

Headline

What Scanner-Panel Data Tell Us about Advertising

Subhead

A Detective Story with a Dark Twist

Authors

Scott Koslow

Waikato Management School (*Te Raupapa*)

University of Waikato (*Te Whare Wānanga o Waikato*)

skoslow@mgt.waikato.ac.nz

Gerard J. Tellis

Marshall School of Business

University of Southern California

tellis@usc.edu

Author Bios

Scott Koslow is associate professor at the Waikato Management School (*Te Raupapa*), University of Waikato (*Te Whare Wānanga o Waikato*), Hamilton, New Zealand. His current research interests include advertising creativity, integration, strategy, and effectiveness. He also researches cross-cultural and international advertising, as well as consumer decision making and research methods. His research has appeared in the *Journal of Advertising Research*, the *Journal of Consumer Research*, the *Journal of Advertising*, the *Journal of the Academy of Marketing Science* and elsewhere. The *Journal of Advertising* gave him their Best Article Award.

Gerard J. Tellis is Jerry and Nancy Neely Chair in American Enterprise and Professor of Marketing, Marshall School of Business, University of Southern California. Gerard Tellis is an expert in innovation, advertising, global market entry, and new product growth. He has published four books and over 100 articles. His research has received 15 awards, including the Bass, Odell, Maynard (twice), and Vijay Mahajan award for lifetime contributions to marketing strategy. Professor Tellis is an Associate Editor of *Journal of Marketing Research* and on the editorial board of *Journal of Marketing*. He is a Trustee of the Marketing Science Institute, Distinguished Professor at Erasmus University and Senior Research Associate at Cambridge University. At USC, he directs the Center for Global Innovation.

Text

ABSTRACT

Scanner-panel data have long been important for understanding advertising's effects. Although the nature of advertising often has been investigated with scanner-panel data, the nature and value of scanner-panel data itself has rarely been considered. Although scanner-panel data is revered—even worshiped—by many, it has measurement issues like any other complex data set. While early estimates of advertising effectiveness from scanner-panel data may have appeared too low, some estimates are not vastly different from other data bases as can be ascertained from recent meta-analyses. But learning about the situations when, where, and how advertising does have large effects is critical and the future development of scanner-panel data does have a way to go to help answer these key questions. To make scanner-panel

data more powerful, we recommend that choice data sets be augmented to correct for their inherent weaknesses.

INTRODUCTION

Round up the usual suspects.

—CASABLANCA

Forget it, Jake. It's Chinatown.

—CHINATOWN

Scanner-panel data and advertising have long had an uneasy relationship. If they were characters in a *film noir* movie, advertising would be the beautiful, young *femme fatale* being investigated by the cynical, hard-boiled detective of scanner-panel data. Although the early years of scanner-panel data seemed to suggest that advertising clearly was an unfaithful wife of the brand—with hardly an advertising effect to be found—more recent evidence shows something more complex. Rather than being a fully objective private eye, scanner-panel data's character flaws make it certainly as complex a character as advertising itself.

Once machine-recorded scanner-panel data were thought of as the stuff dreams were made of—at least for researchers. But when early research failed to show the advertising-sales link, many dismissed advertising's value in favor of sales promotion. Although many researchers focus on the models and how well they estimate advertising effects, this research will focus primarily on the data and what we can learn from it.

Steven M. Shugan, in 2002, raised concerns that researchers often may be inappropriately “worshipping” information like scanner-panel data only because a respected secondary party (Information Resources Inc. [IRI] or Nielsen, for instance) collects it for practitioner use (Shugan, 2002). Instead, one should think of scanner-panel data as a source of a powerful dependent variable—household choices over time—that needs to be paired with other variables and appropriate models to predict those choices.

Occasionally, industry studies are designed to collect such key independent variables as split-cable studies, but other times, the results must be inferred when natural experiments are found in the data. To set the stage, one needs to go back to the origin of this research more than a half-century ago.

SUMMARY OF THE RESEARCH

Enter the Hero: Scanner-Panel Data

Before there was scanner-panel data, researchers still tried to understand and model consumer purchases histories. As early as the 1950s and 1960s, self-reported diaries were used, but the data they reported always seemed to exaggerate the share of the dominant brand at the expense of smaller ones: When consumers were asked to recall brands purchased, if the event was not easy to recall, then brand familiarity was used as a proxy for recalled behavior. These self-reported diaries typically were filled out on the day in the week they were due to be mailed back in, so purchases earlier in the week were the most prone to error. The advent of mechanical scanner readers meant this bias disappeared and led to greater confidence in using this data to understand how to manage brands. As the data evolved into single source data connecting advertising exposure to purchase, a strong understanding emerged regarding the data's strengths and weaknesses (Table 1).

<<Place Table 1 about here>>

The advent of reliable scanner-panel data had an almost revolutionary effect on perceived value of sales promotions versus advertising. The first things researchers began to see in the new scanner data were the pronounced effect sales promotions could have. Previously, brand managers had to rely on warehouse-withdrawal data, which did not pinpoint the timing of sales increases with much accuracy. What the new scanner-panel data showed was that sales spiked dramatically on any kind sales promotion (*e.g.*, cents off at the store shelf; coupons; or even an end-of-aisle display that signaled a temporary price cut) (Guadagni and Little 1983).

By contrast, advertising effects were hard to find. Initially, advertising data weren't even in data set. But, in time, these details became available in a separate file that could be merged into the purchase data. The first researchers couldn't find any advertising effects at all, and many of these early studies never were published. The first effort to find even a modestly significant effect generally was celebrated within the advertising academic community as a significant event (Tellis, 1988a). The *Wall Street Journal*, however, was not impressed with this finding and criticized the advertising community for even suggesting that advertising effects could be so small.

It was too late for such comments, because the lament within the advertising community already had begun. In 1992, David W. Stewart documented the angst and tried to offer consolation. And advertising's reputation continued to suffer. A 1995 analysis of split-cable studies seemed to be the nail in the coffin for advocates of advertising (Lodish *et al.*, 1995a). This work reviewed 389 real-world experiments comparing either a no-advertising condition or a low-advertising condition with a high-advertising condition. In few cases was advertising's effect seen as even statistically significant at the $\alpha=.8$ level; in half of those cases, these effects could have been produced by false positives.

