Commercial Use of UPC Scanner Data: Industry and Academic Perspectives

Randolph E. Bucklin • Sunil Gupta
Anderson School, University of California at Los Angeles, 110 Westwood Plaza, Los Angeles, California 90095, randy.bucklin@anderson.ucla.edu
Graduate School of Business, Columbia University, New York, New York 10027, sg37@columbia.edu

Abstract

The authors report the findings from an exploratory investigation of the use of UPC scanner data in the consumer packaged goods industry in the U.S. The study examines the practitioner community's view of the use of scanner data and compares these views with academic research. Forty-one executives from ten data suppliers, packaged goods manufacturers, and consulting firms participated in wide-ranging, in-person, interviews conducted by the authors. The interviews sought to uncover key questions practitioners would like to answer with scanner data, how scanner data is applied to these questions, and the industry's perspective regarding the success that the use of scanner data has had in each area.

The authors then compare and contrast practitioners' views regarding the resolution of each issue with academic research. This produces a 2 × 2 classification of each question as "resolved" or "unresolved" from the perspectives of industry and academia. Along the diagonal of the 2 × 2, issues viewed as unresolved by both groups are important topics for future research. Issues deemed resolved by both groups are, correspondingly, of lower priority. In the off-diagonal cells, industry and academics disagree. These topics should be given priority for discussion, information exchange, and possible further research.

Practitioners reported that scanner data analysis has had the most success and been most widely adopted for decision making in consumer promotions (i.e., coupons), trade promotions, and pricing. For example, logit and regression models applied to scanner data have revealed very low average consumer response to coupons which has directly led to reduced couponing activity. Managers also reported high levels of comfort with and impact from analyses of trade promotions and price elasticities. While industry views most of the issues in these areas to be resolved, academic research raises concerns about a number of practices in common commercial use. These include price threshold analysis and trade promotion evaluation using baseline and incremental sales.

In product strategy, advertising, and distribution management, practitioners reported that the use of scanner data has had more limited development, success, and impact. In the case of new product decisions, scanner data use has been slow to develop due to the inherent limitations of historical data for these decisions and a heavy reliance on traditional primary research methods. In advertising, scanner data is widely analyzed with models, but confusion among practitioners is very high due to controversies about methods (e.g., what level of data aggregation is best) and conflicting results. In distribution and retail management, scanner data use has tremendous potential but a mixed track record to date. Thus, practitioners view the use of scanner data as unresolved for most issues in product strategy, advertising, and distribution. This view is largely, though not entirely, consistent with academic research, which has only begun to address many of the key questions raised by practitioners.

In light of the large number of unresolved issues and mixed record of scanner data use to date, the authors offer a series of specific recommendations for immediate and long-term research priorities that are likely to have the greatest impact on commercial utilization of UPC scanner data. Topics of immediate priority include price thresholds and gaps, baseline and incremental sales, base price elasticity, competitive reactions, measurement of advertising effects, management of brand equity, rationalization of product assortments, and category management. Long-term priorities include a greater emphasis on profitability versus sales or market share, developing prescriptive models versus descriptive models, and the need for industry standards.
(*Scanner Data; Marketing Research; Marketing Models; Research Priorities*)

0732-2399/99/1803/0247/$05.00
1526-548X electronic ISSN

Marketing Science © 1999 INFORMS
Vol. 18, No. 3, 1999, pp. 247–273

1. Introduction

A major theme of academic research in marketing has been the development of substantive understanding and methodological tools that will help managers make better decisions. Assessments of the actual impact of new approaches on the practice of marketing have, however, been conducted only on an occasional basis (see, e.g., the studies on conjoint analysis by Cattin and Wittink 1982, Wittink and Cattin 1989, and Wittink et al. 1994). While published articles frequently offer managerial implications and often suggest ways in which new models and data analyses may be put to commercial use, it is unfortunate that the literature lacks extensive reporting on the perspectives of practitioners—even in areas in which there is significant ongoing academic activity.

Research on UPC scanner data has been actively pursued by marketing academics since the mid-1980s and many dozens of papers have been published in the major marketing journals since then. Clearly, the body of inquiry collectively known as scanner research has established itself as a major subfield within the marketing science discipline. Much of the intended contribution of the publications in this area has been to give practitioners better tools for understanding their markets (e.g., approaches to analyze consumer response and market segmentation) and for making marketing mix decisions (e.g., setting prices and determining promotion spending). Despite the earnest desire of many academics to contribute to the practical use of scanner data in marketing (see, e.g., Neslin et al. 1994), our understanding of the commercial use of scanner data—and the views practitioners hold regarding its usefulness for decision making—is limited. In this paper, we take a step towards closing this gap.

Our primary objective is to report on the commercial use of UPC scanner data and to provide a comparison of industry practice with academic research. In so doing, we seek to determine where practitioners and academics agree, where they disagree, and why that is so. We will focus our report on a series of key issues that practitioners identified as especially important to them with respect to the use of scanner data. For each of these issues (e.g., trade promotion evaluation), we report on industry practice then compare these practices with those recommended by academic researchers. The agreement (or disagreement) between practice and academia leads naturally to a 2 $\times$ 2 classification of each issue: "resolved" from the perspective of both industry and academia, "unresolved" in the eyes of both, or a disagreement in either direction.

Issues that both groups view as unresolved are important topics for future research because of the potential for significant impact among both academics and practitioners. Issues deemed resolved by both are, correspondingly, of lower priority. In the off-diagonal cells, industry and academics disagree. Issues seen as resolved by one side but unresolved by another may be good candidates for mutual discussion and information exchange (leading to possible further research). For example, practitioners may be unaware of problems identified by academic researchers or, if aware, have good reasons for either viewing the issue as resolved or perceiving the proposed methods as too cumbersome or impractical. Similarly, academics may have elevated certain specific problems to importance levels well beyond what is merited by their practical consequences.

Study Approach

Our approach is exploratory. The views of practitioners have been gathered based on in-depth interviews conducted, in person, with managers from a sample of 10 firms actively engaged in the use and/or analysis of UPC scanner data. Table 1 presents a listing of participating firms and the number of individuals within each firm that were interviewed.[1] We sought to cover

[1]We wish to thank the Marketing Science Institute for providing as-

Table 1 **Participating Companies**

Company	Location	Number of Participants
Information Resources, Inc.	Chicago, IL	7
Nielsen North America	Schaumburg, IL	5
Kraft General Foods, Inc.	Glenview, IL	6
Procter & Gamble, Inc.	Cincinnati, OH	6
The Quaker Oats Company	Chicago, IL	3
Nestle U.S.A.	Glendale, CA	3
PepsiCo	Purchase, NY	2
McKinsey and Company	New York, NY	2
Hudson River Group	Valhalla, NY	3
Media Marketing Assessment (MMA), Inc.	Westport, CT	4

the perspective of packaged goods manufacturers,[2] the two major syndicated data suppliers (Information Resources, Inc. (IRI) and ACNielsen), and third-party consultants. Consulting firms are important in the analysis of scanner data due, at least in part, to the recent downsizing of in-house marketing research capabilities at many consumer products firms. Both IRI and Nielsen also offer extensive consulting and custom-study capabilities and therefore should be considered as consultants as well as data suppliers. We limited the scope of our interviews to the perspectives of packaged goods manufacturers and the consulting firms and data suppliers that serve them directly as clients. Retailers were not included among the interview participants. Thus, the focus of our paper is on the commercial practices with respect to scanner data use for and by *manufacturers.*

We used in-depth personal interviews to generate an interactive discussion with each participant. Unlike the study of the commercial use of a single method (e.g., conjoint analysis), our study investigated the general pattern of use of an extended family of models, tools, and methods that have been applied to secondary data with common characteristics. Consequently, we expected that each interview would take a potentially different path, with participants providing substantial information regarding specific scanner applications with which they were familiar and little, if any, information about others. Since one of our objectives was to investigate areas of disagreement, we needed an approach that permitted detailed follow-up questions and probing. The interviews were conducted from fall 1995 to spring 1996. All interviews were conducted with both authors present (the single exception being the interview with Nestle USA). In the summer of 1997, all participating firms were invited to comment on a preliminary draft of the study report. The input from their responses was then incorporated. We complement the information from our interviews with results from a recent survey-based study by Davidson and Stacey (1997). This telephone survey, conducted in 1996, queried managers from 56 packaged goods firms about the relative importance that their firms attached to a series of 15 marketing issues and the extent to which scanner data was used to address those issues.

We organize our discussion following six areas of the marketing mix: pricing, trade promotion, coupons, advertising, product strategy, and distribution/retail management. For each mix element, we discuss the key issues raised by practitioners, how they use scanner data to address these issues, and how these approaches compare with academic research. This leads to the classification of each issue into one of the four cells in our 2×2 matrix.[3] In our conclusion, we summarize our findings, discuss the factors we believe underlie disagreement between academics and practitioners, and offer suggestions for both immediate and long-term research projects.

2. Pricing

Practitioners considered pricing decisions especially important and scanner data are very widely used in this domain of the marketing mix. The importance of pricing to managers is due to the significant and usually immediate impact of price changes on sales and profits, the ability to implement price changes quickly, and the potential for strong reactions from consumers, retailers, and competitors. Davidson and Stacey (1997) also found that practitioners put pricing at or near the top of the list of business issues. Pricing ranked third in overall importance among the 15 marketing issues in their survey and was cited as "extremely important" by 78% of respondents. The same respondents also rated pricing as the marketing issue for which scanner data analysis and modeling was used more than any other single topic. Sixty-two percent said scanner data modeling and analysis were used "heavily" to address pricing issues.

sistance in securing appropriate contacts at many of the participating organizations. Of the firms we contacted, only one declined to participate in the study.

[2]At each packaged goods company participating, we requested participation from individuals representing brand management, in-house market research and analysis, and field sales.

[3]The degree to which a given issue is seen as resolved by either academics or practitioners lies on a continuum. We adopt the 2×2 classification for expositional ease and it is not meant to imply that each issue appearing in a given cell is precisely the same in terms of its resolution.

Key Questions and Issues

In our interviews, practitioners identified two key pricing issues that they use scanner data to address (or expect to be able to use scanner data to address). These are (1) determining price elasticities, and (2) analyzing price thresholds and determining optimal price gaps with competing brands.

Price Elasticity. Managers are interested in knowing whether a given brand is relatively sensitive (elastic) or insensitive (inelastic) to changes in price. In addition to national-level elasticity figures, managers often expressed a desire to know if price sensitivity varies across regional markets or retail accounts and, if so, how. Practitioners also told us they believed that base or regular-price elasticities differ from elasticities associated with temporary price reductions (TPRs). They want to know both elasticities in order to make decisions about list prices versus price discounts.

Price Thresholds and Price Gaps. If a brand has a price elasticity of -2, a 10% change in its price should change its sales by 20%. But many managers believe that such effects are unlikely to occur unless the price change crosses a threshold. In other words, response to price changes can be "sticky" over certain ranges. Managers want to know where the threshold price points lie so they can fine-tune price setting (e.g., set the maximum possible price without triggering a noticeable loss in share or sales). Managers are also interested in knowing how sales or share will change as the "gap" (i.e., absolute difference in price) between their brand and competing brands widens or narrows. The purpose is to determine how much of a price premium, for example, Marlboro can charge over a private label before the price difference begins to cause a serious erosion in Marlboro's market share.

Price Elasticity

Industry Practices. A typical approach to address price elasticity questions is to run a time-series cross-section regression on store, account, or market-level scanner data. (We will cover the issue of aggregation below.) At IRI and Nielsen, for example, the data set used for the analysis typically includes approximately 100,000 observations covering about 2,000 stores over a one year time period. A typical regression takes the following form (e.g., Wittink et al. 1988):

$$\begin{aligned}\ln(\text{UNITSALES}) = {} & a_0 + b_1 \ln(\text{PRICE}) + b_2\, \text{FEAT} \\ & + b_3\, \text{DISP} + b_4\, \text{FEAT} \cdot \text{DISP} \\ & + b_5\, \text{TPR} + b_6\, \text{SPECPACK} \\ & + b_7\, \text{STORE} + b_8\, \text{WEEK} + \text{ERROR}. \end{aligned} \tag{1}$$

In this equation, the log of the unit sales for a brand or stock-keeping unit (SKU) is a function of the log of price and a series of other independent variables. Here, PRICE refers to the actual shelf price of the brand or SKU. (Note that the price coefficient can be directly interpreted as an elasticity because of the log-log formulation.) Promotion activity is captured by dummy variables (or indices) for feature (FEAT) and display (DISP) activity, as well as whether or not a temporary price reduction (TPR) is being offered.[4] Additional variables include an indicator for special packages (SPECPACK), store dummy variables to control for differences in sales volumes across stores (STORE), and weekly indicator variables (WEEK), to control for seasonality or other special events (e.g., holidays such as the Fourth of July or Thanksgiving).

