Measuring and Allocating Marcom Budgets: Seven Expert Points of View

A Joint Report of the Marketing Science Institute and the University of Michigan Yaffe Center for Persuasive Communication

JANUARY 2003

Marketing Science Institute

The Yaffe Center

This report, prepared with the support of MSI and the Yaffe Center, is being sent to you for your information and review. It is not to be reproduced or published, in any form or by any means, electronic or mechanical, without written permission from the Institute and the author.

The views expressed in this report are not necessarily those of the Marketing Science Institute or the Yaffe Center.

Copyright © 2003 Marketing Science Institute

Contents

Introduction

This joint report of the Marketing Science Institute and the University of Michigan's Yaffe Center for Persuasive Communication contains the points of view of seven experts in the area of marketing communication, responding to the question, How should companies think through the challenge of how to use marketing communication resources most effectively? The seven experts are Paul Farris of the Darden School at the University of Virginia; Dominique Hanssens of the Anderson School at UCLA; Donald Lehmann of Columbia University (also currently MSI's executive director); Leonard Lodish of the Wharton School, University of Pennsylvania; Ambar Rao of the Olin School, Washington University at St. Louis; Don Schultz of the Medill School, Northwestern University; and Gerard Tellis of the Marshall School at the University of Southern California. We thank them all for their insightful contributions.

The idea for this project came from a challenge posed to Rajeev Batra, director of the Yaffe Center, by Jan Valentic, vice president of global marketing at Ford, who sits on the Yaffe Center's industry advisory board, and many of the other Yaffe board members. Ford, like most other companies, is faced with the challenge of deciding how to best allocate its marketing communication dollars across not merely the traditional media of print, television, radio, and outdoor, but also across such media as event marketing, public relations, database and direct marketing, and the Web. If it is hard to measure marcom effectiveness within one medium, it is even harder to do so in many different media, especially in a way that allows comparisons of relative efficiencies and returns, postulating some kind of optimal allocation. How then might companies think through this task? That is, what metrics, tools, and criteria might be the most appropriate in making these allocations?

Since this question involved both marketing metrics and marketing communications efficiency, which are priority topics for MSI, Rajeev approached Dave Reibstein, then MSI's executive director, with the suggestion that MSI and the Yaffe Center launch a joint research initiative in this area. MSI's traditional approach to promoting research in any priority area is to solicit and fund single studies in that topic area. While this is obviously useful, we thought it might be interesting at this early, exploratory phase to try something very different: to approach experts and ask them for their "orienting perspectives" on this broad, somewhat unstructured, question, and to then use their responses to target future research funding more tightly. We invited academic experts with well-established credentials in the areas of marketing communications budgeting practices (Paul Farris), econometric analysis of advertising effectiveness (Gerry Tellis), time-series analyses (Mike Hanssens), scanner data experiments (Len Lodish), meta-analyses (Don Lehmann), field experiments (Ambar Rao), and direct and database marketing (Don Schultz) to give their points of view.

Not surprisingly, they often draw on their technical expertise in their responses. Gerry Tellis, for example, addresses the difficulties of measuring marketing communication inputs and outputs in econometric analyses (the so-called aggregation bias question), and offers recommendations about what kinds of panels, samples, model specifications, etc. might be applied to obtain estimates of the relative efficiencies of spending marcom dollars in different media. Mike Hanssens shows how the gains of marcom-assisted sales over no-marcom baseline sales can be estimated either through experiments or from post-tracking data minus pre-tracking data, then analyzed using vector autoregressive (VAR) multivariate time-series analysis, the results of which can then be fed into simulations to determine optimal spending patterns. Ambar Rao discusses both the key findings and some important limitations of previous field experiments that examined marketing communications spending, pointing out some non-obvious variables (such as price sensitivity) that also require measurement. And Don Lehmann, noting the high cost of individual marketing experiments, suggests how results from many different studies can be combined via meta-analyses, discussing what kind of design, data, and estimation models could be used.

Their four contributions can be considered technical tool-kit reviews, covering the areas of econometric, time-series, experimental, and meta-analytic studies of the measurement of marketing communication effectiveness. The responses of some of the other experts, however, bring up the issue of what we should be measuring—the metrics question—which, logically speaking, must be addressed prior to the question of how to do the measuring. Paul Farris reviews the variety of alternative measures on both the input and output sides and comments on some of the implications of choosing one or another of them. Don Schultz focuses not on the quantitative measures of effectiveness, but rather on the need to understand thoroughly the way in which the targeted customer segments come in contact with the company and its brands. He then addresses the secondary question of which media forms might best enhance that contact—what the best "touch points" are. Len Lodish makes the provocative suggestion that instead of seeking better metrics we should be asking a different question altogether: What are the organizational barriers that are keeping managers from fully utilizing all the data that already exist in companies, and what can we do to remove those barriers?

Clearly, the question of how companies can better allocate their marketing communication dollars across many different media forms is a very complicated one. Readers looking for a simple, canned solution in this collection of points of view will be disappointed. But what you will find are thought-provoking discussions on such questions as, What measures of marcom effectiveness should we employ in the first place? Through what technical methodology should we then attempt to measure, analyze, and optimize them? And, once the technical measurements are over, how do we create organizations that can act on and use these measurements?

We hope you will find these points of view stimulating reading; we hope you will share your reactions and thoughts with us and with the individual authors; and we hope that out of this dialogue will emerge a sharper, more structured set of research questions around which we can commission future research studies. Once again, we would like to thank the seven contributors for their time and for sharing their thoughts with us.

Rajeev Batra
Director, Yaffe Center for Persuasive Communication
University of Michigan

David Reibstein
Executive Director, 1999-2001, Marketing Science Institute
The Wharton School, University of Pennsylvania

Getting the Biggest Bang for Your Marketing Buck: A Point of View

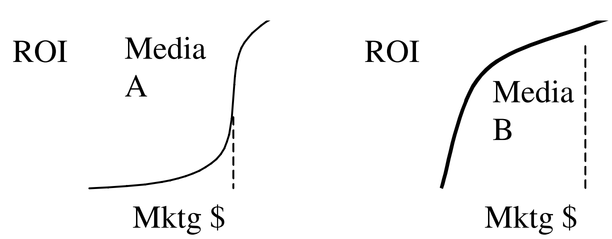
Paul W. Farris, University of Virginia

Introduction

The request: "Please evaluate and recommend study designs and alternative metrics that companies could use to optimize the allocation of marketing communications dollars across different media, including nontraditional media such as Internet websites, events, public relations, database marketing, etc."

As new media vehicles become available, the first questions tend to be overly simplistic. Are banners better than pop-ups? Is a 15-second commercial more efficient than a 30? Should we use vehicle A or B? Only after some initial disillusionment do we begin to grapple with the more difficult, but rewarding, question of how to integrate the new media with the old.

Figure 1. Media Vehicle A or Media Vehicle B?



Anyone who has considered the problem of optimizing marketing communications knows the difficulties. Not only is the range of potential inputs (media vehicles) that might be considered "marketing communications" very large, but there is also no general agreement on the best measures of output (what should be maximized—share, sales, ROI?). Further, it has even been argued that the idea of optimization is potentially destructive, primarily because it will inevitably lead us to narrow our view of the problem to one that is mathematically tractable (even elegant), but probably incomplete in important aspects. Although I believe these criticisms are valid, I also believe that marketers can benefit from applying the logic and metrics represented by a class of model known currently as "media optimizers." In that spirit, I offer some observations on the promise and perils of media models and metrics. The purpose of these media optimizers is to help managers select the individual media vehicles and combinations of vehicles that will maximize the efficiency of their media budget. The optimizer functions by using judgment or data-based inputs on target groups, frequency response, and relative media impact in combination with other data on vehicle reach, audience composition, and costs to evaluate specific combination of vehicles.

Table 1: A Primer on Media Optimizer Terms

Opportunities to See: The "atom" of media vehicle metrics is the vehicle's audience size, typically measured as the number of "opportunities-to-see" (OTS). An OTS is equivalent to an "exposure" or "impression" and is often expressed as a percentage or "rating point."

Target Group Correspondence: Metrics used range from the percentage of vehicle audience "in" the target group to more complex weighting schemes that reflect degrees of target group desirability and the presence of influential opinion leaders, for example.

Reach, Frequency, or Recency: The first media vehicle generates a number of OTS, all of which may be considered reach. Thereafter additional vehicles may increase reach or frequency depending on the overlap of vehicle audiences. Many theories and opinions exist on the circumstances that affect the extent to which one is preferred to the other. Recency is a measure of the time interval between the OTS and the next purchase occasion. Closer, it is argued, is better.

Vehicle Impact/Relevance. The impact of an OTS generated by a banner ad is unlikely to be as great as that of a television commercial during the Super Bowl, but it might be more powerful than the OTS for a brand printed on a speeding NASCAR racer's car. On the other hand, some might argue that the impact of Nike's swoosh logo on Tiger Wood's golf cap as he accepted the winner's cup for U.S. Open was greater than a string of commercials announcing Nike's golf gear.

First, continuous improvement in media allocation is a more reasonable goal than "optimization." However, even this more modest objective requires metrics, benchmarks, and a framework for incorporating learning. It is not enough to somehow test that one media yields a great impact than another (A versus B). As marketers, we need to know why this is so. The explanation needs to be couched in a context that helps us choose between the next set of vehicles (C and D).

The language and structure of media optimizers can help achieve a useful understanding of why B performs better than A. Within the language of media optimizers one media vehicle might be preferred over another for various reasons. For the same dollar expenditure, one vehicle might offer more "opportunities-to-see", a better fit with the target audience, a more relevant exposure context, greater impact from the nature of the media (TV versus billboards) and copy strategy, more efficient trade-offs between reach and frequency, or even the increased ability to deliver exposures adjacent in time to the purchase decision (recency). See Table 1 for a quick review of the metrics used and produced by many optimizers.

Media optimizers generally take the form of estimating the following dimensions of a media plan:

1. Size and composition of vehicle audience (OTS, exposures, impressions)
2. Target group definition and weights (target rating points or weighted exposures)
3. Definition of effective frequency (frequency response functions for individuals, groups)
4. Media vehicle impact/context weighting
5. Timing (recency, seasonality) of exposures

By structuring the media plan output along these dimensions, optimizers can also provide a way of organizing what we need to know. What are frequency response functions? Who are the most responsive targets? Which media vehicles deliver our message with the greatest impact? How important is context and timing of exposures?

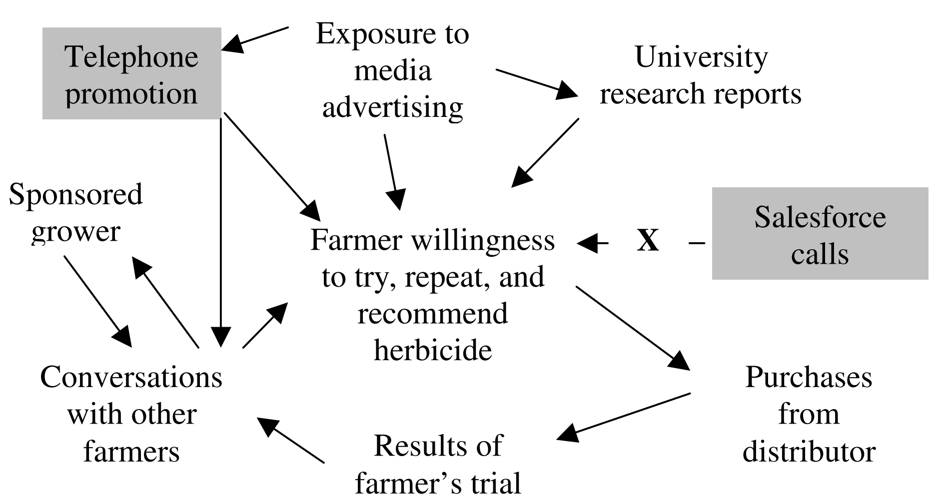
Although there are still real challenges in implementing these models, such as accurate data inputs for number 1, the more significant obstacles to true media optimization are finding reliable, data-based inputs for 2-5. Further, to be useful, optimizers should be adapted to reflect media synergies and interactions—as the following case study will illustrate.

