

THE ROLE OF CONSUMER RESEARCH IN EVALUATING DECEPTION: AN ECONOMIST'S PERSPECTIVE

JANIS K. PAPPALARDO*

I. INTRODUCTION

Consumer research has become a staple of modern life. Newspapers routinely report the results of their own opinion surveys. Nonprofit arts organizations conduct research to help determine which acts to book and how to price tickets. Advertising agencies depend on research to predict and evaluate consumer responses to marketing campaigns. Manufacturers generate consumer research to inform product introduction and pricing decisions. Politicians conduct polls to guide campaign strategies. Why does consumer research pass the market test in so many forums? Imperfect as most research may be, many decision makers apparently believe that it leads to better information and better decisions than expert opinion alone. This should not be surprising: academic studies suggest that experts are often unable to predict consumer opinions.¹

* Economist, Bureau of Economics, Federal Trade Commission. The views expressed in this paper are those of the author. They do not necessarily reflect the opinion of the Federal Trade Commission or of any Commissioner. I thank many people for their thoughtful comments on earlier drafts, including Jonathan Baker, Gerard Butters, Richard Craswell, Theodore H. Hoppock, Pauline Ippolito, James Lacko, Michael B. Mazis, Dennis Murphy, Thomas P. Olson, Paul Pautler, Lee Peeler, Paul L. Yde, participants in a panel discussion of the ABA Section of Antitrust Law, Proving Deceptive Advertising with Surveys, and an anonymous reviewer.

¹ See, e.g., J. Scott Armstrong, *Prediction of Consumer Behavior by Experts and Novices*, 18 J. CONSUMER RES. 251 (1991); Stephen J. Hoch, *Who Do We Know: Predicting the Interests and Opinions of the American Consumer*, 15 J. CONSUMER RES. 315 (1988). Armstrong posed the following research question: "Are those who are familiar with scientific research on consumer behavior better able to make predictions about phenomena in this field?" He reports that: "The practitioners were correct on 58.2 percent of the hypotheses, the students on 56.6 percent, and the academics on 51.3 percent. No group performed better than chance." Hoch asked marketing experts and novices to make predictions about the "activities, interests, and opinions of the American consumer." He reported that:

Predictive accuracy was low overall, and experts were no more accurate than everyday consumers in predicting consumer opinions. This occurred because (1) everyday consumers were much more similar to the target population than were the marketing experts and (2) the experts had difficulty consistently identify-

Given the widespread reliance on consumer research throughout society, one might be inclined to think that consumer research must have an integral and well-defined role in the analysis of consumer deception cases. A key issue in such cases is whether consumers are deceived by marketing practices, and consumer research would seem to be a critical element for making this determination. One might therefore be surprised to learn that recent court opinions and academic scholarship suggest that the role of consumer research in the analysis of consumer deception cases is far from settled.² The catalysts for much of the recent debate are two recent FTC deceptive advertising cases: *FTC v. Kraft, Inc.*³ and *Stouffer Foods Corp.*⁴

In this article I focus on two key questions that lie at the heart of the debate: (1) When should copy tests be used to evaluate deception, and (2) when should copy test evidence be controlled?⁵ These questions are examined from the broad perspective of a social science researcher, keeping in mind first principles of the scientific method and almost two decades of academic literature on deception measurement. I conclude that key questions raised by recent cases were essentially raised and answered long ago.

The finding that core questions under debate today were answered long ago raises an intriguing question: Why hasn't the legal analysis of deception embraced core social science principles? This question is especially intriguing when one compares the role of economic reasoning and econometrics in the evaluation of antitrust cases to the role of economic reasoning and econometrics in the evaluation of consumer protection cases. For example, consumer theory and econometrics are frequently used to predict the competitive effects of mergers and, given

ing other information beyond their own attitudes relevant to the target population. For this task, the experts could not overcome the "information deficit" that accompanies being dissimilar to the typical American consumer.

² For example, in 1995 *The Journal of Public Policy and Marketing* published four articles addressing fundamental research questions raised in *FTC v. Kraft, Inc.* and *Stouffer Foods Corp.*: Jacob Jacoby & George Szybillo, *Consumer Research in FTC Versus Kraft: A Case of Heads We Win, Tails You Lose?*, 14 J. PUB. POL'Y & MKTG. 1 (1995); David W. Stewart, *Deception, Materiality, and Survey Research: Some Lessons from Kraft*, 14 J. PUB. POL'Y & MKTG. 15 (1995); Seymour Sudman, *When Experts Disagree: Comments on the Articles by Jacoby and Szybillo and Stewart*, 14 J. PUB. POL'Y & MKTG. 29 (1995); J. Craig Andrews & Thomas J. Maronick, *Advertising Research Issues from FTC v. Stouffer Foods Corp.*, 14 J. PUB. POL'Y & MKTG. (1995). In June 1996 the Council of Better Business Bureaus sponsored a conference on consumer perception surveys, NAD Workshop VI (1996). In August 1996 the ABA Antitrust Law Section organized a panel on Proving Deceptive Advertising with Surveys.

³ 970 F.2d 311 (7th Cir. 1992), *aff'g* Kraft, Inc., 114 F.T.C. 40 (1991).

⁴ FTC Docket No. 9250, slip op. (1994).