Worse yet, these studies showed that, on average, a one-percent increase in advertising spending returned a .05 percent-increase in sales (Half the amount from the most recent meta-analyses of all estimates of advertising elasticity). With a break-even point generally acknowledged to be between .1 and .15 percent, the report may have been rather negative for advertising.

The revelations from scanner-panel data sent out a ripple in the package-goods industry—and a pronounced shift in spending from advertising to sales promotion (Tellis 1998).

In response, many advertising agencies reacted by diversifying into other types of marketing services, especially sales promotion. A widely circulated 1989 Shearson Lehman Hutton analysis detailed how the ultimate result of this diversification was the merger wave that created the six major holding companies: Omnicom, WPP, Interpublic, Publicis, Dentsu and Havas.

Once the major holding companies were formed, the integrated-marketing communications concept followed as a justification (Sasser, Koslow and Riordan, 2007). Observers conceded that power in package goods distribution had shifted from brands to retailers. Instead of being a loyal wife of the brand—to pick up the *film noir* analogy from the authors' Introduction—scanner-panel data exposed advertising as the *femme fatale*, a deceptively beautiful woman serving her own needs and no one else's. Sales promotions had been transformed as well, from a cheap tramp to a steady, reliable woman who whispers careful warnings in the ear of the detective.

Reactions To The Advertising/Sales Promotion Debate

Many refused to accept the conclusion that advertising's effects could be so small. For example, some analyses went back to the heart of advertising's creative roots and argued that

creative advertising must be effective (Kover, 1995; Kover, Goldberg and James, 1995). Others asked if advertising might have *lost* its creative roots and needed to return to them to be more effective (Reid, King and DeLorme, 1998). Some proposed that creative (and, hence, effective) advertising was something called on by brand managers only under duress (West and Berthon, 1997; West, 1999). In other words: it was possible advertising wasn't effective all the time, but it still could be powerful if clients and agencies would just be more creative.

Yet, many in advertising grudgingly did accept that it was inevitable that advertising had weak effects. These findings were supported by work in the consumer information-processing paradigm. Called the "weak-effects hypothesis," this approach argued that advertising can't be that powerful because most consumers don't think deeply enough about advertising to be profoundly persuaded by it.

Some questioned the new popularity of sales promotions by questioning their strategic basis. Excessive promotions that affect long-term expectations and reference pricing potentially could harm brands in the long run (Jedidi, Mela and Gupta, 1999), despite the fact that such offers could be optimized to maximize retailer (as opposed to brand) returns (*e.g.*, Tellis and Zufryden, 1995). The issues involved, however, concerned the empirical findings of whether or not there were some kind of pre- or post-promotions dips in purchase: Were consumers noticing the timing patterns discounts and strategically stockpiling? Or was there actually increased consumption? Some studies did investigate detailed reasons for dips, but found them varied and complex (Mace and Neslin, 2004).

Despite these concerns, considerable research did eventuate that deepened the understanding of advertising, with the articulation of 13 themes (See Table 2). These views are certainly diverse, some with opposing viewpoints. As Randolph E. Bucklin and Sunil Gupta discussed in 1999, many of the academics who accepted the new logic of the relative influences of promotion and advertising still found many unresolved issues in advertising's subtle effects. Yet practitioners considered many more issues on advertising still unresolved and often based these opinions on their own proprietary data and analysis.

<<Place Table 2 about here>>

Critics of the new advertising/sales promotion logic relied on different kinds of data to show advertising's effects. For example, one study offered a number of analyses of scanner-panel data that showed a more pronounced effect of advertising (Blair and Jones, 1996). The core of the critiques had less to do with the models used, but rather the kinds of data. Both Blair's and Jones' work relied on augmenting scanner-panel data with other types of data.

These issues boiled over in a rather nasty exchange in the normally genteel world of scholars (Blair and Rosenberg 1994; Jones 1995; 1998; Jones and Blair 1996; Lodish 1997; 1998): Blair and Jones argued that Lodish's work showed "poor data hygiene"; Lodish countered that Jones' models were subject to bias. Although this argument seemed to have simmered in the background for several years, the attack on scanner data's reputation as an objective detective was only the first round in what became much wider questioning of "our detective's" character—and, as a result, the value of advertising.

CRITICAL POINTS OF DEPARTURE

Questioning the "Detective": Scanner-Panel Data Revisited

A key to understanding scanner panel's inherent problems started to emerge as expanded computer power and more sophisticated models enabled richer analyses. In the 1980s, modeling the data demanded a sizeable computer. Not until the mid-1990s did researchers have the technology they needed to handle more than one category simultaneously.

An early model to address multiple product categories was created by Andrew Ainslie and Peter E. Rossi (1998). Their work showed that, across five categories, households with sensitivities to price, display, or feature advertising in one category also tended to show the same proclivities in the other categories. And, for the first time, scanner-panel data started to show fairly obvious relationships between household demographics and purchasing behavior.

On the face of it, such findings may hardly seem revolutionary. But, the larger issue related to modeling strategies focused on “unobserved heterogeneity.” Previous modelers were well aware that households were different—often so different that those differences could drive spurious results. In survey research we know that if different subpopulations use different parts of the scale, it could make it look like there were meaningful differences between subpopulations even though it is only response style differences. In scanner-panel data, if one compares different households—like brand loyal households to non-brand loyal ones—if one does not control for these heterogeneous preferences, modeling problems result. What Ainslie and Rossi did was to augment a single scanner-panel data set with information from others and, therefore, partially measure directly what previously had been unobservable.

Several other studies also tried to get at more direct measures of unobserved differences among households. For example, one data-enrichment approach combined conjoint data with scanner-panel data (Swait and Andrews, 2003). Although different households were used in each study, the authors were able to combine the two data sets successfully; their joint model significantly outperformed the pure scanner-panel data model. A key finding is that consumers react to new attributes and new attributes levels in ways that cannot be modeled with historical purchases alone. Consumers also react differently to alternative promotions. In-store promotions and pre-clipped coupons seem to have impact in proportion to the size of their discount, but in-store coupons have a larger effect than the discount alone.