Case Example: Selling Herbicides to Farmers

A relatively new and recently improved agricultural chemical (for weed control) was sold through distributors to farmers. A salesforce that called on large farmers also supported the product. Several companies competed in the market and used a variety of marketing communications to persuade farmers to buy their respective brands. These communications included salesforce calls on individuals, group events to present information to farmers and allow farmers to talk about their experiences with sponsored brands, tours of test plots to view results in the field, and outdoor, radio, and magazine advertising. If the product had been developed last year, the company would surely have had a website too.

The case reviewed the company's media mix, with an emphasis on evaluating a proposal for recruiting farmers to participate in conference-call telephone discussions with other farmers on the merits of the product. Partially because of the need to have high-credibility, neutral-appearing discussion leaders, this technique was much more expensive than traditional media or calls by the salesforce on a cost-per-contact basis (cost was over $200 per farmer). To help decide whether the telephone promotion would be worth the cost, a test was conducted among randomly selected groups of farmers in several areas. To ensure an unbiased test, salespeople were firmly instructed not to contact the farmers participating in the telephone test discussions. Before and after the test telephone promotion, surveys conducted among farmers measured awareness, willingness to try, intention to reuse, and amount of usage anticipated among both test and control groups. The survey data provided comparisons on the variables (including usage and intent to reuse) between the group of farmers receiving the promotion and the control group that did not. The case asks students to use the results of the carefully designed test to recommend a media mix, including the telephone promotion.

Figure 2. Illustrative Map of the Farmer's Decision Process



The data in the case support a moderately successful result for the promotion and it is only after developing the map of the farmer's decision process in Figure 2 that students begin to recognize that the test design was inadvertently biased against the new technique.

In an attempt to eliminate the influence of the salesforce on the test results the nature of what was being tested was changed. After all, was the telephone promotion intended to replace the salesforce or complement it? Or, expressed another way, if the company decided to spend the large sums to conduct the telephone promotion nationally, what would management expect their salesforce to do? Stay away from customers? Of course, put that way, most decide they would want the salesforce to follow up with the farmers participating in the telephone promotions immediately and try to get farmers to purchase from distributors while the product and topic of weed control were still at the top of the participants' minds. In fact, the promotion was implemented, the names of farmers were given to the salesforce for immediate follow-up, and sales exceeded forecasts based on the tests and were limited only by distributor inventories. The telephone promotion was intended to increase the effectiveness of salespeople, not replace them. As the campaign progressed, other synergies became apparent. Farmers who participated in the telephone conversations were used in local radio ads endorsing the product. The sales organization identified influential farmers who would be particularly influential as candidates for the telephone promotion.

The current crop of media optimizers provides a good starting point for developing the architecture of media metrics required to articulate and record learning about why vehicle A works better than B. However, within media optimizers, vehicles are often viewed as independent alternatives to be ranked against each other. Even within the trade-offs of reach and frequency, one vehicle will almost always decrease the effectiveness of another. And the choice is similar to the one depicted in Figure 1. For a given amount of money, which media vehicle offers the greatest incremental impact in sales (or other objective function) for the marketer's money? Where these models tend to fail is in identifying and exploiting synergies between and among different media vehicles.

Media Vehicles: Substitutes or Complements?

An example of a model aimed at capturing marketing synergies is the Gross model. The Gross model was designed to determine how media budgets should be allocated between expenditures for media time and space on the one hand and expenditures to improve the quality of the message execution on the other. Gross's approach was elegant in that he formulated this problem as a combination of substitution and complementary effects. Money spent to improve copy is necessarily money taken away from the purchase of media. But, without effective copy, media funds would be wasted. The real value of the Gross model is that it gives a framework in which complementary effects can be modeled. One might also view the Gross model as a reminder of the importance of "sharpening the saw." Just as it makes sense for a woodcutter to stop and sharpen the saw from time to time, it makes sense to spend money on certain media to enhance the effectiveness of other media. Providing advertising-generated leads to the salesforce is one example. Using offline advertising to increase Web traffic is another. Using billboards to remind viewers of more complex messages in print or TV ads are other examples. These expenditures are not so much choices among independent alternatives as they are choices to spend money on some vehicles to enhance the productivity of other vehicles.

Media Optimizers and Input-Output Metrics

Vehicle Metrics

When considering vehicle economics and ranking their attractiveness, marketers should take into consideration audience size and quality, cost of time or space or exposure (depending on the medium), and impact of message delivery. Cost per thousand impressions (CPM) adjusted for audience quality and message impact is probably the best single measure. If possible, vehicles should be weighted also by ability to deliver contextually relevant messages. These analyses can rank individual vehicle by cost-per-weighted exposures. Unfortunately, many of the key inputs require judgments that few managers are comfortable making. One of the reasons managers are uncomfortable with these inputs is that they too rarely get feedback or an opportunity to learn how to improve the assessments. Even when response data are available that provide some feedback or potential validation of initial judgments, it is rare that the responses can be attributed to a single exposure in a single vehicle. As a result the judgment is often made implicitly instead of explicitly (some vehicles are simple eliminated from consideration).

Media Plan Metrics

Combinations of vehicles will generate patterns of reach and frequency within different target groups that are more difficult to evaluate, because of overlap in vehicle audiences. Further, the vehicles vary in impact, contextual relevance, and ability to satisfy timing objectives (e.g., recency). Reducing these multiple dimensions to a single index of "effective reach" is possible, but it is not at all clear there are rewards for doing so. Unless we can validate the various dimensions and how they are combined, the incentive for "optimizing" plans will be reduced. Instead, the preference is often to make the judgments implicitly by management preference for one combination of vehicles over another combination. At the present, there appears to be a lack of structured research programs for validating the judgmental inputs required by optimizers. Until we develop such programs, the use of optimizers will rely more on managerial faith in the overall logic and the face validity of the individual inputs than hard evidence that the increased efficiency of optimizer-generated plans can be justified by increased marketing productivity. What kind of evidence might justify the application of media optimizers?

One method of validating the optimization engines is to compare plans on the objective functions that are optimized. Typically, for a given dollar amount, this is a form of "effective reach"—exposures weighted by audience correspondence, reach-frequency measures, and vehicle impact/context/timing. Of course, this kind of comparison cannot validate the critical inputs, only the process of forming media plans with those inputs. The global process will need more objective output metrics. In increasing order of the ambition associated with them, I will discuss three levels of these output metrics: transaction, customer, and strategic.

Transaction Metrics

For direct marketers, the first response may be "response." What level of sales and profits can be identified as resulting from a media plan (combination of vehicles)? Using estimates of incremental sales response (whether direct or estimated by some other means) to validate optimizer inputs will be especially tricky when there is a significant marketing communications history that may play a role. In spite of the potential of the history to confound measurement, it would be a significant achievement to make a compelling statistical link between optimizer inputs for frequency response, target group weights, and vehicle impact weights and output metrics of sales.

Customer Metrics

Even more ambitious measures to validate the use of optimizers would include customer economics, in particular, the cost of customer acquisition, lifetime value of customers acquired, retention costs, loyalty rates, and future transactions. The difference in this approach is that the emphasis is much more on the quality of customers and the future revenues they may bring. Further developments might include valuing customers according to their ability to bring other customers into the fold, to pass along the message, or to help with viral marketing efforts. Will the farmers recruited by our telephone sessions be larger farmers who are above average in influence?

Strategic Metrics

The difference in this perspective from the previous two is not so much that it is oriented toward the future, but that it recognizes specific competitors whose actions and alternatives may be affected by our decision. For example, when competitors are expected to follow quickly, the strategic value of some media opportunities may be significant and also difficult to evaluate. With a limited capacity available for the organization implementing the telephone promotion to our farmers, managers believed it was important to deny competitors that competitive weapon. It is often claimed that large beer marketers buy Super Bowl as much to prevent a competitor from stealing a march as any other objective. Clearly, such considerations will remain outside the domain of media optimizers for the foreseeable future.

Summary

First, optimizing the mix of marketing dollars across alternative media requires a system of metrics for both output and input variables. While we have good structure for optimizers, there are a number of inputs that need to be validated if optimizers are to gain widespread acceptance. These include target group definitions and weights, frequency response functions, media vehicle impact, and context weights. Timing of exposures is another. It is difficult to eliminate any of these dimensions as being, in principle, less important than the others.

Second, in the absence of research programs that provide hard data input for the required inputs, optimizers will continue to rely on judgmental inputs. Finding ways for managers to comfortably make and (eventually) validate these judgments is a major research challenge. Shifting frames of reference from transaction, to customer, to strategic economics will only make it more difficult. Since the learning feedback loops—from judgment to plan to implementation to results—are anything but quick and direct, concentrating on the more directly observable transaction economics is probably preferable to more speculative customer equity and strategic considerations. Still, a good many media decisions are probably made with the customer and strategic metrics in mind. Will these decisions make it more difficult to validate the usefulness of optimizers?

Third, it is clear that media vehicles are both potential substitutes for one another and potential complements to one another. Too many separate media planning inputs complicate the process of estimating interactions among different media vehicles and make it more difficult to get the judgments of marketers about the key synergies. In our understandable desire to know more about the relative efficiency and effectiveness of individual media vehicles we should be careful not to let the analytical process ignore or obscure those synergies. In the end, it is those synergies that make it a media "mix" and not a simple ranking of vehicles. I suggest that a map of the customer decision process that marketing is trying to influence should guide the process of modeling interactions. That map should show how the different media are intended to affect the decision process and determine how media substitutes and complements are modeled.

Finally, the basic structure of media optimizers has not changed materially in the last 30 years. While we have developed more sophisticated estimates of reach and frequency and non-linear optimization techniques can now handle more data and media alternatives, the key uncertainties still revolve around the judgmental inputs by the marketers. Is this an indictment of marketing's ability to "get organized" and do the right research to answer these questions? Or, does it perhaps indicate that the development of media models (optimizers) and creative media mixes is a bit like an arms race: more sophisticated media tools simply enable marketers to generate more creative uses of media that are, in turn, more difficult to model? Other hypotheses must include the possibility that marketers understand what optimizers are supposed to do, but don't believe the view will be worth the climb.

Allocating Marketing Communication Expenditures: A Long-Run View

Dominique M. Hanssens, UCLA

Introduction

My point of view on the question of optimal allocation of marketing communication expenditures focuses on the opportunity offered by modern information and modeling technologies. We can now measure communications effects in tangible ways, and therefore use economic principles of resource allocation to guide our spending decisions. I first review the major objectives of marketing communications, and I propose relevant metrics that can be used to measure a firm's progress toward those goals. Next I discuss the implementation of those metrics and the design of marketing studies, with specific reference to the difference between addressable and nonaddressable customers.