⁵ See Jonathan B. Baker, "Hot Topics" at the Federal Trade Commission, Remarks to the ABA Section of Antitrust Law Spring Meeting (Mar. 1996).

recent advances in data availability, sophisticated complete demand system estimation is likely to be used even more frequently in the future.⁶ Although many of the factors that have revolutionized merger analysis (in particular, the availability of scanner data) would seem to apply equally to the analysis of deceptive advertising, there has not yet developed a similar demand analysis revolution for deceptive advertising cases.

Economists may be partly responsible for the relatively elementary state of economics in the realm of consumer protection. For example, although economists have been urging policy makers to adopt an economics-based approach to consumer protection for years,⁷ and some progress has been made toward this goal,⁸ antitrust received serious attention from economists decades before consumer protection did.⁹ It may, therefore, just be a matter of time before consumer protection analysis reaches the same level of economic sophistication, on average, as antitrust analysis.

Another possibility is that economists specializing in consumer protection may be less-effective communicators of economic reasoning than their colleagues who specialize in antitrust. Indeed, although economists are often involved in the evaluation of deceptive advertising cases at the FTC, they have not participated actively in the public debate over the use of copy tests to evaluate deception.

A more systematic application of social science principles to consumer protection may just be a matter of time. The purpose of this article is to contribute to the increasing interest in the application of social science principles to consumer protection. In particular, I hope that it contri-

⁶ Jonathan B. Baker, *Contemporary Empirical Merger Analysis*, 5 GEO. MASON L. REV. 347 (1997); Jerry A. Hausman & Gregory K. Leonard, *Economic Analysis of Differentiated Products Using Real World Data*, 5 GEO. MASON L. REV. 321 (1997).

⁷ Howard Beales et al., *The Efficient Regulation of Consumer Information*, 23 J.L. & ECON. 491 (1981); Pauline Ippolito, *Consumer Protection Economics: A Selective Survey*, in EMPIRICAL APPROACHES TO CONSUMER PROTECTION ECONOMICS 1 (Pauline Ippolito & David Scheffman eds., 1986).

⁸ See, e.g., John E. Calfee & Janis K. Pappalardo, *How Should Health Claims for Foods Be Regulated?: An Economic Perspective* (Bureau of Economics, Federal Trade Commission Sept. 1989) (arguing that the FTC's advertising substantiation standard can be interpreted as being consistent with an expected value standard for regulation); John E. Calfee & Janis K. Pappalardo, *Public Policy Issues in Health Claims in Foods*, 10 J. PUB. POL'Y & MKTG. 33 (1991).

⁹ Baker notes that demand estimates have been used to analyze competition issues since at least 1940. Baker, *supra* note 5. In contrast, economists were rarely consulted on consumer protection issues at the FTC until the early 1970s. And "even economists employed at the FTC spent virtually no time on consumer protection matters before 1974, in contrast to the substantial commitment to competition case work and research." Ippolito, *supra* note 7, at 1.

butes to a conclusion that the time has come for a more explicit incorporation of economic reasoning and social science principles in consumer protection analysis.

II. THE RECENT CONSUMER RESEARCH CONTROVERSIES

Kraft and *Stouffer Foods* are largely responsible for the recent debate over the role of consumer research to evaluate deception. In particular, the cases raise questions about the use of "copy tests." Copy tests are, in essence, a simple type of social science experiment: consumers are shown an ad (the stimulus) and then asked a series of questions to determine the advertisement's likely effects.

Although the concept of a copy test is quite simple, the proper implementation of a copy test requires that many subtle and important research questions be resolved. One of the first tasks is to define the outcome variable. FTC copy tests typically seek to measure the effect of advertising on consumer perceptions. Thus, FTC copy tests typically ask respondents to report what an ad "says or suggests." In other settings, researchers might try to measure the effect of the ad on consumer beliefs, purchase intentions, or, if possible, actual purchases. Researchers must also decide who should be surveyed (sample design), the phrasing and sequence of questions, how the responses should be recorded (coding design), and how the results should be analyzed.

Although *Kraft* and *Stouffer Foods* raise numerous questions about the many details involved in copy test design and analysis, in this section I discuss how these cases pertain to the two central questions examined in this article. The first question is: When should copy tests be used to evaluate deception? As discussed below, the need for a copy test will depend on a number of factors relating to the cost of making a "wrong" policy decision without such a test and the cost of the test itself. The second question is: When should copy test evidence be "controlled?" A "controlled" copy test is one where researchers try to separate the effect of the potentially deceptive advertising claim from the effect of numerous other factors, such as the effects of a respondent's desire to please the researcher, the effects of preexisting beliefs, and the effects of a proposed remedy. I conclude that when additional evidence is deemed necessary to determine whether an ad is deceptive, the default position should be to include several control ad conditions.

A. CONTROVERSY OVER THE NEED TO CONDUCT CONSUMER RESEARCH TO EVALUATE DECEPTION AT THE FTC

Despite the widespread reliance on consumer research to inform marketing and other decisions throughout society, the Seventh Circuit's

ruling in *Kraft* makes clear that consumer research is not necessary for the Commission to prove deception: "We hold that the Commission may rely on its own reasoned analysis to determine what claims, including implied ones, are conveyed in a challenged advertisement, so long as those claims are reasonably clear from the face of the advertisement."¹⁰

Judge Manion's concurring opinion leaves open the question, however, of whether consumer research is desirable as a matter of case selection and policy planning:

While I concur with the opinion of the court, I am concerned that the FTC can avoid extrinsic evidence by simply concluding that a deceptive, implied claim is facially apparent. While the FTC has expertise, consumer surveys provide at least some objective determination of what the purchaser thinks and should be considered since, after all, the consumer is among those we are trying to protect.

