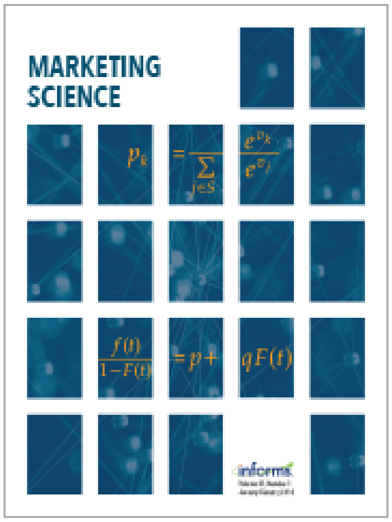


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Practice Prize Reports

Ashish Sinha, J. Jeffrey Inman, Yantao Wang, Joonwook Park, Gerard J. Tellis, Rajesh K. Chandy, Deborah MacInnis, Pattana Thaivanich,

MARKETING SCIENCE
Vol. 24, No. 3, Summer 2005, pp. 351–366
ISSN 0732-2399 | EISSN 1526-548X | 05 | 2403 | 0351

DOI 10.1287/mksc.1050.0159

Practice Prize Report

Attribute Drivers: A Factor Analytic Choice Map Approach for Understanding Choices Among SKUs

Ashish Sinha
School of Marketing and International Business, Victoria University of Wellington,
Wellington, New Zealand, ashish.sinha@vuw.ac.nz

J. Jeffrey Inman
Katz Graduate School of Business, University of Pittsburgh, 356 Mervis Hall,
Pittsburgh, Pennsylvania 15260, jinman@katz.pitt.edu

Yantao Wang
ACNielsen, 2700 Patriot Boulevard, Suite 250, Glenview, Illinois 60025-8025, yantao.wang@acnielsen.com

Joonwook Park
Smeal College of Business Administration, Pennsylvania State University, University Park, Pennsylvania 16802,
joonwookpark@psu.edu

We describe the implementation of Attribute Drivers (AD), a flagship panel product of Information Resources Incorporated, at Campbell Soup Company. AD combines the parsimony of a factor analytic choice map approach with the ability to incorporate the dynamics of choice decisions to understand consumers' choices among stock keeping units (SKUs). A key distinguishing feature of this methodology is its scalability and applicability to large-scale problems.

The application of AD helped Campbell's grow its revenues at twice the category growth rate. This revenue growth was achieved in a climate of high product proliferation, a slow economy, and a five-year decline in unit sales at a category level. Campbell has applied AD in four primary areas: making restaging decisions, identifying potential line extensions and estimating their volume and market share impacts at the brand and category level, performing price gap analysis for new products, and increasing responsiveness to consumers' needs. The model has been used by several other clients, testifying to its transportability.

Key words: product management; choice models; assortment
History: This paper was received September 3, 2004, and was with the authors 2 months for 1 revision; processed by Gary Lilien.

1. Introduction

While consumer packaged goods (CPG) companies enjoy solid balance sheets, healthy profit margins, and a good return on invested capital, many analysts believe that this industry is no longer a growth industry (e.g., McKinsey & Company 2004). Despite continued earnings growth from productivity improvements, many CPG companies have found their stock punished for their failure to deliver robust topline growth. There is consensus among consultants and analysts that the rising retail customer power requires understanding and connecting with the end consumer better than ever before (Rosenbleeth et al. 2002).

Brand managers spend a large portion of their marketing budget on price promotions, displays, and features that increase sales and market share in the short run. The marketing research industry has developed several products, such as ScanPro (ACNielsen) and Mix Drivers (IRI), which seek to provide CPG companies with a better understanding of the impact of marketing mix variables (e.g., advertising, price promotions, sales promotion) on incremental volume (i.e., sales volume beyond what the brand would have sold in the absence of promotion). Since incremental volume is in the range of 20%–30%, it is imperative that managers understand the drivers

Editor Note: This report has benefited from comments from the Practice Prize editor and Committee but has *not* been reviewed by *Marking Science.*

of so-called "base volume," which is defined as the sales volume a brand would sell in the absence of promotion. In order to generate actionable insights, CPG companies are interested in the underpinnings of stock keeping unit (SKU) choice, and product attributes have been identified as important facets in this regard (e.g., Fader and Hardie 1996, Andrews and Manrai 1999).

The focus of this paper is on a commercial product that endeavors to assess the impact of product attributes on consumer choice using scanner panel data—attribute drivers (AD). Soon after AD's introduction in 2001, it became the flagship panel product for Information Resources and has generated over $5 million in revenue. It has been applied to numerous product categories, both slow moving (e.g., cough suppressants) and relatively fast moving (e.g., salty snacks and carbonated soft drinks). Several multibillion dollar companies, such as Procter and Gamble, Campbell, BIC, Pepsico, and Frito-Lay, have used the AD model to obtain insights regarding brand management issues, identify potential line extensions, address category management issues, perform price gap analysis, and assess alternative marketing plans for the launch of a new line item. Table 1 shows some illustrative examples of product categories to which AD has been applied.

The generality of this model and its managerial usefulness are its two most important virtues. AD is the culmination of several years of effort at IRI that stemmed from the application of Fader and Hardie's model (1996) to extremely large categories and the use of store data and multiplicative competitive interaction models (Cooper and Nakanishi 1988) to reveal the drivers of choice. While making important contributions to the practice of marketing research, an important limitation of Fader and Hardie's model (1996) is the so-called "curse of dimensionality" in parameter space, making its application to larger product categories impractical. An alternative, albeit simpler, formulation to Fader and Hardie's model (1996) is the application of the MCI model to store data, enabling the model to be estimated with readily available statistical packages, such as SAS and SPSS, using simple regression. However, the inability of these models to adequately capture market structure led IRI to rethink the implementation of this model. This led to the development of the AD model.

Table 1 Categories to Which AD Has Been Applied

Category	Marketing plans	Brand equity	Price premium	Market structure	Assortment	Line extensions
Toothpaste		X				X
Salty snacks	X	X	X	X	X	
Soups	X	X	X	X	X	X
Shaving products		X	X			X
Frozen foods		X	X	X		

A brief exposition of the AD model that outlines the technical superiority of the system over previous models is an important precursor to the case study. Therefore, in the next section, we provide a brief overview of the model that forms the foundation of AD, followed by the case studies that describe the application of AD to a category management problem for a Campbell's soup brand.

2. The Model

Researchers (Roberts et al. 2004, 2005) have argued that the major reason for the limited impact of marketing science tools is their inability to be adapted to the managerial context. We contend that many marketing science models, particularly in academia, are developed with little consideration of scalability and practical constraints on computational power, time, and resources. In addition, the solutions are often highly situation- and dataset-specific.

These issues become quite pertinent for AD, as the level of complexity is manifold greater than most models estimated on scanner panel data. For one, these models are estimated on choices among SKUs rather than the usual brand-pack-size level. Academic approaches to this area have typically focused on the development of stylized models applied to small data sets (Fader and Hardie 1996, Andrews and Manrai 1999, Sinha et al. 2005). Applying these models in large-scale settings can be quite problematic because the number of SKUs and households even in smaller categories are many times greater than the data set on which these models have been developed. The limited scalability of these models has forced applied researchers and consultants to estimate models at the brand-pack-size level, thereby precluding attribute-level insights.

In order to mitigate these limitations, scalability, generalizability, parsimony, and ease of interpretation were the four guiding principles for the development of the AD model. Our modeling approach is similar to Brownstone and Train (1999), who recommend combining the results from choice models with actual market share data to obtain better forecasts. While their recommendations were for using stated preference models for forecasting the volume of new products, we contend that panel-based models are best suited for understanding substitutability and interaction between products (Bucklin and Gupta 1999). This is due to the fact that panel data overrepresents certain demographic groups, while underrepresenting others. Other researchers (e.g., Hardie et al. 2004) have also noted that managers tend to be highly reliant on

Table 2 Modeling Overview

	Data	Example outcomes
Step 1	Scanner panel data	Understanding market structure, segment-specific parameters, and attribute importance
Step 2	Scanner store data	Market share predictions, volume predictions, profitability, ROI

store scanner data and therefore are reluctant to utilize panel data due to the aforementioned sampling limitations.

Due to the lack of comparability between the market share information provided by store data and panel data, coefficients obtained from the analysis of panel data need to be adjusted. An alternative, simpler method is to first calculate the source of volume for a new SKU in percentage using choice models and panel data. This source of volume provides information pertaining to the degree of cannibalization of the added SKU. Subsequently, the sources of volume in percentages are applied to the SKU market share information obtained from store scanner data to provide a realistic estimate of the market behavior. Hence, once the model has been calibrated on panel data, the estimates can then be applied to store scanner data to obtain more realistic product and alternative strategy forecasts. Our modeling strategy is provided in Table 2 followed by a brief exposition of the scanner panel model.