Ultimately, marketing communications serve the purpose of enabling and supporting the profitable growth of an organization. Media spending should therefore be allocated in conjunction with the firm's quantifiable growth objectives (targets), and in conjunction with past habits, funds available, or competitive norms. While I will discuss some short-term growth targets and their metrics, my emphasis is on the long-term growth that comes from the maximization of customer equity. Customer equity is defined as the discounted sum of profits generated over the lifetime of customers, net of marketing costs. Customer equity is a stock variable, yet most of the firm's observable performance measures are flow variables such as telephone leads, e-mail clickthroughs, or unit sales. By appropriately transforming these short-run flow variables into customer equity, we can achieve the difficult task of measuring marketing's long-run impact and treat communications expenditures as investments.

The customer equity concept is implemented differently for addressable customers than it is for nonaddressable customers. As computing and communications technologies march on, more marketing media are becoming interactive and therefore targeted to addressable customers. For example, the advent of digital television cable service with electronic program guides may create an interactive medium out of this traditional mass medium. This said, the nonaddressable customer category is still the largest in spending and may retain that position as consumers start to resist companies' efforts to personalize marketing communications. I address both types of communication investments.

Implementing the Customer Equity Concept: Addressable Customers

Typical spending categories include direct mail, e-mail, telemarketing, and various forms of permission marketing. Generally speaking, short-run marketing effectiveness is relatively straightforward to assess, at least at the individual-campaign level, so long as direct-response measures are available. However, direct-response measures can overstate marketing effectiveness because they do not readily assess baseline response—the probability that a customer acquired through direct mail would have become a customer anyway, even without the direct-mail campaign.

At the marketing campaign level, the best metrics are response measures that are as close as possible to the point of customer revenue generation, such as online purchases, telephone orders, membership activations, etc. Absent that, good intermediate ("lead") measures include clickthroughs, toll-free-number inquiries, and business cards collected at trade shows. It is important for the marketer to develop benchmarks for rates of conversion of leads into revenue, so that long-term (customer equity) effects of direct-marketing campaigns can be estimated as soon as possible.

In terms of study designs, the experimental approach is a natural for assessing direct-marketing effectiveness and, not surprisingly, it is used relatively frequently in relationship businesses such as banking and telecommunication. Test versus control-group designs may be used to assess all aspects of direct marketing: offer, target, creative execution, and timing. Equipped with good short-run impact intelligence, marketers can maximize campaign effectiveness and allocate media resources according to the simple "marginal cost = marginal revenue" rule.

However, campaign-level response metrics are by definition tactical and short-run in nature. For example, the goal of a campaign may be to hand over as many new target-segment customers as possible to the organization. The more challenging task is to assess long-run direct-marketing effectiveness and to allocate the overall marketing budget across the key activities that generate customer equity: acquisition, retention, up-selling and cross-selling. I have yet to observe an organization that can jointly optimize across those four functions.

The best criterion for budget setting and allocation is customer equity itself (see Blattberg, Thomas, and Getz 2001 for mathematical expressions). Since customer equity is a composite measure, we need multiple response metrics such as acquisition rate and retention rate. For example, in retention marketing we ask such questions as: What is the likely effect of a $150,000-per-year customer hotline service on retention rates? What is the impact of higher customer retention rates on overall customer equity?

For any given set of business and customer response parameters, there is an optimal level of customer acquisition and retention, which translates into optimal acquisition and retention spending levels. In an actual empirical example, acquisition that maximized customer equity was found to be 14 percent of prospects (requiring a spending level of $325,000) and optimal retention was 85 percent of current customers (requiring a $161,000 retention budget). Thus the total marketing budget would be $486,000.[1]

Given optimal spending levels, we can calculate customer-equity elasticities and use them to prioritize marketing investments that enhance the current response parameters. In the earlier example, customer equity was found to be the most sensitive to changes in retention rate (elasticity 2.51) and margins from retained customers (elasticity 1.88). These two areas should therefore be prioritized for additional marketing resources.

In terms of designs of marketing studies, maximizing customer equity requires building a customer-information infrastructure. For example, the results of ongoing campaign experiments should be carefully assembled into a meta-marketing database, to be used to update the key response metrics that drive customer equity. Likewise, changes in revenue generation from existing customers should be captured and added to the database quickly, as they have an impact on optimal spending prescriptions.

Implementing the Customer Equity Concept: Nonaddressable Customers

The typical spending categories include outdoor advertising, print publications, public relations, and traditional electronic broadcasting. For those categories measuring response is more complex, as we cannot readily establish a one-to-one correspondence between marketing and performance. On the other hand, it may be easier to establish baseline performance levels, so long as we have data points for markets or time periods (or both) representing no appreciable marketing communications pressure.

As with addressable customers, response metrics are easier to develop for short-run, tactical objectives. For businesses with short sales cycles, I recommend using transaction-based performance metrics such as orders, unit sales, revenues, and market share. For cases with longer sales cycles, intermediate measures such as awareness, stated preference, and purchase intention can be used, again providing that their connection with subsequent revenue generation can be established.

If experimentation is possible—for example, assigning zero or base-level marketing treatment to regional control markets—short-run marketing effectiveness is readily assessed. If not, then at least a pre- versus post-performance measure should be taken. For example, at a leading toy manufacturer, advertising effects are measured by a simple statistical comparison of average toy sales levels before and after the campaign.

At a more strategic level, media allocation and marketing budget decisions require the development of marketing-mix models. These are typically econometric representations of the marketing, competitive, and environmental drivers of market performance, and they have a rich heritage of analytical development and business applications (see, for example, Hanssens, Parsons, and Schultz 2001). The key response metric in market-response modeling is sales response elasticity, and we now have a body of empirical generalizations of marketing-mix effectiveness expressed as elasticities (see Hanssens, Parsons, and Schultz 2001, chapter 8). Economic principles, beginning with the Dorfman-Steiner theorem, exist that translate those elasticities into prescriptions for optimal pricing and marketing resource allocation.

In terms of study designs, marketing-mix models require the building of an in-house marketing database, though this time it is collected at the aggregate (market) level rather than the individual-response level. While experimentation is possible and should be encouraged, knowledge of marketing-mix effectiveness develops mainly around this historical database. When sufficiently developed, the database becomes the foundation for marketing simulations. Indeed, sophisticated companies in this arena can use marketing-mix models to design and test marketing strategies in much the same way that an airplane manufacturer tests a new aircraft design in a wind tunnel.

In developing and using marketing-mix models strategically, two important considerations arise. First, marketing communications do not operate in a vacuum, but rather in a competitive setting. This calls for a systems approach to marketing-mix modeling, as opposed to a single-equation approach. Second, our focus on long-run marketing effectiveness requires an equivalent measure of customer equity; that is, it requires a metric that transforms a flow variable (sales) into a stock variable. We do not necessarily want to maximize the profits from transitory sales, but rather the profits stemming from a permanent revenue flow that can be boosted by current marketing spending.

Both conditions can be met by developing so-called vector-autoregressive (VAR) models of market response. These are multivariate time-series models in which market performance, marketing, and competitive effort are jointly endogenous, and they require long time series for estimation (for example, two or more years of weekly data or six or more years of monthly data). VAR models are less concerned with period-by-period response elasticities and more with the net (cumulative) effect of a chain of reactions initiated by a marketing or competitive effort, which is captured by the impulse response function. VAR models allow us to measure permanent as well as transitory effects of marketing communications and are therefore useful tools for maximizing the long-run profitability of media spending.

Conclusions

In a nutshell, and at the risk of some oversimplification, marketing communications spending should be driven by the maximization of customer equity. Firms should invest in market response intelligence in order to optimize their media spending. For media with addressable customers, this implies that tactical media decisions are made by successive experiments that maximize short-run direct-response measures. Strategic decisions are based on the relative capability of customer acquisition, retention, and cross-sell spending to increase customer equity.

For media with nonaddressable customers, firms should conduct simple experiments or before and after comparisons to derive short-run marketing effects and make media allocation decisions accordingly. At the strategic level, they should develop multi-equation marketing-mix models that estimate media effectiveness in competitive settings. Long-run response elasticities can now be obtained and used for media allocation decisions that maximize long-term profitability.

Notes

1. An Excel software module that optimizes acquisition and retention spending for maximum customer lifetime value is available from the author upon request.

References

Blattberg, Robert, Jakki Thomas, and Gary Getz (2001), *Customer Equity*. Boston, Mass.: Harvard Business School Press.

Hanssens, Dominique M., Leonard J. Parsons, and Randall L. Schultz (2001), *Market Response Models*, 2nd ed. Boston, Mass.: Kluwer Academic Publishers.

An Approach to Measuring Communication Effectiveness for the Purpose of Allocating Communication Dollars

Donald R. Lehmann, Marketing Science Institute and Columbia University

Optimal allocation of communication dollars is relatively straightforward if we know the effects (both short- and long-run) of different spending patterns (plus cost and margin information and the firm's objective function). Therefore, optimal allocation requires estimation of response patterns. Unfortunately, response patterns are relatively complex, depending on factors related to customers (or audience), product category and brand, and the message (both informational and creative aspects). Other complications include synergy effects (or lack thereof) between different communication and mix elements, and competitive reactions. Since it is impossible to deal completely with all these complexities, much less to do so in five pages, here I focus on two approaches.

One apparently straightforward approach to assessing effectiveness is simply to measure effectiveness for different media. Unfortunately this is not as straightforward as it sounds. First, there is the issue of what measure of effectiveness is relevant: awareness, associations, attitude, intention, sales, share, profit, and so on. Second, media audiences differ, suggesting any differences could be due to audience self-selection in a field study or attention in an experiment rather than the media per se. In addition, different media vary in the senses employed, so in effect copy differs by media, raising the issue of whether differences are due to copy or media. Even if these problems were surmountable, however, another key problem exists.

Put simply, the goal is to predict future effectiveness. Precise measurement of past effects is useful but not sufficient. A number of factors change between the time a study is conducted and the time when its results are needed, even for fast-track studies. While sophisticated modeling procedures that deal with simultaneity, competitive reaction, and heterogeneity are helpful, they do not overcome the shifting effectiveness dilemma. Put differently, knowing exactly the impact of a specific medium and copy combination at time t does not guarantee it will be the same at time $t + k$. As a corollary to this, it makes little sense to assess effectiveness at the level of specific copy or product; a more sensible approach involves type of copy, product, and so on. The question, then, is how best to make effectiveness predictions at this more general level.

One viable approach to producing a baseline forecast involves proactively performing a series of studies according to some experimental design. Given the complexity of the design (including three-way interactions among product, media, and audience), however, the cost would be prohibitive. Therefore we consider an alternative approach, meta-analyses of existing studies, which has been used in many fields, including marketing (Farley and Lehmann 1986) and medicine, among others. In fact, advertising effectiveness has been the subject of a number of meta-analyses (see Lehmann 2000). What is required is a broader meta-analysis that subsumes and extends existing studies.

The keys to such a meta-analysis are (1) development of a design matrix and (2) uncovering past and current studies. Regarding the design matrix, the basic categories would be:

1. Product (i.e., what is being studied), measured on general grounds such as durable versus consumable, consumer versus business-to-business, stage in the life cycle, service versus product, nonprofit versus profit, time period, and so on.

2. Media, defined at a level of detail somewhat more fine-grained than print, television, Internet, etc.

3. Audience characteristics, including demographics, product use, etc.

4. Measure of effectiveness, possibly organized according to a hierarchy-of-effects model.

5. Other (i.e., other media spending), a critical category given the focus on integrated communication, as well as competitive mix, spending, time period, etc.

