Marketing Science

Editorial—It's the Findings, Stupid, Not the Assumptions

Steven M. Shugan,

Please scroll down for article—it is on subsequent pages

Editorial

It's the Findings, Stupid, Not the Assumptions

Steven M. Shugan*

Warrington College of Business, University of Florida, 201B Bryan Hall, P.O. Box 117155,
Gainesville, Florida 32611, steven.shugan@cba.ufl.edu

Observing reality is especially valuable. However, without models, every situation at every time on every variable would be unpredictable. Assumptions allow models and theories to assert constancy. Assumptions distill and simplify reality by dismissing the conspicuous but irrelevant. Criticizing assumptions as unrealistic is absurd. Abstraction is the precise virtue of an assumption. For example, seldom are we prisoners facing interrogation, yet the prisoner's dilemma remains relevant. The adage "A bird in the hand is worth two in the bush" is relevant for more than birds. Unrealistic assumptions that deny current beliefs breed great new theories.

Assumptions are analogous to the basic ingredients in a gourmet recipe. Only the final product of the recipe dictates whether the ingredients suffice. Similarly, assumptions are realistic when they produce good theories, satisfactory predictions, valuable implications, and correct recommendations. Output matters far more than input. Realism is only an issue when creatively diagnosing poorly performing models, not when judging model performance.

Assumptions are the source of value in empirical analyses. If data sets were truly the source of value, empirical research studies would only greatly devalue the raw data by dramatically reducing rich observations to a few meager summary statistics or estimated parameters. Most empirical research makes a contribution by ignoring (assuming away) most information in the data.

We must dramatically shift our attention far away from the hopeless pursuit and sophistry of realistic assumptions to the contribution those assumptions produce. There are scientific methods for evaluating model output (i.e., predictions, findings, implications, recommendation) on criteria such as accuracy, reliability, validity, robustness, and so on. No corresponding objective scientific methods exist for evaluating realism. Realism depends only on personal taste.

Key words: models; mathematical models; realistic assumptions; instrumentalism; empirical research

Why Study Assumptions?

The role of explicit assumptions in building models is widely misunderstood. In fact, observations, from the peer review process at scholarly journals, chronicle widespread misunderstandings, causing mistakes and inconsistencies in the peer review process. Misguided evaluators sadly focus on evaluating assumptions (i.e., the input) rather than the findings (i.e., the output) or the process (i.e., the logic creating the output). Dismissing models based on only the realism of their early assumptions is facile and invalid compared with the correct approach of carefully evaluating the entire model development, the novelty of the findings, and the contribution to the literature. Given that all assumptions are unrealistic per se, the claim of unrealistic assumptions is empty. Assessing the realism of the assumptions might be important for creatively diagnosing a poorly performing model, but it is an inappropriate measure of performance.

Problems with the peer review process alone warrant careful examination of the role of explicit assumptions. Beyond that, however, understanding the role of assumptions is critical for the advancement of knowledge, understanding, and application. Assumptions allow us to focus our modeling effort, to build theory, to extract information from numerical data, and to interpret qualitative observations. Without unrealistic assumptions, most research would add no value to merely reporting raw numeric data or detailing qualitative observations.

We argue that properly evaluating research (and most professional activities as well) requires a focus on output rather than on input. Focusing on input

(the assumptions, the observations, the numerical data, etc.) is wrong. All high-quality research makes explicit and implicit assumptions based on either explicit or implicit past observations. Evaluation of input, without regard to output, is an exercise in personal taste. In contrast, scientific methods exist for evaluating output (e.g., predictions, findings, implications, recommendations) on criteria such as predictive accuracy relative to extant models, reliability, validity, robustness, interobserver agreement, and generality. Moreover, we have criteria for evaluating the scientific research process itself, such as logical validity, objectivity, clarity, meticulousness, replicability, verifiability, and compatibility with scientific procedures.

We also argue that explicit assumptions provide the value from all analyses in all research, whether empirical or theoretical. Models are only useful because they remove the irrelevant aspects of reality. Without models, every situation at every time on every variable would be unpredictable. Assumptions allow models and theories to assert constancy and to predict in new situations.

Finally, for purely empirical research, we require assumptions to reduce a vast quantity of numerical data points to a very small meager set of summary statistics or estimated parameters. It is impossible to add value to a data set by ignoring parts of it, unless ignoring parts of it reveals the predictive or explanatory power of the assumptions. Otherwise, empirical research would simply diminish the value of the raw data by ignoring some of it.