Moreover, the FTC's current procedure threatens to chill nonmisleading, protected speech The Supreme Court has recognized that a free flow of information is indispensable to decisionmaking in the free enterprise system But the FTC jeopardizes this flow by relying on the FTC commissioners' subjective interpretation to determine whether an ad, while literally true, implies a false message.¹¹

Manion further writes that "the FTC would be well advised to take this court's suggestion—apply its expertise and develop a consumer survey methodology that advertisers can use to ascertain whether their ads contain implied, deceptive messages."¹²

The opinion in *Kraft* generated substantial concern over the proper role of consumer research at the FTC, as reflected in an American Bar Association Task Force recommendation that the Commission "respond to the court's invitation in *Kraft* to provide guidance on the use and methods of conducting baseline consumer perception surveys."¹³ The opinion also prompted a series of lively articles regarding the use and evaluation of consumer research in *Kraft*.¹⁴ In particular, the opinion raises the question of whether extrinsic evidence should be applied to case selection and policy planning decisions. It also raises the question of what methodologies should be employed to collect such evidence.¹⁵

¹⁰ *FTC v. Kraft, Inc.*, 970 F.2d 319 (7th Cir. 1992).

¹¹ *Id.* at 327–28.

¹² *Id.* at 328.

¹³ ABA Section of Antitrust Law, Report of the Special Task Force on Competition Policy, reprinted in 64 *Antitrust & Trade Reg. Rep.* (BNA) 1604 (Spec. Supp. Mar. 1993).

¹⁴ See Jef I. Richards & Ivan L. Preston, *Proving and Disproving Materiality of Deceptive Advertising Claims*, 11 *J. PUB. POL'Y & Mktg.* 45 (1992); articles cited *supra* note 2.

¹⁵ For a discussion of the rationale for not setting forth a Commission standard on copy test methodology, see Janet D. Steiger, Address to the Marketing and Public Policy Conference (Mar. 13, 1994).

B. CONTROVERSY OVER THE NEED TO CONDUCT CONTROLLED COPY TESTS

Further questions regarding consumer research methodology were raised in *Stouffer Foods Corp.*, where the Commission stated:

[T]here is no record evidence that, among experts in advertising or consumer research, the use of a control group is considered a *sine qua non* of a valid copy test. In this regard, we note that complaint counsel's expert witnesses testified that the Zinkhan copy test is valid and reliable evidence of what claims the Stouffer ads communicated, without the need for a control group.¹⁶

. . . In any event, there must be evidence of preexisting bias to find that failure to control for such bias is a critical defect.¹⁷

The *Stouffer Foods* decision, therefore, raises questions about the value of controls in copy tests. As discussed further below, those trained in the social sciences, like those trained in the physical sciences, cut their teeth on the concept of hypothesis testing *controlling* for competing hypotheses. The apparent failure of the record in *Stouffer Foods* to reflect this tradition has raised questions among social science researchers.¹⁸

III. AN ECONOMIC APPROACH TO DECEPTION MEASUREMENT

Economists often argue that copy tests and other empirical research should be used to evaluate consumer protection cases whenever such research passes a cost/benefit test. They also tend to argue that such research should be controlled, either statistically or experimentally. To understand why this position is held so often and so vigorously, it is helpful to consider the economists' general approach to hypothesis testing and studies of consumer behavior.

A. WHAT DO ECONOMISTS TRY TO MEASURE WHEN THEY MEASURE "DECEPTION"?

Objects must be defined before they can be measured. That is why it is helpful to specify house plans before gathering construction bids. The same principle applies to the measurement of "deception."

Deception has been defined in numerous ways. One key definition is the widely applied and cited legal definition articulated in the FTC's Policy Statement on Deception:

¹⁶ 1994 FTC LEXIS 196 at 32-33.

¹⁷ *Id.* at 35.

¹⁸ See Andrews & Maronick, *supra* note 12.

First there must be a representation, omission, or practice that is likely to mislead the consumer Second, we examine the practice from the perspective of a consumer acting reasonably in the circumstances. . . . Third, the representation, omission or practice must be a "material" one. The basic question is whether the act or practice is likely to affect the consumer's conduct or decision with regard to a product or service. If so, the practice is material and consumers are likely to have chosen differently but for the deception.¹⁹

Economists, however, often want to know more than whether a case meets a legal standard for deception. From an economic perspective, the critical policy questions are (1) whether a potentially deceptive practice makes consumers, as a whole, worse off, and (2) whether Commission intervention is likely to improve consumer decisions by leading consumers to the choices that more closely approximate those they would have made if they had good information about a product and its alternatives. To economists, the overall objective of government regulation is to promote social welfare, and the only way to estimate whether FTC intervention will improve welfare is by comparing the likely net effect of FTC intervention against the likely net effect of other alternatives, including the possibility of no intervention and the possibility of intervention by other entities, such as local governments, private firms, and industry self-regulatory bodies. Intervention is not likely to be recommended on economic grounds unless the net benefit of the intervention is likely to be positive and higher than the net benefit of other alternatives.²⁰

To operationalize this "net benefit" standard, one must consider a series of questions. First, is a claim likely to be deceptive to some consumers? Second, if a claim is likely to be deceptive, what is the magnitude of injury that is likely to result? If a claim is likely to be deceptive and potentially injurious to some consumers, an analyst must also consider whether the same claim might provide valuable information and benefits to other consumers. A classic example of a claim that could be helpful to some consumers but harmful to other consumers is the "no cholesterol" claim. This claim could be a valuable signal to informed consumers who know how to evaluate it. However, other misinformed consumers, who believe, for example, that a "no cholesterol" claim implies that a product is "fat-free," could potentially be misled and harmed by the claim. Consumers are heterogeneous. A "net benefit" analysis, therefore, requires

¹⁹ FTC Policy Statement on Deception, *reprinted as* appendix to Cliffdale Assocs., 103 F.T.C. 110, 175 (1984).