6. Model form (e.g., linear, multiplicative, or logit).

7. Model specification, including covariates and how heterogeneity is handled.

8. Estimation method, (OLS, multistage, or numerical—for example, Gibbs sampling).

9. Study quality, as measured by sample sizes, or, ideally, variance in estimates of effectiveness as well as source: academic versus industry and lab versus field.

The design should be developed collaboratively (by both academics and practitioners) and iteratively (i.e., the initial design should be modified based on the studies uncovered). Coding should be done by multiple, trained coders.

Regarding selection of studies to include, standard procedures should be employed. First, published literature should be searched using online and hard-copy referencing services, issue-by-issue search of key journals, and reference checking. Second, a search of unpublished academic research and industry files should be conducted. This a potentially rich source, as Batra, Lehmann, Burke, and Pae (1995), Lodish et al. (1995), and Eastlack and Rao (1989) demonstrate. A "public" request could be made through the American Marketing Association, the American Association of Advertising Agencies, the Marketing Science Institute, or similar organizations.

Once the studies are assembled, some thorough and creative work will be needed. While sophisticated meta-analysis procedures are being developed, regression is the appropriate first step. Since the national experimental design would contain empty cells, confounding among variables, etc., a simplification stage would be needed. (An alternative using neural networks can be employed, and though it seems likely to be unstable and hard to interpret, it could provide suggestions for the regression model.) Other issues, such as how to deal with multiple results from the same study (answer: treat them as separate observations and include a study dummy variable to capture some of the idiosyncratic effects) will also need to be dealt with.

In summary, it is not productive to try to rely on a single study to answer the question. To paraphrase Pope, "Whoever seeks a perfect study to design/ seeks what ne'er was nor is nor e'er shall be." Rather, the question suggests a major, and messy, effort to integrate knowledge from the thousands of studies that already exist. By combining the expected result from the meta-analysis with any "case-specific" information (e.g., a specific copy test) that is available (in an informal or formal Bayesian manner), a reasonable forecast can be made. Perhaps equally important, the meta-analysis provides information on the range of results which can be expected. This range can then be used to examine the likelihood that a planned spending program will deliver the results "advertised" in the budget request. For example, advertising elasticities have been found to be between .01 and .03 for increased spending on old campaigns for mature products to about .3 for new messages for new products or new uses for old products. This means that at best a 3% increase in sales can be expected for a 10% increase in advertising spending. Hence, a budget proposal that calls for a 20% increase in sales for a 10% increase in advertising alone is well outside the realm of past experience and hence unlikely to occur. Put differently, arguments that suggest a particular approach is "special" implicitly assume that previous ads were not (and hence that the creative talents and media planning employed in the past were greatly inferior). While this is possible, it is highly unlikely. Thus, a meta-analysis (i.e., an accumulated database of past results) provides not only a baseline forecast of likely impacts which can be used in the absence of any specific data but a sense of what is feasible to expect.

References

Assmus, Gert, John U. Farley, and Donald R. Lehmann (1984), "How Advertising Affects Sales: Meta-Analysis of Econometric Results." *Journal of Marketing Research* 21 (February), 65–74.

Batra, Rajeev, Donald R. Lehmann, Joanne Burke, and Jae Pae (1995), "When Does Advertising Have an Impact? A Study of Tracking Data." *Journal of Advertising Research* 35 (September-October), 19–32.

Eastlack, Joseph O., Jr., and Ambar G. Rao (1989), "Advertising Experiments at Campbell Soup Company." *Marketing Science* 8 (Winter), 57–71.

Farley, John U., John A. Howard, and Donald R. Lehmann (1976), "A 'Working' System Model of Car Buyer Behavior." *Management Science* 23 (November), 235–47.

Farley, John U., Jerrold P. Katz, and Donald R. Lehmann (1978), "Impact of Different Comparison Sets on Evaluation of a New Subcompact Car Brand." *Journal of Consumer Research* 5 (September), 138–42.

Farley, John U., and Donald R. Lehmann (1986), *Meta-Analysis in Marketing: Generalization of Response Models*. Lexington, Mass.: Lexington Books.

Lehmann, Donald R. (1977), "Responses to Advertising a New Car." *Journal of Advertising Research* 17 (August), 23–32.

Lehmann, Donald R. (2000), "Quantitative Empirical Generalizations and Progress Toward Knowledge: Pushing the Meta-Analysis Envelope." Presented to the 2000 Paul D. Converse Symposium, Champaign, Ill., May 6–8.

Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, and Mary Ellen Stevens (1995), "How TV Advertising Works: A Meta-Analysis of 389 Real World Split Cable TV Advertising Experiments." *Journal of Marketing Research* 32 (May), 125–39.

Use of Marketing Metrics: A Different Point of View

Leonard M. Lodish, University of Pennsylvania

There is a more important problem than developing new marketing metrics. Most packaged-goods manufacturers and other marketers are not getting anywhere near full value from the metrics currently available. In this article, I summarize some data on the underutilization of metrics, hypothesize some reasons for this, and describe some steps to be taken to improve firms' performance by using metrics more effectively.

Firms Don't Use Available Metrics Enough

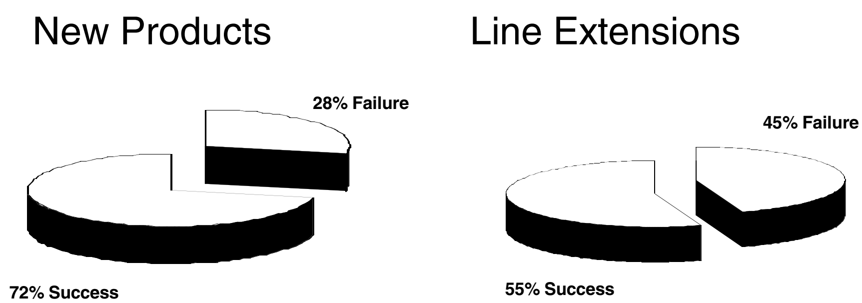
The dot-coms that squandered hundreds of millions of dollars in 1999 and 2000 to "brand" their products and services didn't apply any available measurement devices to improve their performance. If they did use test markets, matched markets, split-cable markets, marketing-mix modeling, etc., it was not evident in their decision making. They squandered literally billions of dollars of venture capital funds on ill-conceived advertising campaigns that did not contribute to their shareholder value. But it is not just dot-coms that are at fault; even packaged-goods marketers, reputed to be the most sophisticated users of marketing data, don't seem to use available metrics as well as possible, though it looks like they may be improving.

There have been numerous metrics available to evaluate new products before they are introduced nationally. Traditional test markets, simulated test markets, and split-cable markets have all been shown to be very effective at reducing the failure rate of new products. The available data shows that packaged-goods marketers are improving in reducing the failure rate of new products, but still have a way to go.

1992 and 2000 Data on New-Product Successes and Failures

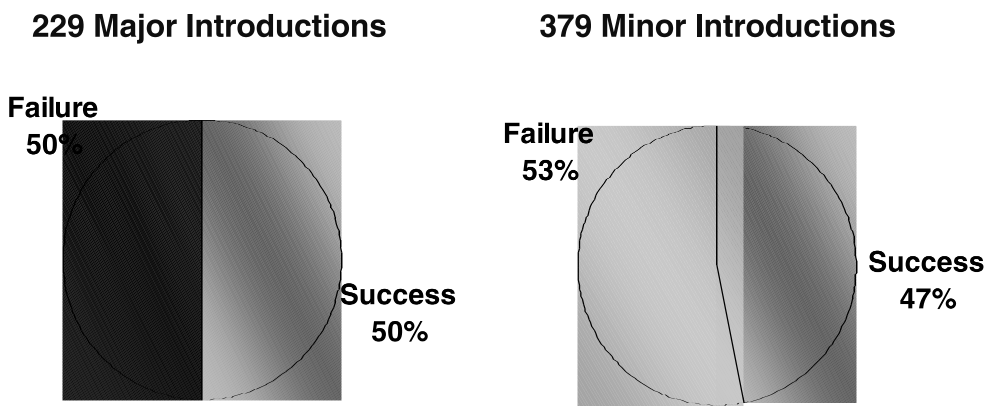
Information Resources Inc. periodically summarizes the success and failure rates of new products in the United States. The data come from their scanner data of food, drug, and mass-merchandise outlets. Failure is defined as a product that either fails to obtain more than one-sixth of the available distribution in food, drug, and mass-merchandise outlets during its first year or a product that loses more than 30 percent of its year one distribution in year two. Figure 1 shows the data for 1992.

Figure 1. New-Product Success and Failure Rates, 1992



For the years 1997-1998 the data were summarized slightly differently into major and minor introductions. Figure 2 shows this data.

Figure 2. Successes and Failures of Major and Minor Introductions, 1997-1998



Source: IRI New Product Trends 2000. Analysis of food/drug/mass distribution and dollar sales for new major and minor introductions launched between October 1996 and October 1998 (looking at two years postintroduction) in 20 categories. Major: new brand or extension into a new segment. Minor: form/flavor extension (not sizes).

There has been an improvement in new-product success in the past seven years. That's the good news. The bad news is that there is still much room for further improvement. Data from the BehaviorScan test-market service shows that improvement is possible, as Figure 3 makes clear.

Figure 3. In-market Testing Significantly Reduces the Risk of Failure

SOURCE: IRI *New Product Trends 2000.*

It is easy to show that the benefits outweigh the costs of using in-market testing. The rest of this article discusses why there are barriers to more profitability using existing metrics and makes some suggestions for improvement.

The Process of Using Metrics: Costs, Value of Information, and Rational Decision Making

The general rational decision-making process is one that MBAs are exposed to more than once in their MBA program. The process is first to establish criteria, pick the best alternative, and then set up feedback mechanisms to continually evaluate the performance of the alternative. The real problem managers continually face is the uncertainty of how the evaluation will go. Most efforts go into trying to minimize the probability of making a decision that turns out not to have been the best option.

Market tests, market research, and experiments can significantly reduce the probability of making a wrong decision. Deciding which market tests, research, and experiments to run (in other words, which metrics to use) involves balancing the costs and the value of the information. If a market test can eliminate a 30 percent probability of losing 10 million dollars, then it has a value close to .3 times 10 million, or 3 million dollars.

Just as decision alternatives need to be rationally considered, so do research, testing, and experiment alternatives. Evaluation of the costs versus the expected value of uncertainty reduction should be carried out for all of them. Once the research, tests, or experiments are executed, their results must be interpreted without bias to best evaluate the decision options that were the subjects of the research.

In order to gain the most advantage from the above process, the decision makers must challenge existing rules of thumb and other mental models to make sure that they are consistent with incoming data and information from the marketplace.

Prerequisites for Using Metrics for More Profitable Learning

The following steps and concepts are not exclusive and may overlap in some instances.

Temper Your Overconfidence

Behavioral researchers have shown over and over that people are generally optimistic. They overrate the chance of good events happening to them and underrate the chances of bad events. They are also overconfident about their relative skills or prospects. For example, Colin Camerer presented data in 1996 that indicated that 90 percent of American drivers in one study thought they ranked in the top half of their demographic group in driving skills.

The way people work in many firms may reinforce these biases toward optimism. It is human nature when making a presentation to management to emphasize the information supporting a new product or advertising campaign and to deemphasize contradictory information. To counter this tendency, some managers will assign staff to be devil's advocates who will effectively present contrary positions.

Management compensation and motivation at many firms also reinforce the optimistic biases. If new-product managers are judged on whether a new product is successful, and if they risk being fired or discharged if the product is withdrawn, they will do almost anything to keep the new product alive as long as possible. If one's career depends solely on the sales performance of a new product, one is not likely to be very rational in viewing information that indicates that the product will not do as well as might be hoped. If, conversely, a new-product team were judged on the overall performance of all the firm's new products, the team would be much more likely to eliminate the losers to concentrate resources on the likely winners.