Panel Data Application

We briefly outline the model estimated on panel data. A more detailed exposition of the model is provided in the appendix. Because most categories are composed of several hundred SKUs, we follow Fader and Hardie's approach (1996) to define preference for an SKU as the sum of the preference for the attribute levels that it possesses, controlling for the effects of marketing mix variables. An implicit assumption of this methodology is that each SKU can be defined by different attributes such as brand, flavor, and size. Further, it presumes that consumers do not form preferences for an SKU; rather, preference for an SKU is the sum of the preferences for the different attribute levels that it possesses. These models can be described as "conjoint analysis meets panel data" attempts to understand the "mind of the consumer."

The proposed model is an extension of the general unfolding methodology proposed by Andrews and Manrai (1999) and Sinha and Sahgal (2002) that incorporates the dynamics of the SKU choice decision. An important advantage of this model is that the total utility that a consumer derives from the consumption of the SKU is decomposed into three components: the utility derived from the marketing mix variables (e.g., price, display, advertising), the utility derived from the attribute levels that are possessed by the SKU (e.g., brand, flavor), and the utility derived from attributes that are uniquely possessed by the SKU (measured by SKU-specific loyalty).[1] The ability of this model to decompose utility that a consumer derives from the consumption of an SKU allows us to understand the effect that the unique attributes have on choice. An assumption of this methodology is that the volume derived from the unique attributes is a measure of the incrementality of the SKU to the category, therefore allowing modelers to estimate the impact on the category of an SKU deletion. In the next section we describe two applications of this model to a "real-world" problem.

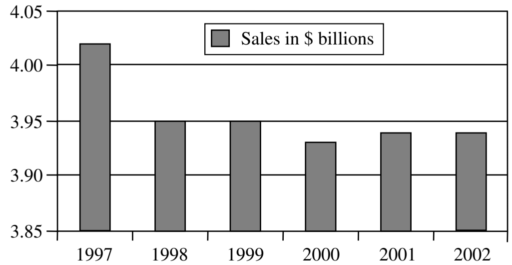
3. Case Study—Identifying Optimal Assortment and Restaging of a Line

In this section we provide an example of how the AD product was applied to measure brand equity and to provide SKU assortment recommendations for Campbell Soup Company. Campbell Soup Company is one of the largest soup companies in the world, operating in over 120 countries, with worldwide revenue of $7.8 billion dollars and a market cap of over $11 billion. It is a key player in the highly fragmented mature soup category, which is characterized by stiff competition from a number of multinational companies as well as high-quality generics and private labels. In addition, changing consumer tastes and needs have highlighted the need to understand the dynamics of SKU choice. In addition, downturn in the economy resulted in flat annual sales at the category level during the last five years. Figure 1 provides sales figures for the soup category from 1997 to 2002, showing that sales in real terms had declined by over 2%.

In sum, changing consumer needs and preferences, competitive pressure, and a mature industry had forced the company to become more innovative. Each year the company introduces several new SKUs to the market. These new SKUs (also called "new news") are considered extremely important to the growth of the company and also for maintaining a healthy position in the marketplace. However, the introduction of these new products, such as the introduction of microwaveable soups and easy-to-open cans, results in a number of strategic and tactical decisions. For one, resource and operational constraints (e.g., retailer shelf space) force the company to identify SKUs that need to be delisted. This delisting is dependent on the

[1] This approach is analogous to the promotional sales bump decomposition reported by van Heerde et al. (2004).

Figure 1 Sales for the Total Soup Category in the U.S. Market



new news introduced that particular year (e.g., the new SKUs may dominate existing SKUs). Therefore, the company needed a delisting strategy, which was termed internal SKU rationalization.

Second, of the many potential candidates for new products, only a few make the cut. While preliminary primary research is used to make these decisions, the information pertaining to the potential success of these products is often incomplete. More accurate estimates, such as concept testing, test markets, and BASES, are expensive and time consuming. Prior to the adoption of AD, new product launch decisions were highly subjective, often depending more on managerial judgment than on facts.

Thirdly, the category is characterized both by product proliferation and a large number of competitive SKU introductions, particularly from low-priced generics and private labels. This phenomenon escalates the pressure on scarce shelf space. Each year the company provides assortment recommendations to key accounts (e.g., Wal-Mart, Kroger). This process involves convincing the account managers of the benefits of stocking the new products that the company has introduced. Account managers are often skeptical of the new product introduction, since new products have no history of success and suffer from a high failure rate. In fact, it is becoming increasingly common for retailers to ask even large national brands to prove their brand's worth to the growth of the category, since low-priced "Me Too" brands often offer quality that is equivalent to the more expensive national brands. Hence, there is a growing need on the part of CPG companies to offer retailers fact-based analysis supporting the new SKUs' potential to increase category profitability.

Lastly, while the category as a whole was characterized by flat annual sales over the last few years, there were certain areas of the market that showed promise, such as ready-to-serve varieties. The ready-to-serve segment had experienced 16% growth between 2000 and 2002, but this success had come at the expense of condensed soups, which saw an 8.8% sales decrease between 1997 and 2002. In response to the growing popularity of the ready-to-serve segment, the company wanted to relaunch the noncondensed line of soups, allowing it to differentiate this line from the condensed soups. Campbell sought insights into the relaunch strategy for their noncondensed line.

Alternative Approaches to Attribute Drivers

A simple approach that CPG companies use is an SKU ranking report, which ranks the SKUs on the basis of sales velocity, defined as the sales in equivalized units (such as pounds) controlling for distributional effects. SKUs that are ranked in the lower quartile are considered prime candidates for delisting. There are three limitations of this method. First, this method does not control for each SKU's marketing support. For instance, certain SKUs have a lower base price and are promoted more often. This may result in their enjoying higher sales, but at the same time they may be more expensive to support than other SKUs. Second, this ranking report does not account for the cannibalization effects of different SKUs. Lastly, SKUs that have high loyalty and low volume are likely to be delisted, even though a high proportion of the volume attributed to the SKU may be incremental to the category (Bucklin and Gupta 1999).

Alternatively, an estimate of the market share for the new product or relaunched product can be obtained by using BASES (ACNielsen), Behavior Scan (IRI), or managerial intuition. Apart from being expensive, the BASES methodology is based on stated preference that often does not reflect what consumers actually do. An alternative to BASES is Behavior Scan, an actual test market that is highly popular for new product introductions and brand relaunches. While Behavior Scan provides accurate results, due to its predictions being based on actual behavior in real test markets, the data requirements, cost, and time needed for this analysis often prove prohibitive.

The AD Solution

Campbell Soup Company wanted a more robust, efficient, and fact-based system for providing a solution to the assortment problem. The company also wanted a simulation system that would help them make decisions pertaining to the relaunch of the noncondensed variety of soup. In addition, Campbell's sought a system that could provide recommendations and evaluations of line extensions, particularly in the microwaveable soup category. In this regard, the company was interested in understanding the category drivers and how these drivers varied across brands in the portfolio and the category. Finally, it required a system that could be used in account level assortment planning.

The AD model was deemed very appropriate for this application. All SKUs were described using seven attributes: *brand*, *vendor* (manufacturer), *type* (e.g., dry

Table 3 Attribute Importance for Brands and Categories

	Category	Chelsea	Savanah	Competition
SKU loyalty	20.6	21.0	23.3	15.2
Vendor	16.9	6.9	13.9	27.8
Package	16.4	16.8	5.8	18.7
Brand	16.2	22.3	7.5	15.9
Flavor	10.9	14.8	20.6	7.8
Eating/cooking	7.1	7.4	11.3	5.1
Calorie	6.2	6.2	9.0	5.3
Type	4.4	3.7	7.2	3.4
Base price	1.3	0.9	1.3	0.7

or wet segments), *subtype* (e.g., eating and cooking), *calorie* (e.g., high or low), *flavor*, and *pack-type*. In addition, SKU loyalty and attribute-specific loyalty terms were included to capture the dynamics of consumer behavior as well as the loyalty toward them that was over and above the attributes that an SKU possessed (see the discussion in the previous section).

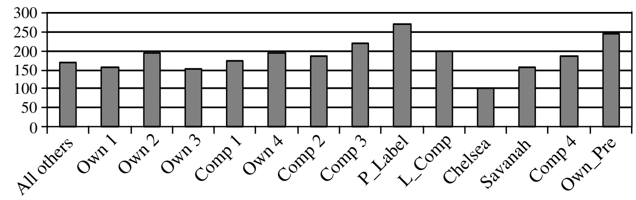
Table 3 provides the importance of each attribute to the consumer at a category and brand level, controlling for the effects of marketing mix variables.[2] It provides Campbell with an understanding of drivers of SKU choice for the category, brands in their portfolio, and for competition. The figure shows that brand, vendor, pack-type, and SKU loyalty account for over 65% of the importance. Hence, these four attributes are the largest drivers of choice, controlling for the effect of marketing mix activities such as price reduction, feature, display, and advertising. Given the importance of brand and package attributes, there appear to be substantial potential gains in realigning the brand and also in introducing new pack-types. Note that the drivers at the brand level are different, with the vendor attribute not being important for Chelsea and Savanah but an important driver for the competition. Also, the variance explained by SKU loyalty is much lower for competition than it is for Chelsea and Savanah. Another insight is that package does not play an important role in the purchase decision for Savanah, implying that the attractiveness of the SKUs for this brand could benefit from attractive new pack-types.