Competitive effects (not shown in Equation (1)) can also enter the model in the form of price or promotional activity for a judgmentally selected set of brands or SKUs. When the regression analysis is done at the SKU level, which is often the case, the complexity of the modeling problem grows substantially. For example, including the price and promotion activity for all competing SKUs as independent variables would rapidly become unwieldy. A common compromise is to include major competitive brands (aggregated across their SKUs) and own-company SKUs "above" and "below" (e.g., one size up and one size down) the target SKU to control for potential cannibalization effects. Competitive promotions are sometimes also rolled into

[4] In this analysis, the price elasticity refers to the shelf price, whose short-run variation is primarily driven by temporary price reductions. In Equation (1), TPR is an indicator variable that controls for the effect of announcing a temporary price reduction via a "shelf-talker" next to the displayed item. We later take up the specific question of whether or not to separate the shelf-price variable into regular price and temporary price discounts.

more aggregate measures (e.g., FEAT, DISP, and TPR may be combined into one variable called promotion). The analysis usually assumes that price and promotion coefficients are the same across all stores, although the equation can be modified to account for known store-specific effects by including additional terms. Due to the robust properties of multiple linear regression and the very large number of observations used, the analysis is almost always conducted using ordinary least squares (OLS). IRI and Nielsen also said that they had virtually no problems with incorrectly signed estimates of own and cross-price elasticities.

In spite of reportedly high degrees of usage and comfort with price elasticity analysis, managers raised a number of issues about the practices in common use. These included: (1) the need for—and difficulties with—analysis and models to take into account the idiosyncrasies of demand and competition in a given category, (2) the appropriate level of analysis (i.e., store, account, market, national) for the elasticity estimation, and (3) problems in estimating base-price elasticity.

A widely held belief among practitioners is that models applied to scanner data must be specifically tailored for the particular industry and product category involved. Kraft executives told two cautionary tales to illustrate their concerns with the use of standardized regression models for pricing decisions:

Idiosyncratic Factors. A few years ago, Kraft experienced declining sales in salad dressings. A regression on historical data suggested that reducing price would significantly boost sales. Looking at the category dynamics in more depth, it turned out that the entire salad dressing category had declined over the previous year because flooding in California had dramatically reduced the supply of lettuce and raised its price. This cut lettuce consumption and led to a drop in demand for salad dressing. Further analysis revealed that there was also a significant time lag in this entire dynamic (i.e., between the flood, the lettuce supply and prices, and the consumption of salad dressing), making it easy to misinterpret regression results. The story highlights the need to include potential idiosyncratic factors in regression models, in this case a strong complementarity effect.

Competitive Reaction and Time Horizon. A category manager was told by market research that the price elasticity for one of his products was -2. The product, however, still had strong sales in spite of several recent price increases. On further investigation, it turned out that competition had also been raising prices, following a short lag. Kraft executives used the story to highlight the question of how to interpret the predictive power of price elasticity. For example, over what time frame should an estimate be considered valid and, moreover, is it useless once competition reacts?

In addition to the complexity of tailoring pricing models to incorporate idiosyncratic factors and competitive reaction, another concern raised by some practitioners (especially consultants) is that the typical analysis can become disconnected with the actual pricing process. Specifically, a regression that provides price elasticity estimates using data across all stores is an analysis that fits best with a brand manager making a pricing decision in a centralized fashion (e.g., a nationwide price discount of 10%). Except in occasional instances (e.g., an across-the-board price cut of 20% by Post cereal), pricing decisions also require attention and management at the account level. This, in turn, leads to a demand for price sensitivity estimates for each account rather than a single estimate that is the same for all. If models are estimated at the account level, however, brand price effects may be overstated if shifts in store patronage are not held constant. A price reduction implemented at one chain but not at others in the same market may attract sales for that brand from competing stores. (In the data, the expected change in brand volume would be augmented by borrowed sales from other stores.) This can lead to larger estimates of price elasticity at the account-level than would be the case had all accounts implemented the price reduction at the same time. As some managers explained, they are therefore unsure whether the resulting elasticities pertain more to the store or to the brand.

When estimating elasticity, many practitioners estimate a single term (using shelf price as in Equation (1)), but others believe that base and TPR price elasticities should be estimated separately because (a) they are different, and (b) they represent two separate pricing decisions. While estimating elasticities for shelf prices or TPRs did not provoke much debate, estimation of base-price elasticity was controversial. Some companies (e.g., IRI) feel comfortable using cross-sectional data (under a no-promotion condition) from several different markets to estimate base-price elasticity. Others question this approach because the model assumes the same base-price elasticity to hold for all markets when, in fact, the market-level base-price elasticities may be very different. Still others (e.g., consultants at McKinsey) believe that there may be insufficient natural variation in base prices to accurately gauge its elasticity with scanner data. Instead of scanner data, they advocate using surveys or choice experiments to determine base-price elasticities.

Academic Perspectives. With respect to the basic determination of a brand's own shelf-price elasticity at the national level, regression analysis on large samples of *store-level data* is likely to be sufficiently robust so that estimates of these elasticities obtained in this manner are likely to possess good properties. A key caveat, however, is the problem of aggregation. As Christen et al. (1997) have shown, aggregating over store-level data to the market level can lead to biases in parameter estimates when the model is nonlinear, as in the log-log form of Equation (1). The same problem can arise if data are aggregated over SKUs (i.e., to form composites or brand-level items) when the movement of price and other marketing variables differs across them. Given this caveat, we consider that the estimation of shelf-price elasticities is largely resolved from the perspective of both academics and practitioners and, accordingly, list it in the upper-left quadrant of Figure 1.

On the other hand, the determination of elasticities *at the account-level* and the estimation of cross-price elasticities are not likely to be handled well by current commercial practice. Recent academic research on price elasticity estimation has emphasized two major themes: (1) there are strong benefits to using prior information (e.g., Bayesian methods) to improve elasticity estimates, especially in cases where sample information is limited (e.g., at the individual account or store level), and (2) price elasticities, particularly in packaged goods, have many empirical generalizations.

Figure 1 **Industry and Academic Perspectives on Commercial Use of Scanner Data**

Industry Perspective	Academic Perspective: Resolved	Academic Perspective: Unresolved
Resolved	• Own-Price Elasticity (TPR) • Coupons • Market Structure Analysis (Shelfspace Management)	• Price Thresholds & Gaps • Baseline Sales (Trade Promotions)
Unresolved	• Account-Specific Elasticity • Cross-Price Elasticity • Advertising Effect Sizes • Micromarketing	• Base-Price Elasticity • Competitive Reaction • Advertising Measurement • Brand Equity • Product Assortment • Category Management

Published academic studies on store-level scanner data that have used classical regression approaches have had a mixed record of producing elasticity estimates with correct signs and stable properties, especially for cross-effects. While Foekens et al. (1994) reported no incorrect signs for either own- or cross-price elasticities, problems have been reported in a number of other cases (e.g., Cooper 1988, Carpenter et al. 1988, Blattberg and Wisniewski 1989). Recognizing the econometric difficulties (e.g., multicollinearity, limited number of observations) often associated with obtaining good estimates for own- and cross-price elasticities for a single account or regional market, researchers have developed some promising new approaches based on Bayesian methods. Working in this type of data setting, researchers have refined methods to impose prior structure or "shrinkage" on the estimates in order to obtain own- and cross-price elasticities with good properties (e.g., Allenby 1990, Blattberg and George 1991, Montgomery 1997). Academic researchers have emphasized the study of regional or account-level elasticity estimation (versus practitioner emphasis on national cross-sections) in part because the data sets available for academic research typically covered only one, or perhaps two, regional markets.

Academic studies also have revealed many regularities in brand price elasticities for packaged goods.

Such "empirical generalizations" can be used as a supplement to regression-based estimates or to provide priors for shrinkage methods. For example, a meta-analysis by Tellis (1988) found price elasticities to average -2.5. Ehrenberg and England (1990) reported that price elasticities for each of four brands across five categories were equal to -2.6 and that this effect size held regardless of whether price changes were large or small. Other researchers have found that price elasticities are related to brand characteristics (e.g., Bolton 1989, Narasimhan et al. 1996, Bell et al. 1999). For example, Bell et al. (1999) reported that 81% of the variance in brand price elasticities can be explained by readily identifiable category, brand, and consumer factors, with category factors (e.g., necessity item, storability) accounting for most of the differences in price elasticities.

With respect to cross-price elasticity estimation, we note that (1) the regressions used in practice do not include estimates for all cross effects, and (2) cross-effect terms are often created by aggregating over both promotional variables as well as SKUs. Moreover, we note that the potential problems are likely to worsen if conventional regression is used to estimate cross-effects at the account-level. With the availability of new estimation approaches, many of these difficulties can be avoided. Latent class analysis (e.g., Kamakura and Russell 1989) can be used to estimate elasticities for clusters of stores or accounts that share common levels of market response parameters. Empirical Bayes methods (e.g., Montgomery 1997) use sample-wide information to "shrink" estimates of account-specific elasticities and cross-elasticities, greatly reducing the occurrence of incorrectly signed elasticies or nonsensical estimates. Bucklin et al. (1998) have shown how cross-price elasticities correspond, via a scaling factor, with brand-switching probabilities. This might be used to take advantage of managers historical comfort with brand-switching data to produce better estimates for own- and cross-price elasticities. Thus, in the estimation of account-specific elasticities and cross-price elasticities, academic researchers can offer a number of new methodologies and useful findings that should be helpful in industry practice. We therefore classify these issues in the lower left quadrant of Figure 1.

Academic research has only just begun to address the specific concerns practitioners raised about competition, confounding store and brand elasticities, and base-price elasticity. Empirical reaction functions (e.g., Bresnahan 1997, Leeflang and Wittink 1996, Jedidi et al. 1999) can be used to assess whether or not predictions of price response will be seriously invalidated by subsequent competitive effects. The empirical reaction functions can also be used to assess the nature and extent of competitive reaction in the marketplace. For example, Leeflang and Wittink (1996) report that competitors generally overreact to price changes. Game-theoretic work also has shed substantial light on the forms of competitive pricing behavior that are commonly found in consumer packaged goods categories and their implications (e.g., Rao et al. 1995, Lal and Padmanabhan 1995). Recently, researchers have used the industrial organization paradigm to estimate demand and reaction functions simultaneously (Kadiyali et al. 1999).

Concerns that brand-price elasticities at the account-level may be inflated due to store switching effects might be addressed with models estimated on panel data. This is because panel data contains information on the specific store visits of each panelist. Thus, if so desired, it is possible to hold the effect of store switching constant (see, e.g., Bucklin and Lattin 1992).

The academic literature reports conflicting results regarding whether or not base-price elasticities differ from separately estimated promotional-price elasticities. For example, Guadagni and Little (1983) report them to be the same while Blattberg and Neslin (1990) show large differences. Part of the confusion stems from the difficulty in completely separating the attention-getting aspect of a temporary price reduction (e.g., the effect of the shelf talker) from the price change itself (e.g., Inman et al. 1990). Intuitively, for price reductions where shoppers are cued that the change is temporary, elasticities should exceed base-price elasticities when the discounts induce consumers to stockpile and switch brands. We note that base-price elasticity estimates from store-level data also might be improved by the use of Bayesian methods, but researchers have not yet explicitly addressed this problem. Due to the nascent academic attention to these issues or to conflicts in the literature, we classify them as unresolved from both perspectives (see Figure 1).