The bias towards optimism causes managers to miscalculate the value of market tests by overstating their odds of making the right decision about a new product or advertising campaign. If managers are estimating the probability of making a wrong decision (which is rarely done explicitly at most firms) as a way of getting at the potential value of research, optimism will bias those calculations. For example, even though typical large packaged-goods marketers see more than 50 percent of their new products fail, most managers who are responsible for new products at those firms would estimate that the probability of failure for their new products is much lower than 50 percent—just as most drivers think they are better than average. If a senior manager would make one person or group responsible for evaluating the specific costs and value of all testing alternatives, and have that person or group evaluate the new-product development group on how well the decisions impacted profitability, it is less likely that the evaluations would be biased.

Challenge the Old Wives' Tales That Are Used as Mental Models

Perhaps the same group that is evaluating the costs and value of testing and research projects should also be in charge of the mental models in use. First, they would ferret out those models, rules, and paradigms. This is a big job, because it takes analysis and probing of decisions to uncover exactly what rules, models, and paradigms hold sway. The group would evaluate the evidence supporting the rules, scrutinize and continuously challenge them, and keep a public record of what was learned. That record would be widely circulated and then become a part of the firm's paradigms.

Don't Overreact to Competition

Many of the old wives' tales perhaps came about because managers have a tendency to overreact to competitive actions in spite of what the market may be telling them. A 1996 article by J. Scott Armstrong and Fred Collopy clearly showed this tendency. In a laboratory study in a simulated environment, when information about competitors' profits were provided, "over 40% of the subjects were willing to sacrifice part of their company's profits to beat or harm their competitor" (p. 188). They also found in a field study of large firms over a half a century that "firms with competitor-oriented (market share) objectives were less profitable and less likely to survive than those whose objectives were directly oriented to profits" (p. 188).

The prevalent use of benchmarking has to be interpreted in light of the above bias. Just because a successful competitor uses so much advertising, or gets a certain television advertising reach or frequency does not necessarily mean that the advertising is responsible for the competitor's success. A better benchmark would be to find the competitor who uses information and market tests the most profitably to get clues about improving your firm's use of tests and experiments.

Keep Time and Competitive Pressure in the Proper Perspective

All too often, time pressures are used as an excuse to avoid doing in-market tests or performing research. George Day quotes one disgruntled manager in a packaged-goods firm: "Concept tests are viewed as obstacles by our product managers. They are rewarded for keeping their products moving ahead" (Day 1994, p. 13). Sometimes when the competitive value of keeping a new product or campaign secret is really large, time and competitive pressures may be legitimate reasons for forgoing tests or experiments. However, the data above show large costs in profitability for forgoing most in-market new-product testing. Again, as above, someone without a stake in the process should be evaluating the tradeoffs using all available objective data and making recommendations to management on what will most likely be most profitable.

A more subtle reason that time pressures get inappropriate emphasis is that many firms promote or move their managers so often that the managers' time horizon precludes the test from helping during "their watch" on the product. If a test will have value for the next three years, and a manager's time horizon left on the brand is nine months, then he or she will be more concerned with how the brand looks after nine months than three years. One way to improve this process is to evaluate managers not just on current revenues and profits, but also on the expected value of tests and experiments under way that will help the brand perform more profitably in the future.

Conclusion

The above suggestions should help firms to improve marketing metrics' contribution to shareholder value, whether the metrics be new or old reliable ones such as in-market testing. I think that sometimes we turn to new technology to help us when we could be even more productive if we used existing technology better.

References

Armstrong, J. Scott, and Fred Collopy (1996), "Competitor Orientation: Effects of Objectives and Information on Managerial Decisions and Profitability." *Journal of Marketing Research* 33 (May), 188–199.

Camerer, Colin (1996), "Progress in Behavioral Game Theory." Pasadena, Calif.: California Institute of Technology, Division of Social Sciences, Working Paper, 228–77.

Day, George (1994), "Continuous Learning About Markets." *California Management Review* 36 (4), 13.

Information Resources Inc. (1999-2000), *IRI New Product Pacesetters* (Annual Report). Chicago, Ill.: Information Resources, Inc.

Advertising Experiments: An Assessment

Ambar G. Rao, Washington University

Introduction

There is a long history of firms trying to evaluate the impact of advertising spending using controlled experiments. Exposing one group of markets to a base level of spending and another group to a test level, and using the difference in sales rates between the two groups to learn the impact of the changed spending, is conceptually elegant and appealing. The advent of instrumented markets (such as those of the BehaviorScan system) and single-source data has simplified the implementation of these experiments. The results of many advertising experiments have been published, enhancing our ability to make some empirical generalizations about the sales response to advertising. In this short article, I first review some of those generalizations. Then I present a critique of the approach to advertising experimentation taken to date. I note that in assessing the impact of advertising in an experiment, we usually (a) ignore other important impacts of advertising on brand profits and (b) frequently treat advertising as a stand-alone expense rather than as an integral part of a marketing strategy that is subject to competitive response. Finally I suggest an approach to ad experiments that may provide a more complete understanding of the impact of advertising. My comments are restricted to mature, well-established brands of consumer packaged goods for which brand awareness and distribution are already high.

Review of Findings

The most widely supported finding with regard to sales response to advertising is that increases in ad spending usually do not result in increases in sales. For example the meta-analysis of BehaviorScan experiments by Lodish et al., (1995a) shows that increases in ad spending result in statistically significant sales increases only about one-third of the time. This finding is consistent with findings from many other ad experiments, including those by Eastlack and Rao (1989), and Aaker and Carman (1982). The recent econometric analysis of P&G's value-pricing strategy, in which ad expenditures were increased 20.3 percent on average, also show that ad elasticities are very low—only .039 on average (see Ailawadi, Lehmann, and Neslin 2001).

The second finding is that when increases in ad spending do increase sales, these increases occur quickly, as reported by Lodish et al. (1995b). Eastlack and Rao (1989) reported that lack of early success accurately predicted lack of eventual success. In addition, when advertising works, it has a significant lagged impact. What we do not know yet is what happens when spending is reduced: When do sales declines begin and how deep are they?

The third finding is the apparent lack of correlation between standard copy-testing measures or tracking-study measures and incremental sales. Again, both Lodish et al. (1995a) and Eastlack and Rao (1989) provide support for this observation.

Finally, a new campaign or the use of an unusual medium almost always results in an increase in sales. Lodish et al. (1995a) and Eastlack and Rao (1989) provide support for this finding.

A Critique

Clearly ad experiments have contributed significantly to our knowledge of how advertising works, but we need to take a fresh and more comprehensive look at various phenomena (in addition to sales volume) associated with advertising in order to increase our understanding of its true impact.

The Value of Increased Advertising Spending

Although the odds are clearly against increases in advertising spending resulting in increases in sales, it is hard to find a brand manager who would turn down an offer of additional advertising funds, or voluntarily accept a reduction in advertising spending! Is this a departure from rationality? I argue that it is not. Marketers offer many reasons for maintaining or increasing spending rates. These can be summarized by saying that marketers believe that advertising nurtures brand equity: Advertising increases brand differentiation, reduces price sensitivity, creates more resistance to competitive price cuts and promotions, and makes it easier to introduce the brand in new markets and extend it to new categories.

What is the empirical evidence supporting these claims? There is a vast amount of support for these ideas from consumer information-processing studies, usually conducted in the lab. These experiments are carefully controlled, theory-based studies, but they might be criticized on the basis that they are conducted under conditions of forced and highly concentrated exposures to novel or hypothetical ads, for novel or hypothetical brands (see Keller 1998 for many references and examples).

Investigations using field experiments or econometric methods have been relatively fewer. Eighteen studies—seven field experiments, ten econometric studies (mostly at the brand or category level), and one lab experiment—have been analyzed by Kaul and Wittink (1995) in an effort to obtain empirical generalizations about the impact of advertising on price sensitivity and price. They conclude that increases in nonprice advertising lead to lower price sensitivity, while increases in price advertising increase price sensitivity and reduce prevailing prices. A study by Ailawadi, Lehmann, and Neslin (2001), also using brand-level data, shows that price sensitivity is lower where category advertising is high. There is empirical support, then, regardless of methodology, for the idea that the proper type of advertising can benefit a brand by making customers less price sensitive, and presumably more resistant to competitive price cuts and promotions as well. Thus, an increase in advertising might be profitable even in the absence of increases in sales, because higher average prices are sustainable. However, we almost never see estimates of the interaction of price elasticity and advertising reported in controlled experiments, although store-level pricing and promotion data are used as covariates in the analysis of BehaviorScan studies (Lodish 1995b). Instrumented markets, from which data at the individual level are collected, would appear to be ideally suited for examining this important impact of advertising.

The Value of Treating Advertising as Part of Marketing Strategy

Advertising is frequently treated as a stand-alone item, not as an integral part of a marketing strategy or marketing budget decision. And yet, the role and type of advertising, both message and medium, clearly depend on strategy and budget decisions. For example, a local-market-oriented strategy might call for a marketing mix light on national television advertising and heavy on local ads, as well as customized pricing, promotions, and public-relations efforts. A mass-market strategy placing little emphasis on market-to-market differences would have more uniform pricing and promotion and a heavy national television presence. Similarly, a substantial change up or down in a marketing budget would call for a very different marketing mix, whereas a marginal change might be accommodated by changes in electronic media expenditure alone. Testing strategies or overall budget levels would require an experiment with a much larger range of ad spending (possibly including zero, or very low levels of spending) and the simultaneous manipulation of other elements of the marketing mix. But it is rare to see such experiments.

An exception is the Navy Enlistment Marketing Experiment (Carroll et al. 1985). This study was the result of congressional interest in establishing appropriate budget levels for U.S. armed services marketing activity in general and for the Navy in particular. The two major elements of the marketing mix for the Navy were advertising (national and local, television and print) and recruiter efforts. For each of the budget levels considered, a marketing and media mix was developed. These mixes (one reduced national television advertising to zero while increasing local and print advertising from previous levels) were implemented in the test markets, resulting in a wide range of expenditure rates for several media and of staffing levels for recruiters. In addition, the marketing activities of the other armed forces and a number of environmental factors were monitored and used as covariates in analyzing the experiment. Implementing such experiments is complicated and often expensive, but the results can more than justify these difficulties.

The Value of Considering Competitive Effects

Most advertising studies reported seem to implicitly assume that the brand studied is a monopolist. Competitive effects are rarely considered—at least in the published literature on ad experiments. Yet, if a brand were to change its advertising strategy or its advertising and promotion mix in a significant way, competitors are likely to react (see Ailawadi, Lehmann, and Neslin 2001 for a recent example). To analyze the impact of potential competitive response we need to estimate the cross-elasticities of the respective marketing-mix variables, to determine both which competitor is most likely to be impacted by our new strategy and how that competitor might respond. Again, the data are available, particularly in instrumented markets, but no analyses of this type have been reported, although it is possible that they may have been done as part of a consulting assignment.

When Advertising Spending Does Increase Sales

Let us turn to the 33 percent of cases in which increases in spending did result in increases in sales. We know very little about the characteristics of these "successful" ads and how they differ from those of ads that did not produce incremental sales. Standard copy-testing methods have not been helpful in this regard. Industry researchers are well aware of this problem, and efforts have been made to validate copy testing methods using a sales criterion (see Haley and Baldinger 1991), but much needs to be done in this area. Questions we should ask include: What type of appeals did the ad contain? What sorts of viewer responses did it create? How do these differ from ads that did not yield sales increases? (MacInnis, Rao, and Weiss [2002] have recently made a preliminary effort to address these issues.) This information could be used to provide guidance to the advertising creative team in the short term, and in the long term it could form the basis of a more theory-driven copy-testing methodology.