²⁰ The economic "checklist" in this section is not new. *See, e.g.*, Ippolito, *supra* note 7, at 2 ("At institutions like the FTC it is now routine for perceived consumer protection problems to be discussed in terms of possible market solutions, the likely effects of alternative regulatory approaches on consumer and firm behavior, and the associated benefits and costs of these interventions.").

consideration of how the magnitude and intensity of harm to one group of consumers from allowing a claim compares to the magnitude and intensity of harm to other consumers from prohibiting a claim (or by imposing remedy requirements so stringent that they would likely lead to the absence of the triggering claim). Thus, there is concern about the potential information “chilling” effect of intervention.

Even if analysts identify a potentially injurious practice, the net effects of intervention cannot be assessed with much certainty without also considering whether a feasible welfare-enhancing remedy exists.²¹ For example, would a proposed information disclosure designed by staff economists and attorneys actually reduce consumer confusion and lead to better decisions? Unfortunately, a substantial literature indicates that remedies often do not work as intended. The Commission is, therefore, often encouraged to test its remedies more routinely.²² The FTC’s Director of the Bureau of Consumer Protection recently focused attention on this problem: “But even after the FTC determines that a law violation has occurred, we will still confront the question of whether the FTC’s proposed remedy—which often involves the disclosure of information—helps the situation, or makes it worse.”²³ A net benefit rule also requires analysts to weigh potential harm from both Type I and Type II regulatory errors. This consideration is particularly important in ad substantiation cases when evidence for a claimed proposition is mixed. In such cases one must consider the tradeoff between harm from allowing claims that are likely to prove unfounded and harm from prohibiting claims that are likely to prove true.²⁴

A focus on the effects of FTC intervention versus other alternatives also encourages analysts to consider whether non-FTC remedies exist that might be more efficient than FTC intervention. For example, in markets where counter-advertising is likely to be effective, the marginal benefit of FTC action might be relatively small. Another consideration is whether private action (for example, through the Lanham Act²⁵) or

²¹ See Beales et al., *supra* note 7.

²² See Jacoby & Szybillo, *supra* note 2; John H. Murphy & Jef I. Richards, *Investigation of the Effects of Disclosure Statements in Rental Car Advertisements*, 26 J. CONSUMER AFF. 351 (1992); William J. Wilkie, *Affirmative Disclosure: Perspectives on FTC Orders*, 1 J. PUB. POL’Y & MKTG. 95 (1982).

²³ Joan Z. Bernstein, *Federal Trade Commission Solicits Consumer Research*, in *ADVANCES IN CONSUMER RES.* 23 (Kim P. Corfman & John G. Lynch, Jr. eds., 1996).

²⁴ Calfee & Pappalardo, *supra* note 8 (arguing that the FTC’s advertising substantiation standard can be interpreted as being consistent with an expected value standard for regulation).

²⁵ 15 U.S.C. §§ 1125(a) et seq.

actions by other government entities are likely to prove more efficient than FTC action.

Finally, an economic analysis of consumer protection regulation requires consideration not only of how intervention might affect consumer behavior in the short run, but also how interventions can change incentives and, hence, consumer and firm behavior in the long run. For example, in markets with emerging technologies one must consider whether restrictions on advertising messages will discourage product innovation by stifling a manufacturer's ability to inform consumers about the innovation.

Although the deception analysis favored by economists might differ in some respects from the legal standard defined in the FTC's Policy Statement on Deception, the definitions are similar in many respects. In particular, both involve consideration of what claims consumers receive and how those claims are likely to affect consumer behavior. Unlike the legal standard, however, the economist's "net benefit" standard to evaluate deception strongly and explicitly reflects concerns raised by Judge Manion regarding the potential harm from chilling the flow of information. The net benefit standard also explicitly reflects concerns about whether a remedy will do more harm than good.

B. TESTING DECEPTION THEORIES

When economists examine any theory, the first question they typically ask is: "What empirical data exists (or might exist) to test the theory against competing theories?" This principle reflects nothing more than an application of the scientific method, as enunciated clearly in Professor Paul Samuelson's classic text (originally published in 1947):

By a meaningful theorem I mean simply a hypothesis about empirical data which could conceivably be refuted, if only under ideal conditions. A meaningful theorem may be false. It may be valid but of trivial importance. Its validity may be indeterminate, and practically difficult or impossible to determine. . . . But it is meaningful because under ideal circumstances an experiment could be devised whereby one could hope to refute the hypothesis.²⁶

To determine whether a deception theory is meaningful, therefore, one must consider what experiment could be devised whereby one could refute the deception hypothesis. The proper "experiment" will depend on how one views deception.

²⁶ PAUL A. SAMUELSON, FOUNDATIONS OF ECONOMIC ANALYSIS 4 (9th ed. 1979).

Two approaches likely to be favored by economists investigating deceptive advertising are discussed below in Parts B.1. and B.2.: the econometric demand approach and the true experimental approach. As discussed in Part B.3., neither approach can be fully implemented without evidence regarding ad interpretation.