What Are Assumptions?

Virtually all scholarly research, benefiting from mathematical models or not, begins with both implicit and explicit assumptions, but the term "assumption" is often ill defined. Although dictionaries define assumptions as something taken as true without formal proof, assumptions need not be true or false. Assumptions can be approximations, limitations, conditions, or merely premises. Most often, assumptions are sufficient conditions that guarantee the validity of the subsequent findings but whose violation by no means necessarily invalidates those findings. Published research often assumes that random error causes unpredicted outcomes, rather than wrong theory (Shugan 2006). Published research often assumes that results are invariant to the time, place, and sample. Most published research assumes continuity when virtually no variables are infinitely divisible. If assumptions could be false, we would dismiss most research.

An Example

Consider, for example, mathematician Albert W. Tucker's famous prisoner's dilemma game (Tucker 1950, Goeree and Holt 2001, Lev 2006). We assume two prisoners, only four possible outcomes, insufficient evidence to convict, guilty prisoners, prisoners acting in their own best interest, prisoners unable to communicate, prisoners knowing four possible outcomes, police who can lessen sentences for confessing, no possible retribution, no possible reputation effects, severe sentences for conviction without confessing, among still more assumptions. These assumptions are sufficient to find that both prisoners will confess, resulting in a worse situation for both than had the prisoners coordinated their decisions.

The assumptions are seldom true (i.e., if truth is even possible). Scholars are seldom prisoners, advising prisoners or advising the police. Nevertheless, there are thousands of references to the prisoner's dilemma. These apparently limiting assumptions reveal practical, remarkably general, and extraordinary powerful implications. It is the generality of the implications that matters, not the realism of the assumptions. Good theory makes good predictions regardless. As Trefil (2003) states "good theory...will do more than incorporate facts already known—it will make predictions about phenomena that haven't been seen before (p. xv)."

Why Make Assumptions?

If the essence of a well-executed competitive strategy is deciding what not to do (e.g., Porter 1996), the essence of mathematical or analytical modeling is deciding what not to model. As it is deceptively easy to add one more strategic activity (e.g., see Porter 1996), it is deceptively easy to add still another variable, relationship, or parameter to a mathematical model. However, the "no free lunch" principle applies. As diversifying competitive activities diverts resources in strategic development, diversifying mathematical models also diverts resources and often hinders the original objective of the entire modeling effort. In this way, assumptions define the research question by dictating what is endogenous (in the research) and exogenous (outside of the research). Assumptions exclude variables, exclude possible relationships, and exclude possible causality. The assumptions are the foundation of proposed models, hypotheses, theories, forecasts, and so on. They dictate which variables to observe, not to observe, and the relationship between them.

Finally, when models become theories, assumptions dictate the processes resulting in causality. For example, we might believe that positive film reviews by film critics cause increases in the sales of better films. Alternatively, we might believe that better films get both greater sales and more positive reviews. Only a deductive causal model derived from

unambiguous assumptions can produce a research design to determine which variables to observe and how to disentangle potential causal relationships. Passively observing the valence of film critic reviews and film box office data disentangles nothing, but it might inspire assumptions that could (Eliashberg and Shugan 1997).

Can Data Speak? If They Do, Assumptions Are the Interpreters

Some authors of research studies erroneously suggest that collecting numeric data eliminates the role of assumptions. That research falsely implies, as they tout, that data can speak. Despite these dubious claims, making and interpreting observations without assumptions is impossible. Certainly, making careful observations and quantifying those observations can create great value and often deserves sincere commendation (or the more preferred reward of publication). Data provide justification and compelling evidence for modeling predictions and correct implications. Moreover, numeric data or qualitative observations, without further adulteration, can provide great value (sans analysis). For example, we should publish raw data, possibly made available by improved measurement technologies, that blatantly contradicts current beliefs or resolves unanswered questions.

Indeed, some scientific disciplines do publish unadulterated raw numerical data or detailed qualitative observations. Sometimes, when, fortuitously, nearly everyone agrees on the assumptions, the observations themselves can become the sole research contribution. For example, when Dutch lens maker Thonius Philips "Antoni" van Leeuwenhoek first observed tiny creatures previously invisible to the human eye (the microorganisms that he termed "animalcules") on a sample of dental plaque, most scientists agreed that the discovery was important because the scientists shared common assumptions. For example, they implicitly assumed that this observation was unexpected, the new lens was not the cause of the new observation, and animalcules were indeed alive.

