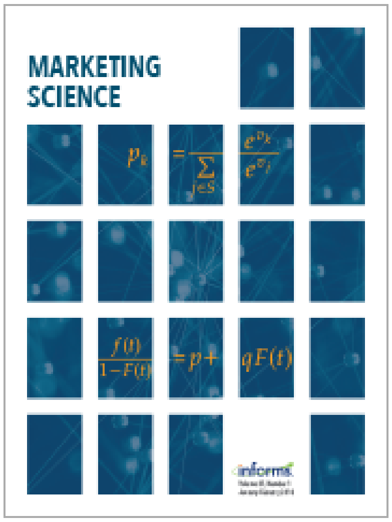

Marketing Science

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

Structural Workshop Paper—Descriptive, Structural, and Experimental Empirical Methods in Marketing Research

Peter C. Reiss,

MARKETING SCIENCE
CELEBRATING 30 YEARS
Vol. 30, No. 6, November–December 2011, pp. 950–964
ISSN 0732-2399 | EISSN 1526-548X | 11 | 3006 | 0950

informs®
http://dx.doi.org/10.1287/mksc.1110.0681

Structural Workshop Paper

Descriptive, Structural, and Experimental Empirical Methods in Marketing Research

Peter C. Reiss
Stanford Graduate School of Business, Stanford University, Stanford, California 94305,
preiss@stanford.edu

What can be learned about marketing phenomena from descriptive, structural, and experimental empirical models? Is structure implicit in a descriptive empirical model? What is a "reduced-form model?" What is a natural experiment, and what can one infer from a study that uses experimental data? Having clear answers to these questions can improve empirical dialog. This paper defines descriptive, structural, and experimental empirical work, provides examples, discusses their similarities and differences, and comments on their strengths and weaknesses. An important theme is that the marketing question and the data available should determine the methods used, and not the other way around. Most of the examples discussed reference linear models that are widely employed in the marketing literature. Many of the points, however, extend to the development and interpretation of cutting-edge nonlinear, dynamic, or nonparametric models used in marketing.

Key words: structural models; experiments; predictive models; instruments; causal effects
History: Received: December 15, 2010; accepted: August 10, 2011; Eric Bradlow served as the editor-in-chief and Jean-Pierre Dubé served as associate editor for this article.

1. Introduction

Marketing researchers currently use a variety of empirical methods to answer research questions. The most common approaches include statistical models that describe patterns and associations, structural models that recover theoretical constructs, and experimental or quasi-experimental studies that measure causal effects. Each of these approaches has strengths and weaknesses when it comes to answering marketing research questions. Not surprisingly, each approach has strong advocates who routinely cite instances in which other approaches are misapplied or misinterpreted. Although debates about the limitations of alternative empirical methods can be useful, they also can be polarizing in ways that do not advance a field's understanding of what each has to offer. Having a clear appreciation of the limits and merits of each approach is important for improving the credibility of all types of empirical research in marketing.

This paper illustrates some basic strengths and weaknesses of descriptive, structural, and experimental empirical models. It does this through examples and discussion of how empirical practice has taken advantage of each approach. It is not meant to be a complete treatment of the many practical issues that arise in implementing these approaches, nor does it attempt to cite all relevant marketing research. Instead, it seeks to stimulate discussion of what each approach has to offer. I emphasize that the questions posed and the data available should determine the modeling approach rather than the other way around. Descriptive analyses, for example, cannot literally test theories. Structural models do not generally provide the best description of data or the best predictive model. Similarly, experimental studies are not assumption-free, and their results may not generalize to real-life marketing applications.

The paper begins with an overview of descriptive empirical models. Descriptive models have long been used in marketing to document interesting patterns, trends, practices, and puzzles (e.g., Kerin 1996, Wilson 2008). Well-executed descriptive empirical work employs easily understood and efficient statistical methods to document generalizable facts. These facts ideally help interpret existing theories or help build new ones. Many concerns about descriptive work center on its generalizability; that is, can insights from descriptive summaries be applied in other contexts? Another set of concerns pertains to attempts by some to draw causal inferences or policy conclusions from simple statistics such as correlations or regressions.

Sections 2 and 3 then provide an overview of structural empirical models. Structural empirical models are a relatively recent development in marketing (e.g., Kadiyali et al. 2001, Chintagunta et al. 2006, Mazzeo 2006), although marketing has for many years relied on sophisticated statistical methods. Structural models combine mathematical, economic, or marketing models of behavior with statistical assumptions to derive estimable empirical models.[1] The "structure" in structural empirical models thus comes from behavioral theories (e.g., marketing, economics, psychology) and statistics. Structural models have the advantage of explicitly linking behavior to data; they also permit researchers to predict how changes in marketing conditions will change consumer and firm behavior. Concerns about structural models often center on their use of simplifying assumptions, lack of fit, and on their not being well connected to marketing institutions.

Section 4 discusses empirical models that seek to infer causation from statistical designs that involve explicit or implicit manipulations of causal variables. I term this the experimental approach; it includes work that uses laboratory experiments, field experiments, "natural experiments," and single-equation instrumental variable techniques. For these studies to be convincing, they must explain how the experimental design and statistical models capture the effect of interest. Absent a clear theoretical model of the experimental effect, it may be unclear what lessons can be extracted from a given experiment or whether those lessons generalize to other settings.

Section 5 contrasts the merits of these different approaches in more detail. In particular, it discusses the misuse of the terms "reduced-form regression" or "reduced-form analysis." It does this by illustrating the difference between a descriptive regression and a reduced-form regression. The blanket term "reduced-form" sometimes leads researchers to misrepresent coefficient estimates as the causal effect of an exogenous variable X on an endogenous variable Y. The section then discusses the proper use of instrumental variables in regression modeling and conditions under which an instrument is valid. It argues that the functional form of the instrument is taken for granted in most applied work and yet it should not be.

The paper concludes with some general observations about empirical work in marketing and issues that arise when assessing what different empirical approaches have to offer.

2. Descriptive Empirical Models

Descriptive empirical work is broadly about the collection, documentation, and interpretation of information. For example, a researcher might document how often stores change prices or use promotional displays. Most current descriptive work in marketing uses data collected from direct observation, surveys, or electronic systems that monitor consumers (e.g., Web-browsing data or point-of-sale information). Sometimes, researchers have the luxury of collecting original data, such as when a researcher collects survey information from consumers. Other times, descriptive research is about characterizing interesting patterns in nonexperimental data. In both cases, good descriptive research is about finding effective and efficient means for summarizing large amounts of information.

It may seem unnecessary to state that good descriptive work is "effective" and "efficient." All too often, however, one hears, "That study had great data, but I don't feel they told me much about it." Effective displays do not happen by accident but are instead deliberately crafted. The most effective displays are self-contained and laid out with an intelligent yet uninformed reader in mind; the analyses answer basic questions the reader may have about how the data were generated, what populations the data represent, and how variables and measures were constructed. Effective displays are straightforward and honest; they should not confuse or deliberately manipulate the reader's impression of the data. A variety of excellent sources illustrate these maxims, beginning with Tukey and continuing with Fienberg (1979), Tufte (1990, 1997), and Wilkinson (2005), to name just a few.

Exemplary descriptive studies in marketing minimally do three things well; they (1) contain clear statements about the facts they document and how those facts inform marketing questions, (2) use relevant and high-quality data that speak directly to the questions of interest, and (3) provide accurate interpretations of the results. Three recent marketing studies that illustrate these principles are Bronnenberg et al. (2009, 2010) and Larson et al. (2005). The first two papers provide intriguing evidence that consumer preferences for brands are highly related to where consumers live. That is, consumers tend to prefer local brands. Moreover, these local preferences are highly persistent—even when consumers move, they tend to carry those preferences with them, although eventually their brand preferences move toward those of the typical consumer at their new locale. These descriptive findings have fostered work examining the sources of local preferences and the competitive implications of preference persistence for nonlocal firms. The Larson et al. paper illustrates how marketers can creatively exploit new technologies to document interesting patterns in electronic tags on shopping carts to follow grocery store customers' movements; these data can then be linked to purchase decisions. Such studies have provided much

[1] The term "structural model" sometimes refers to latent factor or path models. The structure found in those models comes from the latent factor structure imposed on the data. This structure is usually not motivated by an explicit model of behavior.

better information about how consumers respond to store layouts, end-of-aisle displays, checkout lines, etc. These location data have led to richer hypotheses about shopping behavior (see, for example, Hui et al. 2009a, b), and the use of location technologies is being explored in other marketing applications. Similar advances are being made using information collected from the World Wide Web and telecommunications devices.