1. *The Econometric Demand Approach*

From an economist's perspective, deception theories essentially are theories about the effect of advertising and proposed remedies on consumer behavior. When it comes to testing theories about consumer behavior (including purchase behavior, consumption behavior, or time-use behavior) economists tend to rely on models of consumer demand.²⁷

Demand models recognize that many factors—including advertising—are likely to affect purchase decisions. For example, diamond demand is likely to depend on the price of diamonds, the price of substitute goods (such as cubic zirconias or sapphires), the price of complementary goods (such as gold or platinum), income, information about diamonds (for example, the nightly news might report on an anticipated diamond shortage), and consumer tastes and preferences. Because diamond advertising is one potential source of diamond information, it also belongs in the model.

To estimate the effect of advertising on consumer demand, one ideally would like data on the dependent consumer behavior variable (e.g., sales data), as well as data on all of the independent variables (prices, income, advertising, other information sources, and tastes). With such data and proper application of multiple regression techniques, one could theoretically estimate the effects of each independent variable on quantity demanded, holding the effects of the other independent variables constant.²⁸ Thus, one could isolate the effect of advertising on sales.

To test whether deceptive advertising affected sales, one must determine whether an ad caused deceptive interpretations and then estimate how demand under the deceptive condition compared to demand under a nondeceptive condition, all else constant. This might be possible if, for example, a deceptive campaign appeared during one time period and nondeceptive campaigns appeared in different periods. Such estimation might also be possible if geographic differences exist in advertising campaign messages or advertising intensities.

²⁷ See, e.g., HAL R. VARIAN, MICROECONOMIC ANALYSIS (1978).

²⁸ See, e.g., GEORGE G. JUDGE ET AL., THE THEORY AND PRACTICE OF ECONOMETRICS (1980).

The approach typically used to estimate demand models illustrates the principle that by controlling for the effects of various variables statistically, one can rule out alternative hypotheses.²⁹ This principle of control is a fundamental part of the scientific method and is necessary to draw inferences about the effect of one factor on another.

2. *The Experimental Approach*

Another approach to testing the effect of advertising on sales (or other behavior) is the true experimental approach, such as the random control approach required to show that drugs are effective.³⁰ Under an experimental approach, consumers would be randomly assigned to a potentially deceptive ad condition or a control ad condition (analogous to a drug treatment condition and a placebo condition). If both groups are in the same market facing the same prices and both groups have similar tastes, incomes, and other sources of information, then one could attribute differences in behavior between the two groups to differences in advertising. As scanner and cable data technologies improve and become more widespread,³¹ it is increasingly likely that the necessary information will be available to conduct reliable experiments.

3. *Using Copy Tests to Measure Deceptive Ad Interpretations*

Neither approach can be fully operationalized without distinguishing between deceptive and nondeceptive advertising claims—for the question is not how nondeceptive messages affect behavior, but how deceptive messages affect behavior.

Advertising claims fall on a spectrum—some are express and others are barely implied.³² Determining whether an express claim is deceptive is fairly simple (unless the claim involves an area of scientific controversy where the “truth” is murky); a comparison between the literal claim and the “truth” is generally sufficient to determine whether an express claim is deceptive. Among implied claims, some are practically express and can be treated as such. However, a large class of claims are far from express, and this is where issues of ad meaning become controversial.

²⁹ See James J. Heckman & Jeffrey A. Smith, *Assessing the Case for Social Experiments*, 9 J. ECON. PERSP. 85 (1995) (regarding statistical controls in economics).

³⁰ Gary Burtless, *The Case for Randomized Field Trials in Economic and Policy Research*, 9 J. ECON. PERSP. 63 (1995) (regarding experimental controls in economics).

³¹ Dave Kruegel, *Commentary: Television Advertising Effectiveness and Research Innovation*, 5 J. CONSUMER MKTG. 3, 43–51 (1988); Laurence N. Gold, *The Evolution of Television Advertising-Sales Measurement: Past, Present, and Future*, 28 J. ADVER. RES. 19 (1988).

³² See Debra K. Owen & Joyce E. Plyler, *The Role of Empirical Evidence in the Federal Regulation of Advertising*, 10 J. PUB. POL'Y & MKTG. 1 (1991).

From a social science perspective, the proposition that a potentially deceptive implied claim is perceived by consumers is simply another testable hypothesis. This means that alternative hypotheses should be defined and methods applied to control for competing explanations. For example, to show that an ad caused a false impression one would need to consider other likely explanations, such as the effects of prior beliefs. If other likely explanations are not controlled for, then one cannot be certain that the advertising, *per se*, is the cause of the perception.

Borrowing from the drug treatment analogy, the most general way to test for competing explanations is by using a control ad. By comparing the perceptions of those exposed to a control ad versus those exposed to the potentially deceptive ad, one can derive the likely effect of the deceptive ad elements from the likely effects of all other factors, including “yea-saying” and preexisting beliefs. There are many types of control ad conditions available.³³ One can compare the potentially deceptive ad to a similar ad without the potentially deceptive claim. One can also compare the potentially deceptive ad to a “corrected” ad. Another control possibility is a control question; however, this approach typically controls for fewer influences than a control ad.

Failing to control for factors such as preexisting beliefs when assessing advertising is like failing to control for placebo effects when assessing drug efficacy. In essence, the failure is tantamount to giving a drug credit for the placebo effect. Moreover, unless one tests for the placebo effect, one will not know the extent to which a drug’s effect would be overstated absent a placebo effect control. Similarly, unless one tests for the effects of preexisting beliefs, one will not know the extent to which the effects of a potentially deceptive ad are overstated absent a preexisting belief control.