Price Thresholds and Gaps

To investigate price thresholds and price gaps with competing brands, industry practice is to use a simple and intuitive procedure known as *sales velocity analysis.* This procedure essentially creates a cross-tabulation of sales versus price points across stores. It consists of the following steps:

1. For all stores of a common format (e.g., drug stores), collect information on total store sales, sales of the target brand, and its nonpromoted price. (The analysis can be repeated for promoted prices and for other store formats.)
2. Aggregate store and brand sales data across all stores that charge the same nonpromoted price for the target brand.
3. Create a sales rate or sales velocity measure as unit sales of the brand per million dollar store sales. This accounts for differences in brand sales due to differences in overall store sales.
4. Plot the cross-sectional data on sales rate and nonpromoted prices. Visually inspect the plot and make an inference about price thresholds.

A similar analysis also can be conducted for price gaps. In that case, sales rates are plotted against the various price gaps between the target brand and a given competing brand found in the data. The objective in both cases is to use scanner data as a "natural experiment" to help find the best specific price points for a brand.

At PepsiCo, executives recounted the story of how this analysis helped them. A number of years ago, both Coca-Cola and Pepsi were losing market share to private label brands. Management tried different tactics to halt the share loss, but none succeeded. Finally, consultants suggested a price-gap analysis on scanner data. The study revealed price points that were effective at holding share against private labels. These were implemented nationwide and the share erosion stopped.

Practitioners told us that this type of analysis for studying thresholds and gaps is widely used because of its simplicity, ease of use and understanding, and the ability to conduct many such studies in a short period of time. Though sales velocity analysis is appealing to many, executives at Nestle were more cautious. They noted that they found the method useful for diagnostic purposes but believed that, ultimately, price analysis needs to be based on within-store changes that take place over time.

Academic Perspectives. While most practitioners expressed high levels of comfort with the use of sales velocity analysis to determine thresholds and gaps, determining price thresholds (and gaps) is actually a difficult conceptual and statistical problem. We therefore believe that current industry practice has some major limitations. These limitations include the use of cross-sectional data for causal inference (e.g., is it reasonable to combine East Coast data with West Coast data?) and the lack of statistical analysis (e.g., is \$2.29 really a threshold?).

Academic researchers have long recognized that price response functions need not be smooth or even monotonic (e.g., Gabor and Granger 1964, Monroe 1990). Thus, there is agreement with practitioners on the existence of price thresholds as an empirical phenomenon. Academics, however, have taken different approaches to the problem of estimating price thresholds. One stream of research approaches the problem at the individual level, working from the notion that individuals have reference prices against which they evaluate current prices when making a decision (e.g., Thaler 1985). Price threshold effects occur because a latitude of price acceptance (or region of price insensitivity) develops around the reference point. Empirically, researchers have found latitudes of price acceptance in both scanner panel data (e.g., Kalyanaram and Little 1994) and in controlled experiments (e.g., Kalwani and Yim 1992). In the study by Kalyanaram and Little (1994), for example, the latitude of price acceptance was found to be symmetric about the reference point and to average approximately 1.5 times the standard deviation in the brand shelf price. Another stream of work attempts to determine price thresholds from store-level data (or aggregations across stores). A promising approach was recently proposed by Kalyanam and Shively (1998), who use Bayesian methods (specifically Gibbs sampling) in combination with spline regression to estimate irregular price response functions on aggregate-level data. Nevertheless, the estimation problem is challenging in part because each consumer (and therefore each store) may have different zones of price insensitivity.

With respect to price gaps, academic studies have extensively investigated the asymmetric nature of price competition (as captured by cross-price elasticities) between brands in different "price tiers" (e.g., a premium national brand versus private label or private label versus generic). Blattberg and Wisniewski (1989), in a study using store-level scanner data, showed that price cuts on higher price, higher quality, brands had greater effects on the sales of lower price, lower quality brands than vice versa. These effects have also been documented in academic studies based on panel data (e.g., Kamakura and Russell 1989) and modeled in detail with store-level data (Sethuraman 1996). Academics have also investigated some of the implications of price gaps using game-theoretic models. For example, Raju et al. (1990) related the magnitude of price gaps between brands to the equilibrium level of price promotion activity within a product class. In sum, the academic study of price gaps has largely emphasized the competitive implications that may be associated with price gaps, as opposed to the determination of the optimal price gap.

While research has validated the existence of price thresholds (i.e., zones of price insensitivity) and the nature of asymmetries that may exist across price tiers, it has not yet begun to produce methods that can directly provide managers with the answers they expect to be able to obtain from scanner data (e.g., what is the optimal price gap between my brand and a particular rival brand). Given the widespread industry use and comfort with sales velocity analysis, we therefore classify this issue as one that practitioners view as resolved but academics deem unresolved and place it in the upper-right quadrant of Figure 1.[5]

3. Trade Promotion

Trade promotions typically account for the largest portion of the marketing budget at most consumer product companies (approximately 40–50%). Practitioners therefore had a keen interest in understanding the overall effectiveness of trade spending as well as the relative effectiveness of the various components (i.e., TPR, feature, display, etc.). The importance of trade promotion evaluation is underscored by the results from the Davidson and Stacey (1997) survey. They report that managers ranked trade funds management as the most important marketing issue they confront (first among 15), and that they used scanner data analysis extensively for this issue. Specifically, trade funds management ranked third (among 15) in practitioners' reported use of quantitative analysis and modeling, with 54% of respondents indicating heavy usage.

Key Questions

For the managers we interviewed, the critical issue in trade promotions is gauging the effectiveness of a trade deal. Two questions that practitioners seek to answer with scanner data are:

(1) What is the baseline sales volume (i.e., what would sales be if no promotions take place)?

(2) What is the incremental volume[6] (i.e., "lift") due to a trade promotion?

The effectiveness of a trade deal is then assessed by the magnitude of the lift generated by the promotion relative to the baseline sales level. While this type of analysis focuses on sales response, some managers also are interested in extending the assessment of trade promotion effectiveness to profitability. Baseline and incremental volumes are, of course, essential inputs to such a profitability analysis.[7]

Industry Practices

Both ACNielsen and IRI generate estimates of baseline sales using volumes in nonpromoted weeks and an exponential smoothing algorithm (i.e., more recent weeks are given greater weight). In some cases, adjustments are made to the baseline estimate using cross-sectional data (i.e., actual sales in similar stores in the same time period). Although there are some

[5]This is a particularly important research issue because the existence of price thresholds implies that price elasticities are not constant. Thus, inferences about price elasticities made from regression analyses (as in Equation (1)) are subject to this limitation.

[6]Incremental volume, or lift, is defined as the difference between actual volume and estimated baseline volume.

[7]Strictly speaking, TPR, display and feature are actions taken by retailers which often result from trade promotions. Therefore, manufacturers should assess the effectiveness of trade promotions by considering incremental sales *and* retailer pass-through. In this study we focus only on the issue of incremental sales.

technical differences in the two approaches used by ACNielsen and IRI (see Abraham and Lodish 1993 for a description of IRI's approach), they are conceptually very similar.[8]

Trade promotions also can be analyzed with regressions on store-level data. This approach essentially follows the regression template of Equation (1). In trade promotion analysis, however, the simple exponential smoothing procedure is preferred to regression because baseline estimates are readily obtained every week for each market, as well as for every SKU of every brand. Running regressions at this level is not only cumbersome, but also tends to generate unstable parameter estimates. Thus, practitioners have a strong preference for the simplicity and robustness of the smoothing algorithm for obtaining estimates of baseline sales.

Managers and consultants we interviewed had a high degree of comfort with the analysis of trade promotions. Baseline estimates from the smoothing algorithm that are embedded in the on-line data from IRI and ACNielsen are so well accepted that managers often view them as "data" rather than estimates. Trade promotion effectiveness is then analyzed by examing the lift associated with the specific characteristics of the trade deal (e.g., extent of feature, display, and TPR activity obtained at the retail level). Managers then study what types or combinations of deals appear to be most effective, ranking them according to the relative lift (i.e., incremental volume divided by baseline volume) that they generate. Trade dealing can then be adjusted to emphasize the types of deals that generated strong lift over the deals that did not.

Many practitioners consider trade promotion analysis with scanner data to be an industry-wide success story. Consultants at McKinsey said that trade promotion analysis has added "hundreds of millions of dollars to the bottom lines of U.S. packaged goods companies." Nestle executives echoed this view, reporting widespread use and impact. Interestingly, they pointed out that the impact has come not so much from changes in trade promotion spending, but from "redirected" spending, with increases in some areas balanced by decreases in others. Although trade promotion analysis is very useful for marketing managers in budget allocation decisions, it is potentially even more powerful in the hands of the sales force, where it can be used in negotiations with retail accounts. This has led to increased interest in obtaining incremental volume information at the account (or even store level).

Some companies and consulting firms told us that they are extending trade promotion analysis beyond incremental sales to incremental profit (e.g., Kraft and McKinsey specifically indicated that this was an area of emphasis for them). Nevertheless, most companies were focused on sales. The most common explanation for the emphasis on sales versus profitability was the difficulties in assessing the true cost of trade promotions. The difficulties arise because promotions vary on many dimensions (case allowances, off-invoice allowances, scanback pricing, etc.) and because of time lags between invoice and payment. Unless the company has designed a careful accounting system (such as activity based costing), most find it difficult to apportion relevant costs to promotions. This cost allocation problem is further exacerbated when a product is diverted in the channel from one retail account to another. These difficulties in assessing cost may explain why trade promotions continue to thrive even though some studies (e.g., Abraham and Lodish 1990) suggest that almost 84% of trade promotions are not profitable.

While most of the emphasis on trade promotion effectiveness involves the short-run effect of trade deals, an issue that also arose is whether or not trade promotions have long-term effects. On this question, most practitioners responded that they do not find post-promotion dips in their data. This leads many to conclude that promotions do not have long-term negative effects.

Academic Perspectives

Academic researchers have had a significant role in the development of industry procedures for evaluating the effects of trade promotions. The work of Abraham and Lodish (1993) is cited as especially influential (e.g., specifically by executives at IRI). Thus, much of the day-to-day use of baseline and incremental sales analysis

[8]The technicalities have, upon occasion, led to concern at some client companies about discrepancies in baseline and incremental volume estimates from the two data suppliers.

has been developed with the participation of academic researchers. At the same time, however, this research also has recognized many of the limitations inherent in the conventional approach (see, e.g., the discussion of limitations in Abraham and Lodish 1993, pp. 250 and 268). These limitations include the inability of the traditional baseline/incremental method to take into account forward-buying by loyal consumers (i.e., purchase acceleration or stockpiling), store switching by consumers, the cannibalizing effects of sister brand or size promotions, and promotions by competitors.

When we asked managers directly about the limitations of the baseline/incremental approach, most agreed that baseline estimates can have serious problems and that they are probably biased downwards. Downward bias in baseline sales implies that the lift attributed to trade promotions is overstated. This has important implications for assessing the profitability of trade promotion because the reclassification of sales volume from incremental to baseline can have a large negative effect on the economics of a trade deal. Despite these issues, managers told us that they rely on the baseline/incremental approach because (1) it is the best readily available method, and (2) they believe that the *relative* effectiveness of different promotions and promotion vehicles (e.g., TPR, feature and display) is unlikely to change significantly even if the missing factors were to be accounted for. This means that trade deals are effectively sorted by *relative* performance versus evaluated on their *absolute* performance. Relative performance information makes it possible to reallocate trade funds among deal types but does not provide the necessary information for determining the optimal total amount of trade promotion spending. Thus, while there has been enormous positive impact from trade promotion analysis (e.g., the "redirecting" of spending mentioned by Nestle executives), we believe that additional gains in trade promotion productivity can come from using approaches that better assess profitability.