Yet another study tried to understand the effects of sales promotions by including more data (van Heerde, Gupta and Wittink, 2003). Normally, the effects of sales promotions only looked at the brand promoted and they showed that three-quarters of the sales lift came from switches from other brands. These authors, however, expected that, if this phenomenon does hold true, then purchases of other brands must decrease. This strategy, again, augmented the data to provide a richer understanding of household behavior. When modeled in a more sophisticated way, the effects of a sales promotion on brand switching were shown to be far more modest: Brand switching, in fact, accounted for only a third of the lift.

One group of researchers took the research a step farther and used a questionnaire to determine household panel members’ preferences for brands and then integrated that information into a scanner-panel data model (Horsky, Misra, and Nelson, 2006). In doing so, the authors went to the heart of the unobserved-heterogeneity issue and directly measured it. Not only did the model fit better, but the model also suggested that most scanner-panel data models showed inflated sensitivities to prior purchases and price changes and tended to overfit what was observed in the data. If one took a more direct approach and used self-reported measures for brand preference and price sensitivity, the study argued, both have smaller effects than previously had been believed.

The implications are potentially profound. There is a view that coupons or sales promotions can be used to get consumers to try a brand and then consumers will become accustomed to it as they continue to buy it even post-promotion. Such logic may be faulty because some scanner-panel data models may overestimate the price sensitivity of consumers to try a brand and their inertia in continuing to use the same brand after a switch. But the core issue isn’t so much the implications for couponing or sales promotion as it is the underlying nature of scanner-panel data: All data have measurement and modeling challenges.

Character Flaws in Scanner-Panel Data

At first, accusations of flaws were limited to accuracy. But the problems have extended to other problems that are more fundamental. The benefit of hindsight is that Jones and Blair probably were right in pointing out data problems. Lodish and colleagues (Hu, Lodish, and Krieger, 2007) have replicated their analysis with data collected after 1995, which is known to be considerably cleaner. They found a much higher advertising/sales elasticity ratio on the order of .17, which is higher than the .1 found from a recent meta-analysis of all studies (Sethuraman, Tellis and Briesch 2010).

Scanner-panel data sets can only be as accurate as the choices that are in them. It literally is a record of purchase only so if some household purchases are not there for some reason like forgetting an identification card at home or the retailer does not participate in the collection program, those purchases can't be included. To figure out the shelf price of all options the household had at the time of purchase, one can infer this by listing price by UPC by store by date for all buyers in the data set. Normally, these inferences are good ones to make, but others can be more problematic. Coupon availability for a given household also has to be inferred from what others in the data set redeemed (Musalem, Bradlow and Raju, 2008). One also has to infer the consideration set of which products the household actually thought about purchasing before choosing selected brand. Some researchers think this does not introduce bias in estimates but others disagree (Bongers and Hofmeyr 2010; van Nierop et al 2010). Another issue is whether brand loyalty measures inferred from choices alone are good measures (Horsky, Misra and Nelson 2006). Ideally, managers and researchers want to know about the decision making that happened in the supermarket aisle, so as to know what consumers actually looked at and thought about just before they pulled a package off the shelf. The data collected, however, is only what appears at checkout and this is because there are often no real alternative measures from outside the purchase history.

The issue of accuracy became even more problematic when information from several sources—in particular television and print advertising—was combined with purchase data. Today, researchers can look to programs like Project Apollo (Wood, 2009), which accurately integrated several different media and consumer purchases. But the developers of Project Apollo were keenly aware that over a dozen other attempts to integrate data into what is now known as “single-source data” had problems correctly matching purchases with advertising. The trouble was that Project Apollo's accuracy came at a high cost, which apparently many client firms were not willing to underwrite indefinitely and the research program ended in 2008.

Although scanner-panel data collected more recently is more accurate in correctly matching advertisements to purchases than it had been when scanner-panel data was first collected, even if the data were recorded perfectly, the underlying structure of the data presents much more difficult problems. Martin Bongers and Jan Hofmeyr, in a 2010 *Journal of Advertising Research* paper entitled, “Why Modeling Averages Isn't Good Enough: A Critique of the Law of Double Jeopardy” present a strong critique of measurement assumptions underlying some scanner-panel data models. Many models assume that consumers' preferences don't change over long periods of time. For many consumers this is true, but for others, preferences do change and this is what advertisers are hoping to achieve with effective advertising. But because model assumptions don't match real consumers, the effects of advertising may be underestimated.

To go farther into the measurement issues raised by Bongers and Hoffmyer, consider two households. The first household buys Tide in week one. In week two, it buys Tide again. In

week three, Tide is the choice once more. In fact, for 14 weeks in a row, the household continues to buy the same UPC of Tide, and at no time buys any other brand. If one were trying to predict what this household would purchase in week 15, the logical guess would seem to be Tide. Yet if one wanted to understand the motivations that drove the first 14 weeks of purchase, this is not so obvious: There is no variance in the pattern of choice to given any firm clues. One might infer the household likes Tide. But the reasons could be rooted in some sort of deep psychological bond... or something so weak the purchasers weren't even aware of their regular behavior. For some product categories, up to half of all households have completely uniform purchase histories, so the pattern of this first household is not unusual.

Consider another household, however, which buys Tide in week one when it is on a cents-off sales promotion. In week two, the household selects Cheer with a coupon. In week three, it purchases Surf on a sales-promotion special. In week four, the choice is Gain, which is on end-cap display although the sales promotion wasn't supposed to start until the next day. In week five, the household buys Purex with a coupon. In brief: For 14 weeks, this second household continues to always purchase one box per week but only the brand that's supported with some kind of sales promotion. If one wanted to know why the household chose what they bought, the pattern would seem obvious as well—week in and week out, it bought on price. To predict what brand they would buy in the future, however, one would have to know what brand would be on sales promotion. For a researcher predicting sales in a holdout data set, this may not be problematic. But, for a brand manager trying to time her own promotional periods around that of competitors' program, it becomes complicated.

The comparison between these two hypothetical households highlights a unique feature of scanner-panel data: A tradeoff between the ability to predict and the ability to understand causation. The more regular the behavior, the easier it is to make predictions. But it remains difficult to understand why households act the way they do. A high level of variability that allows one to understand causality further obscures the ability to predict. This kind of tradeoff in scanner-panel data is unusual: In most valid data, the better one understands, the better one can predict.