A preference for empirical evidence to test hypotheses about the existence of implied claims is consistent with the economist’s general preference for empirical data to test any hypothesis. This general preference is supported by evidence suggesting that advertising effects are difficult to predict. For example, evidence indicates that even marketing experts are not very good at predicting consumer attitudes and opinions,³⁴ suggesting that they may not be very good at interpreting the effects of ad copy either. This preference is also supported by the observation that survey research passes the market test—advertisers devote sub-

³³ Richard Craswell, “Compared to What?” *The Use of Control Ads in Deceptive Advertising Litigation*, 65 ANTITRUST L.J. 757 (1997).

³⁴ See Armstrong, *supra* note 1; Hoch, *supra* note 1.

stantial resources to communications tests precisely because it is difficult to predict how consumers will respond to advertising.

In addition to a strong preference for controlled data, as discussed further below, economists also tend to evaluate data acquisition decisions from a cost/benefit perspective. This preference translates into a realization that the net benefits of imperfectly controlled data can be positive under some circumstances, as long as the data's limitations are considered and weighted appropriately. For example, if a test that fails to control for preexisting beliefs shows that only 5 percent of consumers receive a false perception, then one can conclude that the upper boundary on deception, whatever the effects of prior beliefs, is relatively small. Moreover, one is inclined to make the most of whatever evidence is available to help evaluate advertising. However, making the most of available evidence is different from deciding how to collect fresh data designed explicitly to evaluate the potential deceptiveness of advertising.³⁵

Finally, even when data exist to estimate demand in an econometric model, such data can be "noisy." Findings of deception within such models, therefore, constitute a sufficient condition to prove deception in an economic sense—not a necessary condition. In most cases data fall far short of the ideal, and economists, like anyone else, must consider how to make the most of whatever data are available. Marketers have specialized in the study of how to evaluate the effects of advertising when behavior data do not exist, and the marketing literature therefore offers insight into the evaluation of deception advertising.

IV. HISTORICAL MARKETING RESEARCH PERSPECTIVE

For decades, marketing researchers have been considering how to measure advertising's effects—including the effects of deceptive advertising. A review of this literature reveals a consistent preference for controlled copy test data to analyze (1) whether an ad is deceptive, and (2) whether a proposed remedy is likely to be effective.

A. WHAT DO MARKETERS TRY TO MEASURE WHEN THEY MEASURE "DECEPTION"?

Marketers generally share the economist's preference for behavioral outcome measures to assess advertising, but such data are often not

³⁵ Control ads, per se, are often not used each time a commercial copy test is conducted. However, marketing researchers often compare advertising effects against a series of "norms" developed over time. See, e.g., Thomas J. Maronick, *Copy Tests in FTC Deception Cases*, 31 J. ADVER. RES. 9 (1991). In this setting, the "norm" becomes the baseline against

available. And even when available, the data are often very "noisy," making it difficult to detect advertising effects. Yet real business decisions about advertising must be made every day. To deal with this problem, marketers have developed several "proxy" measures for behavioral outcomes. These measures are used to define procedures for evaluating deception in the marketing literature. Interestingly, marketers tend not to rely on perception measures to evaluate deception. Instead, they tend to rely on measures of consumer beliefs or consumer purchase intentions.

Consumer beliefs following advertising exposure are often used to define deception, despite limited research testing the validity of belief measures.³⁶ To understand the difference between a belief measure and a perception measure, it is helpful to consider the "Joe Isuzu" advertising campaigns. In these campaigns, the fictional character Joe Isuzu makes outlandish claims about the benefits of purchasing an Isuzu automobile. Thus, if one were to ask what a Joe Isuzu ad "says or suggests," respondents might say that the ad says or suggests that an outlandish outcome would occur. However, if one asked respondents if they believe the Joe Isuzu claims, they might get a different answer because, after all, a reasonable consumer would be unlikely to believe the outlandish claim.

Purchase intentions are another outcome measure that can be used to predict advertising effects. Seymour Sudman recommends that purchase intentions be employed to evaluate deception theories when more direct measures of behavior are not available:

Questions asked after viewing the advertisement should include inquiries concerning buying intentions, which will address issues of materiality, in addition to issues pertaining to perceptions of the advertisement, which measure deception. The issue ultimately is whether the advertisement has an impact on buying behavior, and buying intentions, although imperfect, remain the best predictor of behavior.³⁷

B. HOW SHOULD DECEPTION BE MEASURED?

Marketing researchers have long argued for a consumer research-based approach to deception law. For example, since the early 1970s researchers have argued that the Commission ought to rely on "objective assessment of consumers' understanding of allegedly deceptive advertise-

which new executions are compared; thus, there is no need to use a control ad each time a new campaign is tested for overall effectiveness.

³⁶ Gary M. Armstrong et al., *Defining and Measuring Deception in Advertising: A Review and Evaluation*, in *CURRENT ISSUES & RESEARCH IN ADVERTISING* 17 (James H. Leigh & Claude R. Martin, Jr. eds., 1980) at 27.

³⁷ Sudman, *supra* note 2, at 30.

ments.”³⁸ In addition, since at least the early 1980s, marketing research academics have generally favored a control ad approach to deceptive advertising research. Like Judge Manion, Edward Russo et al. argued for “some equitable, standard procedure to determine whether an ad is misleading.”³⁹ Russo et al. conducted an experiment illustrating their proposed technique. The methodology is important for two reasons. First, the study addresses the question of whether, and how, to use control ads. Second, it considers explicitly the problem of reinforcing or exploiting false beliefs.