Descriptive studies run into problems when they do not make a convincing case along one or more of the three dimensions listed above. Most concerning is work that does not provide an accurate interpretation of research findings, or worse, overinterprets or misinterprets those findings. A familiar example is assigning a causal interpretation to a correlation. On the one hand, it seems clear that merely documenting facts without explaining and interpreting them is not very interesting. On the other hand, there is a temptation for researchers to overinterpret findings in an effort to have them noticed. A familiar example is the belief that for a study to be interesting, it must "test" a formal marketing hypothesis. This belief can go so far as to demand that descriptive studies have testable hypotheses, even if these hypotheses are only loosely connected to a model of behavior. Such an approach narrows the role of descriptive research. Descriptive work is not just about informing existing theories; it also can be about directing attention to new issues and facts worthy of theoretical attention.

On a related note, there is a tendency to have descriptive work reject simple null hypotheses. It is important to understand that not all data are informative. Not rejecting a hypothesis can be just as important as rejecting it. Moreover, rejecting a null hypothesis does not prove a specific alternative hypothesis is correct. As is argued later, descriptive research cannot literally test a theory about consumer or firm behavior—only an econometric representation of the theory can serve as the basis for such a test. This is not to say that descriptive work cannot provide evidence favorable or unfavorable to a theory. Indeed, the comparative statics of a theoretical model might suggest that two marketing variables are positively related. However, just because a descriptive model finds a positive correlation does not prove the theory. This is where descriptive studies sometimes overinterpret what they have established. There are often many competing models that could suggest such a correlation (or not). Moreover, correlations are only one potential measure of the association implied by a theoretical result. Despite this, there is a tendency by some to want to interpret their evidence as causal without cautioning readers that other interpretations of an average, a correlation, a coefficient estimate, etc., may be equally valid.

Occasionally, one will encounter the argument that descriptive studies are superior to structural models or other methods because they do not require assumptions about behavior. Sometimes, this position is signaled with expressions such as, "I've estimated a flexible reduced-form model," "It doesn't matter which way I run the regressions," or "I'm just letting the data speak." The position that descriptive work is preferable because it makes few assumptions is untenable. To interpret data, one always needs assumptions. To improve the credibility of a descriptive study, one needs to be clear about what assumptions have been made to reach conclusions about what the data do or do not show. At a minimum, descriptive work makes judgments about the representativeness of samples, the appropriateness of variables, and the validity of statistical methods. Just because these assumptions may not be spelled out does not mean they are not there. For instance, consider how a researcher might go about describing data on two variables, X and Y. The most general object one could describe is the joint probability density function of all the data: pdf $(X, Y) = f(X_1, X_2, \ldots, X_N, Y_1, Y_2, \ldots, Y_N)$. Without more structure, this object has little practical relevance. For this reason, researchers implicitly or explicitly place structure on the joint density in order to describe it. For example, many researchers assume the pairs $\{X_i, Y_i\}$ are independent and identically distributed. This assumption gives rise to a common bivariate density $f(X_i, Y_i)$, which can be estimated flexibly using nonparametric methods or described graphically. It is important to emphasize that we require statistical structure to give a meaningful interpretation to the nonparametric density estimate or even a scatterplot.[2] This much may be unsurprising. However, it is important to understand that that structure could limit the statements that a marketing researcher can make about behavior. For example, if the data are time-series observations on the behavior of an individual, one might be less comfortable assuming that the joint density of X_i and Y_i is the same over time.

To gain a further appreciation for these issues, it is useful to consider an example. Suppose that Y is an individual's *HEIGHT*, and X is an individual's *WEIGHT*. To describe the heights and weights of a random sample of individuals, I might prepare a scatterplot or nonparametric density estimate. Both would show the two variables are positively related; both have the capacity to reveal finer patterns, such as the bivariate density of *HEIGHT* varies significantly with gender, age, and economic status. Once there are

[2] Consider how a scatterplot of individuals' heights and weights would look for a panel versus what it would look like for a cross section of individuals of the same age.

more than a few discrete conditioning variables, as is a common situation in marketing applications, other methods must be used to describe the data. The most common choice is a regression analysis.

But how does the researcher decide which regression best describes the data? Is it better to regress *WEIGHT* on *HEIGHT*, or *HEIGHT* on *WEIGHT*? Many economists and marketing researchers have told me that it clearly makes more sense to regress *WEIGHT* on *HEIGHT* (and other variables). When pressed to defend this choice, they argue that this is a better regression because *WEIGHT* is a behavioral outcome and *HEIGHT* is reasonably thought of as exogenous; thus, *WEIGHT* should be the dependent variable and *HEIGHT* the independent variable.

In making this argument, these researchers have moved beyond wanting to show that *HEIGHT* and *WEIGHT* are positively correlated (something either regression could reveal) to wanting to interpret the output of the regression as being about behavior. This step is not too surprising, as marketing researchers are focused on describing behavior. However, taking this step overlooks the caution raised above about interpreting the results of a descriptive analysis. The danger in ascribing a behavioral (or causal) meaning to the regression estimate of β in

$$WEIGHT = \alpha + \beta\ HEIGHT + \varepsilon \tag{1}$$

is that I know of no explicit theory that suggests why this linear regression describes the effect of *HEIGHT* on *WEIGHT*. This is not to say that biology might not have theories that would lead to a positive association. Rather, it is to say that it is unclear why biology would justify this linear relation.[3] Indeed, biology might say *WEIGHT* is better related to volume, which may be nonlinearly related to *HEIGHT* in humans.

If a specific behavioral (or biological) theory that leads to a linear relation between *WEIGHT* and *HEIGHT* cannot be cited, what can be said about estimates of this regression? The answer is that a regression provides the best linear predictor of *WEIGHT* given *HEIGHT*. If it is interpreted in this way, then the researcher has made it clear that it should not be interpreted as a causal statement about behavior, but rather a predictive one. The regression simply describes an association, or alternatively, the regression provides a means for predicting *WEIGHT* from observations on *HEIGHT*. Again, problems can arise when a researcher steps beyond these interpretations and seeks to interpret β causally as an estimate of "how much an individual's weight will increase if they grow another inch."

To underscore this point, it is useful to consider what might be said about estimates of the reverse regression,

$$HEIGHT = \gamma + \delta\ WEIGHT + \nu. \tag{2}$$

Some would argue that this regression makes less sense because *WEIGHT* does not causally determine height. But here again, the researcher would be judging this statistical model based on an unspecified behavioral theory rather than judging it for what it does—it summarizes the association between heights and weights in a given sample and delivers the best linear predictor of *HEIGHT* given *WEIGHT*. This interpretation is correct regardless of whether *WEIGHT* causes *HEIGHT* or the converse is true. Which regression we prefer thus should be based on what variable we are trying to describe or predict, not on which variable we think is a better approximation to a behavioral relation. As shown below, we have to do more to establish causality or draw behavioral conclusions from a regression.

Despite these arguments, some readers may still feel that the regression of *WEIGHT* on *HEIGHT* has more meaning than that of *HEIGHT* on *WEIGHT*. Some of this impression may be a by-product of the belief that one of the two relations must represent the true (causal) relation. Such readers might justify *WEIGHT* on *HEIGHT* then by arguing that *HEIGHT* on *WEIGHT* makes no sense from an experimental perspective—it is hard to imagine that fattening a random sample of adults could cause an increase in their overall height. Experimentalists seek to identify causality by conducting lab experiments, finding "natural experiments" in nonexperimental data, or using instrumental variables. Their view is that causation questions can be answered through statistical designs that exogenously manipulate what would otherwise be potentially confounded or endogenous variables. By doing so, experimental work seeks to go a step beyond purely descriptive work in drawing inferences from data. In this example, however, it would be an unusual experimental design where *HEIGHT* could be manipulated independently of *WEIGHT*. Even in situations where it is possible to conduct experiments to interpret the resulting effects as causal, one still needs a theory. This theory minimally says why the treatment sample differs from the control sample in the experiment.