While these flaws in the data may seem problematic, the data are superior to most alternate data used. Also, the strengths of the data far outweigh these flaws (see Table 1).

Coping with Character Flaws

Although *film-noir* detectives typically have character flaws that go unaddressed, the question in marketing research is what to do about the deficiencies (Shugan, 2002). Some conjoint modelers like Jordan J. Louviere refuse to model scanner-panel data. To understand why, compare what can be learned from scanner-panel data to the findings from choice conjoint studies (Elrod, Louviere and Davey, 1993; Oppewal, Louviere and Timmermans 1994; Horowitz and Louviere 1993).

In choice conjoint, researchers present consumers a series of choices specifically designed to identify why they are choosing one product over another. Their designs may involve two levels of five to seven variables and require 16, 32, 64 or more carefully constructed choices to extract their purchase preferences. If it is this hard to dissect purchase behavior under optimal conditions in a laboratory, then inferring preference from field conditions—where fewer choices are observed and those purchases are more regular—has to be even harder. There just isn't that much information that can be extracted from one purchase history to fully understand the behavior of most individual households. In Steven Shugan's 2002

analysis, the data are not inclusive. Instead, one must bring in other information—from heterogeneity assumptions to modeling structure—to extract what one can from this data. And these approaches can present challenges in both under- and over-estimation of effects (Tellis and Weiss, 1995).

A more pragmatic stance: One learns to cope with the data flaws. Five different modeling traditions take this pragmatic stance and have been used on scanner-panel data and the advantages and disadvantages of these are listed in Table 3. Regression has lost favor with many academics, but still is popular among practitioners. Multinomial logit models, while once popular with academics, has been less commonly used lately because of competition from alternative models. Of these newer models, the two major state-of-the-art methods are Bayesian estimation (Rossi and Allenby 2003) and structural models (Chintagunta et al 2006). Specifically:

<<Place Table 3 about here>>

- The value of Bayesian estimation is that it reduces the extent to which one “over-estimates” data with sparse observations, which is one of the problems mentioned above. That is, for households with highly regular purchases, Bayesian estimates “back off” making extreme estimates as well as characterizing the level of uncertainty in making such inferences. Bayesian estimation also is ideal for combining information from choice histories and unmatched questionnaire or experimental data.
- Structural models structure the decision processes of consumers and firms so as to model the outcomes in the data as optimal behavior of various agents. Because of these assumptions, structural models can allow one to extrapolate outside the range of data observed, overcoming the prediction dilemma mentioned above. For example, if some brands in a category never have advertised, structural models might give insights as to whether unadvertised brands might benefit from starting to advertise.

A rule of thumb in statistics should apply in these instances as well: If a data set is large and “clean,” one does not need overly complex models to estimate effects. With clean data, simpler models may perform well. However, scanner panel data are very rich, so simple means and regression can give biased results, because they gloss over effects at the micro level and return the biases from firms targeting of marketing mix (see Table 3). A certain degree of modeling sophistication is essential to recover causality correctly, estimate unbiased effects, and correctly model the disaggregate temporal and cross-sectional detail in the data.

THEORY IN PRACTICE

There are few areas in marketing where there is such an overlap between theory and practice as in the analysis of scanner-panel data. To properly analyze these data, it often takes a PhD to understand and apply challenging choice models. As a result, there is a great deal of overlap between market researchers and academic researchers. Professionals and academics often work together to explore this complex data, and this tradition goes back to some of the first scanner-panel articles (Guadagni and Little, 1983).

While data are richer enough that one can test behavioral hypotheses in scanner panel data (Tellis 1988; MacInnis et al 2002), the complexity and limitations of the data may limit large scale theory testing. Although researchers can find empirical generalizations like those relating to price elasticities (Bijmolt, van Heerde, and Pieters, 2005; Tellis 1998b), there has so far been limited theory testing with scanner panel data. Some do combine the findings

from scanner-panel data models with other types of research to come to definite and valuable contributions and insights about advertising (Sethuraman, Tellis, and Briesch, forthcoming; Tellis 1989; 2005; 2009). Yet to understand marketing theory better and thereby help practice, one normally needs control over the design of studies—and with scanner-panel data, those doing the deciding are practitioners who need to be willing to pay for field experiments. For this purpose, academics need to work closely with practitioners to design field experiments that can generate scanner panel data to test behavioral theories.

THE NEXT STEPS

The Usual Suspects

Character flaws in a detective are not always disastrous—Marlowe still got to the bottom of things—but the more important issue, in this instance, is that scanner data tend to focus one on the usual suspects of research questions rather than address what managers really want to know. Scanner-panel data primarily have been used for pricing issues, particularly for sales promotions. Though researchers and managers alike want to move beyond such a tactical orientation with scanner-panel research, movements in this direction have been tentative at best—and for good reason: The variables with the best accuracy are price variables.

Other “usual suspects” in scanner-panel research include advertising-sales elasticity, sales response curves, advertising decay, and media scheduling.

It still is possible to find a variety of natural experiments imbedded in scanner-panel data. For example, new products appear in these data sets with some regularity and, when they do appear, can be examined (Ataman, Mela, and van Heerde, 2008). Researchers, however, mostly have access to data designed for other purposes and that means that they can use the data to answer marketing and advertising research questions only on an opportunistic basis. Thus, there is a great opportunity to test the numerous theories developed in consumers behavior within the framework of scanner panel data. For this purpose, academics need to design field experiments with the collaboration of practitioners. To do so, academics can make use of the fact that ads of different design can be delivered to matched households under varying experimental situations. The great advantage of these experiments is that they can test theory in real life market contexts.

Going Beyond the Usual Suspects

Marketing researchers do know how to model consumer choice, but they can only do as good a job as the data allow them. And good data need to be inclusive of the effects and that information need to be extractable (Shugan, 2002). Instead of haphazardly relying on finding natural experiments in data, researchers also need to have questions guide research—and then find or collect the data the researchers’ needs. To go beyond the usual suspects, there are four main weaknesses in scanner-panel data that need to be overcome: Content, carryover, media, and markets:

- The creative content of ads may be the most important factor affecting response (Lodish et al 1995a; Tellis, Chandy and Thaivanich 2000). For example, in a 2009 *Journal of Advertising Research* offering, “Short-Term Effects of Advertising: Some Well-Established Empirical Law-Like Patterns,” Leslie Wood showed that the quality of the creative execution has the largest effect on purchase—in fact, more than 10 times more influence than the next closest factor studied.