The methodology endorsed by Russo et al. requires a comparison among beliefs held after exposure to potentially misleading ads, ads without the potentially misleading claim, and ads with a corrected version of the potentially misleading claim. The technique is summarized as follows:

If the level of misleading belief is (statistically significantly) higher for the original group than for the no-ad (control) group, then the ad is found to be incrementally misleading. Exposure to the ad increases the level of false belief.

Exploitative misleadingness occurs when the ad does not increase, but free-rides on, an existing level of misleading belief. If the level of misleading belief is (statistically significantly) higher for the original ad than for the corrected version, the ad is found to be exploitively misleading.⁴⁰

“Misleadingness” is then defined: “An advertisement is misleading if it creates, increases, or exploits a false belief about expected product performance.”⁴¹

Russo et al. recognized that their methodology can sometimes be difficult to implement: “Devising a proper corrected ad requires cleverness and effort. Like its reflection, the control group in experimental science, the correct ad may pose practical difficulties, but at least the goal is clear.”⁴² Nevertheless, they continued to believe that the effort would be justified.

The marketing community built upon the Russo methodology—even after publication of the FTC’s Policy Statement on Deception. For example, Raymond Burke et al. followed the Russo control group tradition

³⁸ Armstrong et al., *supra* note 36, at 18. Also by the early 1970s, behavioral evidence was reportedly considered in roughly half of the FTC’s advertising cases. See J. Edward Russo et al., *Identifying Misleading Advertising*, 8 J. CONSUMER RES. 119 (1981).

³⁹ *Id.*

⁴⁰ *Id.* at 127.

⁴¹ *Id.* at 128.

⁴² *Id.* at 129.

and assessed deception by “comparing consumer responses to the questionable claims against responses to the presentation of no attribute information and true information.”⁴³ As an extension, Burke measured not only the effect of ads on beliefs, but also “affect toward using the brand,” “preference for the brand over competitors,” and “purchase likelihood.”⁴⁴

Widespread preference among marketing researchers for a control ad approach continues today. For example, in his analysis of *Kraft*, Sudman writes:

One of the major strengths of the FTC survey is the presence of a control group. Most of the challenges raised about possible context effects, yea-saying, and question wording can be countered by comparing the control group and the group that saw the disputed advertisement. Using a control group is especially valuable in situations, such as the one discussed by Jacoby and Szybillo and Stewart, in which the disputed advertisement had been running for a long time before the survey was conducted. Again, comparing the control and experimental groups eliminates effects of prior experience.

. . . Advertising studies in malls are typically considered as surveys of the population. However, they are actually more similar to experiments in which the advertisement is the stimulus. Considering them as experiments emphasizes the need for a control group, which is a recommendation for legal studies, as well as for many other mall market research projects.⁴⁵

Similarly, Craig Andrews and Thomas Maronick write:

In any study, it is important to have the ability to control the situation in which the study is being conducted so as to eliminate the role of extraneous forces and competing explanations. A control ad is often used to separate effects due to the challenged ad claim from effects due to external factors associated with such an ad claim. Such external factors include previous exposure to the ad, other nonchallenged executional and copy point elements in an ad, and prior beliefs and knowledge associated with the advertised product.⁴⁶

The more general marketing research literature also reveals a preference for control methods. Recent guidelines formulated for research on marketing communications advocate the use of controls, either through

⁴³ Raymond Burke et al., *Deception by Implication: An Experiment Investigation*, 14 J. CONSUMER RES. 483 (1988).

⁴⁴ *Id.* at 487.

⁴⁵ Sudman, *supra* note 2, at 33.

⁴⁶ Andrews & Maronick, *supra* note 2, at 305 (citations omitted).

experimentation or through statistical corrections.⁴⁷ Similarly, marketing research textbooks warn against the use of “one-shot case studies”:

A useful point of departure for discussing experiments is the one-shot case study A single group of test units is exposed to an experimental variable, and its response is observed once For example, we might interview a convenience sample of those who read a particular trade journal for their reaction to our product. The experimental stimulus here would be the ad.

The one-shot case study is of little value in establishing the validity of hypothesized causal relationships (the ad was responsible for creating a favorable attitude toward our product) because it provides too little control over the extraneous influences. It provided no basis for comparing what happened in the presence of X with what happened when X was absent. Yet the minimum demands of scientific inquiry require that such comparisons be made.

The one-shot case study is more appropriate for exploratory than conclusive research. It is appropriately used to suggest hypotheses; it is not appropriate for testing their validity.⁴⁸

V. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Questions and controversies regarding the use of copy tests and other consumer research to evaluate deception are not new. Consumer research has played a role in FTC analyses for many years. In fact, it appears that consumer research activity at the FTC peaked roughly twenty-five years ago.⁴⁹

Some controversies may never be resolved. However, the economics and marketing approaches are consistent on three key points. First, both fields reveal a strong preference for empirical research to test deception hypotheses. Second, consistent with first principles of the scientific method, both fields place a high premium on controlled data, especially data utilizing control ads. Third, both fields emphasize the need to consider the likely effects of proposed remedies to deception because empirical data suggest that remedies can sometimes do more harm than good.

⁴⁷ Paul N. Bloom et al., *Criteria for Assessing Research on the Effects of Marketing Communications*, Marketing Science Institute Working Paper 94-123, 1995 (available at MSI Cambridge, Mass.).