Much of the academic research on trade promotion has studied market response to promotion variables at the individual-level using scanner panel data. This research has shown that promotions induce consumers to accelerate their purchases and stockpile quantity, not just switch brands (e.g., Neslin et al. 1985). Researchers have also decomposed brand sales elasticity into components due to switching, acceleration, and quantity (e.g., Gupta 1988), finding that brand switching effects are the largest source of sales response, but that significant sales response is often due to purchase acceleration and/or stockpiling. Bell et al. (1999) decomposed elasticities for 173 brands in 13 product categories. On average, brand switching accounted for 75% of total sales elasticity, with a low of 49% and a high of 94% among the categories in their sample.[9] These findings establish that a substantial portion of what appears to be "incremental" sales may be "borrowed" from the future. Unfortunately, the widely used baseline/incremental methods do not make this distinction.

While some practitioners recognized the limitations in the baseline/incremental method, they were also quick to point out severe limitations to panel data studies. Sampling problems are one reason for the reluctance to rely on panel data analysis. For example, many practitioners told us that market shares computed from panel data sometimes move in opposite directions from the shares computed from store-level data. The sampling issue involves both bias (e.g., are scanner panelists representative of the population?) as well as size (e.g., volume and share estimates for small brands or SKUs can be quite shaky). In an important study of the representativeness of panel data, Gupta et al. (1996) reported that the demographic characteristics of a sample of panelists in two data sets did significantly differ from the population and that the brand market shares in the panel data differed from the store data. Nevertheless, *the price elasticities* estimated from the panel and the store-level data differed by only 5–7%.[10] This suggests that panel data may be able to provide key insights into the effectiveness of marketing activity despite the sampling issues. More recently, Silva-Risso et al. (1999–this issue) showed that trade promotion productivity could be improved in a field

[9]A recent study by Van Heerde et al. (1999) uses *store-level* data and distributed lag models to infer that 4–24% of current sales effect is borrowed.

[10]As the authors of the study are careful to point out, this finding is based on the assumption that purchase observations are drawn from the scanner panel data following what they call *a purchase selection* approach and not a *household selection* approach.

application using results from a decision support system based solely on panel data. Clearly, further research is needed to take advantage of the ability of panel data to distinguish incremental from borrowed sales while overcoming sampling problems.

Many industry executives believe that promotions can increase consumption substantially, especially for impulse items such as snack food. Pepsi executives told us that 50% of their incremental volume may come from increased consumption. Some recent academic work has shown this phenomenon to exist in an experimental setting (e.g., Wansink and Deshpande 1994) and it was incorporated into models estimated on panel data (Ailawadi and Neslin 1998). A weakness of scanner data is that it tracks only purchases and provides no specific information on consumption or at-home pantry holdings. This is an important research topic since it affects how much of the incremental sales "bump" represents net new sales of the brand (due to consumption increases) versus sales borrowed from the future (due to pantry loading). Most recently, Bell et al. (1999) reported evidence consistent with consumption increases for 4 out of 13 grocery categories they studied. Specifically, promotions were associated with probable consumption increases in bacon, potato chips, soft drinks, and yogurt.

As noted above, the absence of a dip in sales following a promotion has led many managers to believe that trade promotions have no (negative) long-term effects. The academic perspective is, however, that the lack of a post-promotion dip per se need not suggest anything about the long-term effect of a promotion. Based on simulations, Neslin and Stone (1996) showed that post-promotion dips in sales data will not occur unless consumers are very sensitive to at-home inventory levels—i.e., it is normal not to observe such dips. Over the longer term, if consumers come to expect promotions, overall baselines may become depressed. Again, there need not be any post-promotion dips in the data. Although academic work on long-term effects is limited, initial studies have found that promotions do have negative long-run consequences. For example, Mela et al. (1997) studied data for one product category stretching over an eight-year period. They found that consumers became more sensitive to promotions as promotional activity increased. Using a varying parameter model, Foekens et al. (1999) show that the net incremental sales from a dynamic model can be significantly smaller than the incremental sales based on a static model.

Both practitioners and academics agree that scanner data and trade promotion have been an "industry-wide success story." Perhaps because the positive impact of scanner data has been so strong in trade promotions, practitioners expressed some surprise when we questioned the conventional approaches in widespread use. Nevertheless, we believe that the use of scanner data in trade promotions is not resolved. This is primarily because current methods are likely to overstate both the real sales response to promotion as well as the profitability of those sales. Accordingly we have placed this topic in the upper-right quadrant of Figure 1. In closing, we note that while trade promotion spending has been reallocated among deal types and by retail account, the overall spending on trade promotion has remained at high levels. This may reflect the difficulties involved in assessing the true profitability of these marketing activities.

4. Consumer Promotion

Managerial use of scanner data in consumer promotions is focused primarily on understanding the effects of coupons. Among the managers we interviewed, coupon analysis ranked below pricing and trade promotion in importance, but not necessarily in the expectations for how scanner data could aid decision making. Davidson and Stacey (1997) also report the same pattern. In their study, consumer promotion evaluation ranked seventh out of 15 issues in terms of importance, but fourth in the use of scanner data analysis and modeling. The major issues in consumer promotions are similar to those in trade promotions: understand incremental sales and profits due to coupons. At the broadest level this will help managers decide whether to increase or decrease spending on coupons. Some detailed decision issues involve the face value and frequency of coupon drops, expiration dates, and the potential advertising role of coupons (e.g., how important is coupon design and copy).

Industry Practices

In our interviews, we found that practitioners employed three approaches to the analysis of coupons: (1) judgment, (2) regression, and (3) logit models. With respect to the first, managerial judgment, several managers expressed high levels of comfort in relying on judgment as the primary means to decide on coupon related issues. They also noted that they want to be able to include such judgments along with other analytical techniques. For example, the average purchase cycle of the product (which can be measured with scanner panels or by survey) is considered a driving factor in deciding coupon frequency. As one manager explained, it does not make sense to him to drop coupons every 30 days if the average purchase cycle of your product is two months. Competitive intensity of coupons is also an important factor in deciding coupon face value and frequency. Similarly, expiration dates of coupons are a function of purchase cycle, the urgency that managers wish to create, and logistics factors for coupon clearance and payment to relevant parties.

A simple analytical tool used by some practitioners is to treat coupons as another promotion variable in the store-level regression model and assess its effectiveness in driving sales (see Equation (1)). A limitation of this approach is that it is unable to distinguish when coupons are used primarily by loyal consumers or by switchers (panel data is required for that). Thus, while the method is useful for obtaining general results on coupon response, it has limitations for assessing borrowed versus incremental sales.

A more sophisticated approach used by some companies (via custom consulting studies conducted by IRI and others) is the household-level logit model to assess incremental sales due to coupons (Guadagni and Little 1983, Little 1994). This approach, known at IRI as CouponScan, fits a logit (or nested logit) model to panel data and then simulates brand sales to determine what the sales would have been if coupons had not been dropped.[11] Incremental sales are then computed as the difference between actual and simulated sales. Thus, the principle is analogous to that of the baseline sales approach used in trade promotion. The major difference is that baseline sales estimates are produced from the logit model based on panel data, not an exponential smoothing algorithm based on store-level data.

By setting up the analysis in the baseline/incremental fashion, the CouponScan approach gets around the problem of the lack of information on coupon *availability*. While data are usually collected on coupon *redemption*, i.e., we know which panelists redeemed a coupon, there is usually no information on when a panelist had a coupon available but decided not to use it. Executives at Procter and Gamble expressed a very high level of comfort with the logit modeling approach used in CouponScan. Said one, "If we get a result out of the logit model that sounds weird, we will challenge the data first."

We found a strong and consistent response from all parties about coupons: they are generally regarded as relatively ineffective. Managers at Quaker reported that the studies they commissioned on marketing mix effectiveness (e.g., by ACNielsen) for their products revealed that coupons had the lowest impact of any marketing mix element. They also noted that their brands' heaviest users were also the heaviest redeemers of coupons. Consultants at the Hudson River Group reported that they found in their work an average elasticity for coupon spending of 0.07 (based on regression analysis of aggregate-level data). They noted that this places coupon response below that of media advertising (see, e.g., Assmus et al. 1984, Lodish et al. 1995). As one of the consultants put it, "We love to shoot coupons." At McKinsey, consultants called coupons, "a lousy investment."

Several factors were cited as contributing to the strong conclusions about coupons. These were low redemption rates, the large proportion of coupon redemptions by loyal consumers as revealed by panel data analysis, low levels of lift as revealed by regression and logit modeling (especially when compared with trade promotion), lack of enthusiasm from retailers in handling coupons, and the cost of clearing and payments. The consensus regarding the relative ineffectiveness of coupons has led many packaged goods

[11]Despite the claims of some academic researchers that probit is a superior choice model to logit, probit models remain unappealing to practitioners because of the complexities associated with their estimation. Said one executive at IRI, "Probit versus logit was never a debate here; probit would be too costly."

manufacturers to reduce their spending on coupons. Executives at both Quaker and Procter and Gamble, for example, said that scanner-based findings led management to cut coupons. Indeed, P&G recently conducted tests in upstate New York where they eliminated the use of coupons (e.g., Narisetti 1996). At Nestle, coupons are now used primarily in association with new product launches. Still others have moved to more targeted use of coupons (e.g., the check-out couponing system from Catalina Marketing, Inc.). In sum, the use of scanner data has had a major impact on how managers view coupon promotions in the packaged goods industry. This is because the data have clearly shown that coupons tend to have relatively small effects on sales and because coupons are redeemed heavily by loyal users. Our discussions surfaced little in the way of unresolved issues from practitioners.

Academic Perspectives

Many of the issues in evaluating the response to and profitability of coupon promotions parallel those in trade promotion. Both require that sales be decomposed into those incremental to the promotion versus those that would have been obtained without the promotion (i.e., baseline). The CouponScan system, for example, is engineered to do exactly this. Researchers have also stressed the need to determine the proportion of coupon redemptions that deliver incremental sales to the brand versus providing a discount to buyers who would have purchased the brand without the coupon (e.g., Blattberg and Neslin 1990). Analogous to trade promotion, there are major implications for the profitability of couponing. For example, Neslin (1990) reports that for the category he studied only about 40% of coupon redemptions can be classified as incremental sales for the brand. Based on typical coupon cost estimates, he argues that a much higher proportion is probably needed in order for couponing activity to be profitable. This reinforces the need to examine coupon promotions with panel data (as practitioners are indeed doing) to assess the extent of redemptions made by loyal buyers. We note that most of these studies have evaluated coupon profitability assuming that regular shelf prices remain constant. If coupons are used as a price discrimination device (Narasimhan 1984), it may be desirable to increase regular price while increasing the frequency and/or depth of coupons.

Another issue that has begun to be examined by academic researchers involves the potential for coupons to have an advertising effect on brand sales. Srinivasan et al. (1995), for example, found that coupons in free-standing inserts can have advertisement value to some consumers. The potential for this effect to improve the economics of a coupon promotion has been incorporated into a model of coupon profitability by Leone and Srinivasan (1996).

While there are clearly opportunities for additional research in coupons (e.g., the advertising effect of coupons, refining market response and profitability models), the perspectives of practitioners and academics are better aligned with respect to coupons than most other areas of the marketing mix. In particular, it is critical that practitioners have recognized that a significant portion of the temporary sales increase that a coupon may generate is likely to represent borrowed sales. Accordingly, we classify the evaluation of coupons as largely resolved from both the academic and practitioner perspectives.

5. Advertising

Advertising is extremely important to practitioners in packaged goods, but the expectations for how scanner data can help improve its effectiveness are more modest. Much of this comes from the role that creative factors play in advertising strategy and the widespread use of survey-based testing and tracking measures. The Davidson and Stacey (1997) study reports that advertising effectiveness is ranked fifth in importance but ninth in the use of scanner data analysis and modeling among the 15 issues they investigated. Thus, the relative importance of advertising as a business issue is greater than the current use of scanner data to help address it.

The practitioners we interviewed highlighted three key questions that they would like to be able to address with scanner data: (1) the short-term effect of advertising on brand sales, (2) the long-term effect of advertising, and (3) the indirect effects of advertising (e.g., the effects on the health or equity of a brand). Answers to these questions have implications for advertising spending, allocation of resources between promotion and advertising, and evaluation of advertising content.

