Overcontrol in Advertising Experiments

Paul W. Farris and David J. Reibstein

Market A is selected as the test city. Three levels of advertising "frequency" are to be tested by beaming corresponding numbers of commercials to three groups of randomly selected households. Sales to the three groups are to be measured with either diary panel data or scanner data. Since market A has been determined to be representative of the total area under consideration, the three levels of purchase probability associated with the levels of advertising are to be used in constructing an advertising-sales response function for budgeting purposes.

According to many (Simon and Arndt, 1980, for example), such an experiment manifests the state of the art for assessing the effects of advertising on sales. This paper argues that such state-of-the-art experiments overcontrol the environment, and hence may mismeasure the effects of advertising. More specifically, a single-market experiment may control "away" effects of advertising. The net result is a bias which underestimates the effects of advertising. Introducing a clear bias in an experimental design, it seems to us, can only be defended if one has an idea of how serious the bias is and how to correct for it.

---

**External Validity versus Control**

Some noted authorities on experiment design have expressed a concern that "overcontrolling" can undermine the external validity of a study (Campbell, 1957; Campbell and Stanley, 1966). "Overcontrol" results when the design of the experiment cuts off naturally occurring feedback and second-order effects.

These threats on external validity apply to macro-measurement in marketing. In particular, we are concerned about the inferences drawn from field experiments about advertising's impact.

With the advent of split-cable systems, advertising field experiments have increased in popularity (Aaker and Carman, 1982). In a simplified form, the split-cable design implies Model I which is represented in Figure 1A. That is, advertising exposures lead to consumer demand which leads to sales. By measuring the effects of various exposure levels (advertising frequency) on consumer demand, it is assumed that one can approximate the effect of advertising on sales in order to establish advertising budgets. This is the most common use of the split-cable system according to Aaker and Carman (1982).

Model 2, represented in Figure 1B, depicts a slightly different view of advertising's effects which include the role of the retailer, the effects of advertising on distribution, and distribution's effect on sales. It assumes that increased consumer demand will lead to increased availability and that still another increase in sales can be expected from increased availability. (Availability, as will be explained...
shortly, includes more than the number and size of outlets carrying a brand.

Model 2 reflects two assumptions that the authors believe are valid, but which will probably be most applicable to products with intensive distribution strategies and to new products which have not yet reached full distribution potential. Marketing literature provides substantial support for the assumptions advanced.

**Assumptions.** Assumption 1: Increases in distribution (availability) of the product will be associated with increases in average individual purchase probability (e). Simply stated, if a product is difficult to find, fewer people will purchase it and those purchasing it will purchase it less often. Distribution and availability are loosely defined here, but include the number of stores carrying a product, the number of stockouts, the number of sizes carried, and the prominence of shelving and display. (Retail advertising of a product and/or feature pricing are probably also relevant, but are not essential considerations.)

This first assumption is most likely to hold true for “convenience” goods, which consumers like to find available in many outlets. Even for specialty or shopping goods, for which consumers are apt to expend more effort in finding, extremely low levels of distribution will be associated with lower purchase probabilities.

The increases in sales from increased distribution need not come only from additional “impulse” purchases, but may also come from brand-choice decisions in the product category. The POPAI/DuPont study indicates that up to 50 percent of purchases are planned by category, but not brand. The average for all products investigated was 22.5 percent. Only 41 percent of all purchases were planned with respect to both category and brand (Advertising Age, 1982).

In a review of 12 studies on this topic, Leone and Schultz (1980) conclude, “the empirical research in this area is in general agreement that there is a positive relationship between shelf-space and unit sales.” The number of outlets carrying a brand has also been shown to be positively related to sales by the work of Parsons (1974) and Lambin (1972).

It should be noted that even the leading brands in convenience-goods categories rarely if ever attain full “availability.” When one considers distribution, shelving, and stockouts for all sizes and product variations (colors, flavors, etc.), there is usually room for growth.