Proposed Metrics and Designs for Advertising Experimentation

The critique above addresses some areas in which ad experiments need to be extended in order to obtain a more comprehensive understanding of the impact of advertising expenditure. Here are some suggestions for the design, analysis, and assessment of advertising experiments:

1. An often unreported benefit of advertising is the reduction of price sensitivity. If an increase in advertising helps us support a higher average price, it may be justified even if there is no impact on sales. Thus we must analyze the interaction of advertising weight with own- and competitive-price and promotion variables. If instrumented markets are used to implement the experiment, individual-level data can be used to test for the interaction of advertising and price or promotion. Recent advances in choice modeling should make such analyses relatively straightforward to implement. (In fact, the IRI database used for the Lodish et al. 1995a meta-analysis provides an excellent starting point to check out these ideas). Even if tests are done at the market level, pricing and promotion data are available in great detail, so analysis is feasible, although potentially less reliable because of concerns regarding aggregation bias.

2. Treat advertising as part of a total marketing strategy and test alternative strategies. This may require manipulating other marketing-mix variables and testing wider ranges of spending, including very low levels of expenditure. For example, one could compare two strategies—one for brand-heavy advertising and low promotion and one for low advertising and heavy promotion—plus a control strategy with current spending levels. The individual impacts of advertising and promotion plus their interaction effect could be estimated from this design. And reasonable estimates of cross-elasticities may also be feasible so that alternative competitive-response scenarios could be evaluated. It is interesting to speculate whether P&G would have implemented their value-pricing policy had they first experimented with it in test markets and done the type of analysis suggested here.

3. Analyze the executional cues of and viewer responses to successful and unsuccessful ads to determine which creative elements differentiate them. Viewer responses could be collected in the course of the normal tracking-study interviews that are part of the standard market research activity for most brands. Use this information to guide the development of ad creative platforms.

Points 1 and 2 above are a call for more thorough analysis of data from experiments, together with a more serious consideration of the totality of marketing strategy and competitive response that has heretofore been the norm. I am calling for a melding of experiments and econometric analysis, and in my third point I am calling for the integration of concepts from the consumer behavior literature into the quantitative framework. Perhaps it is a tall order, but it certainly would help both the theory and practice of marketing if we moved in this direction.

References

Aaker, David A., and James M Carman (1982), "Are You Over-Advertising?" *Journal of Advertising Research* 22 (4)(August/September), 57–70

Ailawadi, Kusum L., Donald R. Lehmann, and Scott A. Neslin (2001), "Market Response to a Major Policy Change in the Marketing Mix: Learning from Procter and Gamble's Value Pricing Strategy." *Journal Of Marketing* 65 (January), 44–61.

Carroll, Vincent P., Ambar G. Rao, Hau L. Lee, Arthur Shapiro, and Barry L. Bayus (1985), "The Navy Enlistment Marketing Experiment." *Marketing Science* 4 (4), 352–74.

Eastlack, Joseph O., and Ambar G. Rao (1989), "Advertising Experiments at the Campbell Soup Company." *Marketing Science* 8 (1), 57–71.

Haley, Russell I., and Allan L. Baldinger (1991), "The ARF Copy Research Validity Project." *Journal of Advertising Research* 31 (April-May), 11–32.

Kaul, Anil, and Dick R. Wittink (1995), "Empirical Generalizations about the Impact of Advertising on Price Sensitivity and Price." *Marketing Science* 14 (3) G151–60.

Keller, Kevin Lane (1998), *Strategic Brand Management: Building, Measuring, and Managing Brand Equity*. Upper Saddle River, N.J.: Prentice Hall.

Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, and Mary Ellen Stevens (1995a), "How Advertising Works: A Meta-Analysis of 389 Real World Split Cable TV Advertising Experiments." *Journal of Marketing Research* 32 (May), 125–39.

Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, and Mary Ellen Stevens (1995b), "A Summary of Fifty Five In-Market Experimental Estimates of the Long Term Effects of Advertising." *Marketing Science* 14 (3), G133-40.

MacInnis, Deborah J., Ambar G. Rao, and Allen M. Weiss (2002), "Assessing When Increased Media Weight Helps Real World Ads Generate Incremental Sales." *Journal of Marketing Research* (forthcoming)

Marketing Communications Allocation: Asking the Right Questions

Don E. Schultz, Northwestern University

The Marketing Science Institute has requested a limited number of experts in various fields to submit proposals in response to the following request: "Please evaluate and recommend study designs and alternative metrics that companies could use to optimize the allocation of marketing communications dollars across different media, including nontraditional media such as Internet websites, events, public relations, database marketing, etc."

I have been asked to respond with a point of view. Further, I have been asked to suggest research methodologies that might be useful in addressing this challenge. For the sake of ensuring a range of perspectives, I have been asked to consider the request from the point of view of a direct or database marketer.

The Challenge and Some Assumptions

The request for a point of view on the optimization of the allocation of marketing communication dollars assumes a number of issues have already been resolved. For example, it assumes a target market has been selected, specific communication or behavioral goals have been set, relevant time frames for the execution of the communication program have been established, and that appropriate dollars or other funding have been made available. It also assumes that a wide variety of media forms are available to the planner, no specific return-on-investment requirements are demanded, and so on. I will take these as givens in my recommendation. Otherwise, the hypothetical nature of the question limits any type of useful response.

The major assumption made in my proposal is that the marketing communication allocations are being made primarily against end users or consumers rather than against channels, shareholders, employees, or other relevant stakeholders. Further, I assume the product categories for which the media allocation is being made are primarily consumer goods being sold through some type of distribution system. In other words, the products or services are not being sold directly to the end user. These assumptions are made even though I have been asked to take the view of a direct or database marketer. Were the products or services being sold directly to customers or end users, my answer to the question would likely be quite different. With those caveats in place, my suggestions and recommendations follow.

A Point of View on Media Allocation

Inherent in the request is the assumption that different media forms will likely provide different results or provide differing values when used in a communication plan. Therefore, different combinations of media forms will produce different responses and provide different values to the marketing organization. Thus, the purpose of the media allocation process is to attempt to identify which media forms or which combinations of media forms will provide the best marketplace impact and thus the greatest returns to the marketing organization.

The problem is that the question and the inherent assumption embedded in the question disregard one of the most important issues in marketing communication. That issue is whether the marketing communication investment is being made to purchase media to deliver messages and incentives to a maximum number of people or whether it is being made to communicate with a targeted group of the best customers or prospects for the product or service. In either case, is the goal to maximize distribution of messages and incentives, or is it to optimize returns from the investment? Acknowledging that, from the perspective of corporate value, marketing communication investments are not really for purchasing various marketing communication media vehicles but rather are for attempting to influence the behaviors of various customers or prospects changes dramatically the question of what the best media allocation is.

Given that new understanding, the first question now becomes: What customers or prospects or customer or prospect groups should we target with our finite corporate resources in the form of marketing communication programs? That question is much more important than what media form should be purchased. Thus, the planner's first decision should be a customer or prospect decision, not a media decision.

While this might appear to be only a semantic argument, it is really the major challenge for the marketing communication industry today. It changes the marketing manager's primary question from What do we as marketers want to do? to What do customers want done to them?

In the MSI question, the selection of customers or prospects to reach is not raised. As it happens, in most cases the media planner has only very broad descriptions of "product or service suspects." For example, most media-planning information is based on broad-scale demographics such as age, sex, income, education, geography and the like, perhaps enhanced by broadcast daypart or past-section viewing or reading. That's because that is essentially all the information the media generally can provide about their audiences. Because they often have only sketchy audience information or data, media planners are often forced to evaluate media forms in the very broadest terms and then try to relate that to media costs. This commonly results in the now infamous CPM or CPR syndromes that still drive media models.

By taking my proposed customer-first approach, the question of marketing communication media allocation becomes secondary to selecting or identifying the customers or prospects or customer or prospect groups to be reached and determining the outcome desired as a result of those contacts. Rather than selecting the media forms that could or might be used to deliver the most or the most efficient number of outbound messages, the challenge is now identifying what media forms, or better said, what "brand touch points" the preselected customer or prospect has with the marketing organization. For example, what media forms does the customer or prospect use? When or how does the customer or prospect come in contact with the marketing organization, its brands, its distribution systems, its products and services, its employees, its advocates, its detractors, and so on? The marketing organization controls some of these brand contacts or touch points but by no means all of them. Consequently, the marketing or media planner may well determine that in some cases it is no longer important what outbound communication forms he or she selects or would like to use. Instead, the greater concern becomes what media forms or customer touch points the prospect has with the organization. The media or marketing manager begins to recognize that he or she is involved in an interactive, reciprocal, and customer-response-driven marketplace, not one in which the marketing organization controls or attempts to control customers by sending out volumes of marketing communication messages using the most efficient media forms. The challenge is no longer figuring out what is best for the marketing organization; it is figuring out what is best for the development or maintenance of customer-marketer relationships. Questions such as, What will build long-term and ongoing responses? not What will maximize short-term organizational transactions? become the key questions for the media planner.

This approach acknowledges the growing impact and effect of the new electronic and relational forms of media and the interactive nature of the Internet, World Wide Web, and mobile and wireless forms of communication. At the same time, it also recognizes that traditional forms of marketing communication such as advertising, public relations, direct marketing, and the like will not disappear. Instead, they will be integrated and aligned with the media usage of customers and prospects rather than existing as separate, stand-alone activities. The challenge of marketing communication allocation is not just to figure out how much should go to traditional marketing communication forms and how much to new media, but how to mix and match all forms of communication or "touch points" to build the desired ongoing customer relationships and generate the desired behavioral response.

Further Complications Arising from Reframing the Question

By thinking of marketing and marketing communication investments and expected returns in terms of whom the media investments should be made against, the question becomes more one of investment and return than of media allocation. One then must ask, What type or form of communication do those customers or prospects either use or prefer, and to which would they be most responsive for this particular product or service in this particular category?

But simply reframing the question does not make it easier to answer. In fact, it makes it more difficult. The media allocator, hitherto an expert in various forms of marketing and communication message and distribution systems, must now become knowledgeable about customers and prospects. But that is difficult in the consumer products arena, in which, in many cases, the marketer is separated from the final user by several levels of channels, all of whom influence the behaviors of customer or prospects to varying degrees and in varying ways. Those circumstances make analysis and decision making all the more difficult, but that is the situation that I believe media planners face in the 21st-century marketplace.

As for which customers or which prospects the organization should target with its marketing communication dollars, obviously it should target those who will likely provide the greatest return, immediately and over time. From the view of the direct or database marketer, this question is generally easily solved. Direct and database marketers should invest in those who have the greatest current or potential value or those who have exhibited the highest response rate in the past, or the lowest cost per order, or some similar favorable behavioral measure. To generate new customers, one should model off one's best customers and find more like them. The common approach to making this decision is to use database and historical-purchase records. This is complemented by standard and analytical techniques such as measurement of recency of purchase (that is, who has purchased from the marketer in the most recent time frame), frequency (who has purchased the greatest number of times over a given period or who has purchased in the greatest volume during the period), and monetary value (those who provide the greatest returns based on their purchase activity). This last valuation is commonly determined at the contribution margin line. While there are other variables that might be added into this analytical framework, such as share of requirements, advocacy, and the like, for the most part, and for most direct marketers, these elements provide a useful if rough first-cut at target customer identification.