Short-Term Effect of Advertising

Two broad approaches are used by practitioners to assess advertising effectiveness: an experimental approach and a regression-based approach. An example of the experimental approach is the split-cable method, where the media weight, creative, and scheduling all can be manipulated across different households in a city and their response measured through sales data captured by scanners. A detailed discussion is provided in Lodish et al. (1995), who describe a meta-analysis of 389 split-cable experiments conducted by IRI. The second method for analyzing the effectiveness of advertising is to include advertising spending measures or Gross Rating Points (GRPs) as independent variables in the regression model as described in Equation (1).

Considerable disagreement exists among practitioners regarding short-run advertising effect sizes. Some executives we interviewed stated that their analyses showed no short-term effect of advertising on sales (e.g., PepsiCo). To these managers, such results were neither surprising nor problematic since they believed the purpose of advertising is to positively impact the brand in the long run. Another view holds that advertising is a "sacred cow," not to be sacrificed at any cost, and that any advertising is an absolute good (e.g., P&G). A middle-ground view taken by many others holds that advertising has a small, but significant short-term impact on sales. Several practitioners told us that they have found elasticities to average about 0.10 to 0.12 for mature products and 0.20 to 0.40 for new products. A study by Frito-Lay found the advertising impact to be significantly greater for small brands, and for brands that have some "news" to offer (Riskey 1997).

Academic studies are generally consistent with the middle-ground—small but significant short-run elasticities. For example, based on a meta-analysis, Assmus et al. (1984) found average short-term advertising elasticities for established products to be about 0.15. Based on 389 real-world split-cable TV ad experiments, Lodish et al. (1995) concluded that the average advertising elasticity is 0.05 for established products and 0.26 for new products. We emphasize that these numbers are *averages*. For example, although the average elasticity for established products in the Lodish et al. study was 0.05, they also found that these effects were not statistically significant for 67% of the products. In other words, while academic studies tend to agree on the *average* short-term effect sizes of advertising, there is still considerable uncertainty regarding *specific* products or campaigns.

This conflict is represented by several studies. Lodish et al. (1995) are convinced that increasing advertising budgets in relation to competitors does not increase sales in general. In contrast, some academic studies (e.g., Pedrick and Zufryden 1991) as well as industry studies (e.g., research by ASI Market Research and Media Marketing Assessment Inc. (MMA)) show strong evidence of advertising weight or GRP on sales. It is difficult to say whether these conflicting findings arise due to (a) differences in methodology (split-cable for Lodish et al., logit model for Pedrick and Zufryden, and regression-based models for ASI and MMA), (b) differences in product categories studied, or (c) differences in the level of aggregation—both time (e.g., Lodish et al. use annual data, while Pedrick and Zufryden use weekly data) and geographic aggregation (e.g., Pedrick and Zufryden use panel data, while MMA typically uses market-level data).

A major source of controversy is the appropriate level of aggregation at which to study advertising effects with regression analysis. Many companies, such as Nielsen and IRI, argue in favor of store- or account-level analysis. They claim that this is consistent with the analysis conducted to assess the effectiveness of price and promotions (which vary by stores/accounts). Consequently, it is also appropriate to assess advertising effects after controlling for sales effects due to price and promotions. Others (MMA in particular) argue strongly in favor of market-level analysis. They suggest that since advertising decisions are made at the market level—not at the store or account level—differences in advertising weight or copy can only be observed at the market level.

While Nielsen and IRI have store- and account-level data, third-party consultants such as MMA usually conduct their analysis on the market-level data to which they are given access. Consequently, it is difficult to disentangle the self-interest of the respective parties from the technical merits of the approaches. Though their paper is not about advertising per se, Christen et al. (1997) have shown that the aggregation of store-level data to the market level can create biases in estimated effect sizes of marketing variables (e.g., feature and display). A careful comparative study of the estimation of advertising effect sizes with store-level versus market-level data would be of substantial interest to the wide variety of practitioners who come into contact with this issue.

Another controversy involves experimental approaches to assessing advertising effect sizes. Recently, John Philip Jones (1995) proposed an approach to measure the effect of advertising using a method called STAS (Short Term Advertising Strength). Using AC-Nielsen single-source panel data, the STAS procedure first estimates baseline and stimulated STAS. The baseline STAS is a brand's share of purchase occasions among households who have not been exposed to television advertising for it during the seven days before purchase. The stimulated STAS is the brand's share among households who have had at least one television ad exposure for the brand in the seven-day period. The difference between the baseline and stimulated STAS is a measure of the short-term impact of advertising. The STAS procedure has been criticized by Lodish (1997) for lack of random assignment of households to advertising conditions. He advocates the use of a split-cable approach, such as IRI's BehaviorScan. This continuing controversy (see, e.g., Jones 1998 and Lodish 1998) provides another example of the importance of resolving the measurement issues pertaining to the assessment of advertising effects.

Long-Term Effects of Advertising

Three issues arise in the study of long-term effects of advertising: (a) how large is the long-term effect, i.e. the effect size (if the long-term effect exists), (b) how long does the advertising effect last, i.e. the duration, and (c) what are the sources of long-term effect. Although most practitioners agree that advertising should have a long-term effect on sales, few know how to measure it. Managers told us that they consider the long-run effects of advertising so difficult to quantify that they make little attempt to measure them. Some research companies, such as Millward Brown International, are an exception. Based on its client work and proprietary model, Millward Brown suggests that the long-term effect of advertising is about two times the short-term effect, and in some cases these effects could be as large as eight times the short-term effect (Hollis 1997, Scott and Ward 1997).

Academic studies have typically used a distributed-lag (e.g. Koyck) model to assess the short- and long-term impact of advertising on sales. Assuming a geometric decay of advertising effect, the Koyck model can be written as:

$$S_t = \alpha + \beta A_t + \lambda S_{t-1} + \epsilon_t \qquad (2)$$

where S_t is the sales at time t, A_t is advertising at time t, β is the short-term effect of advertising, and λ is the lag coefficient or carry-over effect of advertising. It is then easy to show that in this model, the long-term effect of advertising is $\beta/(1 - \lambda)$.

Assmus et al. (1984) conducted a meta-analysis on 128 models reported in 22 studies and reported an average λ of 0.468, which suggests the long-term effect to be $1/(1 - 0.468) = 1.88$ times the short-term effect (β). Using four product categories, Givon and Horsky (1990) estimated λ to be between 0.25 to 0.82 with an average of 0.59. Based on three categories, DeKluyver and Brodie (1987) estimated λ to lie between 0.50 and 0.80 with a mean of 0.70. Using an eight-year stretch of panel data, Mela et al. (1997) estimated the long-term effect to be between 1.5 to 2 times the short-term effect. Collectively, all these academic studies suggest the long-term effect to be between 1.3 to 5.6 times the short-term effect with the average close to 2.0.

Lodish et al. (1995) used an experimental approach to assess the long-term effect of advertising. They conducted 42 in-market split-cable experiments on established brands where the advertising weight was changed for one year, and its sales effect (compared to a control cell) was monitored over the next two years. They found that, on average, the long-term effect (i.e., two years following the test year) of advertising on sales was about two times its short-term (or test year) effect. In sum, all the academic studies corroborate the Millward Brown work, showing that, on average, the long-term effects of advertising are about twice the short-term effects. These effects, however, could be significantly larger for some brands and products.[12]

[12]Millward Brown finds that, in general, large brands have small short-term but large long-term effects.

The second issue related to the long-term effect of advertising is the duration or length of the effect. There are conflicting results in this area. In its split-cable experiments, Frito-Lay found that most of the effect of advertising occurs within six months (Riskey 1997). In contrast, Millward Brown finds the long-term effects of advertising have a half-life of well over six months (Hollis 1997). Within academia there is also little consistency on this issue. The classic study by Clarke (1976) found the 90% duration interval of advertising (i.e., the length of time it takes to realize 90% of the effect of the advertising) to last 6–9 months. But more recent studies by Lodish et al. (1995) and DeKimpe and Hanssens (1995) both indicate that advertising effects may last over several years.

Lastly, there are few studies devoted to understanding and quantifying the sources of long-term effects of advertising. Nevertheless, there is some consistency in the results among the work that has addressed this issue. For example, both academics (Lodish et al. 1995) and practitioners (von Gonten and Donius 1997) agree that the long-term effects of advertising are more attributable to increases in consumer buying rates (i.e., more sales to existing customers) than to higher penetrations of the brand into nonusers.

Indirect Effects of Advertising

Many managers we interviewed said that they, and especially their advertising agencies, focus on a number of dependent variables other than sales when assessing advertising effectiveness. For example, many executives feel that the role of advertising is not to influence sales directly but to create awareness and act as a reminder at the time of purchase. The actual sales are then a function of price, promotion, product quality, and distribution. Since our focus is scanner data, which does not include scores on awareness, recall, and persuasion, we will not discuss most of these indirect effects.

One indirect effect that has been examined with scanner data is the effect of advertising on brand equity. While the long-term effect of advertising may capture some aspects of this effect on brand equity, it is far from a complete representation. There is an ongoing debate within both industry and academia about how to measure brand equity and advertising's effect on it. We found many executives concerned about brand equity issues but few had any sound methodology to address them.[13] Some executives suggested using the ratio of own-price to cross-price elasticity as a measure of brand equity (in this approach one can either make brand equity relative to each competing brand, or use cross-price elasticity of a generic brand as a benchmark). A few firms offer models of brand equity (e.g., BrandBuilder by The NPD Group) but the acceptance of these models and approaches is limited.[14] Academic research on this topic is also far from conclusive. Scanner-based research has considered the "intercept" in the logit model as a possible surrogate for brand equity (Kamakura and Russell 1993). Using this surrogate, Jedidi et al. (1999) found that advertising has a significant positive impact on brand equity while long-term promotion has a negative impact on brand equity. While these preliminary results are intuitively appealing, more research is clearly needed.

Another indirect effect of advertising is on consumers' current and future purchase behavior. For example, many academic studies have shown that nonprice advertising reduces consumers' price sensitivity (Kaul and Wittink 1995). Practitioners with whom we discussed this issue also believe this to be true and several companies rely on analysis by Nielsen and IRI to measure this effect. A remaining issue is how this indirect effect should be captured when attempting to assess the overall (i.e., both direct and indirect) effectiveness of advertising.

Summary for Advertising

There is remarkable consistency between the academic and "middle-ground" industry views regarding the average size of short-run advertising elasticities (0.05–0.15 for mature products and 0.20–0.40 for new products). Moreover, most of the academic as well as the industry studies on the subject agree that, on average, the long-run effect of advertising is approximately two

[13]Corroborating this, Davidson and Stacey (1997) report that measurement and management of brand equity ranked ninth out of 15 marketing issues in importance, but fourteenth out of 15 in use of scanner data analysis and modeling.

[14]As is often the case with proprietary industry models, it is impossible for outsiders to evaluate them.

times the short-run effect. Both practitioners and academics also agree that ad copy makes a significant difference in the impact of advertising on sales, the source of long-run effects is due more to increasing buying and not increased penetration, and that advertising decreases price sensitivity. Nevertheless, practitioners view the assessment of all advertising effects with scanner data to be highly unresolved, perhaps due to the difficulties involved in measuring them in their *own specific businesses.* This is highlighted by the controversies surrounding both the regression-based approaches (i.e., the appropriate level of data aggregation) and other approaches (i.e., experimental as in split-cable versus nonexperimental as in STAS). There is also some debate about the correlation between recall and persuasion scores and brand sales, the duration of advertising effects, and how to capture the impact of advertising on brand equity. Accordingly, we place the average size of advertising effects in the lower-left quadrant of Figure 1 (resolved for academics, unresolved for practitioners), while the specific measurement issues appear in the lower-right quadrant (unresolved for both groups).

6. Product

For the managers we interviewed, the most important issue in product policy and strategy is the rationalization of product lines and limiting product proliferation. As companies added new SKUs to their lines without deleting older items, significant product proliferation occurred to the point where many categories, such as toothpaste, now have as many as 1,200 separate UPC items. Several studies have shown (e.g., Cummings et al. 1990) that assuming the 80/20 rule (i.e., 80% of the profit is accounted for by 20% of the items), such product proliferation not only adds significant cost to the system but it also creates confusion among consumers. In addition, retailers are demanding that manufacturers rationalize product lines in order to obtain and hold supermarket shelf space. Recent industry trends, such as the movement to Efficient Consumer Response (ECR), also highlighted this issue for most of the practitioners we interviewed.