Applications of AD

Relaunch. Figure 2 illustrates the attractiveness scores for the different brands in the category. The attractiveness scores are obtained by taking the exponents of the brand coefficients and normalizing them, such that Chelsea has a score of 100 (i.e., benchmarking them against the Chelsea brand). This table shows that Chelsea is the least preferred brand, and the premium own brand and private label are the two most preferred brands. Given the importance of the brand

[2] Brand names have been disguised to maintain confidentiality.

Figure 2 Attractiveness Scores of Different Brands*



*Own_Pre = own premium brand, L_Comp = large competitor brand, Own = own brands, Comp = competitors.

as a driver of purchase, particularly at a category level, the management decided to relaunch a portion of the Chelsea line under a different brand name, as Chelsea's brand attractiveness was the lowest of all the brands in the category.

Volume forecasts and the sources of volume for this introduction were of interest to Campbell's management. Because the new brand name did not exist in the market, past primary research and managerial intuition guided its attractiveness score estimate. Management believed that the relaunched brand name would be more attractive than the original Chelsea. It was also felt that the attractiveness of the new brand would lie somewhere between the attractiveness of Chelsea and Savanah. Two attractiveness scores were chosen for the analysis: 50% (an attractiveness score of 129) and 75% (score of 144) of the way between Chelsea and Savanah. The simulation consisted of delisting ten Chelsea SKUs and reintroducing nine same or improved SKUs under the new brand name. Table 4 provides an estimate of the net volume gain of introducing the new line, which shows an increase of 2.3 million (equivalized) units represented in terms of pound sales of equivalized volume. Therefore, our analysis shows that rebranding of the Chelsea brand will lead to an increase of 2.3 million units sold.

Assortment. The assortment component of the project involved the delivery of a decision support system that predicted the brand and category volume as well as profitability changes from delisting an existing SKU or from introducing a new SKU in an account. Table 5 shows that the most incremental SKUs to a brand do not necessarily have the greatest volume.[3] Clearly, the ranking of the SKUs depends upon the criterion that is being used to evaluate an SKU (e.g., profitability, volume, incrementality, or some combination of these measures). In addition, it shows that SKUs that have low volume are not necessarily the best candidates for delisting, which is the usual practice in the industry. Hence, consistent with our earlier contention, as well as that of other

[3] Incrementality is defined as the volume lost to the competition if a particular SKU is deleted.

Table 4 **Volume Impacts due to Relaunch of Noncondensed Line of Product**

	% Vol incremental to portfolio	% Vol cannibalized to portfolio	Sales lost/gained (equivalized units)
Delisting	−59.1	−40.9	−47.7 M
New brand launched	60.5	39.5	50.0 M

Table 5 **SKU Incrementality to the Brand**

SKU No.	Segment	Volume	Percentage incremental	Incrementality ranking
1	Condensed	2,938	72.5	1
2	RTS soup	46	70.9	4
3	RTS broth	13,012,908	70.3	6

research (Bucklin and Gupta 1999), a simple sales velocity approach to assortment decisions is biased at best and completely erroneous at worst.

The analysis was used for two purposes. First, it provided a basis for company-wide SKU rationalization decisions. Second, it helped the company acquire scarce shelf space at an important account. In terms of SKU rationalization, the least incremental SKUs were prime candidates for the national delisting strategy. Hence, AD helped the company become more cost efficient. On the other hand, the most incremental SKUs required more careful management of marketing support. In particular, low volume but highly incremental and profitable SKUs were identified as possible opportunities. One reason for highly attractive SKUs to have low volume is that a simple sales measure previously identified these SKUs as poorly performing, leading to their being inadequately supported. AD provides a method for identifying attractive SKUs and for understanding the degree of support (e.g., distribution, price discounts, displays) required to make the product succeed in the market. This enabled the company to deploy its resources more effectively.

The decision support system was also used to help Campbell's acquire scarce shelf space at a major retail account. Using the assortment tool, the effect of removing 8% of SKUs of a store brand from an important account was assessed. The simulation showed that while there was no loss to the category volume (for these SKUs, no significant volume was attributable to SKU loyalty), a significant proportion of the volume would go to Campbell's with approximately a third of the volume being retained by the store brand. This provided Campbell's the space to place twenty new items in the account. *This was considered as a huge win by management, and the ability to use models and simulations to show retailers the impact of delisting competitor's SKUs was a critical aspect of this analysis.*

4. Implementation and Transportability

Based on our model results, the following strategy was recommended. First, we recommended the relaunch of the line of Chelsea brand under a different brand name. The rationale for that has been provided in Table 3, which shows that the model predicts an increase of 2.3 million units of sales from the renaming of this line of soups. Second, the use of microwaveable containers rather than styrofoam cups in the ready-to-serve segment was also recommended to the management. In addition, given that pack-type was an important driver in the purchase decision, it was recommended that the "easy-to-open" pack-type, rather than regular cans, be the preferred package. Lastly, while the recommended national assortment was identified on the basis of profitability, we note that the management used this recommendation as one of the many inputs for making this decision. Specifically, the optimal assortment was identified based on profitability calculations that required estimates of SKU incrementality to the brand obtained from the model, national sales in dollars obtained from scanner sales data, and the cost of the SKU.

Besides using the assortment tool used for national level delisting decisions, this tool has also been used for providing guidance and recommendations at the regional level. In fact, this decision support system was rolled out to more than thirty accounts, and decisions at each of these accounts are being made using it. In several situations, the managers were also able to show higher profitability from stocking Campbell's SKUs versus those of competitors, thereby helping the company make a case for increased shelf space for new product introductions. Furthermore, simulations for twenty new products were also provided as a part of this project. Hence, this tool and the findings from this project are being used for making both strategic decisions, as in the case of restaging of a brand, to more tactical decisions, such as account level assortment decisions.

Beyond the organization-wide impact of this project, there have also been several financial gains. For one, the company increased its sales by over 2% in the year 2003, outpacing category revenue growth of 1% and declines in volume. Because most companies run multiple research projects concurrently, particularly in situations of brand and product launches, the results from these studies are often used in conjunction with each other rather than in isolation. Therefore, one project can rarely claim to be the sole driver of improved company performance. Rather, we believe that the company used inputs from several studies and that most of the recommendations that we made had a bearing on the eventual decisions. The financial benefit of running this approach, calculated in terms of ROI, is in the vicinity of 3,300%.

The Campbell's story is just one illustration of how this approach can provide client insights and benefits. There are many other examples of the transportability of the AD model. For instance, the model has been used in the shaving category to understand the price gap between the client's brand and the premium brand in the category, thereby helping the client make more effective pricing decisions. Further, the AD model has been used in the salty snacks category in making assortment decisions and in understanding the brand equity of all the brands in the client's portfolio. Similar to the Campbell case, the analysis was used in this category for the addition and deletion of SKUs. The model's transportability is evinced by the fact that it has been applied to twenty two client situations across twelve companies by some of the largest CPG companies in the world. The significant repeat business from existing clients also attests to its success.

5. Summary

We have described IRI's Attribute Drivers, a scanner panel data-based choice model that attempts to understand the drivers of choice at an SKU level. It uses a principal components specification, coupled with Guadagni and Little (1983) loyalty terms to specify loyalty at the attribute and SKU level, for a parsimonious solution to the problem of the so-called "curse of dimensionality" in parameter space. Apart from having the advantage of being highly parsimonious, this model also allows managers to understand the drivers of choice at different levels of geography and segments. While providing a brief background of the model, we also described the implementation of the system in a multibillion dollar soup company. Readers interested in the details of the methodology should refer to Sinha et al. (2005).

The system is being used for predicting the success and failures of new product introductions and identifying attribute levels for further investigation and also for tactical and strategic decisions such as brand restaging and assortment decisions. While the company spent under $300,000 on the project, the benefit derived from this was in excess of $10,000,000. It is important to note that several other insights that this model provides, such as price gap analysis, pricing recommendations for new products and brand equity, to name a few, have not been discussed at length here. Instances and examples of the use of these insights can be obtained from the authors.

Acknowledgments
The authors thank Information Resources Incorporated for their help and assistance in the preparation of this paper.