In reality, from a database or direct marketer's perspective, most sophisticated marketing organizations already know the value of their customers or prospects—their current value, share of requirements, potential for cross-sells or up-sells, migration, and the like. The reason they know this is that they know their customers' and prospects' behaviors. Thus, in many cases they can identify the ways in which marketing communication investments should be made. For example, if a customer has purchased through a direct-mail offer in the past, the marketing manager knows that making additional direct-mail offers are likely much more relevant and will be more productive than developing some type of marketing event or investing in a sponsorship of some sort. Thus, for those types of organizations that have direct customer access or have ongoing, interactive experiences with their customers and prospects, the challenge posed in the initial question is not a challenge at all: Those firms have generally answered the question through data analysis and sophisticated predictive modeling.

The challenge, of course, still remains for those organizations that do not have direct customer contact or who have little or no data on a major portion of their customers or prospects. That is the situation for many consumer product organizations. Yet while the direct or database marketer does have substantial amounts of information on customers and prospects, or can obtain various levels of customer or prospect information, some major challenges still remain. And, they are the same challenges that face the consumer product marketer. These are discussed in the next section. They begin to frame the research question and potential forms of solution.

Not What Markers Want to Do; What Customers Want Done

At this point, it should be clear that in the scenario just described, the real question of media communication allocation is how customers and prospects would like to acquire information or material about the marketer's products and services, rather than what methods or distribution systems the marketer would like to use. The marketer must therefore try to determine the form, format, and system customers prefer and then use those media forms.

Again, we return to the question of what it is that customers and prospects want. That is an easy question to ask but a hard one to answer. It often makes decisions more difficult for the marketing organization and the media communications allocator. That is why it is often ignored or shunted aside. But it is only by answering that question that we can figure out how best to allocate media communication resources.

While some advocates of one-to-one marketing might argue for an individualized, customer-by-customer approach to finding out what customers want, for most organizations, particularly in the consumer product field, there are simply too many customers and prospects, too many manufacturing challenges, and too many distribution variables for that to be a realistic solution. Instead, because the consumer product marketer commonly achieves relatively low margins on the products being sold, the media manager must find more efficient and affordable—yet still effective—ways of communicating with customers and prospects. Typically, those will continue to be advertising, public relations, sales promotion, and so on, with the addition of the Web, Internet, and other new and useful electronic and interactive methods. Those media may not always be forms of mass communication, but they must have some level of targeting and personalization capability to be useful and effective.

Identifying Customers in Whom to Invest and Communication Vehicles to Use

Since it is not feasible to ask every single customer or prospect what media forms they prefer when it comes to receiving product information, media allocators must estimate or model their customers' or prospects' marketing communication choices or media preferences. In some cases, allocators must identify the media forms that customers or prospects may not say that they prefer but to which they have responded over time. If that can be done, then a methodology can likely be developed that will accommodate those choices.

At this point, the research questions to be resolved are:

1. When is the marketing communication most relevant to the selected customers or prospects for the particular product or service?

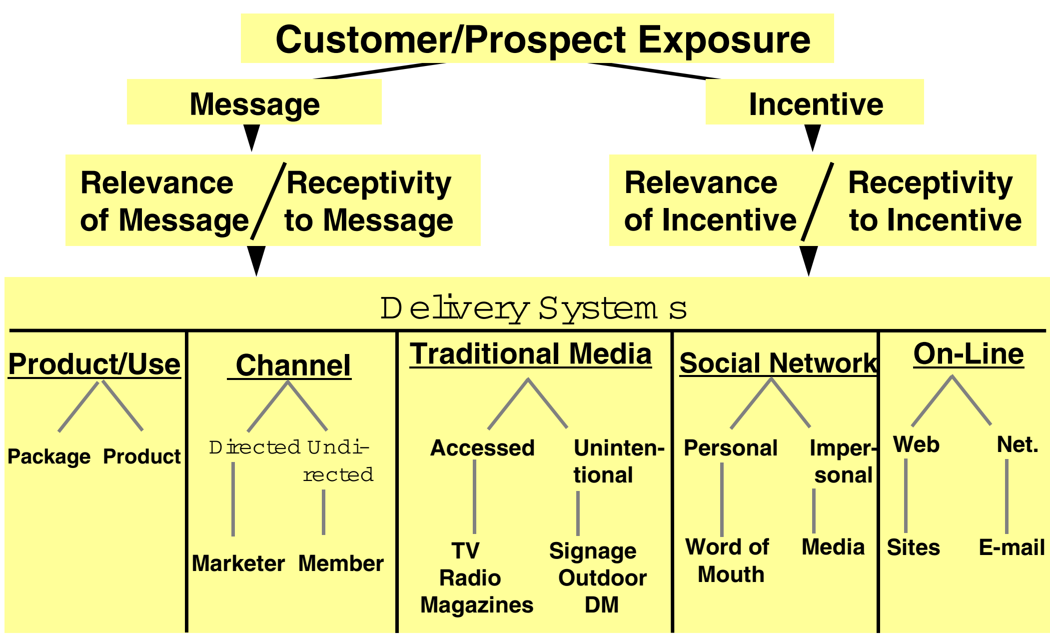
2. When would the customers or prospects be most receptive to receiving the marketing communication?

3. Through what media forms would the customers or prospects most like to receive the information?

One unique feature of the model above is the use of messages and incentives for planning purposes. The reason for using these terms and not such typical media forms as advertising, sales promotion, new media, and the like is simple. Most of the terms and terminology used in media planning have to do with the delivery form. For example, advertising has to go through the media, direct mail through the postal service, sales promotion has to be linked to a price reduction and the like. The real goal of the marketer should be first to determine whether the goal is to deliver (a) a message, which for this paper I define as a communication unit designed to be absorbed, retained, and used by the customer or prospect over time, or (b) an incentive, which is an activity that attempts to influence short-term customer or prospect behavior—typically some type of purchase or purchase-related behavior. By first determining whether the goal is to deliver a message or an incentive, basic media decisions can be reached without all the traditional encumbrance of media forms and distribution systems.

Based on work I have done in this area, I have developed a conceptual model on message and incentive relevance and receptivity. That is shown as Figure 1.

Figure 1. Model of Delivery System Selection



As shown, this model attempts to determine to what media forms customers and prospects currently expose themselves. That is the first cut at proper media allocation. The second step is to attempt to determine when the marketing communication message might be most relevant to customers or prospects. Also, at what point might they be most receptive? That then provides a broad array of marketing communication alternatives, some of which are shown in the illustration. From this decision calculus, a media allocation process, focused on available customers or prospects, could be developed.

The model above is crude and needs refinement. But it does have value if properly implemented. It certainly puts marketing communication in a different light. It poses the marketing communication question as, What do customers or prospects want? not as, What do marketers want? It subsumes the question of media efficiency and instead inserts the question of communication effectiveness. Most of all, it recognizes the shift in marketplace power away from the marketer, who instead of deciding what media forms are most efficient for him or her must now learn first which marketing forms customers want or will attend to and then which of those generate effective results. In short, this approach refocuses the media planner's efforts on outcomes as opposed to mere outputs.

To restate my earlier three questions:

1. How can we identify customer relevance, or how can we identify which messages and incentives should be delivered for the specific product or service to the specific target group of customers or prospects?

2. How can we determine receptivity, or when the customer or prospect might be most interested in receiving or accessing the message or incentive?

3. What media forms do customers and prospects prefer, or to which do they respond?

If the marketing media allocator can find ways to answer those three questions, the industry will have taken a giant step toward resolving the ongoing question of media resource allocation. We will also likely write a new chapter in actually achieving the goal of the marketing concept, which is providing for the wants and needs of customers and prospects, not just trying to maximize returns for the marketing organization.

Econometric Estimation of Media Effectiveness and the Optimal Allocation of Media Expenditure: A Proposal

Gerard J. Tellis, University of Southern California

Introduction

Advances in data collection, research tools, and analytical methods provide us with a great opportunity to research the effectiveness of advertising more precisely than ever before. Such research is useful for optimally allocating advertising resources. However, the success of this task depends critically on our definition and use of appropriate metrics and research designs. This essay first covers the issues involved in metrics and econometric design. It then proposes a research design to test the effectiveness of media expenditures using the econometric approach. Finally, it suggests a method for the optimal allocation of media resources based on the estimated econometric model.

Metrics

Under metrics, we need to consider two issues: measure of independent variable and choice of dependent variable.

Measure of Independent Variable

Advertising can be evaluated by a number of different measures, such as advertising dollars, frequency, reach, gross rating points, and exposures. These measures differ systematically in terms of the level of aggregation. Going from the most disaggregate to the most aggregate, these measures can be ranked as an individual consumer's exposure to an ad, frequency of exposure of various segments reached, gross rating points of an ad or a campaign, and advertising dollars spent. Of these metrics, use of the most disaggregate measure—an individual's exposure to an ad—is probably the most appropriate for several reasons.

First, advertising works because an ad reaches consumers through one or more of their senses, and causes some change in thought, attitude, or behavior. Every exposure causes some change, however small. Indeed, individual exposures might relate to one another, leading to a buildup or wear-out in advertising effectiveness with frequency of exposures. The advertiser has to plan and be responsible for every single exposure. Thus the most reasonable level at which to measure the ad's effect is at the point at which individuals are exposed to it. As soon as researchers choose a more aggregate measure than that, they lose valuable information about the mode, depth, and intensity of consumer response. Such information may be critical for buying time or space and optimally placing ads.

Second, in the current age, media buys can occur at fairly disaggregate levels. For example, an advertiser can choose to position an ad in a particular issue of a magazine, and in one or more articles or pages of the issue. The same holds for various shows on television, or various pages or sites on the Internet. If good decisions must be made at this level, then it is imperative that the researcher measure advertising and its effects at the level of specific programs, articles, shows, or sites of the various media used.

Third, for several media, advertisers can choose to time their ads at various hours of the day. For example, television, radio, Internet, and telephone ads can be varied by times of the day. Ads for different products may vary in effectiveness for different segments at various times of the day. Aggregation, even up to the daily level, causes an irretrievable loss of information. In order to better understand such effects and design ads for them, researchers need to measure the effects of ads by various times of the day.

Fourth, the effects of advertising are not instantaneous. Because of their level of conviction, need to discuss with peers, need for the product, or their personal purchase occasion, consumers' behavioral response to an ad may take place some time after exposure. Temporal aggregation of data leads to a bias in the estimation of this carryover effect. In previous decades, this topic has been researched and debated under the topic of the data aggregation bias. Researchers seemed to have reached a solution to the problem by developing various methods to "recover" the true effects of advertising even with aggregate data. However, all those methods assume that the researcher knows the true data level.

Today, researchers face abundant data at various levels of aggregation. The problem is not so much to recover the true effects of advertising, but to identify and analyze data at the right or true level of temporal aggregation. What is this true level of aggregation? Some authors think it is the inter-purchase time or the time between exposures. I can prove mathematically that the appropriate level of aggregation that is free from temporal bias is the inter-exposure time. This temporal period is the smallest time between exposures. Any higher level of aggregation will cause temporal bias.

Interestingly, the need for this level of aggregation further justifies the use of measures such as advertising exposures at various hours of the day, and at various pages, sites, programs, and articles of those media.

Measure of Dependent Variable

The field of advertising research is currently split between behavioral researchers, who focus on cognitive and affective measures of advertising effectiveness, and econometric modelers, who focus on purchase measures of advertising effectiveness. Which approach is correct?