Similarly, research articles in anthropology might provide descriptions of recently discovered ancient remains. Research articles in virology might report descriptions of recently discovered viruses. Research articles in many disciplines might report observations made available by more creative or more sophisticated measurement. Moreover, one could argue that experimentation and market surveys produce new data which, without further manipulation, is a publishable contribution.

However, most social science disciplines, and many other disciplines of study, seldom, if ever, publish complete original raw data in their unmodified form. In fact, journal editors and authors alike often limit the number of purely descriptive data tables, preferring to emphasize tables that summarize various findings and surmise relationships. Of course, constructing these tables involves myriad assumptions. Going from observations to unambiguous conclusions usually requires many assumptions, at least, about what we are observing and why we are observing it.

Only when nearly everyone adopts common assumptions can mere observations be sufficient. Otherwise, assumptions add the value. The data tests subsequent predictions, implications, recommendations, theoretical explanations, methodological improvements, and so on. If data do speak, assumptions are the interpreters.

Hence, for most research, the assumptions are the primary source of value added by the researcher. Without assumptions, publishing raw data (qualitative observations or unadulterated numeric observations) is preferable to publishing research studies that manipulate and summarize the data.

All Empirical Analyses Use Assumptions to Add Value

Certainly, empirical research is essential for the advancement of theory, models, methods, and technology. However, most published empirical studies are not simply tabulations of unadulterated raw numerical data or detailed qualitative observations. Many empirical research studies claim that their numeric data provide their primary contribution while their assumptions play little or no role in the value added by their so-called empirical research.

This dubious claim is logically inconsistent. To understand, consider a typical empirical study with a grand database consisting of copious numbers reflecting the values of a variety of potentially interesting variables, possibly at different points in time (e.g., sales transactions by all competitors for each item at each point in time in each location). Most research studies then reduce this rich reservoir of numbers to a few meager numbers (e.g., average sales, aggregate market share, a few estimated parameters). Thousands of numerical observations become less than fifty numbers. For example, copious amounts of sales, advertising, and pricing data reduce to a few meager cross-elasticity estimates or a simple linear relationship that fails to recover exactly any of the original data.

Conceptually, if the data were the primary source of value, reducing many numbers to a few would only remove information and diminish the value of the data, rather than enhance it. If the original raw data were the source of primary value, ignoring parts

of it would only lessen its value. At best, reducing the number of data points by ignoring or summarizing would create nothing more. At worst, discarding data by summary or estimation would decrease available information. It almost appears that models squander the information in the original raw data.

An alternative argument yields the same conclusion. If assumptions played little role in the contribution of empirical research, then given exactly the same data, all analyses should come to exactly the same conclusions. However, different analyses apparently come to different conclusions, at least, so researchers claim. Barring incompetency, different conclusions must result from different assumptions and not from the data.

Consider still another argument. Modelers generally believe that meticulously applying the same methods to the same data would replicate the same findings (at least, that is the assumption of all published research). However, few believe in reverse replication, that is, being able to always recover the same data from the reported findings. This lack of reversibility again suggests that the research studies ignore (assume away) some information in the original data. The assumptions dismiss the part eluding recovery.

Beyond the scant value created by doing the mechanical computations, it is the assumptions that add value to the data. Assumptions concerning which variables to observe, not to observe, their relationships, possible causality and so on, reduce what we might call the entropy in the data. The assumptions create order from disorder. Only the assumptions can reduce the entropy of the real world to the world of the model. Only the assumptions can create order in apparently complex and bewildering systems. When a research study starts with data or observations, the assumptions are the sole source of value from empirical analyses.

Of course, not every assumption adds value. Assumptions can certainly create bad models that fail to predict and provide flawed implications.

Good Theory Requires Unrealistic Assumptions

Unrealistic assumptions, often considered ludicrous when first proposed, spawned truly great discoveries, theories, technologies, and conceptualizations. For example, consider the following unrealistic assumptions—certainly considered absurd when first proposed. Theoretical physicist Albert Einstein assumed acceleration and gravity were the same force. English mathematician Edmund Halley assumed, among other remarkable assertions, that different comets sighted in different years were indeed the

same comet. German mathematician Johannes Kepler assumed planetary orbits were ellipses rather than circles, without the aid of a telescope. Dutch mathematician Christiaan Huygens, beyond writing the first book on probability theory, assumed that light was also a wave rather than merely a particle (termed "corpuscles" by Isaac Newton). English mathematician Charles Baggage assumed machines could do computations. English physicist James Prescott Joule assumed that energy could not be destroyed. Sir Timothy John Berners-Lee assumed that everyone on earth could be electronically networked using a simple common language leading to the World Wide Web. Some physicists claim that Einstein's theory of relativity is widely misunderstood not because of its complexity but because the assumptions contradict common sense. For example, Teller et al. (1991) state that "if I say something that you think is absurd the automatic reaction is that your earflaps come done and you stop listening...[but] things that are 'obviously' wrong; in fact, are true (p. 2)."