Appendix

Scanner Panel Data Model. We assume that the probability that the ith individual will choose the jth-brand in the tth time period, given that the individual belongs to the sth segment, can be expressed as follows:

$$p_{ijt/s} = \frac{\exp(V_{ijt/s})}{\sum_{j'} \exp(V_{ij't/s})}, \tag{A1}$$

where $V_{ijt/s}$ is the deterministic component of the utility that the ith individual derives from the jth SKU in the tth time period, given that the ith individual belongs to the sth segment. The most general form of the deterministic component is represented as follows:

$$V_{ijt/s} = \alpha_{ijt|s} + \sum_n \sum_k \beta_{iknt|s} * X_{knj} + \sum_m \kappa_{sm} * Z_{mjt}, \tag{A2}$$

where $\alpha_{ijt|s}$ is the time varying intercept term for the jth SKU, given that the ith individual belongs to the sth segment; $\beta_{iknt|s}$ is the time varying preference that the ith individual has for the kth attribute level of the nth attribute, given that the ith individual belongs to the sth segment; $X_{knj} = 1$, if the kth attribute level of the nth attribute is possessed by the jth SKU, $X_{knj} = 0$, otherwise; κ_{sm} is the parameter that measures the effect of the mth marketing mix variable for the sth segment; and Z_{mjt} is the value for the mth marketing mix variable of the jth SKU at the tth time period.

In summary, $\alpha_{ijt|s}$ is a time-varying, individual-specific term that measures the effect of attributes that are uniquely possessed by an SKU, and $\beta_{iknt|s}$ measures the effect of the attribute levels possessed by the SKU, while κ_{sm} controls for the impact of marketing mix variables on consumer choice. Strictly speaking, Equation A2 is not identified and, hence, inestimable, as both $\alpha_{ijt|s}$ and $\beta_{iknt|s}$ parameters vary over time and individuals. We impose restrictions on these parameters for identification reasons, which are discussed below.

It was deemed important for the model to incorporate market structure, both at a brand and attribute level, in order to avoid the independence from irrelevant alternatives (IIA) property and to provide an understanding of the interactions between SKUs. In addition, the model also needed to incorporate dynamics of consumer choice, particularly for categories in which consumers were expected to exhibit variety-seeking behavior (i.e., negative-state dependence). A choice map approach (e.g., Sinha et al. 2005) is applied to obtain a parsimonious solution to this problem. In this approach, each segment is represented as a linear transformation of any one of the segments. For example, if an attribute (e.g., brand) has ten levels, then a one-dimension, three-segment solution will only have $(10-1)+2$ parameters rather than $(10-1)*3$ parameters in the unconstrained Fader and Hardie model (1996). A major advantage of this method is that in addition to parsimony, it provides an understanding of the interaction between SKUs, thereby providing accurate sources of volumes for SKU additions and deletions. Therefore, $\beta_{ikn|s}$ is reparameterized and restricted as follows:

$$\beta_{iknt|s} = \beta_{kn|s} = \sum_f \{l_{fkn} * w_{sfn}\}, \tag{A3}$$

where l_{fkn} denotes the location of the kth attribute level of the nth attribute along the fth latent factor (benefit), and w_{sfn} is the weight that the sth segment attaches to the fth factor (benefit) of the nth attribute.

Elrod and Keane (1995) showed that a simple choice map approach (a principal component approach) is often insufficient to capture consumer choice behavior, as consumer responses are also driven by factors that are idiosyncratic to an attribute level. One could interpret this as loyalty that consumers have towards an attribute level that is not captured by factors that are common to all attribute levels. A Guadagni and Little (G&L)-type loyalty variable (1983) is used to capture this component of behavior. Note that this variable also helps to capture the dynamics of consumer behavior, particularly for categories in which consumers exhibit variety-seeking behavior. A negative parameter for attribute-specific loyalty variables indicates that the preference for an attribute level in the current time period is reduced if it was bought in the last time period. Hence, this model also captures the temporal changes in the preference structure for an attribute level (Inman 2001):

$$\beta_{iknt|s} = \sum_{f} \{ l_{fkn} * w_{sfn} \} + \mu_{sn} * GL(iknt \mid \delta_n), \tag{A4}$$

where $GL(iknt \mid \delta_n) = GL_{iknt-1} * \delta_n + d_{iknt-1} * (1 - \delta_n)$ is the Guadagni and Little (1983)-type loyalty variable that the ith individual has for the kth attribute level of the nth attribute at the tth time period; $d_{iknt-1} = 1$, if ith individual has chosen the kth attribute level of the nth attribute at $t-1$ time period; and δ_n is the smoothing parameter for the nth attribute (see Swait and Erdem (2002) for a similar specification).

Lastly, there may be consumer loyalty for a particular SKU beyond the preference and loyalty towards the attribute levels (e.g., a combination of attribute levels that is unique and idiosyncratic to the SKU). This SKU loyalty is captured by a G&L (1983)-type loyalty variable defined at the SKU level. In line with the above argument, the parameter $\alpha_{ijt|s}$ is parameterized as follows:

$$\alpha_{ijt|s} = \lambda_s * GL(ijt \mid \nu), \tag{A5}$$

where $GL(ijt \mid \nu)$ is the G&L (1983)-type loyalty variable that the ith individual has for the jth SKU at the tth time period, and ν is the smoothing constant. One can view our model as a highly parsimonious factor-analytic solution applied to SKU choice behavior, in which the choice map captures drivers that are common to all attribute levels within an attribute. The idiosyncratic component of preference for an attribute level captured by G&L (1983)-type loyalty variable (μ_{sn}), while the unique component of the SKU is also captured by a G&L (1983)-type loyalty variable (λ_s). An important advantage of separating out the total preference for an SKU into these components (controlling for marketing mix variables) is that it allows the total sales volume for the SKU to be decomposed into volume that is transferable to other SKUs in the category and nontransferable volume, defined as the volume that would be lost from the category if the SKU were removed from distribution. The nontransferable component of volume is a measure of the incrementality of an SKU to the category. Clearly, the greater the intrinsic preference for an SKU (measured by the parameter α_{sjt}), the greater the nontransferable volume and incrementality of the SKU to the category.

An implicit assumption of our model is that the incremental volume contribution of the SKU to the category is largely dependent on the unique attributes that the SKU possesses.[4] Therefore, a methodology was developed in conjunction with the model that provided an estimate of the volume contribution of the different variables, such as common and unique attributes and marketing mix variables. None of the earlier models of SKU choice are able to provide this insight, so this represents a key contribution of the AD model.

Scanner Store Data Correction. An important element of our methodology is the application of the choice probabilities to the store scanner data. For instance the impact of deleting the rth SKU on the market share of the jth SKU is estimated by first calculating the aggregate market share of the jth SKU using Equation A1. This is given as follows:

$$MS_j^P = \sum_i \frac{1}{T_i} \left(\sum_t \sum_s (p(d_{ijt} \mid s) * \varphi_{is}) \right) * o_i, \tag{A6}$$

where MS_j^P is the aggregate level market share for the jth SKU calculated using panel data, and o_i is a weighting factor used for correcting under- or overrepresentation of the demographic segments; φ_{is} is the posterior probability of the ith individual belonging to the sth segment; and T_i are the total number of observations for the ith individual.

As this market share is calculated by summing probability of choice over individuals, time periods, and segments, as per Equation A1, this aggregate market share estimate does not suffer from IIA. Therefore, this is an important advantage of the panel-based approach. Due to the sample limitations of scanner panel data, this panel-based market share value is corrected for by multiplying this estimate to the actual volume of the rth SKU obtained from scanner store data. Therefore, the volume gained by the jth SKU is as follows:

$$\Delta V_j^{ST} = MS_j^P * V_r^{ST}, \tag{A7}$$

where ΔV_j^{ST} is the store level data estimate for the gain in the volume of the jth SKU, and V_r^{ST} is the actual store level data value for volume of the rth SKU.

We note that a similar analysis can be conducted for addition of an SKU.

References

Andrews, Rick L., Ajay K. Manrai. 1999. MDS maps for product attributes and market response: an application to scanner panel data. *Marketing Sci.* **18**(4) 584–604.

Brownstone, David, Kenneth Train. 1999. Forecasting new product penetration with flexible substitution patterns. *J. Econometrics* **89**(1/2) 109–123.

Bucklin, Randolph E., Sunil Gupta. 1999. Commercial use of UPC scanner data: industry and academic perspectives. *Marketing Sci.* **18**(3) 247–273.

[4] An alternative model formulation is to use a two-stage purchase incidence/SKU choice model rather than a single-stage SKU choice model such as that used in this paper. While the two-stage model would provide a more theoretically sound method for understanding the impact of an SKU on the category volume, the weekly data requirements of this model precluded this course.

Cooper, Lee, Masao Nakanishi. 1988. *Market Share Analysis*. Kluwer Academic Publishers, Boston, MA.

Elrod, Terry, Michael P. Keane. 1995. A factor-analytic model for representing the market structure in panel data. *J. Marketing Res.* **32**(February) 1–16.

Fader, Peter S., Bruce G. S. Hardie. 1996. Modeling consumer choice among SKUs. *J. Marketing Res.* **33**(November) 442–452.

Guadagni, Peter, John D. C. Little. 1983. A logit model of brand choice. *Marketing Sci.* **2**(1) 203–238.

Hardie, Bruce, Leonard Lodish, Peter Fader, Alistair Sutcliffe, William Kirk. 2004. Attribute based market share models: methodological development and managerial applications. Working paper 98-009, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Inman, J. Jeffrey. 2001. The role of sensory-specific satiety in attribute-level variety seeking. *J. Consumer Res.* **28**(June) 105–120.