A potential limitation of descriptive work that is sometimes shared with structural and experimental work is the issue of generalizability. To what extent are findings relevant beyond a particular study? This question often is not explicitly answered and instead

[3] For example, biology might suggest that nutrition and genetics play a large role in producing a positive association, or that the relationship between *WEIGHT* and *HEIGHT* is nonlinear, or that the reverse causation holds when malnutrition is a problem.

left for readers to evaluate. In general, marketing researchers who do descriptive work face a dilemma when it comes to the potential generalizability of their study. If the researcher narrows the focus of his or her study to a particular market, firm, or set of consumers, he or she often can hold constant many nonessential factors that could also affect the variables of interest. This focus simplifies statistical models and can result in a more compelling description of the data. By narrowing the focus of the study, however, the researcher will tend to raise questions about the generalizability of his or her findings. To illustrate, consider a study that wishes to describe how consumers shop online, what information they pay attention to, and how they respond to different displays of information. To do this, a researcher could analyze data from a particular website. By focusing on one website, the researcher can take advantage of special ways that the website displays information and directs consumers to different Web pages. This focus can ultimately be a disadvantage because it raises concerns about whether the website and consumers on it represent all websites and consumers. These types of issues are familiar to descriptive researchers. They do not reduce the value of descriptive work, but they may temper the types of inferences that can be drawn from descriptive analyses.

3. Structural Models

The previous section characterized descriptive empirical work as summarizing the joint distribution of data or features of that distribution. Descriptive summaries can be parametric or nonparametric, or Bayesian or non-Bayesian. They may contain simple descriptive statistics, regressions, and maximum likelihood estimates, or more complicated nonparametric estimates of joint densities, conditional densities, or conditional moments. Structural models, on the other hand, are empirical models of individual or firm behavior that give rise to a particular joint density of the data, i.e., $f(X, Y)$, or properties of this joint density, such as the conditional density $f(Y \mid X)$ or conditional expectation $E(Y \mid X)$.

Structural models rely on two main ingredients (see Reiss and Wolak 2007). The first is a formal model that mathematically relates endogenous decision variables Y to exogenous variables X. An example would be a demand equation derived from a utility maximization model that links items and quantities consumed to a consumer's income, advertising, prices, etc. The second key component is a stochastic specification that maps the theoretical model to a joint density of data. This stochastic specification recognizes that mathematical models of behavior do not fit data perfectly. For example, a Cobb–Douglas demand function will not perfectly fit the consumption decisions of a random sample of households. A structural modeler therefore introduces heterogeneity in consumer preference parameters, consumer decision-making errors, or measurement errors to explain why a mathematical theory does not perfectly explain data. The fact that structural models contain errors raises the issue of how well the structural model ultimately fits the data. This is the same question that descriptive work faces and is one for which there is a well-developed set of strategies for reporting in- or out-of-sample measures of model fit.[4]

Marketing researchers have become interested in structural models for three main reasons. First, structural models allow researchers to estimate useful theoretical objects. These include behavioral constructs (e.g., degree of risk aversion), decision rules (e.g., how much to bid in an online auction), or parameters (e.g., a constant price elasticity or a marginal utility of income). Second, once estimated, structural models can answer marketing policy questions. For example, if two channel partners merge, what will happen to equilibrium prices? In performing these counterfactuals, the structural modeler alters key parts of the model while assuming other parts do not change. For instance, when the two partners merge, one might assume that consumer preferences do not change. Third, structural models allow the researcher to assess what aspects of behavior constrain the data. For example, a researcher can explore how the distribution of bids in an online auction might be different if bidders were risk neutral as opposed to risk averse. Or, one might ask how uncertainty about competitor's costs might affect the prices a competitor charges. Additionally, structural models allow researchers to compare different theories based on how they fit the data in sample or out of sample. Here, it should be noted that any tests or comparisons of models are not general tests but are instead tests that rely on the assumptions that were used to develop the competing models. For example, a study testing whether a collusive or a Bertrand-Nash pricing model better explains prices might assume a linear demand and constant marginal costs. In rejecting one pricing model in favor of the other, the study does so only under these specific assumptions about demand and costs.

Finally, it is useful to comment on the use of structural model counterfactuals in marketing. Sometimes, marketers express the concern that structural models are value-free descriptions of optimizing behavior (e.g., what price to set to maximize profits), whereas many marketers are ultimately concerned with normative questions (e.g., how to change prices

[4] See here also the cautionary comments on predictive performance versus policy relevance discussed in Bronnenberg et al. (2005).

to increase profitability). The concern continues that if a structural model presumes optimal behavior (e.g., firms have set prices to maximize profits), then it is not useful because there is no room to improve behavior. Most structural modelers would disagree this need be true. Structural models describe behavior subject to institutional and practical realities. How one understands and incorporates these realities impacts what types of positive or normative conclusions one can draw. To illustrate, one might observe a firm setting the same price in two different markets. One structural model might treat this behavior as a constraint (heuristic, etc.), estimate demand and supply parameters, and then calculate the profit consequences of charging one versus two different prices. (See Chintagunta et al. 2003.) This structural model therefore has the potential to make normative statements. To do so, however, the marketer will first have to address the question of why the constraint (or heuristic) is there initially. Similarly, one could instead formulate a structural model in which the firm optimally picks a different price in each market, but these prices turn out to be the same. In such a model, there is little room to make normative recommendations for changing prices. On the other hand, as a positive model of pricing, there is room to explore (and potentially test) just what it is about demand and supply in both markets that makes the same price optimal. Yet other modeling strategies can be employed here to introduce alternative positive or normative dimensions into marketing policy evaluations (e.g., Bronnenberg et al. 2005). The key point is that an advantage of structural models is that they enforce the discipline of being clear about what behavior is being modeled, in what sense it is optimal, and in what sense it might be changed.[5]

3.1. An Extended Example

To illustrate further the sources of "structure" in structural models, it is useful to revisit the demand and supply models found in most econometrics textbooks. These models view observed prices and quantities as equilibrium outcomes corresponding to the intersections of market demand and supply curves. The market demand curve is obtained by aggregating the demands of individual consumers. These individual demands can be built up from utility-maximizing models of consumer behavior. The market supply curve is derived similarly. For example, in a competitive market, market supply is the sum of individual firm supply curves. A complete structural demand and supply model consists of three (or more) equations: the market demand equation, the market supply equation, and a market clearing or equilibrium condition. With linear demand and supply curves, for instance, one would have

$$\begin{aligned} &\text{(Demand)} && Q^D = \beta_{10} + \beta_{11} x_1 + \gamma_{11} P, \\ &\text{(Supply)} && Q^S = \beta_{20} + \beta_{22} x_2 + \gamma_{21} P, \\ &\text{(Market clearing)} && Q = Q^D = Q^S. \end{aligned} \tag{3}$$

The first two equations are "behavioral" in that they represent how much quantity market participants would demand at price P and how much quantity participants would supply at price P, respectively. The structural modeler is interested in recovering consistent estimates of the demand and supply parameters β and γ.

The third equation does not contain any parameters and is seemingly trivial, yet it plays an important role in the structural model. To see this, consider what would happen if one instead imposed the rationing assumption $Q = \min(Q^D(P), Q^S(P))$, as is common in disequilibrium models (e.g., Fair and Jaffee 1972, Maddala 1986). Economically, this assumption says price does not adjust to equilibrate demand with supply; instead, at the given price, demand or supply decisions limit the quantity that is sold. This change in "structure" affects predictions about prices and quantities. Under the usual market-clearing assumption, prices adjust so that $Q^D = Q^S$. Under the assumption $Q = \min(Q^D(P), Q^S(P))$, there may be two different prices that can correspond to a given quantity.[6]

These different theoretical predictions could easily be tested if one knew the market demand and supply curves. This is because with six or more independent observations of the variables $\{P, Q, x_1, x_2\}$, one should be able to recover the six conditional mean parameters $\{\beta_{10}, \beta_{11}, \beta_{20}, \beta_{22}, \gamma_{11}, \gamma_{21}\}$ exactly.[7] Of course, in practice, theoretical models never fit marketing data perfectly. The absence of a perfect fit does not mean the theory is fundamentally wrong, but rather that the theory is a simplification of the process that generated data. In reality there are random shocks that affect behavior, behavior is heterogeneous in unobservable ways, decision makers make mistakes, and outcomes are measured with error (see Reiss and Wolak 2007). As an example, one might add normally distributed, mean-zero errors to the demand and supply equations and claim that they reflect measurement errors in prices and quantities.