Assumption 2: Retailers react to increases in consumer demand by being more likely to carry the product, display it prominently, keep stockouts down, and carry more sizes (f). They may also react to the level of a product’s advertising (independent of the advertising’s effect on demand) in the same ways (d). Thus, product distribution and availability might be affected by advertising in two ways. The

Paul W. Farris is professor of business administration at The Colgate Darden Graduate School of Business Administration, University of Virginia, where he teaches advertising management and marketing. He holds a B.S. in business economics from the University of Missouri, a M.B.A. from the University of Washington, and a D.B.A. from the Harvard Business School. His work experience includes assignments in account management for the Lintas advertising agency and product management for Unilever.

first effect of advertising would result from increases in demand caused by the advertising and the second from demand anticipated by the retailer, because of the level of advertising being done. For the purpose of this article either of the two effects caused by advertising would increase demand, although the second might occur more quickly. (In some cases, when the consumer advertising campaign is promoted to the trade, such effects could occur before the campaign has been implemented.)

Haines and Silk (1967) state that advertising in consumer media affects retailers' decisions to stock the advertised product and to promote it themselves, e.g., by giving it favorable shelf space. The notion that consumer advertising has such effects is frequently embodied in statements of firms' advertising objectives. Haines and Silk also note empirical studies of situations in which advertising was able to gain some resellers for the brand who had previously refused to carry it. This and similar effects are documented by Gordon (1961); Henderson et al., (1962); and Montgomery (1975).

Heller, Kearney, and Mehaffey's (1973) model for predicting new-product acceptances by supermarkets used advertising at the time of introduction as the most important variable. Their conclusion is supported by several other studies (Graf, 1968; Borden, 1968; Progressive Grocer, 1978). The following quotations from drugstore buyers indicate that the phenomenon is not confined to supermarkets (Inhorn, 1981):

'It's super for us when we know an OTC has been pre-advertised. The manufacturer's done all the work for us. Then, it's just a matter of taking the product on and promoting it in our circular. Even if we set it at a really reduced price, it still makes a profit when it's backed by consumer advertising.

The main thing any buyer looks for is whether a new product is going to be backed with sufficient consumer advertising... Buyers want heavy and continual consumer advertising.'

Summary of assumptions: Changes in advertising can cause changes in distribution; changes in demand can cause changes in distribution; and changes in distribution can change sales.

---

**Implications for Advertising Testing**

Retailer behavior is obviously the key to translating consumer demand into actual sales. Unfortunately, direct and indirect effects of advertising on retail behavior will be difficult to measure within a single retail-market area. In other words, with single-market, split-cable experiments, individual retailers will perceive an "average" level of demand and advertising for the entire market, and will base their decisions on that "average." The net effect of the additional control imparted by a single-market, split-cable design may be a bias. Table 1 illustrates the nature of this bias.

Since the trade perceives only one level of advertising and demand in the single-market (the average of low, medium, and high), the retail reaction will be the same for each level of advertising tested. For the three-market test, retail reactions will presumably vary according to the overall level of advertising and demand in each of the three markets. Figure 2 is a graphic illustration of the same effects for two conceivable response functions.

As a matter of practice, researchers sometimes ensure that distribution in

<table>
<thead>
<tr>
<th>Advertising levels</th>
<th>Consumer effect</th>
<th>Trade effect</th>
<th>Total effect*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>low</td>
<td>avg</td>
<td>low + avg</td>
</tr>
<tr>
<td>med</td>
<td>med</td>
<td>avg</td>
<td>med + avg</td>
</tr>
<tr>
<td>high</td>
<td>high</td>
<td>avg</td>
<td>high + avg</td>
</tr>
</tbody>
</table>

* The total effect represents the joint effect of both the trade and consumer effects. It does not intend to imply additivity.

* As would be produced from a three-market test

---

David J. Reibstein is an associate professor of marketing at The Wharton School, University of Pennsylvania. Prior to joining the Wharton faculty, he was an assistant professor at Harvard Business School. He received his Ph.D. from Purdue University.

the test market is established at an artificially high level so that advertising effects are not obscured by lack of availability. This practice strongly implies that the researcher believes an interaction exists between the two variables. Otherwise there would be no reason for this control. Indeed, if the effects were additive, not interactive, the practice of establishing distribution at an artificially high level would tend to dampen the relative differences in total demand between advertising levels and leave the absolute differences unaffected. If the true relationship is interactive, then establishing artificially high levels of availability and not allowing availability to vary among advertising treatments would compound the errors of improper specification with the elimination of feedback and second-order effects.

Relative differences would be underestimated, but absolute differences might be either under- or overestimated. For example, if "real world" maximum distribution is 20 percent for a product, but the test establishes it at 100 percent, then no effect of advertising on distribution can be observed. The relative differences between low and high advertising levels will be underestimated, but the absolute difference in sales might be lower or higher. Percentaging the differences between the low and high treatment will still result in underestimation of advertising effects, however. Table 2 shows an example of this result, as does Figure 3, for two alternative response functions.