Tellis, Chandy and Thaivanich (2000) developed a model to show how individual creatives affected advertising response. The authors worked with a category that used 70 different creatives over different cities. While their analysis was not based on scanner data, the model can easily be adapted. For this purpose, keeping track of individual creatives that are used in scanner panel data would be very important. Furth, Tellis, Chandy, MacInnis, and Thaivanich 2005 analyzed what aspects of ad content were most important. Their instrument for rating ad content can be applied to ads in scanner data. Such analyses can be very useful when designing future ads.

For this purpose, when brand managers evaluate advertising in a split cable test, they also need multiple campaigns to compare why some advertising content works better. Instead of using only one campaign in a split-cable test, more than one might be needed. To develop advertisements with the largest effects—presumably the most creative ones—brand managers also need to be more open to exploring a wider range of creative work (Koslow, Sasser and Riordan 2006).

- Researchers have now proposed models to evaluate the roles of media and time of day (e.g., Tellis Chandy and Thaivanich 2000). A more fundamental understanding is needed on how media interact with one another and with time of the day to influence choice. Joan Fitzgerald, in “Evaluation Return on Investment of Multimedia Advertising with a Single-Source Panel: A Retail Case Study” (2004), tied media exposure to retail visits.

Although some data sets like Project Apollo’s has detailed information on multiple media (Wood, 2009), the split-cable data in general does not provide as detailed information on the timings to examine media schedules. More detail on timings needs to be collected and appropriately modeled.

- Carry-over effects often are difficult to assess due to a variety of technical problems (Tellis and Weiss 1995). Most importantly, aggregate data are known to inflate advertising effects. Yet many researchers still use aggregate temporal data estimate advertising effects because of the ease of working with aggregate data. The greatest loss is when researchers take aggregate already disaggregate scanner data for convenience.
- Analysis is limited to only a markets: a few advertisements and test sites. While this may be good from the perspective of experimental design, there often are generalizability issues because these test sites frequently are complex environments. Even worse, Wal-Mart, most notably, never has participated in scanner-panel data collection by the major research firms, so households who purchase—in full or in part, at Wal-Mart—have escaped analysis. To have truly generalizable findings, the “Wal-Mart effect” needs to be understood.

CONCLUSION

Film noir has an undeserved reputation for unhappy endings. But there is no reason why the relationship between scanner-panel data and advertising needs to be negative. The rough initial relationship between scanner-panel data and advertising has matured into a more realistic one. A third of a century’s experience with scanner-panel data suggests that researchers will find that sometimes advertising does not work, but sometimes it does.

Scanner-panel data, however, are merely one set of players in a larger drama of consumer decision-making. And although scanner-panel data are good sources of an excellent dependent realistic variable, household choices over time, the full story still requires a full cast of independent variables.

Initially, the most popular independent variables were pricing related. But the extent to which one can augment this data to include variables like media, creatives, ad content, and theoretically design conditions makes scanner-panel data much more powerful. But to go further will require researchers to spend much more time augmenting the data. Single-source data like Project Apollo, which tried to pull together advertising and choice measures, obviously are valuable, but the cost was prohibitive. Researchers—academics, especially—need to make their own opportunities.

The lesson: Advertising may still at times play the *femme fatale*, but other times she can become the heroine as well—and those are the times we need to understand better.

Table 1

Advantages and Limitations of Single Source Data

Advantages	Limitations
<ul style="list-style-type: none">• Observed data: eliminates problems that plague survey data like social desirability or demand effects.• Temporal disaggregation: can track effects on several levels, by the week day, minute and even second, which allows one to separate out current effects from carry-over effects as well as estimate wear-in and wear-out.• Cross-sectional disaggregation: can compare different types of households and control for their effects, for example, heavy versus light users, brand loyals versus non-loyals, etc.• Purchase disaggregation: can break down a single purchase into four components: timing choice, store choice, brand choice and quantity choice.• Market structure: can reveal how households switch between brands and if these purchases are for the same uses, then which brands compete most intensely.• Segmentation analysis: can show how households differ in patterns in brands purchased or in responses to the marketing mix.	<ul style="list-style-type: none">• Household representation: despite efforts to keep panels representative of aggregate USA demographics, households that tend to be more cooperative tend to show up more in panels, <i>e.g.</i>, more widows, fewer young, mobile singles, and more households sensitive to the discount incentives use to reward panel members.• Store representation: some discount stores refuse to cooperate with the major scanner-panel data collections firms, <i>e.g.</i>, Wal-Mart, Costco and Sam's Club.• Advertising representation: purchases are usually tied to either TV meter data or print in the form of store feature advertising.

Table 2
Themes in Advertising Research Using Scanner-Panel Data

Theme	Selected Citations
Advertising's short-run effects are on average small—but highly variable.	Tellis (1988a) Pedrick and Zufryden (1993) Lodish et al (1995a) Tellis and Weiss (1995) MacInnis, Rao and Weiss (2002)
Advertising's long-run effects may be larger.	Givon and Horsky (1990) Jones (1995) Lodish et al (1995b) Mela, Gupta and Lehmann (1997)
Advertising's effects are smaller than promotions' in the short run, but larger than promotions' in the long run.	Jedidi, Mela and Gupta (1999) Pauwels, Hanssens and Siddarth (2002) Ataman, van Heerde and Mela (2010)
Price discounting of various forms may be helpful in the short run, but lead to reduced loyalty and higher price sensitivity in the long run—problems that need to be fixed by advertising or new product activity.	Papatla and Krishnanurthi (1996) Jedidi, Mela and Gupta (1999) Ailawadi, Lehmann and Neslin (2001) Ailawadi, Gedenk, Lutzky and Neslin (2007) Ataman, van Heerde and Mela (2010)
Unlike promotion, advertising does not seem to increase consumption rates.	Neslin and Shoemaker (1989) Deighton, Henderson and Neslin (1994)
Advertising can interact with other marketing mix variables.	Balachander and Ghose (2003) Lemon and Nowlis (2002) Steenkamp et al (2005) Zhang (2006)
Advertising seems to work by expanding the number of users or by building on loyalty.	Tellis (1988a) Aliawadi, Lehmann and Neslin (2001)
One role of advertising is to reinforce quality and justify a higher price.	Kanetkar, Weinberg and Weiss (1992) Mela, Gupta and Lehmann (1997) Erdem and Sun (2002) Erdem, Keane and Sun (2008) Mehta, Chen and Narasimhan (2008)
Media plans matter.	Pedrick and Zufryden (1991) Fitzgerald (2004) Taylor, Kennedy and Sharp (2009)
Maintaining share of voice in advertising is important.	Pedrick and Zufryden (1991) Danaher, Bonfrer and Dhar (2008)
The effectiveness of advertising should be viewed on a segment-by-segment basis.	Zufryden, Pedrick and Sankaralingam (1993) Garretson and Burton (2003) Zhang (2006)
Over time, advertising effects measured with scanner-panel data have increased in size, while the effectiveness measures taken with other methods have decreased; as a result, current advertising elasticity measures using either method have come to have greater agreement and seem about the same across methods.	Jones and Blair (1996) Ainslie and Rossi (1998) Hu, Lodish and Krieger (2007) Hu et al (2009)