[5] See also the papers on marketing policy evaluation in the *Journal of Marketing Research* (February 2005) issue on econometric models for marketing decisions.

[6] To see this, fix a quantity below the quantity that equates demand and supply. To achieve this quantity, demanders may face a high price (thereby rationing suppliers) or suppliers may face a low price (thereby rationing demanders).

[7] This assumes the model is identified in the traditional sense and the observations are linearly independent.

The stochastic terms of a model, together with the functional form of the demand and supply equations and the market clearing condition, deliver the structural model. This structural model implies a joint density of market prices and quantities given x_1 and x_2—$f(P, Q \mid x_1, x_2, \beta, \gamma, \Sigma)$.[8] The functional form of this conditional density and its dependence on the "structural" parameters β and γ is what distinguishes structural models from descriptive models. A structural model has content to the extent that it allows us to recover consistent estimates of β and γ. The descriptive approach, on the other hand, has no behavioral parameters and is limited to summarizing statistical features of the density $f(P, Q \mid x_1, x_2)$.

For the structural model to be more than a descriptive statistical model, one must be able to recover consistent estimates of some or all the structural parameters. This difference is best understood by returning to the demand and supply model (3) and adding errors

$$\begin{aligned} &\text{(Demand)} \quad Q^D = \beta_{10} + \beta_{11}x_1 + \gamma_{11}P + \varepsilon_1, \\ &\text{(Supply)} \quad Q^S = \beta_{20} + \beta_{22}x_2 + \gamma_{21}P + \varepsilon_2, \\ &\text{(Market clearing)} \quad Q = Q^D = Q^S. \end{aligned} \tag{4}$$

To obtain the conditional density $f(P, Q \mid x_1, x_2, \beta, \gamma, \Sigma)$, one can solve for price and quantity as a function of the conditioning variables x_1, x_2, and the demand and supply errors. The resulting set of equations is

$$\begin{aligned} Q &= \pi_{10} + \pi_{11}x_1 + \pi_{12}x_2 + v_1, \\ P &= \pi_{20} + \pi_{21}x_1 + \pi_{22}x_2 + v_2. \end{aligned} \tag{5}$$

These equations describe how the density of price and quantity, $f(P, Q \mid x_1, x_2, \pi, \Omega)$, depends on the conditioning variables x_1 and x_2 and linear combinations of the demand and supply errors. For example, when the demand and supply errors are jointly normally distributed conditional on x_1 and x_2, the conditional density will also be joint normal.

Estimating the conditional density $f(P, Q \mid x_1, x_2, \pi, \Omega)$ could well be the object of a descriptive study. For example, one could use nonparametric methods to estimate it. One could also choose to estimate just the conditional mean of the conditional density using nonparametric regression. Estimates of the conditional means do not deliver estimates of the behavioral demand and supply parameters, however. These structural parameters only come from knowing how to recover the demand and supply parameters β and γ from information contained in the conditional distribution. This information is the "structure" in the structural model.

Here, the process of recovering the structural parameters involves relating the structural parameters to the reduced form. Historically, the term "reduced form" was used to denote the set of Equations (5) that expressed the endogenous variables in terms of exogenous variables and unobservables. By this definition, a reduced form exists only to the extent that it can be rationalized by a structural model. Thus, the fact that the reduced form (5) is linear in x_1 and x_2 and the demand and supply errors is a consequence of starting with linear demand and supply equations. This linearity is what makes π_{12} interpretable as the partial derivative of the conditional expectation of quantity given x_1 and x_2 (i.e., $\pi_{12} = \partial E(q \mid x_1, x_2)/\partial x_2$). Absent this specific structure, π_{12} instead becomes simply a coefficient in the best linear predictor of quantity given x_1 and x_2. Although this distinction may at first appear minor, it goes to the heart of the difference between a structural model and a descriptive or predictive model.

Today, many empirical researchers use the term "reduced-form analysis" to describe a regression of endogenous variables on exogenous variables. In doing so, they sometimes mistakenly give the regression coefficients causal interpretations. As we have just seen, a reduced form does not exist without an underlying structural model. Absent a structural model, all we have is a linear regression that can be interpreted as delivering best linear predictors. To illustrate, suppose that in contrast to (4), the structural demand and supply equations have the form

$$\begin{aligned} &\text{(Demand)} \quad \ln Q^D = \beta_{10} + \beta_{11}x_1 + \gamma_{11}\exp(P) + \varepsilon_1, \\ &\text{(Supply)} \quad \exp(Q^S) = \beta_{20} + \beta_{22}x_2 + \gamma_{21}P + \varepsilon_2, \\ &\text{(Market clearing)} \quad Q = Q^D = Q^S. \end{aligned}$$

The reduced form for this system is not available in closed form. Nevertheless, a researcher might proceed by estimating the "reduced-form" regressions

$$\begin{aligned} Q &= \pi_{10} + \pi_{11}x_1 + \pi_{12}x_2 + v_1, \\ P &= \pi_{20} + \pi_{21}x_1 + \pi_{22}x_2 + v_2. \end{aligned}$$

It should be clear that these regressions do not have a causal interpretation, such as how much expected quantity would change when x_1 or x_2 changes. Instead, these are descriptive regressions that, when estimated by ordinary least squares (OLS), yield the best linear predictor of price and quantity given x_1 and x_2. There is no obvious way to recover the β and γ parameters from estimates of the πs. In general, one will have to work with the joint density $f(P, Q \mid x_1, x_2, \beta, \gamma, \Sigma)$ for this nonlinear system to determine whether the structural parameters can be recovered from estimates of π and Ω.

[8] Here, Σ represents the variance–covariance matrix of the normally distributed demand and supply errors. The variance–covariance matrix of the errors in the reduced form is denoted by Ω.

Without being clear on the structural model, best linear predictors cannot reveal much about demand and supply behavior. The best linear predictor coefficients do, however, have meaning. For example, the quantity regression estimate of π_{12} can be interpreted as follows:

> Consider two observations of $\{Q, x_1, x_2\}$ drawn from the population for which the demand and supply model is relevant. Suppose that these two observations have the same x_1 value, and their x_2 values differ by one unit. Then the best (in a mean squared error sense) prediction for the difference in prices is π_{12}.

The danger in labeling such regressions as "reduced-form regressions" is that it suggests a structure that is known. A better label is "descriptive regression" or "predictive regression." These labels make it clear that caution should be used in giving causal interpretations to the results, just as one would hopefully do when in describing the results of a regression of *HEIGHT* on *WEIGHT* (or vice versa).

3.2. Caveats

It often is difficult and time consuming to develop and estimate structural models. This is because the structural model must be flexible enough to fit data while at the same time simple enough to estimate. Often, simplicity comes at the cost of realism. This is because it is hard to formulate simple theoretical models that respect the institutional realities of marketing applications and yet are estimable given the data available. For example, to estimate a model of ready-to-eat cereal manufacturer competition, one might assume that each cereal manufacturer produces only one brand of cereal. Whereas this assumption reduces the number of prices the researcher needs to model, it potentially ignores cross-brand price promotion opportunities. Compelling structural models attempt to avoid these types of assumptions when they significantly limit the model's practical relevance.