Ironically, the most vocal advocates for downplaying assumptions are those researchers most concerned with application and practicality. For example, Friedman (1953) argues that it is impossible to test descriptive theories from the realism of their assumptions, in part, because assumptions are unable "to determine the circumstances for which a theory holds...(p. 19)." In fact, Friedman (1953) argues that "truly important and significant hypotheses will be found to have assumptions that are widely inaccurate descriptive representations of reality, and, in general the more significant the theory, the more unrealistic the assumptions...(p. 14)."

These arguments were never intended to justify naïve descriptive theories or abstract analyses. Rather, Friedman (1953) argues that focusing on assumptions hinders, if not thwarts completely, the proper and more onerous task of properly evaluating descriptive theories based on the quality of what the theory produces. Friedman (1953), for example, concludes "the question whether a theory is realistic enough can be settled only by seeing whether it yields predictions that are good enough for the purpose in hand or that are better than predictions from alternative theories (p. 41)."

Note that the ideas of Friedman (1953) seem consistent with the philosophy of instrumentalism in contrast to realism (Boland 1979, Wible 1982, Boland 1984). Instrumentalism views descriptive theories as "valuable for predicting observable events but not to be viewed as genuine statements whose truth or falsity may be significantly investigated (Crilly 1974, p. 112)." In contrast, realism views theories as true or not and considers some theoretical entities as real (Crilly 1974, 1975; Popper 1985). There

has been considerable debate on whether Friedman (1953) was advocating instrumentalism (Mayer 1993), just being pragmatic (Pelloni 1996), ignoring causality (Helm 1984, Hicks 1979), practicing what he preachers (Caldwell 1980), voicing Dewey's view of pragmatism (Dewey 1938, Wible 1984), and rejecting the inclusion of all relevant background information (e.g., Blaug 1980). Still others have taken Friedman's arguments further. For example, Stigler and Becker (1977) advocate dismissing almost everything unobservable, including unobservable predictions.

However, the debate regarding the ideas of Friedman (1953) often concerns Friedman's ultimate use of these arguments, i.e., to further Stigler's (1949) criticism of monopolistic competition. Moreover, accepting the realism viewpoint still requires the evaluation of models on their observable predictions, rather than on only a priori predilection or distaste for the model (or the assumptions). We need not resolve the issue of whether these models are causal theories representing true causal relationships or merely good approximations to argue that observations should not contradict the predictions of good theories. As Barr (2006) states: "An incomplete and inexact description of a situation may be sufficient to convey a completely true insight into that situation. Otherwise, we could never learn anything (p. 61)."

Descriptive (Positive) Models vs. Prescriptive (Normative) Models

Unlike Friedman (1953), who focused on traditional scientific methods for evaluating descriptive (or positive) models and theories (e.g., see Keynes 1891 for distinctions), Little (1970, 2004) considered evaluation of prescriptive (or normative) models intended to aid decision makers. Little (1970) advocates specific criteria for evaluating prescriptive models and evaluates several specific mathematical models for aiding decisions on those criteria.

Little (1970, 2004) and many other researchers seem to imply that prescriptive models, unlike descriptive models, should have realistic assumptions. Perhaps, one reason is that prescriptive models are often tools or technologies designed to outperform alternatives only under specific situations. Users seek to know when each technology is superior. For example, the relative performance of durable (Bass 1969, Urban et al. 1993) and nondurable (Silk and Urban 1978) product forecasting models might depend on the frequency of repeat purchases. The relative performance of logit and discriminant statistical models seems to depend on whether specific distributional assumptions are met (Maddala 1991, Press and Wilson 1978, Lachenbruch and Goldstein 1979).