⁴⁸ GILBERT A. CHURCHILL, JR., *MARKETING RESEARCH METHODOLOGICAL FOUNDATIONS* 187 (5th ed. 1991).

⁴⁹ See Maronick, *supra* note 35; Mary G. Jones, *Marketing Academics at the FTC: Reflections and Recommendations*, in *MARKETING AND ADVERTISING REGULATION: THE FEDERAL TRADE COMMISSION IN THE 1990s* 216-220 (Patrick E. Murphy & William L. Wilkie eds., 1990); Patrick E. Murphy, *Past FTC Participation by Marketing Academics*, in *id.* at 205-15.

Given the strong theoretical preference to test deception theories with consumer survey data—particularly controlled data—one might wonder why controlled consumer research studies are not performed routinely. There is one obvious explanation. As discussed above, the Commission has not been required, as a matter of law, to produce consumer survey evidence to prove deception theories. Nor is the Commission required to employ particular control measures when it does collect fresh copy test evidence.

This leaves unanswered several important policy questions: Should consumer research be used more routinely, as a matter of case selection and policy planning? Should deception laws be changed to require a more systematic application of marketing and economic research principles? Should standards be adopted to guide consumer research protocols whenever the Commission finds it necessary to gather fresh evidence to evaluate deception theories?

Unfortunately, there are no easy answers to any of these questions. Consumer research can be costly, and controls substantially increase research costs.⁵⁰ From an economic perspective, data acquisition and analysis costs ought to be scrutinized under the same cost/benefit microscope as anything else. Ultimately, the net benefits of consumer research will depend on the potential harm from making “wrong” policy decisions in the absence of such research. Consistent with the economic deception analysis described above, costs for solid testing are more likely to be justified the greater the potential harm from chilling truthful information, the less certain we are that a remedy will fix a problem, and the further away claims are from the “express” end of the “express claim/ implied claim” continuum.

In cases where analysts and decision makers agree on likely reasonable ad interpretations, where there is little risk of over-deterrence, and where there is strong reason to believe that a remedy will work as intended, one can understand why consumer research is not conducted routinely. Given how often consumer research is used to make decisions throughout society, I am inclined to believe that decisions regarding deception might benefit from more reliance on survey data. This, however, is a testable hypothesis that would benefit from further research.

In cases where consumer research is deemed necessary to evaluate ad interpretation and de novo data are collected, the economic and marketing literature indicates that the default position should be to include controls. To be fair, there are some practitioners who question

⁵⁰ See Sudman, *supra* note 2.

the value of controls even in these cases. One argument against controls is that they can be difficult to devise.⁵¹ This is sometimes a real problem, as anyone who designs copy tests will verify. However, if copy test data are important to a case, the lack of an ideal control ad does not necessarily imply that controls should not be used. A better approach would be to use a "sensitivity analysis," whereby several reasonable controls are employed and the results of the different control conditions are compared. Such an approach seems consistent with the general recommendation that "whenever possible, associative and causal studies should use different methods and scales to measure the constructs under investigation."⁵²

Another argument against the use of controls arises in cases where deceptive ad interpretations arise from preexisting false beliefs, which may have been caused by previous advertising campaigns. Some practitioners argue that use of a control in such cases would allow marketers to avoid responsibility for deception that they caused. Although such cases require particular care regarding the choice of controls, they do not necessarily render controls useless. As suggested by Russo, remedy controls are particularly useful in these circumstances. For example, if consumers are shown corrected ads, how does this change their perceptions and likely purchase behavior? Of course, determining the extent to which misperceptions are caused by previous advertising versus other factors is still undeniably difficult. However, remedy controls would be helpful to evaluate the extent to which intervention is likely to improve consumer welfare.

As a matter of science, there is no question that controlled data (controlled statistically or experimentally) is preferred to uncontrolled data. As a practical matter, however, the magnitude of bias resulting from a lack of controls is unclear. A more precise estimate of the costs of failing to conduct copy tests and failing to use controls—in terms of identifying the existence of deceptive perceptions or the effects of a proposed remedy—would give policy analysts and the private bar a better way to estimate when the net benefits of conducting controlled tests are likely to be positive. Perhaps researchers could compare outcomes in published litigation tests with the outcomes that would have been observed if criticized aspects had been corrected. If it turns out that responses to various criticisms substantially change outcomes, then the fuss over

⁵¹ See Russo et al., *supra* note 38; Michael B. Mazis, *Copy-Testing Issues in FTC Advertising Cases*, in *MARKETING AND PUBLIC POLICY CONFERENCE PROCEEDINGS* 122-30 (1996).

⁵² Bloom et al., *supra* note 47, at 23.

controls will be of more than theoretical interest and will likely gain more attention from practitioners and policy makers alike.

Disagreements regarding the appropriateness of alternative outcome measures would also benefit from further research. Should copy tests focus on consumer perceptions, beliefs, purchase intentions, or other measures? Unless policy participants agree on which broad concepts should be measured, even good measures of a construct could be irrelevant to decision makers. Research validating various approaches would be a welcome addition to the public policy and marketing literature and could help to bridge gaps that sometimes exist between the economic, marketing, and legal perspectives.

The information revolution is already changing the face of antitrust analysis. As we head into the Twenty-First Century, this would seem to be an opportune time to consider how improvements in consumer research technology might improve consumer protection analysis. Surely, a consumer protection revolution cannot be far away.