Corroborating the views of the practitioners we interviewed, the Davidson and Stacey (1997) study ranked product assortment/SKU optimization sixth in importance and fifth in the use of scanner data analysis and modeling. Their study also included new product introductions as an issue, and it ranked very close to the assortment/SKU optimization issue: fourth in importance and sixth in the use of scanner analysis and modeling. Many of the issues involved in new product introduction (especially for line and brand extensions) overlap extensively with product assortment and SKU optimization. With respect to launching new products, many practitioners said that they rely on traditional survey-based methods for concept testing.

Industry and Academic Perspectives

A few industry studies have been conducted recently to examine the impact of SKU reduction on category sales. The Food Marketing Institute reported that reducing the number of SKUs in six test categories in three retail chains resulted in no significant loss in category sales (Kurt Salmon Associates 1993). A similar result was reported by *Progressive Grocer* (Krum 1994). The executives we interviewed shared the belief that many categories could be trimmed of SKUs with little effect on sales. Nevertheless, there remains a reluctance to cut SKUs in categories that are highly competitive or where supermarket shelf space is scarce.

Academic studies that have addressed this issue have reached roughly the same conclusion. Dreze et al. (1994) customized shelf facings of SKUs while deleting approximately 10% of the less popular items. They found that this change resulted in a 4% *increase* in sales. Based on lab experiments, field tests, and surveys, Broniarczyk et al. (1998) concluded that eliminating as many as 50% of the items had no significant impact on consumers perceptions or purchase behavior as long as category shelf space was held constant and most consumers could find their favorite items. On the other hand, a number of studies have argued that consumers demand a wider assortment of varieties when they buy large quantities of a product (Simonson 1990, Simonson and Winer 1992, Walsh 1995). Behaviorally, this phenomenon results from consumers' uncertainty about future consumption preferences. If this were a large effect in the marketplace, it would provide a counterargument for retaining large product assortments. Examining the phenomenon in the yogurt category, Bucklin et al. (1998) show that the size of assortment effect is modest—even with a 50% increase in total quantity purchased. Taken together, these studies are consistent with the industry perspective that a reduction in the number of SKUs is unlikely to affect category sales.

Given the agreement on the need to rationalize product assortment and reduce the number of SKUs, the next question is how to decide which items to eliminate. Our discussion with the practitioners suggested that they follow a simple, and somewhat naive, procedure of deleting, say, the bottom third (in terms of sales or profits) of the items in a category. Some retailers have adopted this approach unless they obtain specific guidance from manufacturers. Since this approach ignores market structure issues (i.e., where do sales go when an item is deleted), it does not necessarily provide the retailer or manufacturer with the best reduced assortment of items.

One promising approach to this problem is due to Fader and Hardie (1996). Their approach combines the features of the logit model with certain aspects of conjoint analysis and is currently in widespread use at IRI. Traditional logit models (e.g., Guadagni and Little 1983) estimate a constant for each item, brand-size, or SKU. This model has severe limitations when the number of SKUs is very large, a situation which is now typical in most packaged goods categories. Most academic studies have dealt with this problem by either eliminating low share items or by aggregating over them in some fashion. But this approach is often incompatible with the needs of practitioners because their focus is on understanding and enhancing the position of specific SKUs. Fader and Hardie (1996) suggest a solution by characterizing an SKU as a combination of its product attributes (e.g., brand name, package size, flavor, variety, type, etc.). This respecification significantly reduces the complexity of the logit model for the product category. For example, if there are eight characteristics, each with two levels (e.g., package size: large and small), then the traditional logit model will estimate $2^8 - 1 = 255$ item constants. (Indeed, many of these may not be estimable if certain combinations do not exist in the market place.) The Fader and Hardie approach will need to estimate only eight parameters and, as in conjoint analysis, also can forecast the share for a new combination of product characteristics. Although this approach assumes that the product characteristics only have main effects, and researchers may want to test this assumption, the approach is noteworthy because it is one of the first to be able to use scanner data to provide direct guidance on product strategy.

Traditional approaches to new product development and testing (e.g., concept testing) still dominate the decision-making process in the new products domain. The Fader and Hardie (1996) model represents an important step in bringing the use of scanner data and models into product policy and strategy. Nevertheless, because scanners provide behavior and not intentions data, it may not be possible to forecast the potential share of a truly new concept with scanner data alone. In sum, we found little disagreement between the academic and practitioner perspectives on product-related issues. Both agree that reducing the number of SKUs in a category may not damage sales, and that more study is needed on the best way to go about optimizing a particular assortment.

7. Distribution and Retail Management[15]

Executives in our survey raised three key issues regarding distribution and retail management. These are, in order of importance, (a) category management, (b) account-specific insights, and (c) managing shelf space. The most significant of these issues for manufacturers and retailers is category management. This concept has gained substantial favor since the Kurt Salmon Associates (1993) study on efficient consumer response. (The study recommended four major areas for improvement, one of which focused on category management.) The primary implication of category management is a shift in retailers' attitudes from a buyer orientation (where money is made based on how the product is bought), to a buyer and merchandiser orientation (where the focus is on *category* profitability). Davidson and Stacey (1997) report that category management was the second most important marketing issue for the practitioners they surveyed, with 77

[15]Note that we did not interview retailers. This discussion focuses on manufacturers' perspectives on how to create "win-win" strategies with retailers.

percent indicating that it was extremely important. Scanner data use, however, is only moderate with the issue ranking seventh out of 15 on this dimension.

A second key issue is the desire of the sales force for account-specific insights. Simply focusing on national or even market-level price elasticities is considered meaningless by most sales people. Skepticism about market-level results comes from the strong belief that "my market or my store is different and unique." Davidson and Stacey (1997) reported that micromarketing/store cluster analysis ranked twelfth out of 15 issues in importance, but was eighth in terms of the use of scanner data analysis and modeling.

With respect to the issue of shelf space, manufacturers and consultants both noted that retailers are growing more interested in the competitive market structure of products (e.g., tree-structured representations of consumer switching behavior) which can help them better organize their displays. Indeed, manufacturers say that retailers often expect them to provide this information and related analyses. For example, retailers want to know whether it is better to group all Coca Cola and all PepsiCo soft drink SKUs together, versus grouping the diet and nondiet SKUs together. The analyses needed to answer these questions must often be conducted at the disaggregate level (either with scanner panel data or by survey). In the Davidson and Stacey (1997) study, shelf-space allocation was ranked eleventh in overall importance and in the use of scanner analysis and modeling.

Category Management

Most manufacturers perceive retailers' sophistication in category management to vary widely and that some retailers still focus on cost and gross margins. Nevertheless, our interviewees reported that retailers are looking closely at the role of a category in the success of a store, its traffic, customer penetration, and growth. This leads to questions about the role that a particular manufacturer's brand has in enhancing the overall profitability of the category. At Procter and Gamble, executives said that they welcomed the change because it gives sales people and retailers an objective way to develop win-win strategies. The fact that most P&G brands are market leaders—or near-leaders—in their categories helps to reinforce this view.

Even though executives consider category management to be an important issue, few told us that they rely on quantitative models to address this topic. Current practice is characterized by the simple approach of comparing winners and losers. In this approach, a manufacturer attempts to convince a retailer to make its brand the "category captain" by contrasting the category sales of a store where its brand is the "category captain," with another store where its brand is not the category captain.

Several academic studies have explicitly addressed the category management issue. Studies that decompose brand sales into category volume and brand share/choice are also useful in addressing category management issues (e.g., Bell et al. 1999, Dillon and Gupta 1996). Hoch et al. (1994) conducted field experiments in 26 product categories in a grocery chain with 86 stores and found that while an every day low price (EDLP) strategy provides a small win for manufacturers (average 3% increase in unit sales), it reduces retailer profits by 18%. As retailers continue to move toward category management, these strategic issues are likely to grow in importance for manufacturers as they attempt to design "win-win" strategies. The decision support system developed by Silva-Risso et al. (1999–this issue), for example, takes one step in this direction by optimizing the manufacturer's promotion plans subject to the fulfillment of "win-win" constraints imposed by the retailer.

Micromarketing and Account-Specific Insights

Manufacturers report that their sales people want decentralized decision making because they are closer to the market and know the idiosyncrasies of each account. Procter and Gamble is using an approach where headquarters decides on national and international strategies (e.g., the budget allocation across advertising versus promotion), but leaves micro decisions to individual sales teams. Managers said this could be thought of as follows: the responsibility of the marketing department is to manage baseline sales (e.g., brand equity), while the responsibility of the sales force is to manage incremental sales. Implementing this approach required equipping the sales force with, and training it in the use of, more advanced analytical tools.

Similar to category management, current practice in account-specific management typically involves the comparison of winners and losers to search for causal factors. For example, a company may match consumer and store demographics to see what types of consumers shop where, which stores grew, which stores did not grow, and what they did differently in terms of merchandising, store layout, etc. Another example of this simple analysis includes an SKU productivity comparison (across SKUs and across stores), where shelf audit data are used to assess sales per square foot, share of shelf for own brand, "fair share" of shelf, and over/underdeveloped SKUs. These approaches provide account-specific recommendations (e.g., account A should promote more of item X) and also aid in the design of win-win strategies. Some manufacturers (e.g., P&G) are providing their retail accounts with analyses of how the promotion of their brands can add to the store's overall sales based on the type of buyer that the brand attracts to the store. For the sales force, the appeal of these approaches is in their simplicity, ease of understanding and communicability to retailers, and their ability to provide major insights for large number of SKUs and stores in a short period of time.

A second approach is an innovative procedure pioneered by practitioners. To develop better promotion plans for each retail account, one manufacturer, with the help of a data supplier and an outside consulting firm, designed a decision support system based on a creative combination of panel and store-level data. This approach consisted of five steps:

1. Use panel data to create preference segments of consumers. For example, 30% of consumers may be loyal to brand A, 20% may be very price sensitive and frequently switch among brands, 10% may be loyal to brand B, and so on.
2. Use panel data to find where consumers shop.
3. Assess the percent of each type of consumer (e.g., loyal to A) that makes up a store profile. For example, store 1 may have a consumer profile of 10% loyal to A, 5% loyal to B, and 85% switchers, whereas store 2 may get mostly brand loyal consumers.
4. Cluster stores based on the similarity of their consumer preference profile. For example, one cluster of stores may consist mainly of loyal consumers, another cluster of stores may consist mainly of switchers, and so on.
5. Conduct analysis for each cluster of stores separately. For example, a pricing analysis would be done separately for cluster one, cluster two, etc. The implicit assumption is that stores in the same cluster will have a similar response function since the brand preference profile of consumers is similar within each cluster.

Note that a store-by-store analysis may be infeasible due to the limited number of observations for each individual store, while an aggregate analysis may mask important differences across stores. Thus, the above procedure provides one way of overcoming these problems. An alternative approach, advanced by ASI Market Research, uses "modified pooled" models where the data are pooled across accounts or markets unless the data suggests otherwise (Richardson and Dratfield 1997). Markets that deviate from the pooled results are reported separately. This approach also attempts to strike a balance between one model for all stores/markets versus one model for each store.

Significant progress has been made in the academic literature on this issue. For example, using weekly scanner data for 18 categories for a chain of 83 supermarkets, Hoch et al. (1995) correlate store-specific price elasticities to consumer demographics and competitive variables. They find that consumer demographics explain a large proportion of the variance in price elasticities across stores. These results can be used to design more effective everyday and promotional pricing strategies that exploit store-level differences in price sensitivity. Although academic studies have not specifically looked at the issue of store segmentation, approaches that employ latent class analysis for consumer segmentation could be used to segment stores based on store-specific price and promotion responsiveness (e.g., Wedel and Kamakura 1998). Similarly, Bayesian models could also be used to address the problem of determining elasticities for a specific store.