The content of advertising is paramount.	Blair and Rosenberg (1994) MacInnis, Rao and Weiss (2002) Blair and Kuse (2004) Mehta, Chen and Narasimhan (2008) Wood (2009)
--	---

Table 3
Advantages and Limitations of Models Estimating Advertising Effects in Scanner-Panel Data

Modeling Tradition	Advantages	Limitations
Regression	<ul style="list-style-type: none"> • Simple and easy to understand. • Frequently used by industry. 	<ul style="list-style-type: none"> • Dynamics are ignored, <i>e.g.</i>, advertising's effect on quality or price inferences. • Can't deal with the direction of causality, that is, does advertising cause sales or sales cause advertising. • Hard to use to make strategy implications because most predictions are outside the range of data.
Time Series	<ul style="list-style-type: none"> • Can cope with dynamics in the data like carryover effects, wear-in or wear-out. • Can avoid reverse causality problems. 	<ul style="list-style-type: none"> • Can get computationally complex, especially with many lagged variables. • To measure current advertising effects require very short time interval, but this reduces the size of advertising effect. • When time intervals are very short, the data and computational burden can be enormous.
Multinomial Logit	<ul style="list-style-type: none"> • Allows for cross-sectional disaggregation minimizing biases from data aggregation. • Allows for temporal aggregation analysing choice at the moment they are made. • Accounts for differences in household advertising exposures. • Market structure analysis can be performed using cross-advertising elasticities of brands. • Segmentation can be performed using advertising response elasticities. 	<ul style="list-style-type: none"> • Data are hard to set up with significant, painstaking data cleaning required. • More complex estimation than regression or time series models. • Subject to the IIA assumption that choice alternatives are independent, which rarely happens. • Strategy implications are difficult because the models assume only minor shifts in strategy will occur—one is typically outside the range of the data for major strategy shifts.
Structural Models	<ul style="list-style-type: none"> • Add in realistic economic theory to models, <i>e.g.</i>, promotion increasing one brand's sales should cause competing brand sales to fall. • Can infer some unobserved issues like marginal cost or competitive conduct. 	<ul style="list-style-type: none"> • Computationally intensive requiring sophisticated optimization techniques. • Although these models aim to make few assumptions, in practice they may make more, overly simple assumptions. • Hard to standardized across product categories.

	<ul style="list-style-type: none"> • Can deal with endogeneity concerns like rational consumers who anticipate promotions exploit them for stockpiling. • Good for out of sample predictions, in the future or after major strategic changes. 	<ul style="list-style-type: none"> • Hard to use outside the situation under investigation because theoretical assumptions can break down in practice.
Bayesian Methods	<ul style="list-style-type: none"> • Allow for household level estimates of responses to price, promotion and advertising. • Can incorporate prior information, especially where there are fewer observations. • Can estimate difficult models involving high dimensional integration when close-form likelihoods functions cannot be obtained. • Allow flexibility in model formulation when there are many parameters and limited data. 	<ul style="list-style-type: none"> • Computationally intensive, and in some cases estimates are not reliable. • Selection of priors can be abused by the researcher, especially when priors are subjective.

References

- Ailawadi, Kusum L., Donald R. Lehmann and Scott A. Neslin (2001), "Market Response to a Major Policy Change in the Marketing Mix: Learning from Proctor & Gamble's Value Pricing Strategy," *Journal of Marketing*, 65 (January), 44-61.
- Ailawadi, Kusum L., Karen Gedenk, Christian Lutzky and Scott A. Neslin (2007), "Decomposition of the Sales Impact of Promotion-Induced Stockpiling," *Journal of Marketing Research*, 44(August), 450-467.
- Ainslie, Andrew and Peter E. Rossi (1998), "Similarities in Choice Behavior Across Product Categories," *Marketing Science*, 17(2), 91-106.
- Ataman, M. Berk, Harald J. van Heerde and Carl F. Mela (2010), "The Long-Term Effect of Marketing Strategy on Brand Sales," *Journal of Marketing Research*, 47(October), 866-882.
- Balachander, Subramanian and Sanjay Ghose (2003), "Reciprocal Spillover Effects: A Strategic Benefit of Brand Extensions," *Journal of Marketing*, 67 (January), 4-13.
- Blair, Margaret Henderson and Karl E. Rosenberg (1994), "Convergent Findings Increase Our Understanding of How Advertising Works," *Journal of Advertising Research*, (May/June) 35-45.
- Bodapati, Anand V. (2008), "Recommendation Systems with Purchase Data," *Journal of Marketing Research*, 45 (February), 77-93.
- Bongers, Martin and Jan Hofmeyr (2010), "Why Modeling Averages Isn't Good Enough: A Critique of the Law of Double Jeopardy," *Journal of Advertising Research*, 50(3), 323-333.
- Bucklin, Randolph E. and Sunil Gupta (1999), "Commercial Use of UPC Scanner Data: Industry and Academic Perspectives," *Marketing Science*, 18(3), 247-273.
- Chintagunta, Pradeep, Tulin Erdem, Peter E. Rossi and Michel Wedel (2006), "Structural Modeling in Marketing: Review and Assessment," *Marketing Science*, 25(6), 604-616.
- Danaher, Peter J., Andre Bonfrer and Sanjay Dhar (2000), "The Effect of Competitive Advertising Interference on Sales of Packaged Goods," *Journal of Marketing Research*, 45(April), 211-225.
- Deighton, John, Carline M. Henderson and Scott A. Neslin (1994), "The Effects of Advertising on Brand Switching and Repeat Purchasing," *Journal of Marketing Research*, 31(February), 28-43.
- Elrod, Terry, Jordan J. Louviere and Krishnakumar S. Davey (1992), "An Empirical Comparison of Ratings-Based and Choice-Based Conjoint Models," *Journal of Marketing Research*, 29(August), 368-77.
- Erdem, Tulin and Baohong Sun (2002), "An Empirical Investigation of the Spillover Effects of Advertising and Sales Promotion in Umbrella Branding," *Journal of Marketing Research*, 39(November), 408-420.
- Erdem, Tulin, Michael P. Keane and Baohong Sun (2008), "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality," *Marketing Science*, 27 (6), 1111-1125,
- Fitzgerald, Joan (2004), "Evaluation Return on Investment of Multimedia Advertising with a Single-Source Panel: A Retail Case Study," *Journal of Advertising Research*, (September), 262-270.
- Garretson, Judith A. and Scot Burton (2003), "Highly Coupon and Sale Prone Consumers: Benefits Beyond Price Savings," *Journal of Advertising Research*, June, 162-172.
- Givon, Moshe and Dan Horsky (1990), "Untangling the Effects of Purchase Reinforcement and Advertising Carryover," *Marketing Science*, 9(2), 171-187.