Structural modelers must also confront the reality that theory rarely delivers a complete empirical model. Typically, structural modelers must add variables, parameters, functional form assumptions, and other elements to make the theory operational. For instance, choice theory does not spell out what a car shopper's utility function looks like, what attributes the consumer cares or knows about, or how the consumer budgets for the car. To make progress formulating a structural choice model, the researcher must make assumptions about these elements (and others). Although these elements often permit a much better match of the theoretical model to the data, they can appear ad hoc, or even arbitrary. It is therefore useful for structural modelers to explore the sensitivity of their conclusions to these additional elements.

Recently, structural modelers and econometricians have begun to explore more systematically the role of supplemental assumptions used to complete models. In particular, the literatures on nonparametric identification and partial identification have developed methods for analyzing the sensitivity of estimates or inferences to alternative economic and statistical assumptions. The literature on nonparametric identification takes as given a base set of assumptions and then asks what functions or function properties are recoverable from the joint distributions of data. As an example, many marketing studies have estimated panel or cross-section discrete choice models that rely on parametric specifications of consumer utility and how utility differs across consumers. Many of these choice models assume that consumer utility is extreme value or normally distributed, conditional on unobserved consumer preferences (e.g., random coefficients) that are normally distributed. These distributional assumptions typically have no economic defense but are rather viewed as conventional or computationally convenient. A number of recent papers have sought to examine whether it is possible to identify nonparametrically the distribution of consumer heterogeneity or other properties of tastes from the joint distribution of product attributes and product choices.[9] Beyond showing that it is possible to relax standard parametric assumptions, these papers suggest what it is about the data and economic choice model that identifies taste heterogeneity. (See also Matzkin 2007.)

Although nonparametric identification arguments have the potential to establish that functional forms can flexibly be recovered from joint distributions of data, the practical reality is that it may not always be possible to demonstrate that a model is nonparametrically identified. In this case, identification arguments often rely on showing what information in the joint distribution of the data *uniquely* identifies the chosen parametric functional forms.[10] Such arguments are standard in econometrics (e.g., identification of the parameters in the linear simultaneous equations model). The literature on partial identification takes a broader approach to the identification of structural parameters. It treats model assumptions as potentially yielding bounds on parameters or functions rather than unique values. An early example of such a bounds analysis is Frisch (1934); he showed that although the regression slope in a classical measurement error model was not identified, bounds could be

[9] See, for example, Briesch et al. (2010), Fox and Ghandi (2009), and Bajari et al. (2009).

[10] Considerable attention is also being devoted to identification in semiparametric models that combine parametric and nonparametric assumptions.

placed on it.[11] Since then, partial identification strategies have been used to estimate a variety of empirical models. Many of these models rely on inequalities implied by optimizing behavior. For instance, the economics of discrete strategic interaction models, such as those being explored in marketing, readily lends itself to bounds analyses. (See, for example, Ellickson and Misra 2011, Manski 2003, Manski and Tamer 2002, and Berry and Reiss 2007.)

4. The Experimentalist Approach

A third style of empirical work in marketing is the experimental approach, which includes not only lab experiments run by behavioral marketing researchers but also randomized field experiments, "natural experiments," and instrumental variable methods that correct nonrandom assignment problems. The general logic of the experimental approach is to find data that have exogenous variation in variables that would otherwise be endogenous or confounded because of behavior, omitted variables, selection effects, etc.

To illustrate the logic of the experimental approach, consider a marketer who is interested in measuring how consumer purchases change in response to a price change. If nonexperimental price and quantity data were available, one might be able to use structural modeling techniques to estimate a consumer demand function. These estimates would, of course, require functional form and other assumptions. A key issue in the estimation would be how to handle the likely endogeneity of price. There could be many reasons for this endogeneity. For example, the error could include unobserved factors correlated with price, such as advertising, competitor prices, or discounts. Or, firms could have decision rules that, say, raise price when the unobserved error ε is high. The experimentalist seeks to circumvent these issues by developing data where price is manipulated independently of these factors or by finding data where price can reasonably be regarded as being exogenous to these factors. In some cases, it might make sense to do laboratory experiments with consumers as a means of generating data. This approach is most desirable when the researcher wants to establish a causal connection that otherwise could not be established from nonexperimental data.

Whereas laboratory experiments provide the researcher with a controlled environment to demonstrate the internal validity of their causal hypothesis, such experiments may have difficulty demonstrating external validity. Perhaps for this reason, marketing researchers are turning to field experiments as a more realistic means for gauging consumer behavior. Returning to the demand example, a field experiment might attempt to replicate a controlled lab environment by having a company vary prices while holding other marketing variables fixed.[12] Although such experiments have greater external validity, a potential problem with them is that it may be difficult to hold constant all factors that matter (such as competitor prices, discounts, advertising, etc.). Experimentalists have also relied on other approaches, including "natural experiments," instrumental variables, or propensity scores. These approaches do not automatically solve the problems noted above. In general, they too require support from theory to clarify what they can and cannot measure.

Natural experiments occur when someone or something other than the researcher has exogenously manipulated the causal variables of interest. In some cases these manipulations occur when subsets of the data arbitrarily receiving a single "treatment." In other cases there might be a multivalued index that ranks the varying strength of treatments. Causal effects are most often inferred by comparing the different subsets treated. For example, in a demand study, one might have data from two different geographic areas that are identical apart from a much higher per-unit tax rate in one area. In this case, the tax authorities might be regarded as running the "experiment." However, note that absent an explanation (or theory) about the determinants of tax rates, it would be unclear whether the variation in tax rates is exogenous to the errors in demand.

To gauge the slope of demand using the differences in taxes, a researcher might estimate the coefficient β_1 in the regression

$$Q = \beta_0 + \beta_1 P + \varepsilon \tag{6}$$

using instrumental variables. This would result in the population estimate

$$\beta_1 = \frac{\mathrm{cov}(Z, Q)}{\mathrm{cov}(Z, P)} = \frac{E(Q \mid Z=1) - E(Q \mid Z=0)}{E(P \mid Z=1) - E(P \mid Z=0)},$$

where Z denotes a 0–1 dummy instrumental variable equal to 1 in the high tax state. The first piece of this formula is recognizable as the formula for the indirect least squares instrumental variable estimator. The second piece decomposes the covariances using the fact that the instrument Z is a binary indicator.

[11] Most researchers are familiar with the lower bound, which is the same as that obtained by noting that the estimated regression slope coefficient in a bivariate regression with measurement error is biased downward.

[12] Anderson and Simester (2003, 2004) report three interesting experiments in this vein. There are other examples of experiments done with traditional advertising (see Lodish et al. 1995) or modern media (Lewis and Reiley 2009). Other interesting field experiments have been done with product attributes (e.g., Levav et al. 2010). For other possibilities, see Stout (1969).

Replacing the population unknowns with their sample counterparts results in a consistent estimator of the demand slope, provided the (discrete) instrument Z is correlated with price (meaning the mean of P differs with Z), and the error ε is mean independent of Z. Verifying these assumptions can be difficult in practice, as will be shown.

The search for these types of natural experiments, or, more generally, instrumental variables, has become a popular way of estimating causal effects. The main attractions of this approach seem to be its operational simplicity (find the right data or instrument) and its conceptual simplicity (it appears not to require the overhead of developing a structural model with lots of assumptions). In fact, even in situations where there are natural experiments or seemingly valid instruments, the experimental approach requires additional assumptions to interpret the resulting estimator. This point has been made in a variety of economic applications (e.g., Heckman 1997, Rosenzweig and Wolpin 2000, Reiss and Wolak 2007, Keane 2010) and applies equally in marketing applications.

5. Contrasting the Approaches

Having separately discussed three widely used approaches to empirical work in marketing, this section compares them using one application so as to better understand what each can accomplish. Consider again a marketer with price and quantity information who contemplates running the simple regression:

$$Q = \beta_0 + \beta_1 P + \varepsilon. \tag{7}$$

This regression looks like the structural demand specification (4) because quantity is on the left-hand side and price is on the right. Is this relation a demand equation? If it is not, how should a researcher interpret regression estimates of this equation?

The answers to these questions are clear if we are solely interested in descriptive questions, such as how well quantity can be predicted using price. Under general conditions, ordinary least squares provides consistent estimates of the best linear predictor of quantity given price. This statement is true whether this is a demand equation or not.