McKinsey & Company. 2004. Consumer packaged goods, executive insights. http://www.mckinsey.com/clientservice/consumerpackagedgoods/insight.asp.

Roberts, John H., Pamela D. Morrison, Charles J. Nelson. 2004. Implementing a pre-launch diffusion model: measurement and management challenges of the Telstra switching study. *Marketing Sci.* **23**(2) 186–191.

Roberts, John H., Charles J. Nelson, Pamela D. Morrison. 2005. A pre-launch diffusion model for evaluation of market defense strategies. *Marketing Sci.* **24**(1) 150–164.

Rosenbleeth, Mitch, Chris Dallas-Feeney, Stephen S. Simmerman, Tom Casey. 2002. Capturing value through customer strategy. Booz Allen Hamilton, www.boozallen.com.

Sinha, Ashish, Joonwook Park, J. Jeffrey Inman, Yantao Wang. 2005. A factor analytic choice map approach to modeling attribute-level varied behavior among SKUs. Working paper, Victoria University of Wellington, Wellington, NZ.

Sinha, Ashish, Anna Sahgal. 2002. A general unfolding model of consumer choice among SKUs. *Australian and New Zealand Marketing Acad. Conf. (ANZMAC)*, Adelaide, Australia.

Swait, Joffre, Tulin Erdem. 2002. The effects of temporal consistency of sales promotion and availability on consumer choice behavior. *J. Marketing Res.* **39**(August) 304–320.

van Heerde, Harald J., Peter S. H. Leeflang, Dick R. Wittink. 2004. Decomposing the sales promotion bump with store data. *Marketing Sci.* **23**(3) 317–334.

Practice Prize Report

Modeling the Microeffects of Television Advertising: Which Ad Works, When, Where, for How Long, and Why?

Gerard J. Tellis
Marshall School of Business, University of Southern California, Los Angeles, California 90089, tellis@usc.edu

Rajesh K. Chandy
Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455, rchandy@csom.umn.edu

Deborah MacInnis
Marshall School of Business, University of Southern California, Los Angeles, California 90089, macinnis@usc.edu

Pattana Thaivanich
Hawaiikiki LLC, 10220 31st Avenue N, Plymouth, Minnesota 55441, pthaivanich@yahoo.com

Most past research has focused on how aggregate advertising works in field settings. However, the information most critical to managers is which ad works, in which medium or vehicle, at what time of the day, at what level of repetition, and for how long. Managers also need to know why a particular ad works in terms of the characteristics (or cues) of its creative. The proposed model addresses these issues. It provides a comprehensive method to evaluate the effect of TV advertising on sales by simultaneously separating the effects of the ad itself from that of the time, placement (channel), creative cues, repetition, age of the ad, and age of the market. It also captures ad decay by hour to avoid problems of data aggregation. No model in the literature provides such an in-depth and comprehensive analysis of advertising effectiveness. Applications of the model have saved millions of dollars in costs of media and design of creatives.

Key words: advertising response; wear-in; wear-out; carry-over effect; long-term effect; ad creative; ad cues
History: This paper was received August 19, 2004, and was with the authors 3 months for 2 revisions; processed by Gary Lilien.

Introduction

Firms spend billions of dollars each year on media advertising, yet much of this expenditure is made with limited testing of how or whether these expenditures will pay out in terms of sales and profits. Published studies on advertising's effectiveness have generally studied the aggregate effects of advertising on sales or market share. Most have focused on technical issues involved in efficiently capturing the

This report has benefited from comments from the Practice Prize editor and committee but has *not* been reviewed by *Marketing Science*.

unbiased effects of advertising using field data. The consensus from this research is that advertising does affect sales, though its elasticity is small and difficult to estimate (e.g., Sethuraman and Tellis 1993, Tellis 1988, Tellis and Weiss 1995).

One reason the elasticity is small could be that it reflects an average of many factors, including the medium, timing, repetition, age of the market, ad age (wear-in and wear-out), and ad creative cues (defined here as ad execution or content elements). Some of these factors might contribute to strong effects, while others might be weak or have no effect at all. Knowledge of the impact of these individual factors would greatly help managerial decision making. In particular, managers today need to know which particular ad works, in which medium or vehicle, at what time of the day for broadcast media, at what level of repetition, for how long, and in which market. Managers also need to know why a particular ad works and which aspects of its creative cues need to be changed to make it more effective. These questions are important because they reflect the way managers create and schedule advertisements, and hence they are critical to advertising creative and to media decisions. The need is all the more acute given increasing pressure on managers to show tangible performance outcomes from ad expenditures. No model or study has so far has addressed all or even most of these issues simultaneously.

The model described in this paper is designed to provide insight into all these issues simultaneously. The model was initially developed and applied in the context of a leading provider of toll-free medical referral services called Futuredontics (see Chandy et al. 2001, Tellis et al. 2000). It has since been successfully applied to a number of other contexts. The model and approach also have important research implications because they join two substantively relevant yet largely independent streams of academic research: consumer behavior and econometric modeling. The consumer behavior stream has focused on advertising creative content, testing with experiments that creative cues or what content elements of an ad achieve a better impact on consumer behavior. The econometric modeling has focused on estimating the effect of ad intensity (in terms of dollars, gross rating points (GRPs), or exposures) on sales.

The sections below outline the existing approaches to the problem in the application context and the field. We then describe our model with particular focus on its implementation, novel insights, and impact on marketing practice. Next, we highlight some illustrative outputs and analyses made possible by the model. We end with a discussion of the model's scope of application.

Existing Approaches to the Problem

In the section that follows, we provide a context for our model's impact, first describing the approach to creative, media decisions, and ad testing taken at Futuredontics, a medical referral service for which our model was initially developed. We further underscore its contribution by describing the two main approaches typically used in the advertising industry to test advertising effectiveness (split-cable single-source experiments and call tracking). We articulate the disadvantages of these approaches and illustrate how our approach avoids the disadvantages.

Approach at Futuredontics

Futuredontics operates a dentist referral service (1-800-DENTIST™) that advertises in more than 60 major markets in the United States. Their multimillion-dollar advertising budget includes more than 5,000 TV ad exposures per month. The company also runs some radio and billboard ads. Currently, Futuredontics receives more than three million calls per year to its toll-free referral service. Callers are connected to a call center in Santa Monica, California, that employs more than 80 operators. The operators collect information on the preferences of each caller and seek to match these preferences to the profiles of dentists in the company's member database. In the event of a match, the caller is connected to the office of the matching dentist. Such a connection is called a referral. The referral is free to the caller. Dentists listed in the database pay a monthly membership fee.

Our involvement with Futuredontics began in 1996. At that time, and consistent with traditional advertising practices, managers at Futuredontics used a combination of GRPs, Nielsen ratings, promotional offers from media outlets, and their own experience, intuition, and informal analyses (described later) to determine which ads to run, when, where, and how often. To minimize costs while increasing ad exposures, Futuredontics often sought "floating" spot ads, whereby TV vehicles would provide price discounts to the firm, provided they retained some discretion on exactly when the company's ads would run.

This simple approach at Futuredontics suffered from at least four problems also prevalent in the two more-sophisticated state-of-the-art approaches used in the advertising field and described below.

Approach of Split-Cable Single-Source Experiments

The split-cable single-source approach is used by IRI and Nielsen and described in a series of articles by Lodish and his colleagues (e.g., Lodish et al. 1995a, b). The approach consists of the following steps:

1. Split a panel of cable subscribers in a city so that they receive different ads.

2. Collect cumulative ad exposure and sales from households in these separate conditions.

3. Check differences in sales responses by level or type of advertising.

Notably, this approach suffers from one or more of the following limitations, which our approach avoids:

• *There is no mechanism to evaluate the creative content of the ad*. The system uses only a small set of ads. There is neither a scale nor a method by which specific elements of an ad (e.g., celebrities, arguments, music) can be included to assess which element(s) best drive sales. Moreover, the system does not evaluate how frequently one can use an ad before wear-out sets in.

• *There is no mechanism to evaluate the role of media and time of day*. The split-cable approach does not separate the effects of TV media channels or times of the day from that of the creative characteristics or repetition of the ad. For example, how much of the effect of advertising on sales is due to the program (e.g., *Friends*), its prime-time placement, creative aspects of the ad, or ad frequency? The ability to tell exactly what drives sales can help in the appropriate design, choice, placement, and scheduling of ads.

• *It is difficult to assess the carryover effects of advertising*. It is well known that advertising has carryover effects. However, it is also well known that such effects are biased upward with the use of aggregate data. Because split-cable tests involve aggregate data, the effects of advertising are potentially biased upward.

• *Analysis is limited to a few ads and test sites*. Only a few ads are tested in a few cities. The analysis is rarely conducted over the whole library of ads, population of cities, respondents, and regions.