- Guadagni, Peter M. and John D.C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2(3), 203-238.
- Horowitz, Joel J. and Jordan J. Louviere (1993), "Testing Predicted Choices against Observations in Probabilistic Discrete-Choice Models," *Marketing Science*, 12(3), 270-279.
- Horsky, Dan, Sanjog Misra and Paul Nelson (2006), "Observed and Unobserved Preference Heterogeneity in Brand-Choice Models," *Marketing Science*, 25(4), 322-335.
- Hu, Yu, Leonard M. Lodish and Abba M. Krieger (2007), "An Analysis of Real World TV Advertising Tests: A 15-Year Update," *Journal of Advertising Research*, 47(September), 341-353.
- Hu, Yu, Leonard M. Lodish, Abba M. Krieger and Babak Hayati (2009), "An Update of Real-World TV Advertising Tests," *Journal of Advertising Research*, June, 201-206.
- Jedidi, Kamel, Carl F. Mela and Sunil Gupta (1999), "Managing Advertising and Promotion for Long-Run Profitability," *Marketing Science* 18(1), 1-22.
- Jones, John Philip (1995), "Single-Source Research Begins to Fulfill Its Promise," *Journal of Advertising Research*, (May/June), 9-16.
- Jones, John Philip (1998), "Point of View: STAS and BehaviorScan—Yet Another View," *Journal of Advertising Research*, (March/April), 51-53.
- Jones, John Philip and Margaret H. Blair (1996), "Examining 'Conventional Wisdoms' about Advertising Effects with Evidence from Independent Sources," *Journal of Advertising Research*, (November/December) 37-59.
- Kanetkar, Vinay, Charles B. Weinberg and Doyle L. Weiss (1992), "Price Sensitivity and Television Advertising Exposures: Some Empirical Findings," *Marketing Science*, 11 (4), 359-371.
- Kover, Arthur J. (1995), "Copywriters' Implicit Theories of Communication: An Exploration," *Journal of Consumer Research*, 21 (March), 596-611.
- Kover, Arthur J., Stephen M. Goldberg and William L. James (1995), "Creativity Vs. Effectiveness? An Integrative Classification for Advertising," *Journal of Advertising Research*, 35(November/December), 29-38.
- Koslow, Scott, Sheila L. Sasser and Edward A. Riordan (2006), "Do Marketers Get the Advertising They Need or the Advertising They Deserve? Agency Views of How Clients Influence Creativity," *Journal of Advertising*, 35 (Fall), 81-101.
- Lemon, Katherine N. and Stephen M. Nowlis (2002), "Developing Synergies Between Promotions and Brands in Different Price-Quality Tiers," *Journal of Marketing Research*, 39(May), 171-185.
- Lodish, Leonard M. (1997), "Point of View: J.P. Jones and M.H. Blair on Measuring Advertising Effects—Another Point of View," *Journal of Advertising Research*, (September/October), 75-79.
- Lodish, Leonard M. (1998), "STAS and BehaviorScan—It's Just Not That Simple," *Journal of Advertising Research*, (March/April), 54-56.
- Lodish, Leonard M., Magid Abraham, Stuart Kalmenson, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson and Mary Ellen Stevens (1995a), "How T.V. Advertising Works: A Meta-Analysis of 389 Real World Split Cable T.V. Advertising Experiments," *Journal of Marketing Research*, 32(May) 125-139.
- Lodish, Leonard M., Magid M. Abraham, Jeanne Livelsberger, Beth Lubetkin, Bruce Richardson and Mary Ellen Stevens (1995b), "A Summary of Fifty-Five in-Market Experimental Estimates of the Long-Term Effect of TV Advertising," *Marketing Science*, 14(3), G133-G140.