An experimentalist might claim that with the right experimental data, Equation (7) can be interpreted as a demand equation. To estimate it one would need exogenous variation in price or an instrumental variable to get a consistent estimate of the demand slope β_1. This argument implicitly treats (7) as the correct conditional expectation of quantity given price. Recall that the best linear predictor of quantity given price is not the same as the conditional expectation of quantity given price, $E(Q \mid P)$. The best linear predictor solves the statistical prediction problem

$$\min E(Q - \beta_0 + \beta_1 P)^2.$$

The population solution to this problem is

$$\beta_0 = E(Q) - \beta_1 E(P),$$
$$\beta_1 = \operatorname{cov}(P, Q)/\operatorname{var}(P).$$

These relations hold even when price is endogenous and the expectation of quantity given price is nonlinear in price. The solutions for the coefficients make clear what information about the distribution of price and quantity is summarized in the coefficient estimates.

Whereas regression estimates of (7) always have a descriptive interpretation, it may also be possible to give them a structural interpretation. To give the coefficient estimates of this equation a deeper structural or causal interpretation, we must first address why we have regressed Q on P and not P on Q. Additionally, we must justify the use of a linear relation and explain what the error term represents. This challenge confronts both experimentalists and structural modelers.

To answer the first question, structural modelers appeal to theory. Economic theory treats prices and quantities as either market-determined variables (e.g., in competitive markets) or choice variables (e.g., in oligopoly markets). In a market setting, demand curves can either be expressed as demand

$$Q = h(P, X)$$

or inverse demand

$$P = h^{-1}(Q, X).$$

Similarly, supply curves can either be expressed as supply

$$Q = g(P, W)$$

or inverse supply

$$P = g^{-1}(Q, W).$$

These expressions show that from a theoretical perspective, simply putting Q on the left-hand side does not guarantee that we are estimating a demand curve. The only way we know whether we are estimating a demand curve is by observing what variables other than price and quantity are included or excluded from the equations of interest. A demand equation is one that includes X and excludes W. As it stands, Equation (7) does not have any X variables, meaning that for it to be a demand specification, we have to argue that there are no observable X variables and that we can account for any unobserved X variables in the estimation (see below). In general, a structural modeler can remove ambiguity about structural regressions such as (7) by specifying both the demand and supply equations.

The error terms are also an important part of any structural model. If Equation (7) is to be interpreted as a demand equation, then it is important to understand the sources of the error ε. A common concern raised in demand estimation is that price is "endogenous." Because of this, ordinary least squares will deliver inconsistent estimates of the linear demand coefficients. This concern about endogeneity is sometimes raised by those estimating descriptive regressions even though the best linear predictor interpretation of a descriptive regression is correct when there is endogeneity. Taking the additive linear structure of (7) as given, the conditional expectation of quantity given price is

$$E(Q \mid P) = \beta_0 + \beta_1 P + E(\varepsilon \mid P),$$

implying that we can write regression (7) as a conditional mean plus a new error:

$$Q = \beta_0 + \beta_1 P + E(\varepsilon \mid P) + \xi.$$

In general, $E(\varepsilon \mid P)$ may vary with price, making the interpretation of β_1 as a "price effect" difficult without knowing the structure of $E(\varepsilon \mid P)$.

By placing structure on $E(\varepsilon \mid P)$, it may, of course, be possible to provide a clearer interpretation of β_1. A familiar way of placing structure on ε and its correlation with price is to find an instrumental variable. The most common practical definition of an instrumental variable is "a variable correlated with the endogenous variable (here price) but uncorrelated with the error." In symbols the researcher seeks a variable Z that satisfies the following two conditions:

$$\text{(Relevance condition)} \quad E(ZP) \neq 0,$$

$$\text{(Exogeneity condition)} \quad E(Z\varepsilon) = 0.$$

Together, these two conditions say that we have found a variable Z that is correlated with the linear part of the conditional mean of quantity ($\beta_0 + \beta_1 P$) and uncorrelated with the residual that potentially depends on price ($E(\varepsilon \mid P) + \xi$).

A critical practical question is, how did the researcher come up with this instrument absent knowledge of the correlation structure between the residuals and the regressors? Without knowledge or assumptions about the source of the error term, it is difficult to argue that the exogeneity and relevance conditions are satisfied. In addition, it turns out that the exogeneity and relevance conditions are necessary but not sufficient conditions for instrumental variables estimators to produce consistent estimates. This latter point is particularly important for researchers who choose instruments independently of a model.

To appreciate this last point, suppose that we are trying to estimate the quantity equation:

$$Q = \beta_1 P + \beta_2 X + \varepsilon,$$

and we have a second "supply" equation:

$$P = \pi_1 X + \eta,$$

where ε and η are contemporaneously correlated, there is one X variable, and X is independent of either error. As the model stands, the β coefficients in the quantity equation are not identified because there is no excluded exogenous variable (or instrument) W. Absent a theory about what might be an appropriate instrument, a researcher might instead appeal to the two necessary conditions—relevance and exogeneity—to find an instrument. Suppose, for example, I construct a new instrument Z by taking X and adding random noise to it—i.e., I set $Z = X + \omega$, where ω is a mean-zero error that is independent of X, ε, and η. This new variable Z satisfies the exogeneity and relevance conditions: it is correlated with price (because X is correlated with price) and uncorrelated with ε (because ω is independent by construction).

My hope in creating this example is that the reader is uneasy about this approach to finding an instrument for price. How can adding noise to a valid instrument suddenly create a new valid instrument? And yet, this instrument satisfies the standard conditions researchers believe deliver a valid instrument. Understanding why it does not create an additional instrument is important for understanding instrumental variable methods and the sense in which they can estimate a causal effect of interest.

To see why the new instrument Z does not work, consider the two moments that define the instrumental variable (IV) estimator (which, in this just-identified case corresponds to indirect least squares):

$$\begin{bmatrix} \sum_{1=1}^{N} x_i p_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{1=1}^{N} z_i p_i & \sum_{i=1}^{N} z_i x_i \end{bmatrix} \hat{\beta}_{\mathrm{IV}} = \begin{bmatrix} \sum_{1=1}^{N} x_i q_i \\ \sum_{1=1}^{N} z_i q_i \end{bmatrix}.$$

For $\hat{\beta}_{\mathrm{IV}}$ to be a consistent estimator, we require the probability limit of the matrix on the left-hand side to converge to a positive definite (invertible) matrix when appropriately scaled by functions of the sample size. Suppose that the first matrix is of order N. Then the probability limit of the first matrix is

$$\operatorname{p\,lim} \frac{1}{N} \begin{bmatrix} \sum_{i=1}^{N} x_i (x_i \pi_1 + n_i) & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} (x_i + \omega_i)(x_i \pi_1 + n_i) & \sum_{i=1}^{N} (x_i + \omega_i) x_i \end{bmatrix}$$

$$= \begin{bmatrix} Q\pi_1 & Q \\ Q\pi_1 & Q \end{bmatrix},$$

where $Q = \text{p}\lim((\sum_{i=1}^{N} x_i^2)/N) > 0$. This matrix is of deficient rank and is hence not invertible. This means that the randomly generated instrument Z fails to identify the coefficient vector β. The reason why the relevance and exogeneity conditions are not sufficient to identify β is that the rank condition for identification is not satisfied.

Upon seeing this example, some researchers dismiss it as absurd. No self-respecting marketing researcher would propose using an instrument based on computer-generated noise. Instead, some researchers who have seen this example propose using the square (or square root, etc.) of X.[13] Their argument is that if one can assume that the error term ε is mean independent of X, i.e., $E(\varepsilon \mid X) = 0$, then any function of X should work as an instrument. The problem is that these functions do not work either. For example, with X^2,

$$\text{p}\lim \frac{1}{N} \begin{bmatrix} \sum_{i=1}^{N} x_i(x_i\pi_1 + \xi_i) & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i^2(x_i\pi_1 + \xi_i) & \sum_{i=1}^{N} x_i^3 \end{bmatrix} = \begin{bmatrix} Q\pi_1 & Q \\ P\pi_1 & P \end{bmatrix},$$

where $P = \text{p}\lim((\sum_{i=1}^{N} x_i^3)/N)$. This matrix is singular, as in the previous case, meaning β is not identified.