Managing Shelf Space

Because of requests from retailers for information about "how consumers shop the category," the class of techniques known in the academic literature as *market structure analysis* is attracting renewed attention from practitioners. The analysis seeks to reveal the underlying substructure of a product category, usually in terms of the degree of brand and/or item substitutability. Most approaches involve panel-data based models such as Hendry (e.g., Kalwani and Morrison 1977), tree models (e.g., Rao and Sabavala 1981), or overlapping clusters obtained via latent class analysis (e.g., Grover and Srinivasan 1987, Kamakura and Russell 1989). Both IRI and Nielsen are well acquainted with the academic literature and methods in this area and regularly provide custom studies for their clients on this topic.

Interestingly, these developments coincide with the introduction of frequent shopper programs by major supermarket chains in the U.S. Indeed, some practitioners foresee that a key benefit to retailers will be the panel-level information on transactions generated by these programs and, hence, the ability to better understand the shopping behavior of consumers in their stores. In the meantime, manufacturers reported that retailers are increasingly using such measures as DPP (direct product profitability) and sales/profit per square foot to allocate shelf space in their stores. As in category management, comparing winners and losers is the most widely used approach.

Summary for Distribution and Retail

While academics have recently begun to develop models to address category management (e.g., Silva-Risso et al. 1999—this issue), these are in the early stages and have not yet found extensive use in practice. We classify category management as unresolved from both perspectives. Academics have made good progress in the area of account-specific management or micromarketing but, apart from a few innovative approaches, practitioners are generally using a simple comparison of winners and losers. Consequently, we classify micromarketing as resolved from the academic perspective but unresolved from the practitioner perspective. In the area of shelf-space management, both industry and academia are in agreement on the benefits of using market structure analysis, and we therefore classify its use as resolved from both perspectives.

8. Conclusion

Our goals have been to investigate, distill, and present the practitioner community's perspective on the use of UPC scanner data and to contrast this with the perspectives represented in the academic marketing literature. For each of the issues raised by managers, we reviewed the academic literature to assess whether or not academics have developed models and/or analytical procedures that have or readily could settle the issue. This leads to a 2×2 classification of the various issues as resolved or unresolved from the perspective of both academics and practitioners (see Figure 1). Given the desire that many marketing academics have for their research to acquire a broad impact, we hope that seeking out and reporting back "the voice of the practitioner" and comparing it to the academic marketing literature will help in guiding future research, both in the short and long run. We now turn to a summary of what we believe to be the most important immediate and long-term research needs in this area.

Immediate Research Needs

With respect to immediate needs for more academic research, we follow the unresolved items on the right-hand side of Figure 1 and propose the following:

1. **Price Thresholds and Gaps.** Academic researchers could extend existing models of price thresholds that are based on store-level and panel-level data to address practitioners' questions regarding optimal price points and price gaps versus competing brands. For example, this could be done by incorporating response models into decision support systems.

2. **Baseline and Incremental Sales.** Research is needed to develop simple, robust models that will produce better estimates of promotional sales that are truly incremental for the manufacturer, not borrowed from the future, from another store, or from a sister brand. These models will also need to take into account the effect that promotions may have on consumption in some product categories (e.g., stocking up on soda means drinking more soda).

3. **Base Price Elasticity.** Studies are needed to determine whether or not base-price elasticities can be reliably estimated from scanner data and, if so, what is the extent of natural price variation or number of observations needed. These estimates could also be compared with and tested against those obtained from survey-based methods such as discrete choice analysis.

4. **Competitive Reactions.** Omission of competitive reactions may lead to biased estimates of market

response to a change in price (or other marketing variables). Two potential approaches to this problem are to build game-theoretic models or to incorporate empirical reaction functions into models of market response.

5. Advertising Measurement. Research is needed to resolve aggregation issues in assessing advertising effects. Determining the best approach to assess advertising effects with scanner data could resolve some of the methods and data aggregation debate in this area. The vast majority of practitioners we interviewed expressed high levels of frustration with the inability to get clean, consistent answers about advertising effectiveness from scanner data. As one manager at Nestle put it, "Somebody needs to solve this."

6. Brand Equity. More research is needed into what can be learned about brand equity from the analysis of scanner data. Choice models, for example, could be extended to capture consumer valuation of product features as opposed to brand names, following the approach developed by Fader and Hardie (1996). Academics also will need to develop clear definitions regarding precisely what brand equity represents and how it should be measured.

7. Product Assortment. Manufacturers (and retailers) could benefit from methods to determine the costs and benefits of broader versus narrower product assortment (e.g., number of flavors and/or varieties).

8. Category Management. Manufacturers need simple models and decision support systems to automate elements of the category management function.

Even though research on scanner data has been actively pursued by academics for well over a decade, the length of the above list shows that there are significant gaps in our knowledge and numerous opportunities for research projects of high academic and commercial impact.

Long-Term Needs

In addition to the topics described above, we believe that there are also several important long-term needs for research and action. These include the following:

1. From Tactics to Strategy. Most of the scanner data analysis to date has focused on short-term tactical issues (e.g., deciding whether the price should be $2.99 or $2.79). We believe that for this type of analysis to gain greater attention with senior management, it must also address strategic issues such as brand equity, customer equity, competitive reactions, category expansion, and everyday pricing (EDLP) versus promotional pricing (Hi-Lo).

2. From Sales/Share to Profit/Industry Surplus. Both academic studies and industry practice focus almost exclusively on sales and share. Academics have not studied the impact of marketing activities on profits partly because suitable data has not been available.[16] Interestingly, practitioners told us that they had similar problems: marketing managers seldom evaluate profit impact because it is hard to assess and allocate promotion costs. We believe that it is important to move beyond sales analysis and to carefully study the impact of price, promotions, and other mix elements on short- and long-run profits, and the implications for how total industry surplus is shared among manufacturers and retailers.

3. From Descriptive to Prescriptive Models. Most scanner data models are descriptive in nature. Although these models provide broad guidelines (e.g., if your price elasticity is low in absolute terms, a price reduction may not be advisable), they still leave managers uncomfortable about making specific decisions. A typical complaint along these lines is, "You are telling me that the price elasticity for my brand is -2, but what do I do with it?" To aid managers in making these decisions, we need to develop prescriptive models that will incorporate the key aspects of descriptive models as well as assess competitive reactions, impact on profits, and long-run health of the business.

4. From Method Confusion to Industry Standards. We found packaged goods companies to be bombarded with a variety of methods from third-party

[16] The Marketing Science Institute attempted to assemble a "brand-level data base" in the early 1990s that would enable researchers to study the impact of marketing tactics and strategy on profitability. It reportedly ran into resistance from packaged goods companies regarding the release of proprietary information on costs and profitability and the project was dropped.

consultants, the details of which are often not disclosed to clients or outsiders. This creates methods confusion and makes it impossible to compare results and resolve controversies. As we discussed above, this problem is most acute in the area of advertising. We believe that it may be quite helpful for an industry council (or similar forum) to actively promote open discussion and debate to help establish methods standards for scanner data analysis.

Accelerating the Adoption of Scanner Analytics

A series of barriers to the diffusion of new analytical approaches were noted by practitioners. First, managers must have confidence that scanner analysis can offer tangible advantages over judgment and other methods. At PepsiCo, one executive said that the prevailing perception in the organization was that the scanner analysis did not necessarily improve the "win rate" for marketing decisions. Second, many manufacturers commented on what they perceived to be the high costs for scanner data and custom studies. At Procter and Gamble, executives said they viewed scanner analysis as a cost and that cheaper solutions are constantly sought out. At Kraft, however, executives said that they view scanner analysis as an investment. A third barrier, at least from the perspective of the data suppliers, was a concern that introducing a new analysis technique may be met with unrealistic client expectations. An executive at ACNielsen summarized the problem: "As soon as you make something new available, all of the worst problems line up first." Practitioners also noted that the quantitative orientation of key decision makers (e.g., senior vice-president of marketing) is a predictive factor in a company's use of scanner data analysis (i.e., going beyond a basic scorecard function). Finally, large organizations are thought to be bigger users of more advanced analysis techniques. In the words of one IRI executive, "Size correlates with [analytical] sophistication."

Academics can help to speed diffusion in many ways. By stepping back and arriving at fundamental substantive findings, academics can offer practitioners the power of good social science to build confidence in results from scanner data analysis. They can also determine the generalizability of empirical results (e.g., meta-analyses on price elasticity or advertising elasticity; see Tellis 1988 or Lodish et al. 1995) which helps provide reasonable bounds when estimating models. Methodologically, academics can work to make models more robust and more parsimonious. Though more complex models can sometimes offer technical advantages, these must be carefully weighed against the negative effect that complexity has on commercial adoption. Academics can also show how management judgment can be incorporated into the modeling process (e.g., Blattberg and Hoch 1990) and they can provide impartial assessments of different approaches, helping to resolve controversies and technical debates. Perhaps most importantly, in their teaching roles they can provide learning experiences that give current and future managers confidence that using analytical methods indeed leads to better decision making.[17]

References

Abraham, Magid M., Leonard M. Lodish. 1990. Getting the most out of advertising and promotion. *Harvard Bus. Rev.* **68** (May–June) 50–60.

——, ——. 1993. An implemented system for improving promotion productivity using store scanner data. *Marketing Sci.* **12** (Summer) 248–269.

Ailawadi, Kusum, Scott A. Neslin. 1998. The effect of promotion on consumption: buying more and consuming it faster. *J. Marketing Res.* **35** (August) 390–398.

Allenby, Greg M. 1990. Hypothesis testing with scanner data: the advantage of Bayesian methods. *J. Marketing Res.* **27** (November) 379–389.

Assmus, Gert, John U. Farley, Donald R. Lehmann. 1984. How advertising affects sales: a meta analysis of econometric results. *J. Marketing Res.* **21** (February) 65–74.

Bell, David R., Jeongwen Chiang, V. Padmanaban. 1999. The decomposition of promotional response: an empirical generalization. *Marketing Sci.* **18** (4) forthcoming.

Blattberg, Robert C., Stephen J. Hoch. 1990. Database models and managerial intuition: 50% model, 50% manager. *Management Sci.* **36** 887–899.

——, Edward I. George. 1991. Shrinkage estimation of price and promotional elasticities: seemingly unrelated equations. *J. Amer. Statist. Assoc.* **86** (June) 304–315.

——, Scott Neslin. 1990. *Sales Promotion: Concepts, Methods, Strategies.* Prentice-Hall, Englewood Cliffs, NJ.

——, Kenneth J. Wisniewski. 1989. Price-induced patterns of competition. *Marketing Sci.* **8** (Fall) 291–309.

[17]The authors thank the Marketing Science Institute for funding and the practitioners at participating firms for their time. We also thank the editor, area editor, and two anonymous reviewers for their comments and suggestions.

Bolton, Ruth N. 1989. The relationship between market characteristics and promotional price elasticities. *Marketing Sci.* **8** (2) 153–169.

Bresnahan, Timothy. 1997. Testing and measurement in competition models. David M. Kreps, Kenneth F. Halls, eds., *Advances in Economics and Econometrics: Theory and Applications*. Cambridge University Press, Great Britain.

Broniarczyk, Susan, Wayne D. Hoyer, Leigh McAlister. 1998. Consumers' perceptions of the assortment offered in a grocery category: the impact of item reduction. *J. Marketing Res.* **35** (May) 166–176.

Bucklin, Randolph E., Sunil Gupta, S. Siddarth. 1998. Modeling the effect of purchase quantity on consumer choice of product assortment. *J. Forecasting* **17** (June/July) 281–301.

——, James M. Lattin. 1992. A model of product category competition among grocery retailers. *J. Retailing* **68** (Fall) 271–293.

——, Gary J. Russell, V. Srinivasan. 1998. A relationship between market share elasticities and brand switching probabilities. *J. Marketing Res.* **35** (February) 99–113.

Carpenter, Gregory S., Lee G. Cooper, Dominique M. Hanssens, David F. Midgley. 1988. Modeling asymmetric competition. *Marketing Sci.* **7** (Fall) 393–412.

Cattin, Phillipe, Dick R. Wittink. 1982. Commercial use of conjoint analysis: a survey. *J. Marketing* **46** (Summer) 44–53.

Christen, Markus, Sachin Gupta, John C. Porter, Richard Staelin, Dick R. Wittink. 1997. Using market-level data to understand promotion effects in a nonlinear model. *J. Marketing Res.* **34** (August) 322–334.