There are several points to be made here. First, assumptions are virtually always only sufficient conditions. More general sufficient conditions do not guarantee a more general tool. Weaker sufficient conditions fail to guarantee weaker necessary conditions. For example, when $A > 1$ and $B > 9$, it is still possible that $A > B$ always. One notable example is the phenomena of overfitting, where predictive accuracy diminishes when increasing the flexibility (i.e., relaxing the assumptions) of a statistical model. Hence, knowing whether stronger sufficient conditions yield stronger or weaker implications is difficult. Sometimes, stronger assumptions yield extraordinarily general conclusions (e.g., the prisoner's dilemma). Sometimes, apparently weak assumptions produce rarely observed situations (e.g., consider Sir Robert Giffen's, famous but rare, Giffen Good). Relaxing assumptions to create realism works only when relaxation produces better output (e.g., predictions, implications, recommendations, etc.). Hence, we are unable to conclude that more realistic assumptions yield more generally applicable tools.

Second, although Little (1970, 2004) failed to use the terms input and output, Little's criteria for evaluating prescriptive models focus on the model's output (always sensible predictions, predictions conditional on control variables, predictions change as new information becomes available, predictions are easily understood) rather than input (i.e., realistic assumptions, interesting data, extensive observations).

Third, in practice, the apparent scrutiny of the assumptions often becomes an evaluation of the model output. For example, when analyzing the sensitivity of the findings to particular assumptions (i.e., robustness), the focus remains on the objective stability of the findings to variations in the assumptions rather than the more ambiguous realism of the assumptions. When demanding explicit, easily understood assumptions, the focus remains on whether the user knows when to trust the model's predictions. Hence, sensitivity analysis scrutinizes output conditional on input, rather than focusing merely on realistic assumptions. For example, when evaluating a model of adaptive control of promotional spending, Little (1966) judges the outcome of the sensitivity analysis by observing the model output. Of course, there are many recent uses of sensitivity analysis for this purpose (e.g., Sriram et al. 2006, Ho et al. 2006, Lewis et al. 2006, Naik 2005, Chintagunta 2005, Mitra and Golder 2006). Mazzeo (2006) advocates an analogous sensitivity analysis when employing descriptive structural models.

Fourth, researchers, who explicitly or implicitly advocate realistic assumptions, often ultimately test their assumptions based on model output rather than on the realism of the input. For example, Hauser et al.

(2006) strongly advocate the testing of assumptions (and hypotheses) of new product models and, in particular, models of network externalities (Hauser et al. 2006). However, they cite research using output measures for evaluation.

Radas and Shugan (1998) show that modeling seasonality by changing the speed at which time advances has analytical advantages, but their empirical evaluation of the model depends on the observed fit and prediction of the box office of motion pictures. Although Bergen et al. (1996) propose a theoretical model implying that diminishing retail competition with branded variants (minor variations of the product) will enhance the manufacturer's distribution by encouraging retail adoption, their empirical test is based on the observed number of items that retailers carry. Pasa and Shugan (1996) propose a theoretical model implying that firms gain from marketing expertise more when the environment is changing, but their empirical test is based on the observed backgrounds of chief executive officers in different industries. Ehrman and Shugan (1995) provide a theorem proving that, when forecasts influence outcomes, unbiased forecasts will appear downward biased. However, their empirical test employs observed forecasts.

Finally, we must remember that prescriptive models need not be sufficiently realistic or comprehensive to replace decision makers (Shugan 2004a). These models need only aid decisions. Analogously, a tax accountant might provide useful and accurate tax advice regarding an investment, but other concerns might have a greater influence on the investor. Models can predict enhanced profits from adding features, changing profit margins, or increasing rebates. Despite the value of these predictions, decision makers might have strategic reasons (such as channel relationships and image considerations) that trump the model's recommendations in specific situations. Division of labor suggests using assumptions to limit the scope of prescriptive models. Hence, complete realism is seldom sought.

Evaluating Assumptions Is Personal Taste, Not Science

Judging the realism of an assumption, in isolation from the findings, is an arbitrary unscientific exercise in personal taste rather than an objective step in any scientific process or methodological development. We could easily criticize assumptions as unrealistic when they dismiss our favorite variable, contradict the past literature, or conflict with current beliefs. Consequently, all breakthrough research would require unrealistic assumptions.

We can effectively argue that myriad aspects of reality have influence (to some extent). These aspects include advertising, bounded rationality, brand loyalty, branding, category management, channels, competitive dynamics, consumer response, coupons, diffusion, equilibrium behavior, externalities, firm entry, heterogeneity, supplier incentives, inventory, learning, the operations interface, packaging, physical distribution, prior beliefs, product attributes, product entry, product positioning, product-line concerns, promotions, quality, rationality, risk aversion, salesforce decisions, targeting, uncertainty, and so on. Researchers, who research different research questions, have different preferences about which variables create realism. Those preferences dictate which variables to exclude and include.