Note that these limitations relate to the specific methods and models used to design and analyze split-cable experiments. They are not intrinsic to single-source data, so our model could also be used by firms such as IRI or Nielsen.

Approach of Call Tracking

Two common approaches to call tracking are used to evaluate responses to advertising in the toll-free advertising market:

1. Track calls during and immediately after the airing of each ad. Prior to our involvement, Futuredontics was informally using this approach.

2. Track calls from multiple 800 numbers—a different number for each vehicle—to identify which vehicle yields more calls. Many infomercials use this approach; they forgo the use of mnemonic numbers in order to be able to track calls more closely.

However, these approaches suffer from one or more of the following problems, which our model avoids:

• *There is no mechanism to evaluate the role of the creative characteristics or cues of the ad*. There is neither a scale nor a method by which various characteristics of an ad can be included in the model. As such, the creative executional elements that drive sales are unknown.

• *It is difficult to build brand equity*. A single number (especially a mnemonic number) is far superior to scattered generic numbers in its ability to enhance the future effects of advertising and to increase familiarity with the number and service.

• *It is hard to isolate all drivers of sales volume*. For example, the approach does not allow one to assess how much of the advertising effect is due to the vehicle (e.g., NBC) or to the ad itself. The ability to tell exactly what drives call volume can help spot problems quickly. Perhaps the problem lies in a weak creative, or the problem could lie in the vehicle used. Perhaps a new media plan is needed to draw new groups of consumers.

• *It prevents an accurate assessment of carryover effects*. This problem is similar to the use of simple models and aggregate data.

Again, these problems are due to the method or model of call tracking in use. They are not intrinsic to the form of the data per se.

Our Approach

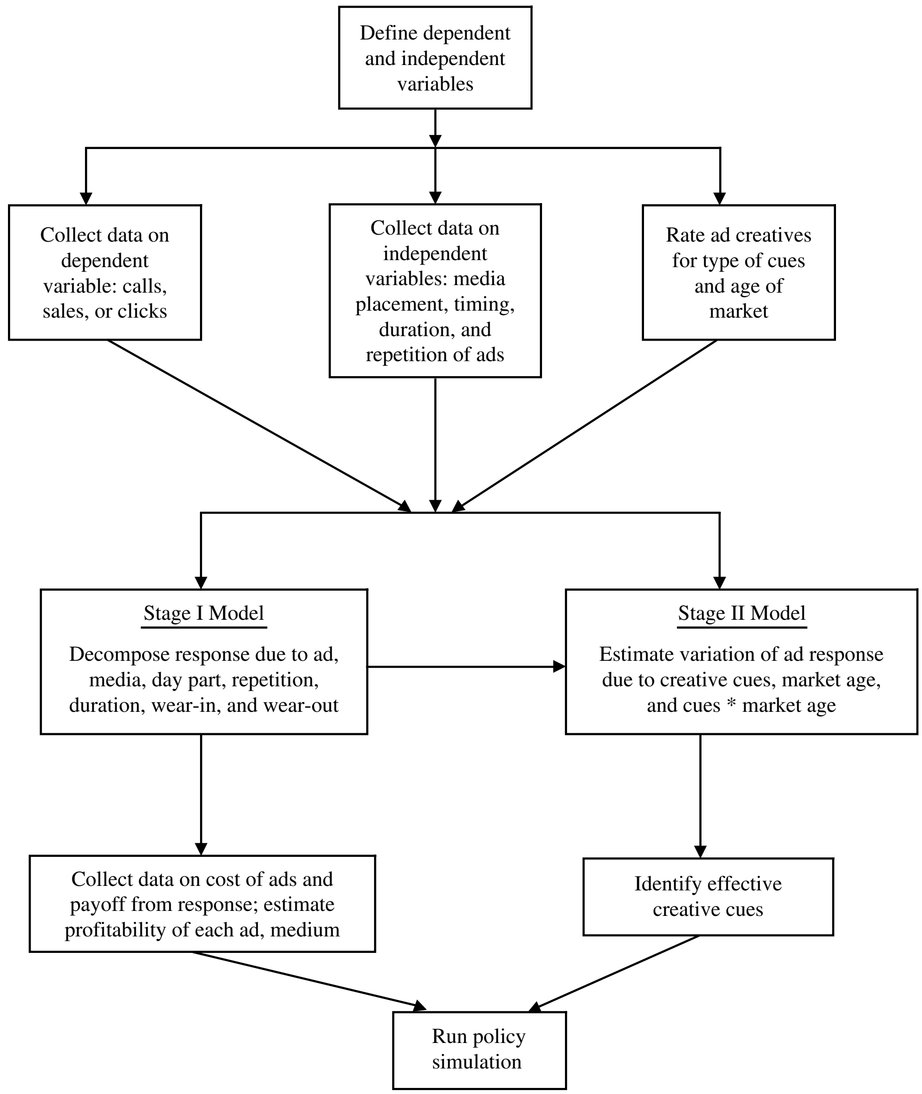
Our approach addresses each of the problems that limit the above approaches. It allows an advertiser to build brand equity around a single distinctive name. At the same time, it allows for a systematic, comprehensive, and reliable assessment of which ad works, when, where, and how often, how the advertising effect decays over time, and how repeated use of an ad wears out over time. Importantly, it can tell which creative elements in an ad work most effectively so the advertiser can develop better ads in the future. Our method is unique in this regard.

We describe our approach in four sections: Data Preparation, Model, Analysis 1, and Analysis 2. Figure 1 provides a schematic overview of the approach.

Data Preparation

Our initial task at Futuredontics involved compiling data in a manner suitable for statistical analysis. To compile the history of calls and referrals made through the toll-free number, it was necessary to first obtain the telephone logs for the number and identify the time of each call and the area codes associated with each call. The information from the telephone logs was then used to further identify the hour in which each call was made and the designated market area (DMA) from which it originated. This process resulted in a count of the number of calls and referrals made in each hour for each market.

Figure 1 **Flow Diagram for Micromodeling of Ad Response**



The data on the level of usage of each creative, vehicle, and media type were available from the billing invoices from each vehicle. The invoices, stored by the company in paper form, specify which ads aired on a particular vehicle during the period covered by the invoice, the time at which the ads were aired, and the cost associated with airing the ad. To compile the history of ad placements in each market, it was necessary to convert this information to electronic form. It was also necessary to match each vehicle with its DMA, so that ads on the vehicle could be linked to the referrals received from that DMA.

The above steps yielded a database that contained the number of referrals received and the specifics of the media execution of the company for each market in each hour over a relatively long period of time. This database permitted the modeling of a number of important issues on advertising effectiveness.

Model

The model uses a two-stage hierarchical design (Chandy et al. 2001). The first stage estimates the effectiveness of various ads across markets.

$$R = \alpha + \big(\mathbf{R}_{-l}\lambda + \mathbf{A}\beta_A + \mathbf{A}_M\beta_M + \mathbf{S}\beta_S + \mathbf{SH}\beta_{SH} + \mathbf{HD}\beta_{HD} + \mathbf{C}\beta_c + \big)O + \varepsilon_t, \quad (8)$$

where

$R =$ a vector of referrals by hour, $\mathbf{R}_{-l} =$ a matrix of lagged referrals by hour, $\mathbf{A} =$ a matrix of current and lagged ads by hour, $\mathbf{A}_M =$ a matrix of current and lagged morning ads by hour, $\mathbf{S} =$ a matrix of current and lagged ads in each TV station by hour, $\mathbf{H} =$ a matrix of dummy variables for time of day by hour, $\mathbf{D} =$ a matrix of dummy variables for day of week by hour, $O =$ a vector of dummies recording whether the service is open by hour, $\mathbf{C} =$ a matrix of dummy variables indicating whether or not an ad is used in each hour, $\alpha =$ constant term to be estimated, $\lambda =$ a vector of coefficients to be estimated for lagged referrals, $\beta_i =$ vectors of coefficients to be estimated, and $\varepsilon_t =$ a vector of error terms, initially assumed to be I. I. D. normal.

The second-stage analyzes the effectiveness of the ads as a function of the measured ad creative characteristics (emotion, argument, etc.).

$$\begin{aligned}\beta_{c,m} = {} & \varphi_1 \text{Argument}_c + \varphi_2(\text{Argument}_c \times \text{Age}_m)\\ & + \varphi_3 \text{Emotion}_c + \varphi_4(\text{Emotion}_c \times \text{Age}_m)\\ & + \varphi_5 800\,\text{Visible}_c + \varphi_6(800\,\text{Visible}_c \times \text{Age}_m)\\ & + \varphi_7 \text{Negative}_c + \varphi_8(\text{Negative}_c \times \text{Age}_m)\\ & + \varphi_9 \text{Positive}_c + \varphi_{10}(\text{Positive}_c \times \text{Age}_m)\\ & + \varphi_{11} \text{Expert}_c + \varphi_{12}(\text{Expert}_c \times \text{Age}_m)\\ & + \varphi_{13} \text{NonExpert}_c + \varphi_{14}(\text{NonExpert}_c \times \text{Age}_m)\\ & + \varphi_{15} \text{Age}_m + \varphi_{16}(\text{Age}_m)^2 + \mathbf{\Gamma}\,\mathbf{Market} + v, \quad (9)\end{aligned}$$

where

$\beta_{c,m} =$ coefficients of creative c in market m from Equation (1),
Age = market age (number of weeks since the inception of service in the market),
Market = matrix of market dummies,
$\mathbf{\Gamma} =$ vector of market coefficients,
$v =$ vector of errors,

and other variables are as defined in Equation (4).