- Mace, Sandrine and Scott A. Neslin (2004), "The Determinates of Pre- and Postpromotion Dips in Sales of Frequently Purchased Goods," *Journal of Marketing Research*, 41(August), 339-350.
- MacInnis, Deborah J., Ambar G. Rao and Allen M. Weiss (2002), "Assessing When Increased Media Weight of Real-World Advertisements Helps Sales," *Journal of Marketing Research*, 39(November), 391-407.
- Mehta, Nitin, Xinlei (Jack) Chen and Om Narasimhan (2008), "Informing, Transforming, and Persuading: Disentangling the Multiple Effects of Advertising on Brand Choice Decisions," *Marketing Science* 27 (3), 334-355.
- Mela, Carl F., Sunil Gupta and Donald R. Lehmann (1997), "The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice," *Journal of Marketing Research*, 34(May), 248-261.
- Musalem, Andres, Eric T. Bradlow and Jagmonhan S. Raju (2008), "Who's Go the Coupon? Estimating Consumer Preferences and Coupon Usage from Aggregate Information," *Journal of Marketing Research*, 45(December), 7-15-730.
- Neslin, Scott A. and Robert W. Shoemaker (1989), "An Alternative Explanation for Lower Repeat Rates After Promotion Purchases," *Journal of Marketing Research*, 26(May), 205-13.
- Newstead, Kate, Jennifer Taylor, Rachel Kennedy and Bryon Sharp (2009), "The Total Long-Term Sales Effects of Advertising: Lessons from Single Source," *Journal of Advertising Research*, (June) 207-210.
- Oppewal, Harmen, Jordan J. Louviere and Harry J. P. Timmermans (1994), "Modeling Hierarchical Conjoint Processes with Integrated Choice Experiments," *Journal of Marketing Research*, 31 (February), 92-105.
- Papatla, Purushottam and Lakshman Krishnamurthi (1996), "Measuring the Dynamic Effects of Promotions on Brand Choice," *Journal of Marketing Research*, 33(February), 20-35.
- Pauwels, Koen, Dominique M. Hanssens and S. Siddarth (2002), "The Long-Term Effects of Price Promotions on Category Incidence, Brand Choice, and Purchase Quantity," *Journal of Marketing Research*, 39(November), 421-439.
- Pedrick, James H. and Fred S. Zufryden (1991), "Evaluating the Impact of Advertising Media Plans: A Model of Consumer Purchase Dynamics Using Single-Source Data," *Marketing Science*, 10(2), 111-130.
- Pedrick, James H. and Fred S. Zufryden (1993), "Measuring the Competitive Effects of Advertising Media Plans," *Journal of Advertising Research*, (November/December), 11-20.
- Reid, Leonard N., Karen Whitehill King and Denise E. DeLorme (1998), "Top-Level Agency Creatives Look at Advertising Creativity Then and Now," *Journal of Advertising*, 27(2), 1-16.
- Rossi, Peter E. and Greg M. Allenby (2003), "Bayesian Statistics and Marketing," *Marketing Science*, 22(3), 304-328.
- Sasser, Sheila L., Scott Koslow and Edward A. Riodan (2007), "Creative and Interactive Media Use of Agencies: Engaging and IMC Media Palette of Implementing Advertising Campaigns," *Journal of Advertising Research*, (September), 237-256.
- Sethuraman, Raj, Gerard J. Tellis and Richard A. Briesch (forthcoming), "How Well Does Advertising Work? Generalizations from Meta-analysis of Brand Advertising Elasticities," *Journal of Marketing Research*.
- Shearson Lehman Hutton (1989), *Consumer Advertising Report: Diversification Begets Integration*. (New York: American Express Publication).

- Shugan, Steven M. (2002), "In Search of Data: An Editorial," *Marketing Science*, 21 (4), 369-377.
- Stewart, David W. (1992), "Speculations on the Future of Advertising Research," *Journal of Advertising*, 21(September), 1-18.
- Stewart, David W. and David Furse (1986), *Effective Television Advertising: A Study of 1000 Commercials*. New York: Lexington Books.
- Stewart, David W. and David H. Furse (1984/1985), "Analysis of the Impact of Executional Factors on Advertising Performance," *Journal of Advertising Research*, (December/January), 23-26.
- Stewart, David W. and Scott Koslow (1989), "Executional Factors and Advertising Effectiveness: A Replication," *Journal of Advertising*, 18(3), 21-32.
- Swait, Joffre and Rick L. Andrews (2003), "Enriching Scanner Panel Models with Choice Experiments," *Marketing Science*, 22(4), 442-460.
- Tellis, Gerard J. (1988a), "Advertising Exposure, Loyalty, and Brand Purchase: A Two-Stage Model of Choice," *Journal of Marketing Research*, 25 (May), 134-44.
- Tellis, Gerard J. (1988b), "The Price Elasticity of Selective Demand: A Meta-Analysis of Economic Models of Sales," *Journal of Marketing Research*, 25(4), 331-341.
- Tellis, Gerard J. (1989), "Point of View: Interpreting Advertising and Price Elasticities," *Journal of Advertising Research*, (August/September), 40-43.
- Tellis, Gerard J. (2005), "Advertising's Role in Capitalist Markets: What Do We Know and Where Do We Go from Here?" *Journal of Advertising Research*, (June), 162-170.
- Tellis, Gerard J. (2009), "Generalizations about Advertising Effectiveness in Markets," *Journal of Advertising Research*, 49(2), 240-245.
- Tellis, Gerard J. and Doyle L. Weiss (1995), "Does TV Advertising Really Affect Sales? The Role of Measures, Models and Data Aggregation," *Journal of Advertising*, 24 (3) 1-12.
- Tellis, Gerard J. and Fred S. Zufryden (1995), "Tackling the Retailer Decision Maze: Which Brands to Discount, How Much, When and Why?" *Marketing Science*, 14(3), 271-299.
- Van Nierop, Erjen, Bart Bronnesberg, Richard Paap, Michel Wedel and Philip Hans Franses (2010), "Retrieving Unobserved Consideration Sets from Household Panel Data," *Journal of Marketing Research*, 47(February), 63-74.
- West, Douglas C. (1999), "360 of Creative Risk," *Journal of Advertising Research*, 39(January/February), 39-50.
- West, Douglas C. and Pierre Berthon (1997), "Antecedents of Risk-Takeing Behavior by Advertisers: Empirical Evidence and Management Implications," *Journal of Advertising Research*, 37(September/October), 27-40.
- Wood, Leslie (2009), "Short-Term Effects of Advertising: Some Well-Established Empirical Law-Like Patterns," *Journal of Advertising Research*, June, 186-192.
- Zhang, Jie (2006), "An Integrated Choice Model Incorporating Alternative Mechanisms for Consumers' Reactions to In-Store Display and Feature Advertising," *Marketing Science*, 25(3), 278-290.
- Zufryden, Fred S., James H. Pedrick and Avu Sankaralingam (1993), "Zapping and Its Impact on Brand Purchase Behavior," *Journal of Advertising Research* (January/February), 58-66.