The problem with both of these examples is that even though the statistical relevance and exogeneity conditions are satisfied, neither condition recognizes the second structural equation. Put another way, the problem is that there are no additional exogenous variables that affect the conditional mean of price but do not enter the quantity equation. This example thus illustrates the danger of not using theory to identify a causal effect of interest.

This last point is particularly important for the marketing literature. Many marketing researchers see a regression estimated by least squares and almost reflexively raise concerns about endogeneity and the need for instrumental variables. This happens even when a researcher is estimating a descriptive regression and treating the results as such. Somehow, they believe that even descriptive prediction equations need to be concerned with endogeneity. Such concerns confuse the purpose of descriptive versus structural regressions. Regressing quantity on price under very general conditions produces consistent estimates of the best linear predictor of quantity given price. What it does not necessarily do is produce estimates of a demand or supply curve. To obtain those behavioral relations, we need a structural model of both endogenous variables (price and quantity). Although we can regress quantity on price using instrumental variables without specifying a second equation, we may still be subject to the identification problem raised above.

To illustrate how these issues arise in practice, it is useful to turn to an applied example. Epple and McCallum (2006) have developed a canonical data set with which to explore issues that arise in estimating demand and supply models. Their data consist of annual data on the quantities and prices of broiler chickens sold in the United States. One of their specifications is a log-log demand specification,

$$\ln Q = \beta_0 + \beta_P \ln P + \beta_I \ln I + \varepsilon, \tag{8}$$

where I is real income. Imagine three researchers, each of whom believes this equation represents a demand curve. They wish to identify the price elasticity of demand, β_P, which gives the percentage change in aggregate broiler demand with a 1% increase in price.

Suppose each of our researchers decides to use an instrumental variable approach to estimating β_P. Researcher 1 describes his choice of instrument as follows: "I used a [log] time trend as an instrument because prices are upward trending during the period and time is clearly an exogenous variable." Researcher 2 states, "I used the price of corn as an instrument. Corn is used in chicken feed and therefore likely something that would affect the marginal cost [supply] of producing chickens." Researcher 3 proposes something close to Researcher 2: "I used the log price of corn as an instrument…because the demand equation is formulated in natural logarithms."

Which, if any, of these researchers, has a valid instrument (or at least has provided a solid defense of his instrumental variable)? None of the three has explicitly specified a supply equation, although Researchers 2 and 3 appear to have selected their instrument with a supply equation in mind. The difference between Researcher 2's and Researcher 3's choice raises an intriguing question: Does the functional form of the instrument matter? Or should it matter? This issue is rarely explored in applied work, yet without an explicit supply equation, it seems difficult to choose between the instrument being the corn price in levels, natural logarithms, or some other function.

One potential way in which researchers judge an instrument is to see whether the results "make sense" and are "robust." By this, they mean, do the proposed instruments result in IV estimates that are different from the OLS estimates by the magnitude/sign one might expect, and do the IV estimates make economic sense? Table 1 contains the OLS estimate for Equation (8) and IV estimates for each researcher's instrument.

[13] This assumes that X is not a 0–1 indicator variable.

Table 1 Broiler Demand Equation Estimates

Estimated parameter	OLS	ln(Time trend)	IV ln(Corn price)	Corn price
β_0	−4.86	−2.70	6.42	−1.34
	(0.67)	(0.94)	(7.16)	(1.44)
β_P	−0.28	−0.52	−1.52	−0.67
	(0.07)	(0.10)	(0.79)	(0.16)
β_I	0.87	0.65	−0.27	0.51
	(0.07)	(0.10)	(0.72)	(0.14)
ρ^2	0.98	0.98	0.85	0.97
t^2	—	61.9	2.9	22.0

Notes. Source: Author's calculations using Epple and McCallum's (2006) data. Coefficient standard errors are in parentheses. The log time trend model uses the natural logarithm of the authors' trend variable. The row ρ^2 is the squared correlation coefficient between log quantity and predicted log quantity. The row t^2 is the square of the t-statistic for the instrument in the reduced form for log price.

Based on these results, all three researchers might claim success in that their instrument makes the demand curve appear more elastic, which is consistent with the bias one might expect from estimating demand using ordinary least squares. Nevertheless, the elasticity estimates vary considerably in economic magnitude and statistical significance. For example, using the log of the corn price, we get an estimate that is about three times that produced by the other two instruments. The income elasticity estimates tell a different story. The log of corn price IV specification has a negative income elasticity, suggesting that broilers are an inferior good. On the other hand, the other three specifications suggest broilers are a normal good (with a relatively high income elasticity). Based on our economic priors, we may have reason to prefer the log corn price specification if we believed broilers are relatively price elastic and an inferior good. On the other hand, perhaps we have another reason to believe that broilers are less price elastic and a normal good.

If the researchers in this example were made aware of each other's results, what would they make of the differences? Gauging by the squared correlation between the actual and fitted quantities (the ρ^2), there is little reason to choose among the models (although the log corn price model does have a worse fit). The square of the t-statistic on the log corn price is small and suggests, according to Stock and Yogo (2005), that the log corn price is a weak instrument. This might lead one to choose one of the other specifications.

There is, of course, a deeper issue as to whether their choice of instrument is appropriate for the endogeneity problem they envision. In some cases, it may be possible to use statistical tests for endogeneity and overidentification to check assumptions about the instruments.[14] For example, Reiss and Wolak (2011) have developed tests to assess instruments based on functional form assumptions. In the end, a researcher's first defense of instruments should come from economic theory, because it is that theory that helps make the case that the effect of interest is truly identified. Implicitly stated, each of the researchers has a different theory (here, a second equation) that rationalizes why his estimator is consistent. Knowing these differences is what allows researchers to make progress in understanding potential economic or marketing reasons for the differences in their findings. This again illustrates the value of being explicit about the assumptions that go into estimation and the usefulness of exploring alternatives. (See Reiss and Wolak 2011 for more discussion.)

6. Concluding Remarks

Empirical research in marketing has embraced a variety of different empirical approaches. Each of these approaches has an important role to play in marketing research. Understanding both their advantages and disadvantages for framing and addressing research questions is essential for encouraging and improving empirical research in marketing.

Descriptive models in marketing are primarily about using statistical methods to characterize features of the joint distribution of data. Descriptive analyses ideally uncover interesting facts, trends, practices, and puzzles that can help shape existing marketing theories or prompt new models. Descriptive analyses can also be an integral component of structural or experimental analyses. For example, structural papers are often strengthened by descriptive work that documents the features of the data that the structural model is attempting to fit. Additionally, structural modelers can benefit from descriptive analyses that show whether the structural model adequately explains the data. Similar points can be made about descriptive data analyses in experimental studies.

Descriptive researchers lessen their credibility when they misinterpret or simply overinterpret results. Such missteps occur when researchers treat statistical objects as behavioral constructs or attempt to draw policy conclusions without reference to a model. The use of the term "reduced-form analysis" to describe a descriptive regression is one such example in which such problems can arise. Whereas a true reduced form is derived from a structural model, this term is now routinely used to describe descriptive (linear) regressions. To avoid confusion, researchers should only use the term "reduced form" to refer to relations derived

[14] See Hahn and Hausman (2002) and Stock and Yogo (2005).

from a structural model. The terms "descriptive" or "predictive regression" are much more accurate representations of what most researchers do when they regress Y on X.

This advice extends to applications of instrumental variables. One can regress quantity sold on price and advertising mix variables and interpret this regression as a best linear prediction equation. Some believe if we recognize that price and advertising mix variables are potentially correlated with the error, that instrumenting for these variables suddenly gives the regression coefficient estimates a causal interpretation. Such casual interpretations, however, rely on functional form and distributional assumptions that are best developed through a complete structural model. Additionally, as was seen in §5, even if we are sure about functional form, it is still useful to have a structural model that can motivate the researcher's choice of instruments.