Equations (1) and (2) are estimated sequentially. The estimated coefficients are unbiased because the measurement error in the dependent variable in Equation (2) can be absorbed in the disturbance term of the equation and ignored (Greene 2003, p. 84). The next section summarizes two analyses that address a number of research questions of interest to advertisers (see Tellis et al. 2000, Chandy et al. 2001). We focus here on key results and impact on practice.

Analysis 1: Which Ads Work, When, Where, and for How Long?

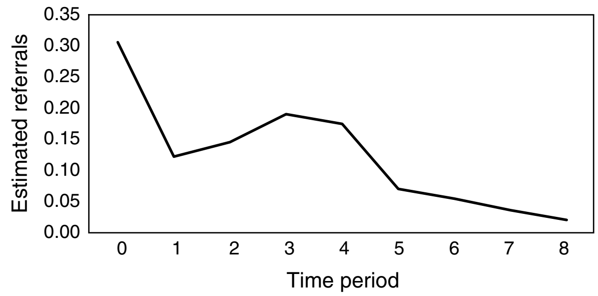
This analysis was based on the first stage of the two-stage hierarchical approach (Model 1). Its goals, results, and impact are as follows.

Goals. The goals of study 1 were to determine the following specifics of advertising effectiveness:

- Which particular ads were effective?
- When or at what times were they effective?
- Where or on which vehicle were they effective?
- For how long were they effective or how long was the time for wear-in and wear-out?

The entire analysis was at a highly disaggregate (hourly) level. We used a distributed lag model with a highly flexible carryover effect that incorporated nonlinear decay over hour of the day (Tellis et al. 2000).

Figure 2 Nonmonotonic Ad Carryover in Sacramento

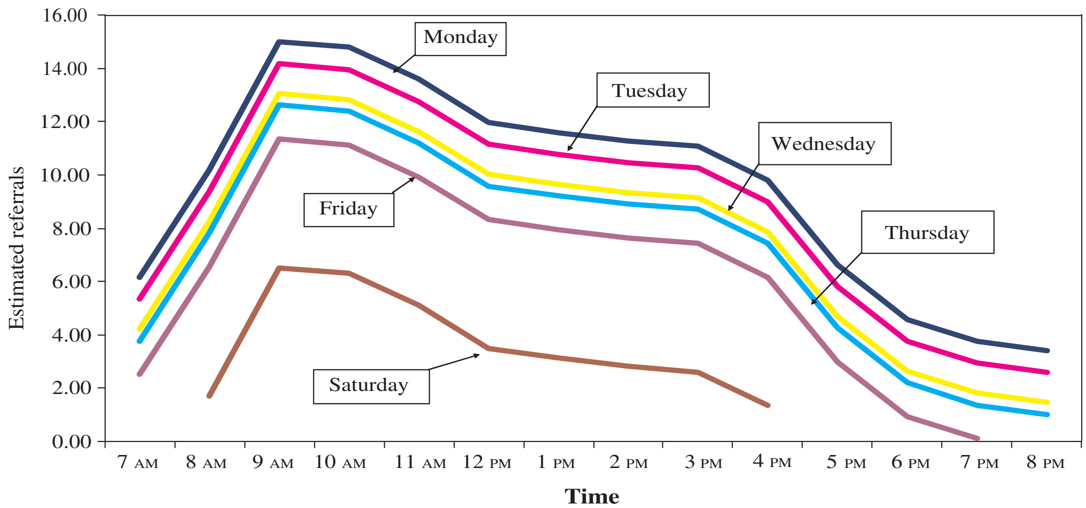


Note. Advertising carryover decays rapidly and mostly dissipates within eight hours.

Results. The estimation of the first-stage model led to the following key results. First, advertising carryover decays rapidly and mostly dissipates within eight hours (Figure 2). Additional analysis reveals that the peak of the carryover effect generally occurs in the current hour for daytime advertising but in subsequent hours for morning advertising. Thus, daytime advertising decay generally follows an exponential pattern, whereas morning advertising decay follows an inverted U-shaped pattern. Second, the effect of advertising is over and above a baseline referral, which follows distinct patterns by time of day and day of week (Figure 3). Third, when effective, new creatives are effective immediately (that is, wear-in is immediate). However, creatives also begin to wear-out rapidly. The most rapid wear-out occurs in the first few weeks of exposure. Fourth, ads vary greatly in effectiveness and profitability, with many ads being unprofitable.

Impact on Practice. Results from the study had the following impact on practice at Futuredontics. The firm used the results for baseline sales and advertising carryover to assess call center staffing and to schedule call center operators in a manner that minimized wait times. The firm also adopted an advertising schedule that focused on heavier advertising in the early part of the week, a Sunday to Tuesday schedule or a Monday to Wednesday schedule, and completely dropped advertising on Thursdays, Fridays, and Saturdays. Michael Apstein, CEO of Futuredontics at the time this research was conducted, estimates that this shift provided a savings of 30%–35% of media expenditures. Based on decay pattern results, Apstein noted, "We realized that we could go completely off air for one week out of four months." This finding resulted in savings of more than $1 million on media each year. Results also indicated that advertising effects differ substantially by TV channel and ad creative. Many vehicles and creatives did not contribute to an increase in referrals. The firm used this

Figure 3 **Baseline Referrals in Chicago by Hour of Day**



Note. Baseline referrals follow distinct patterns by time of day and day of week. The peaks in the curves indicate the hours in which customers are most likely to respond to the service. Knowledge of this pattern helps in planning advertising, and in staffing operators at call centers during different hours of the day. When extended to nonservice contexts, it could also help inventory planning.

information to reduce weight on the least effective creatives and channels. Apstein indicated that prior to the model the company developed creatives by a "go by your gut approach" coupled with expensive ad testing. "Once the research was done and the creative elements within the ad could be clearly identified, we became much more effective in the production of new creatives," said Apstein. The result was significant savings on the creative budget. Results also indicated that even when advertising was effective, it might not have been profitable. The firm used this information to increase weight on the most profitable vehicles, ads, and times of day. Apstein noted, "We were able to develop creatives that really addressed the questions of consumers. Those particular creatives really drove a significant increase in call volume to the company very rapidly once they were put on the air."

Ideally, modelers of ad competition, ad response, and optimal strategy (e.g., Banerjee and Bandopadhyay 2003, Dukes and Gal-Or 2003, Pauwels 2004, Vakratsas et al. 2004) should incorporate these more detailed aspects of advertising response into their models.

Policy Simulation. To facilitate use of the results for future planning, we developed Excel-based programs to simulate alternate advertising schedules and evaluate the results in terms of costs and benefits. The simulation enables the analyst to conveniently describe the results in each market and to suggest managerially actionable alternatives.

The simulation also helps managers plan media outlays and placements based on results from the model. Insights from these analyses can help in designing an advertising strategy that maximizes revenues, while minimizing costs.

Analysis 2: Why Ads Work

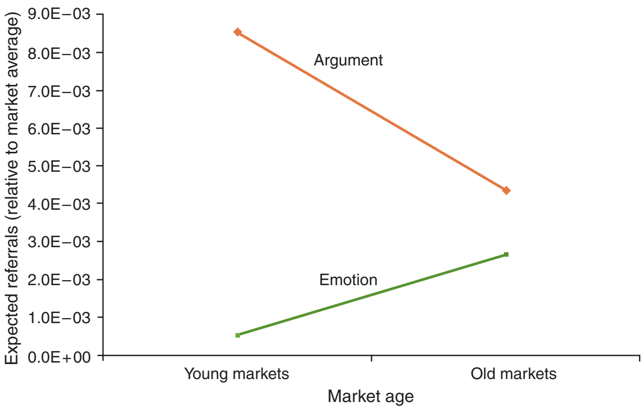
This analysis used Model 2. Its goal was to determine the creative characteristics that made specific ads effective. We describe the design, results, and impact of the study.

Design. Relying on theory in consumer information processing, we developed an instrument incorporating a rich set of creative cues, such as emotion, information, endorsers, etc. (see Chandy et al. 2001, MacInnis et al. 2002). Two trained assistants coded each ad's creative cues with this instrument. Over the years, Futuredontics developed a bank of 70 ads with varying creative content, which it scheduled without any particular pattern across 60 markets (cities) of varying ages. To determine how the effectiveness of ads varies by the age of markets and the ad's creative characteristics, we examined the interaction of age $\times$ creative cues.