Neither. The ideal measures of advertising effectiveness should take into account both mental processes and behavior. Focusing only on mental processes ignores whether those measures really account for how advertising works in the marketplace, when faced with distracted consumers, various promotions, and intense competition. Focusing on behavioral measures can only identify what intensity and timing of advertising is most effective. It does not reveal in what other ways the advertising could be wrong or could be improved, such as various creative aspects or timing aspects.

When dealing with the behavioral measures, researchers face a question similar to the one they face when measuring the independent variable: What is the appropriate measure and the appropriate level of aggregation? We can identify several measures, including brand choice, brand switching or repurchase, quantity purchased, timing of purchase, sales, and market share. Of these, the first three are relatively disaggregate measures, while the latter two are relatively aggregate. Analysis at the level of choices, quantities, and timing provide a very rich analysis of how advertising really works. For example, ads for toilet tissue may cause brand switching, those for cookies may cause increased consumption, while those for a dental referral service may cause a change in timing. However, for some products, use of such disaggregate data may enable the researcher to learn whether the advertising causes changes in all three of those metrics.

In terms of the temporal level of aggregation, the time between choices or sales purchases is not critical. What matters even here is the level of aggregation of the independent variable. The appropriate level of the independent variable—advertising—determines the appropriate level of the dependent variable: choice or sales. That is because we are interested in the precise effect of each exposure. Any aggregation among exposures clouds what this particular effect is, especially if the effect is not instantaneous. Thus, we need to measure the data at the level of the inter-exposure time. For example, even if sales occur round the clock, if the advertising is only once a day, then all variables need to be measured at the daily level. On the other hand, even if sales occur one a week, if advertising occurs once a day, then all variables need to be measured at the daily level. In the second case, sales will be, of course, zero for six of the seven days.

Research Design

Goals

The goals of this study are to determine the effectiveness of alternate advertising media and to ascertain optimal advertising dollars to be allocated to each of these media. In particular, the study will determine the effectiveness of television, magazine, and online advertising, in as natural a field (market) setting as possible. The design will adopt the disaggregate econometric approach outlined above.

Sampling Design

This section describes the design of the sample of markets, categories, consumers, and time periods.

Sample Markets. The study proposes to sample one or two small to medium-sized U.S. cities, preferably one that is already hooked up for the collection of syndicated single-source data, and whose demographics relative to the U.S. population at large is well known.

Sample Categories. The study will cover at least the following categories. More categories can be added depending on the interests of the advertiser, the availability of data, and the type of design being used for panelists.

- ❒ Shampoo (or other frequently purchased branded toiletry)
- ❒ Ketchup (or other frequently purchased branded edible product)
- ❒ Television (or other fast-moving electronic product)
- ❒ Computer (or computer accessory)

Differences, if any, in these categories will provide an idea of the generalization of the results, especially across low-priced nondurables and high-priced durables.

Sample Panelists. The study proposes one of two alternate designs depending on the cooperation of syndicated marketing research firms such as Nielsen or IRI.

Design 1 assumes access to a sample of consumers who are already in a single-source data environment such as that supplied by Nielsen or IRI. That is, they already belong to a panel of consumers who make purchases with a card and have their television hooked up with a meter to record their channel viewing. From among such consumers, recruit a subsample that are willing to have their Web surfing fully monitored and to have their purchases on the Web and their purchases of sample nongrocery products also monitored. They also must agree to fill out a questionnaire, which will provide exposure to magazine ads in the form of both subscriptions and reading. Consumers must agree to make nongrocery purchases in the sampled categories with a debit card or to retain the receipts for their purchases. They will be reimbursed 20 percent for small items in the sample and 10 percent for big-ticket items.

If access to the sample described for Design 1 is not available, then consider developing Design 2, as outlined below:

Recruit 500 consumers in one or two cities. These consumers must be willing to have their televisions hooked up to a meter to monitor exposure to television ads. They must be willing to fill out a questionnaire that ascertains various demographic information, including exposure to magazines. They must be willing to have their online access monitored with a permanent software program. In return, the researcher agrees to pay for their online access. Consumers agree to make purchases in the sampled categories separately with a debit card or retain the receipts for their purchases. They will be reimbursed 20 percent for items purchased in the low-priced categories and 10 percent for items purchased in the high-priced categories in the sample.

Sample Time. The study should take one to two years, depending on urgency and the availability of resources. If two years are used, the first year can be used as a calibration period that will cover consumer and market behavior over all the seasons taking place in the year. On the other hand, if only one year is being used, then some of the effects due to seasonality may not be separated from those due to the independent variables.

If only one year of data is used, the test period should be autumn to ensure that the sample covers the time when sales of durables peak.

Cost. Total cost for the sample will be about $20,000 per year.

Econometric Model

Model Specification

The model being used for analysis will be a choice model using the framework of Tellis (1988). However, it will have the lag structure and the inter-media effects of the framework of Tellis, Chandy, and Thaivanich 2000. Thus the model will be a disaggregate choice model allowing for multiple media effects, which each have a multiperiod carryover effect at the hourly level. In a further refinement, individual variables will also be included for the interaction effects of television, magazine, and online advertising. The lag structure of the interaction effects will also be determined.

Key Features of the Model

- ❒ There will be a calibration period of the first half of the sampling time and a test period of the second half of the sampling period. The calibration period will be used to estimate consumers' baseline preferences in terms of brand loyalty, quantity purchased, consumption, and stocking, as outlined in Tellis (1988).
- ❒ Individual-level panel analysis will allow for the separation of loyalty effects outlined above from panelists' true response to marketing mix.
- ❒ Tracking of advertising on television, in magazines and over the Internet together with the channels, articles, or sites, respectively, in which those ad appeared will allow for the decomposition of ad effects in terms of the various media and vehicles that brands use.
- ❒ The use of interaction effects of the three media enables us to test for the effectiveness in overlap of reach of various media, if any.
- ❒ The analysis of lag structure will allow for the determination of the precise day of advertising within day, week and over time, if any. This can also be done within each medium, as well for interaction, if any, among the media.

Estimation of Empirical Ad-Response Function

Once the data are available and cleaned, one can estimate the model specified in the previous section. Using the beta coefficients derived from this estimation we can derive the advertising response function of each medium over the panel of subjects in the sample. This response function will be S-shaped, following the logit model.

Optimal Allocation of Media Expenditures

This exercise can be broken into the estimation of sales response function, estimation of the profit function, identification of the optima, and allocation of media expenditures.

To analyze profits, we need to choose a level of aggregation that does not involve too much loss of the refinement of the disaggregate analysis, yet is sufficiently aggregate to enable one to apportion costs of advertising. The weekly level should be suitable for this purpose.

Determining the Sales Response Function. Weekly sales can be obtained as follows. First, if the response function is likely to differ by usage or demographic segments, then determine those segments. Then estimate the model separately for each segment. Also estimate the mean weekly sales for each segment. Estimate the media response function for each segment in terms of probability of purchase as a function of media exposures, using the econometric model specified and estimated above. Estimate sales as a function of the media exposures by multiplying average sales per week into probability of purchase as a function of media exposures, for each sample.

If strong seasonality exists and ad response is likely to vary by seasons, then this entire analysis must be done separately for each season.

If the sample is not representative of the population in the city, the sampled sales response curve must be corrected for sampling difference relative to the population.

Determining the Profit Function. Now Profits = Revenues – Costs (all measured at the weekly level). Revenues are sales per week into average price per week. The key costs that we must consider are the media costs. The media costs are the total advertising dollars spent to reach the panelists for each week.

Identification of Optima. Given that the underlying probability of purchase is S-shaped, the sales and revenue functions will also be S-shaped, because sales are obtained from probability of purchase by the multiplication of a constant (average sales per week for the segments). Similarly, revenues are obtained from sales by the multiplication of a constant, average price per week.

However, if the revenue function is S-shaped, the profit function will be an asymmetric bell shape, because when sales increase first slowly and then rapidly, profits will increase first slowly and then rapidly, since advertising costs increase at a constant rate. However, when sales taper off and remain constant at the upper end of the S-shaped curve, costs continue to increase at a constant rate. Thus profits will fall at an increasing rate.

Such a profit function will have a unique optimum for each medium used and for each segment and season, if estimated separately.

Optimal Allocation of Media Budget. The optimal allocation will have to be done sequentially. For initial allocation, once the profit function has been obtained, the optimal allocation of media expenditures is relatively easy. The optimal advertising for each medium during a week is the level at which profits for that week are at a maximum. Then there are iterative adjustments. Will the optimal allocation of media expenditures remain the same?

The answer depends on consumer behavior. There are two possibilities. On the one hand, consumer response itself may vary depending on the level of advertising. Thus, to achieve continuous optimal allocation, the analysis needs to be continued on a regular basis via tracking analysis on a limited sample of subjects using the above framework.

With repetition of the analysis, one can determine how and to what extent the response function itself is sensitive to the level of advertising. The optimal allocation in the subsequent periods will have to be changed based on the response function of the previous period. Over a period of such adjustments, the researcher may determine a pattern in allocation that will allow for some sort of global steady state optimal.

On the other hand, if consumer behavior is not sensitive to the level of media allocation, then the estimation in the first stage above will arrive at the global advertising. In the latter case, changes, if any, will be merely fine tuning of the above optima after periodic repetitions of the above analysis, to estimate small changes in the basic response functions.

Conclusion

The optimal allocation of advertising resources is a perennial problem facing managers. This paper argues that new data that are currently available offer a promise to achieve such goals far more efficiently than ever before. The key to such good analyses is to use appropriate metrics to measure advertising and its effects. In particular, the use of highly disaggregate measures of advertising for highly disaggregate time periods is more precise, less biased, and more insightful. This paper proposes a research design to carry out such an analysis and to use the findings for the optimal allocation of media expenditures.

References

Tellis, Gerard J. (1988), "Advertising Exposure, Loyalty and Brand Purchase: A Two Stage Model of Choice." *Journal of Marketing Research* 15 (May), 134-44.

Tellis, Gerard J., Rajesh Chandy, and Pattana Thaivanich (2000), "Decomposing the Effects of Direct Advertising: Which Brand Works, When, Where, and How Long?" *Journal of Marketing Research* 37 (February), 32-46.

About MSI

MSI was established in 1961 as a not-for-profit institute with the goal of bringing together business leaders and academics to create knowledge that will improve business performance. The primary mission was to provide intellectual leadership in marketing and its allied fields. Over the years, MSI's global network of scholars from leading graduate schools of management and thought leaders from sponsoring corporations has expanded to encompass multiple business functions and disciplines. Issues of key importance to business performance are identified by the Board of Trustees, which represents MSI corporations and the academic community. MSI supports studies by academics on these issues and disseminates the results through conferences and workshops, as well as through its publications series.

About the Yaffe Center

A joint venture of the School of Business and the School of Art & Design, this University of Michigan initiative is a collaborative, cross-disciplinary center for the study of persuasive mass communication—both explicit (content) and implicit (design and style), in all its forms and media. Its research agenda includes the synthesizing of existing knowledge about persuasive communications that today exists in fields as diverse as advertising, film and video, law, politics, and religion, as well as the creation of new knowledge on visual persuasion and other fields. It conducts courses for students as well as conferences and workshops for practitioners and scholars. For more information please see the Center's web-site at www.yaffecenter.org.

MSI

MARKETING SCIENCE INSTITUTE

1000 Massachusetts Avenue
Cambridge, MA 02138 USA
Phone: 617.491.2060
www.msi.org