Structural models are making inroads in marketing applications largely because they enable researchers to perform counterfactual predictions about how changes in marketing variables will affect consumer and firm behavior. Although structural models provide this and other benefits, the challenge marketing researchers face is to develop structural models that provide realistic descriptions of the environments in which firms market products and services. As Chintagunta et al. (2006) note, many structural marketing models currently rely heavily on generic economic models. Moving these models toward marketing applications where consumers develop tastes for goods and those tastes are in turn influenced by marketing activities will lead to more credible marketing research. In the process of doing so, marketing researchers must come to grips with the fact that theory is rarely complete. Researchers will have to add functional form assumptions, variables, and errors to these models to capture important sources of variation in marketing data. Adding these elements poses risks. For instance, it is possible to add structure that, in fact, delivers an undesired result. (Witness the homogeneous cross-price elasticities in the multinomial logit.) Moreover, in complicated models, it is sometimes unclear whether the results are unduly affected by strong assumptions. This makes it essential for structural modelers to explore whether their results are sensitive to their modeling assumptions.

My hope is that this paper prompts further reflections on the role of empirical research in marketing. Each type of empirical approach has a role to play in the development of the field, and each type should be part of a serious empiricist's toolkit.

Acknowledgments

This paper expands on a presentation given at the Workshop on Quantitative Marketing held at the Fuqua School of Business, Duke University, August 2010. The author is particularly indebted to his colleague Frank Wolak. Many of the ideas presented here are products of their prior collaborations and ongoing discussions. The author is also grateful to conference participants, Eric Bradlow, an area editor, and two referees for their comments. The author alone is responsible for this paper.

References

Anderson, E. T., D. I. Simester. 2003. Effects of $9 price endings on retail sales: Evidence from field experiments. *Quant. Marketing Econom.* **1**(1) 93–110.

Anderson, E. T., D. I. Simester. 2004. Long-run effects of promotion depth on new versus established customers: Three field studies. *Marketing Sci.* **23**(1) 4–20.

Bajari, P., J. Fox, K. Kim, S. Ryan. 2009. Identification and estimation of random utility models. Working paper, University of Minnesota, Twin Cities, Minneapolis.

Berry, S., P. Reiss. 2007. Empirical models of entry and market structure. M. Armstrong, R. Porter, eds. *Handbook of Industrial Organization*, Vol. 3. Elsevier Science, Amsterdam, 1845–1886.

Briesch, R., P. Chintagunta, R. Matzkin. 2010. Nonparametric discrete choice models with unobserved heterogeneity. *J. Bus. Econom. Statist.* **28**(2) 291–307.

Bronnenberg, B. J., S. K. Dhar, J.-P. H. Dubé. 2009. Brand history, geography, and the persistence of CPG brand shares. *J. Political Econom.* **117**(1) 87–115.

Bronnenberg, B., J.-P. Dubé, M. Gentzkow. 2010. The evolution of brand preferences: Evidence from consumer migration. NBER Working Paper 16267, National Bureau of Economic Research, Cambridge, MA.

Bronnenberg, B. J., P. E. Rossi, N. J. Vilcassim. 2005. Structural modeling and policy simulation. *J. Marketing Res.* **42**(1) 22–26.

Chintagunta, P., J.-P. Dubé, V. Singh. 2003. Balancing profitability and customer welfare in a supermarket chain. *Quant. Marketing Econom.* **1**(1) 111–147.

Chintagunta, P., T. Erdem, P. E. Rossi, M. Wedel. 2006. Structural modeling in marketing: Review and assessment. *Marketing Sci.* **25**(6) 604–616.

Ellickson, P. B., S. Misra. 2011. Estimating discrete games. *Marketing Sci.* **30**(6) 997–1010.

Epple, D., B. McCallum. 2006. Simultaneous equation econometrics: The missing example. *Econom. Inquiry* **44**(2) 374–384.

Fair, R. C., D. M. Jaffee. 1972. Methods of estimation for markets in disequilibrium. *Econometrica* **40**(3) 497–514.

Fienberg, S. 1979. Graphical methods in statistics. *Amer. Statistician* **33**(4) 165–178.

Fox, J., A. Gandhi. 2009. Identifying heterogeneity in economic choice models. Working paper, University of Chicago, Chicago.

Frisch, R. 1934. *Statistical Confluence Analysis by Means of Complete Regression Systems*, Publication 5. University Institute of Economics, Oslo, Norway.

Hahn, J., J. Hausman. 2002. A new specification test for the validity of instrumental variables. *Econometrica* **70**(1) 163–189.

Heckman, J. 1997. Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations. *J. Human Resources* **32**(3) 441–462.

Hui, S., E. Bradlow, P. Fader. 2009a. Testing behavioral hypotheses using an integrated model of grocery store shopping path and purchase behavior. *J. Consumer Res.* **36**(3) 478–493.

Hui, S., E. Bradlow, P. Fader. 2009b. The traveling salesman goes shopping: The systematic deviations of grocery paths from TSP optimality. *Marketing Sci.* **28**(3) 566–572.

Kadiyali, V., K. Sudhir, V. Rao. 2001. Structural analysis of competitive behavior: New empirical industrial organization methods in marketing. *Internat. J. Res. Marketing* **18**(1–2) 161–186.

Keane, M. 2010. A structural perspective on the experimentalist school. *J. Econom. Perspect.* **24**(2) 47–58.

Kerin, R. 1996. In pursuit of an ideal: The editorial and literary history of the *Journal of Marketing*. *J. Marketing* **60**(1) 1–13.

Larson, J., E. Bradlow, P. Fader. 2005. An exploratory look at supermarket shopping paths. *Internat. J. Res. Marketing* **22**(4) 395–414.

Levav, J., M. Heitmann, A. Herrmann, S. Iyengar. 2010. Order in product customization decisions: Evidence from field experiments. *J. Political Econom.* **118**(2) 274–299.

Lewis, R., D. Reiley. 2009. Retail advertising works! Working paper, University of Arizona, Tucson.

Lodish, L. M., M. M. Abraham, J. Livelsberger, B. Lubetkin, B. Richardson, M. E. Stevens. 1995. A summary of fifty-five in-market experiments of the long-term effect of TV advertising. *Marketing Sci.* **14**(3, Part 2 of 2) G133–G140.

Maddala, G. S. 1986. Disequilibrium, self-selection, and switching models. Z. Griliches, M. Intriligator, eds. *Handbook of Econometrics*, Vol. 3. Elsevier Science, Amsterdam, 1633–1688.

Manski, C. 2003. *Partial Identification of Probability Distributions*. Springer-Verlag, New York.

Manski, C., E. Tamer. 2002. Inference on regressions with interval data on a regressor or outcome. *Econometrica* **70**(2) 519–546.

Matzkin, R. 2007. Heterogeneous choice. Working paper, Northwestern University, Evanston, IL.

Mazzeo, M. J. 2006. Marketing structural models: "Keep it real." *Marketing Sci.* **25**(6) 617–619.

Reiss, P., F. Wolak. 2007. Structural econometric modeling: Rationales and examples from industrial organization. J. J. Heckman, E. E. Leamer, eds. *Handbook of Econometrics*, Vol. 6A. Elsevier Science, Amsterdam, 4277–4415.

Reiss, P., F. Wolak. 2011. How can instrumental variable methods be defended? Working paper, Stanford Graduate School of Business, Stanford University, Stanford, CA.

Rosenzweig, M., K. Wolpin. 2000. Natural "natural experiments" in economics. *J. Econom. Literature* **38**(4) 827–874.

Stock, J., M. Yogo. 2005. Testing for weak instruments in linear IV regression. J. Stock, D. Andrews, eds. *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg*, Chapter 5. Cambridge University Press, Cambridge, UK.

Stout, R. 1969. Developing data to estimate price-quantity relationships. *J. Marketing* **33**(2) 34–36.

Tufte, E. 1990. *Envisioning Information*. Graphics Press, Cheshire, CT.

Tufte, E. 1997. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT.

Wilkinson, L. 2005. *The Grammar of Graphics*, 2nd ed. Springer, New York.

Wilson, A. 2008. Marketing research. M. Baker, S. Hart, eds. *The Marketing Book*, Chapter 8. Butterworth-Heinemann, Oxford, UK.