Clarke, D. G. 1976. Econometric measurement of the duration of advertising effects on sales. *J. Marketing Res.* **10** 250–261.

Cooper, Lee G. 1988. Competitive maps: the structure underlying asymmetric cross elasticities. *Management Sci.* **34** (June) 707–723.

Cummings, Peter, David White, Stefan Wisniowski. 1990. Strategic simplicity. *McKinsey Quart.* (3) 80–90.

Davidson, Timothy A., E. Craig Stacey. 1997. 1996 marketing analysis and modeling survey. Summary Report, Prevision Corporation, Wellesley, MA.

DeKimpe, Marnik G., Dominique M. Hanssens. 1995. The persistence of marketing effects on sales. *Marketing Sci.* **14** (Winter) 1–21.

DeKluyver, C. A., R. J. Brodie. 1987. "Advertising-versus-marketing mix carryover effects: an empirical investigation. *J. Bus. Res.* **15** 269–287.

Dillon, William R., Sunil Gupta. 1996. A segment-level model of category volume and brand choice. *Marketing Sci.* **15** (1) 38–59.

Dreze, Xavier, Stephen J. Hoch, Mary E. Purk. 1994. Shelf management and space elasticity. *J. Retailing* **70** (4) 301–326.

Ehrenberg, A. S. C., L. R. England. 1990. Generalizing a pricing effect. *J. Indust. Econom.* **34** (September) 47–68.

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

Foekens, Eijte W., Peter S. H. Leeflang, Dick R. Wittink. 1994. A comparison and an exploration of the forecasting accuracy of a loglinear model at different levels of aggregation. *Internat. J. Forecasting* **10** (September) 245–261.

——, ——, ——. 1999. Varying parameter models to accommodate dynamic promotion effects. *J. Econometrics* **89** (March–April) 249–268.

Gabor, Andrew, C. W. J. Granger. 1964. Price sensitivity of the consumer. *J. Adv. Res.* **4** (December) 40–44.

Givon, Moshe, Dan Horsky. 1990. Untangling the effects of purchase reinforcement and advertising carryover. *Marketing Sci.* **9** (2) 171–187.

Grover, Rajiv, V. Srinivasan. 1987. A simultaneous approach to market segmentation and market structuring. *J. Marketing Res.* **24** (May) 139–153.

Guadagni, Peter M., John D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2** (Summer) 203–238.

Gupta, Sachin, Pradeep Chintagunta, Anil Kual, Dick R. Wittink. 1996. Do household scanner data provide representative inferences from brand choices: a comparison with store data. *J. Marketing Res.* **33** (November) 383–398.

Gupta, Sunil. 1988. Impact of price promotions on when, what and how much to buy. *J. Marketing Res.* **25** (November) 342–356.

Hoch, Stephen J., Xavier Dreze, Mary E. Purk. 1994. EDLP, Hi-Lo, and margin arithmetic. *J. Marketing* **58** (October) 16–27.

——, J., Byung-Do Kim, Alan M. Montgomery, Peter E. Rossi. 1995. Determinants of store-level price elasticity. *J. Marketing Res.* **32** (February) 17–29.

Hollis, Nigel S. 1997. Looking longer-term: how best to identify longer-term advertising effects. Presentation made at *The Advertising Research Foundation Conference*, New York, June.

Inman, J. Jeffrey, Leigh McAlister, Wayne D. Hoyer. 1990. Promotion signal: proxy for a price cut? *J. Consumer Res.* **17** (June) 74–81.

Jedidi, Kamel, Carl F. Mela, Sunil Gupta. 1999. Managing advertising and promotion for long-run profitability. *Marketing Sci.* **18** (1) 1–22.

Jones, John Philip. 1995. *When Ads Work: New Proof that Advertising Triggers Sales.* Simon and Schuster, New York.

——. 1998. Point of view: STAS and BehaviorScan—yet another view. *J. Adv. Res.* **38** (March/April) 51–53.

Kadiyali, Vrinda, Naufel Vilcassim, Pradeep Chintagunta. 1999. Product line extensions and competitive market interactions: an empirical analysis. *J. Econometrics* **89** (March–April) 339–364.

Kalwani, Manohar U., Donald G. Morrison. 1977. A parsimonious description of the Hendry system. *Management Sci.* **23** (January) 467–477.

——, Chi Kin Yim. 1992. Consumer price and promotion expectations: an experimental study. *J. Marketing Res.* **29** (February) 90–100.

Kalyanam, Kirthi, Thomas S. Shively. 1998. "Estimating irregular pricing effects: a stochastic spline regression approach." *J. Marketing Res.* **35** (February) 16–29.

Kalyanaram, Gurumurthy, John D. C. Little. 1994. An empirical analysis of latitude of price acceptance in consumer package goods. *J. Consumer Res.* **21** (December) 408–418.

Kamakura, Wagner A., Gary J. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* **26** (November) 420–434.

——, ——. 1993. Measuring brand value with scanner data. *Internat. J. Res. in Marketing* **10** (March) 9–22.

Kaul, Anil, Dick R. Wittink. 1995. Empirical generalizations about the impact of advertising on price sensitivity and price. *Marketing Sci.* **14** (3) G151–G160.

Krum, Franklin. 1994. Quantum leap. *Progressive Grocer* (January) 41–43.

Kurt Salmon Associates, Inc. 1993. *Efficient Consumer Response.* Food Marketing Institute, Washington, DC.

Lal, Rajiv, V. Padmanabhan. 1995. Competitive response and equilibria. *Marketing Sci.* **14** (3) G101–G108.

Leeflang, Peter S. H., Dick R. Wittink. 1996. Competitive reaction versus consumer response: do managers overreact? *Internat. J. Res. in Marketing* **13** (April) 103–120.

Leone, Robert P., Srini S. Srinivasan. 1996. Coupon face value: its impact on coupon redemptions, brand sales, and brand profitability. *J. Retailing* **72** (Fall) 273–289.

Little, John D. C. 1994. Modeling market response in large customer panels. Robert C. Blattberg, Rashi Glazer, John D. C. Little, eds., *The Marketing Information Revolution.* Harvard Business School Press, Boston, MA.

Lodish, Leonard M. 1997. Point of view: J. P. Jones and M. H. Blair on measuring advertising effects—another point of view. *J. Adv. Res.* **37** (Sept./Oct.) 75–79.

——. 1998. STAS and BehaviorScan—it's not just that simple. *J. Adv. Res.* **38** (March/April) 54–56.

——, Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetlein, Bruce Richardson, Mary Ellen Stevens. 1995. How T.V. advertising works: a meta-analysis of 389 real world split cable TV advertising experiments. *J. Marketing Res.* **32** (May) 125–139.

Mela, Carl F., Sunil Gupta, Donald R. Lehmann. 1997. The long-term impact of promotion and advertising on consumer brand choice. *J. Marketing Res.* **34** (May) 248–261.

Monroe, Kent B. 1990. *Pricing: Making Profitable Decisions.* McGraw-Hill, New York.

Montgomery, Alan. 1997. Creating micro-marketing pricing strategies using supermarket scanner data. *Marketing Sci.* **16** (4) 315–337.

Narasimhan, Chakravarthi. 1984. A price discrimination theory of coupons. *Marketing Sci.* **3** (Spring) 128–147.

——, Scott A. Neslin, Subrata K. Sen. 1996. Promotional elasticities and category characteristics. *J. Marketing* **60** (2) 17–30.

Narisetti, Raju. 1996. P&G ad chief plots demise of the coupon. *Wall Street J.* April 17, B1.

Neslin, Scott A. 1990. A market response model for coupon promotions. *Marketing Sci.* **9** (2) 125–145.

——, Caroline Henderson, John Quelch. 1985. Consumer promotions and the acceleration of product purchases. *Marketing Sci.* **4** (Spring) 147–165.

——, Greg Allenby, Andrew Ehremberg, Steve Hoch, Gilles Laurent, Robert Leone, John Little, Leonard Lodish, Robert Shoemaker, Dick Wittink. 1994. A research agenda for making scanner data more useful to managers. *Marketing Letters* **5** (4) 395–412.

——, Linda G. Schneider Stone. 1996. Consumer inventory sensitivity and the postpromotion dip. *Marketing Letters* **7** (January) 77–94.

Pedrick, James H., Fred S. Zufryden. 1991. Evaluating the impact of advertising media plans: a model of consumer purchase dynamics using single-source data. *Marketing Sci.* **10** (Spring) 111–130.

Rao, Ram C., Ramesh V. Arjunji, B. P. S. Murthi. 1995. Game theory and empirical generalizations concerning competitive promotions. *Marketing Sci.* **14** (3) G89–G100.

Rao, Vithala R., Darius J. Sabavala. 1981. Inference of hierarchical choice processes from panel data. *J. Consumer Res.* **8** (June) 85–96.

Raju, Jagmohan S., V. Srinivasan, Rajiv Lal. 1990. The effects of brand loyalty on competitive price promotional strategies. *Management Sci.* **36** (March) 276–304.

Richardson, Kevin, Simon Dratfield. 1997. Bringing new capabilities to marketing analysis. Presentation made at *The Advertising Research Foundation Conference,* New York, June.

Riskey, Dwight R. 1997. How T.V. advertising works: an industry response. *J. Marketing Res.* **34** (May) 292–293.

Scott, Doug R., Kirk Ward. 1997. But which half is working? Measuring the impact of marketing activity on business results. Presentation made at *The Advertising Research Foundation Conference,* New York, June.

Sethuraman, Raj. 1996. A model of how discounting high-priced brands affects the sales of low-priced brands. *J. Marketing Res.* **33** (November) 399–409.

Silva-Risso, Jorge M., Randolph E. Bucklin, Donald G. Morrison. 1999. A decision support system for planning manufacturers' sales promotion calendars. Marketing Sci. **18** (3) this issue.

Simonson, Itamar. 1990. The effect of purchase quantity and timing on variety-seeking behavior. *J. Marketing Res.* **27** (May) 150–162.

——, Russell S. Winer. 1992. The influence of purchase quantity and display format on consumer preference for variety. *J. Consumer Res.* **19** (June) 133–138.

Srinivasan, Srini S., Robert P. Leone, Francis J. Mulhern. 1995. The advertising exposure effect of free standing inserts. *J. Adv.* **24** (Spring) 29–40.

Tellis, Gerard J. 1988. The price elasticity of selective demand. *J. Marketing Res.* **25** (November) 331–341.

Thaler, Richard. 1985. Mental accounting and consumer choice. *Marketing Sci.* **4** (Summer) 199–214.

Von Gonten, Michael F., James F. Donius. 1997. Advertising exposure and effects: new panel-based findings. Presentation made at *The Advertising Research Foundation Conference,* New York, June.

Van Heerde, Harald, Peter S. H. Leeflang, Dick R. Wittink. 1999. The estimation of pre- and postpromotion dips with store-level scanner data. Working Paper, Yale School of Management, New Haven, CT.

Walsh, John W. 1995. Flexibility in consumer purchasing for uncertain future tastes. *Marketing Sci.* **14** (2) 148–165.

Wansink, Brian, Rohit Deshpande. 1994. Out of sight, out of mind:

pantry stockpiling and brand-usage frequency. *Marketing Letters* **5** (January) 91–100.
Wedel, Michel, Wagner A. Kamakura. 1998. *Market Segmentation: Conceptual and Methodological Foundations.* Kluwer Academic Publishers, Boston, MA.
Wittink, Dick R., Michael Addona, William Hawkes, John Porter. 1988. SCAN*PRO: the estimation, validation and use of promotional effects based on scanner data. Working Paper. AC Nielsen, Schaumburg, IL.
——, Phillipe Cattin. 1989. Commercial use of conjoint analysis: an update. *J. Marketing* **53** (July) 91–96.
——, Marco Vriens, Wim Burhenne. 1994. Commercial use of conjoint analysis in Europe: results and critical reflections. *Internat. J. Research in Marketing* **11** (January) 41–52.

This paper was received December 17, 1997, and has been with the authors 10 months for 1 revision; processed by Dick Wittink.