Results. Ads differ substantially in effectiveness due primarily to variation in an ad's creative cues. Importantly, the effectiveness of creative cues is moderated by market age. Argument-based appeals, expert sources, and negatively framed messages are particularly effective in newer markets. In contrast, emotion-based appeals and positively framed messages are more effective in older markets. Figure 4 illustrates the results for arguments vs. emotions in young vs. old markets.

Impact on Practice. Futuredontics used these results to tailor ads that fit the age of specific markets, with different ad creative characteristics for markets of different ages. In addition, the firm designed entirely new ads with creative characteristics that

Figure 4 Effects of Argument vs. Emotion by Market Age



Note. Argument-based appeals are more effective in young markets. Emotion-based appeals are more effective in old markets. (Adapted from Chandy et al. 2001.)

better fit the age of each market. Prior to this analysis, a wild card creative had been produced quarterly just to test different creative cues. According to Michael Apstein, former CEO of Futuredontics, the elimination of the earlier approach resulted in a savings of approximately 25% of the creative budget. The firm scheduled current ads in age-relevant markets in which they would be more effective.

Cost Effectiveness Analysis

We assess cost effectiveness by conducting a cost-per-call analysis. We compare the cost of each element during the analysis period with the number of calls due to the same element during the same period.

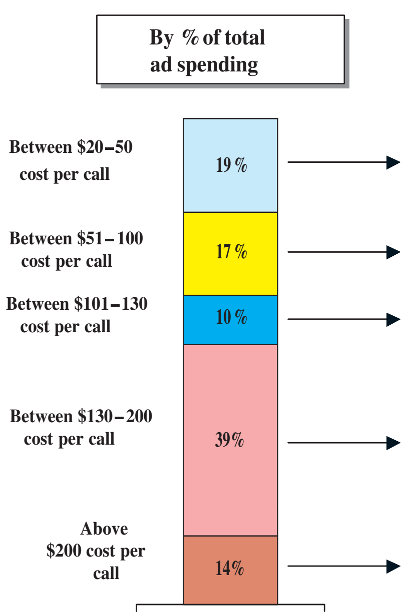
Figure 5 provides a comprehensive look at estimated costs per call and other relevant statistics for each vehicle by hour of day. (This table is illustrative and based on masked data.) The information from this analysis can be used to guide future media buying and scheduling. The rows with the lowest "estimated cost per call" are the most cost-effective. These day-parts might be worthy of increases in ad spending. On the other hand, the rows with the highest "estimated cost per call" are the least cost-effective media placements. These day-parts might be ripe for dropping, trimming, or aggressive price negotiations aimed at reducing costs.

Generalizing to Other Contexts

As a result of our model, Futuredontics was able to reduce advertising expenditures substantially while maintaining the number of calls received applied the model in all its major markets. The company has applied the model to another service category (1-800-BEAUTIFY™) with similar results. The model has also been applied successfully at 1-800-PLUMBER™, a toll-free plumber referral service. Apstein views the model as "tremendously helpful" to these contexts and as particularly critical to start-ups.

The model has also been applied by the Best Buy Corporation, the leading electronics retailer in the United States, to assess the impact of its Sunday newspaper insert advertising. Finally, the model is now the basis for an entrepreneurial start-up venture. This firm is currently in the process of further developing the model and applying it to other toll-free advertising contexts.

Figure 5 Detailed Analysis of Cost Effectiveness



Network	Time	Average cost per ad ($)	Estimated call per ad	Estimated cost per call ($)	Total usage during this period
NBC	7 AM–8 AM	106	5.0	21.2	89
WCOO	7PM–8 PM	92	3.0	30.7	95
NBC	1 PM–2 PM	172	5.0	34.4	101
NBC	9 AM–10 AM	73	1.5	48.7	85
KZYX	5 AM–6 AM	81	0.7	52.7	102
NBC	6 AM–7 AM	110	0.7	75.3	32
KXXX	7 AM–8 AM	150	0.5	81.2	50
ABC	9 AM–10 AM	370	0.3	100.5	50
ABC	10 AM–11 AM	263	0.4	110.2	41
FOX	11 AM–12 AM	233	0.5	120.5	15
NBC	12 PM –1 AM	350	0.4	128.5	22
FOX	1 PM–2 PM	547	4.2	130.3	31
NBC	2 PM–3 PM	438	3.2	137.0	31
FOX	3 PM–4 PM	183	1.3	140.5	15
ABC	3 PM–4 PM	213	1.5	141.7	20
CBS	6 PM–7 PM	580	3.2	181.2	15
CBS	11 PM–12 AM	287	1.5	191.5	20
FOX	6 PM–7 PM	1,606	8.2	195.9	20
CBS	4 AM–5 AM	118	1.7	201.2	20
NBC	6 AM–7 PM	350	0.6	222.3	11
CBS	7 PM–8 PM	600	0.4	230.5	21
CBS	4 PM–5 PM	233	1.1	250.9	51

Note. This chart provides a comprehensive look at cost-effectiveness by vehicles and day-parts.

A question we sometimes face is whether this model is applicable to other classes of products (e.g., packaged goods). Some readers point out that responses in our data are by phone, whereas consumers do not buy packaged goods by phone. We believe that our model is indeed applicable. The key issue is to track response by highly disaggregate time intervals, such as hour or minute. Whether that response is by phone, Internet, or checkout counter is immaterial. Thus, when porting our model to other contexts, all the analyst has to do is to replace the dependent variable, Referrals in the current context with order, sales, or clicks, as the case might be. The analyst will need to enhance the data with information on related characteristics (independent variables) describing their advertising, such as the time-of-day, the channel of placement, the duration of usage, a rating of the content of the ad, or age of market. Information on these latter elements is additive and separable. Thus, if information is available on some independent variables and not the others, the model can then be used for those variables on which required data should be available to advertisers.

Indeed, contrary to perception, scanner data and single-source data are even richer than the data we have used so far, because the former contain individual household information. Thus, the analyses can be done at the individual consumer level, allowing for variation in ad response by individuals. Such analyses will allow for more fine-grained segmentation.

In general, we believe the model has fairly wide applicability—all it requires is highly disaggregate data. In the current realm of abundant data collected via electronic media and stored on computers, the restriction is not data access but sound models. Thus our model is highly relevant to managers, can be readily operationalized with disaggregate data, and is generalizable across contexts.

Conclusion

We propose a comprehensive model to evaluate the effects of TV advertising on sales, which simultaneously separates the effects of the ad itself from that of the time, placement, length of usage, repetition, creative cues of the ad, and type of market in which it is shown. It also captures ad decay by hour to avoid problems of data aggregation. There is no other model in the literature that currently does such an in-depth and comprehensive analysis of advertising effectiveness.

References

Banerjee, Bibek, Subir Bandyopadhyay. 2003. Advertising competition under consumer inertia. *Marketing Sci.* **22**(1) 131–144.

Chandy, Rajesh, Gerard Tellis, Deborah MacInnis, Pattana Thaivanich. 2001. What to say when: Advertising execution in evolving markets. *J. Marketing Res.* **38**(November).

Dukes, Anthony, Esther Gal-Or. 2003. Negotiations and exclusivity contracts for advertising. *Marketing Sci.* **22**(2) 222–245.

Greene, William. 2003. *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ.

Lodish, Leonard M., Magid Abraham, Jeanne Livelsberger, Bruce Richardson, Mary Ellen Stevens. 1995a. A summary of fifty-five in-market experimental estimates of the long-term effect of TV advertising. *Marketing Sci.* **14**(3) 143–140.

Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson, Mary Ellen Stevens. 1995b. How TV advertising works: A meta-analysis of 389 real world split cable TV advertising experiments. *J. Marketing Res.* **32**(May) 125–139.

MacInnis, Deborah J., Ambar Rao, Allen Weiss. 2002. Assessing when increased media weight of real-world advertisements helps sales. *J. Marketing Res.* **39**(November) 391–407.

Pauwels, Koen. 2004. How dynamic consumer response, competitor response, company support, and company inertia shape long-term marketing effectiveness. *Marketing Sci.* **23**(4) 596–610.

Sethuraman, Raj, Gerard J. Tellis. 1991. An analysis of the tradeoff between advertising and pricing. *J. Marketing Res.* **31**(2) 160–174.

Tellis, Gerard J. 1988. Advertising exposure, loyalty, and brand purchase: A two-stage model of choice. *J. Marketing Res.* **15**(2) 134–144.

Tellis, Gerard J., Doyle Weiss. 1995. Does TV advertising really affect sales? *J. Advertising* **24**(3) 1–12.

Tellis, Gerard, Rajesh Chandy, Pattana Thaivanich. 2000. Which ad works, when, where and how often? Testing the effects of direct TV advertising. *J. Marketing Res.* **37**(February) 32–46.

Vakratsas, Demetrios, Fred M. Feinberg, Frank M. Bass, Gurumurthy Kalyanaram. 2004. The shape of advertising response functions revisited: A model of dynamic probabilistic thresholds. *Marketing Sci.* **23**(1) 109–119.