The biases which result from the split-cable design take two forms: (1) interactions and (2) feedback and second-order effects.

Interactions. It is not sufficient to hold interaction variables constant in the test. One must be able to hold them constant at the level which will make the results projectable to the population in question, or one must know how to account for their effects (have a properly specified model for interpreting the experimental results). In an excellent article in the *Journal of Consumer Research*, Lynch (1982) focused on the potential of background factors to interact with treatment factors in a way that distorts the external validity of experimental results. "The external validity of experimental findings depends upon whether "background" factors (e.g., subject or setting factors) that are held relatively constant over the cells of an experimental design, interact in nature with the manipulated variables. If they do so, the relationships observed in experimental data would not be observed if an attempt were made to replicate the study while holding these background factors constant at a different level."

A more precise expression of this problem is the following:

\( S = f(A, D) \)

where,

- \( S \) = sales in market \( n \)
- \( A \) = advertising in market \( n \)
- \( D \) = variable which interacts with \( A \), in this case distribution, such that \( dS/dA = g(D) \)
Specifying any particular level of distribution across test markets, as is done in the split-cable design, could result in either overestimating or underestimating the effects of A, depending on the level of distribution which is representative of all markets in the population of interest.

**Feedback and Second-order Effects.**

If changes in the treatment variable, advertising, affect the dependent variable, sales, through consumer demand, C, we can measure the effect of advertising on sales by allowing only C to vary. If, however, there is a fourth variable D, distribution, which is affected by advertising, and distribution affects sales, we just allow D also to vary. Otherwise the total effect will be underestimated. For example, if the true process is

\[ A \rightarrow C \rightarrow S \]

\[ A \rightarrow D \rightarrow S \]

a model which does not allow D to vary will "overcontrol" the process and underestimate \( dS/dA \). That is, in the above structure

\[ dS/dA = \frac{dS/dC}{dC/dA} + \frac{dS/dD}{dD/dA} \]

An experimental design which holds D constant and equates \( dS/dC \) with \( dS/dA \) is clearly incorrect for the purposes of setting budgets.

### Discussion

A recent article by Aaker and Carman (1982) concludes that advertisers may be generally overspending. Much of the support for this contention derives from experimental results measured by the AdTel split-cable system. Forty-two of forty-eight studies investigated only the effects of increased advertising budgets and 70 percent of these "heavy-up" budgets failed to justify higher budgets with increased sales levels. Aaker and Carman use this evidence to support the conclusion that many marketers may be spending too much money for advertising. As this article has shown, however, such experiments may systematically underestimate advertising effects.

The potential biases discussed are not restricted to experimental measures of advertising effects. Econometric models, which introduce distr...
distribution as a variable in the equation explaining sales or purchase probability, may create similar biases. If distribution were an exogenous variable, one not affected by marketing budget decisions, the variability it accounts for might reasonably be removed when analyzing marketing spending. However, distribution can usually not be assumed to stay constant or independent of marketing spending.

Of course, a change in advertising, sales-force expenditures, trade promotion, or other marketing decisions may be so small that distribution can safely be assumed to stay constant. One must question, however, whether investigating the effect of such minor changes is worth the bother and money. One can also find brands that enjoy such a high degree of distribution and availability that further increases are no longer feasible. In the latter cases one would worry just as much about methods of measuring marketing effects that ignore the worth of the distribution and availability already obtained at least for downside decisions.

There are probably variables other than availability and distribution which are subject to the same sort of "over-control." We have chosen to focus on distribution and availability because (a) they are not allowed to vary normally by the very nature of the split-cable systems and (b) there is clear evidence that advertising does affect distribution and that distribution affects sales. Advertising experiments and econometric models of advertising effects which neglect the impact of advertising on prices (Farris and Reibstein, 1979) can similarly underestimate the effects of advertising on dollar sales and profits.

References


Conclusions

Conclusive evidence exists that advertising affects distribution and distribution, in turn, affects sales for many products. As such, experimental designs which ignore or exclude these relationships are likely to yield biased results.

Frank and Massy (1965) stated, "It is impossible to predict the effects of a change in manufacturers' promotional strategies upon sales and profits without considering the reaction of retail middlemen." Unfortunately, as new research tools evolve, we may sometimes forget old lessons.