Realism is a function of personal taste. Consider models that assume consumers act in their own best interests. Those who find the assumption distasteful enumerate numerous observations of apparently altruistic, self-destructive, or dysfunctional behavior. Those who find the assumption appealing enumerate numerous observations of apparently blatant, egocentric, power-hungry behavior or creatively reinterpret self-interest to include latent or unobservable self-interest (e.g., feeling good, conforming to societal indoctrination). Unlike the relatively objective procedures for evaluating model performance, modeling practices consider the creation and evaluation of assumptions to be an art dependent on the creativity, skills (e.g., Aris 1994, p. 222), prejudices, and personal tastes of the modeler.

We should be extremely skeptical, if not entirely dismissive, of the condemnation of modeling assumptions as unrealistic without an explicit explanation of how those assumptions adversely impact the subsequent findings, predictions, or implications. Be critical of assumptions only when different assumptions improve model output. Assumptions are only wrong when mutually inconsistent. Otherwise, assumptions are merely building blocks of a model or a theory (i.e., models with causality). If assumptions produce a model, theory, method, or conceptualization that achieves the desired objective, then we should consider the corresponding assumptions to be realistic.

In sum, assumptions about what to model and what is realistic relies completely on personal taste. We should scrutinize the validity of findings, not the realism of assumptions.

Do Unrealistic Assumptions Devalue Understanding?

It might appear that diminishing the importance of realistic assumptions is devaluing the deductive part of research. Saying deduction is superfluous is equivalent to saying that explanations are

superfluous (Keynes 1891). Staelin (1998) stresses that, "there is little managerial value in constructing a model that reproduces some set of interesting phenomena unless the model…helps one better identify and understand the underlying factors driving the solution…(p. 300)." These views do not necessarily conflict with judging whether the understanding is correct based on the model output rather than the purported realism of the assumptions. Although evaluating models based on assumptions is wrong, creatively linking assumptions to reality might help diagnosis of the cause of poor model predictions.

Hence, linking assumptions to reality might be valuable for inspiring future research, rather than evaluating the research at hand. Understanding which assumptions are creating predictive problems, for example, might allow future research to modify those assumptions. Understanding which causal assumptions are creating predictive problems might allow future research to propose better explanations. The diagnostic task of linking assumptions to reality might be a critical step in the creative evolution of theory.

What Are Strong Assumptions?

The term "strong assumption" is often left ill-defined and expresses many meanings. Certainly, one assumption could be stronger than another. For example, assuming that random variables are normally distributed is stronger than only assuming that variables are random (e.g., Mood et al. 1974, Chapter 3). Assuming a subgame perfect equilibrium is stronger than only assuming a Nash equilibrium (e.g., Fudenberg and Tirole 1991, pp. 72–74).

However, beyond relative comparisons, absolute meanings are ambiguous. For example, strong assumptions could represent assumptions that prior research refutes, fail to hold for many situations, poorly approximate most situations, contradict popular beliefs, are stronger than past assumptions, are stronger than needed, or are merely disliked by those favoring different assumptions. Perhaps, strong assumptions are assumptions that dramatically change the findings. Strong assumptions that produce new, explicable, easily verified, and valid findings are the sole source of value in all great research studies.

Is a Sliver Enough?

When evaluating models whose findings depend on values of assumed parameters, scrutinizing the parameter range is wrong. Some evaluators criticize a model, theory, or technology because the sufficient conditions hold only for what the evaluators call a sliver of the range of possible parameter values. This fallacy is equivalent to the realism fallacy. Again, a sliver may or may not represent what we observe. Astronomy suggests that only planets spanning a minute range of possible parameters (i.e., the stable habitable zone) can support life as we know it, yet we have life. The price ranges in all product categories span only a small sliver of all possible prices. The lifetimes of business are only a small sliver of all possible lifetimes. We observe nearly all market data within only a small sliver of their unbounded values. Focusing on input produces misleading conclusions. We are unable to assess whether a particular range for a parameter is realistic by comparing the range with all feasible values. Whether a sliver of the parameter space is realistic depends on whether the model explains what we observe and do not observe.

Objectives Trump Assumptions

When models are truly models and lack any pretense of being a causal theory, then the research objectives completely trump the so-called realism of an assumption. For example, we are unable to judge the realism of photographic pictures without knowing the objective. It might be sufficient to show only one side of a product on the packaging (i.e., assume away the other side).

Similarly, if a map fails to include local streets (i.e., assume away local streets), the geographic map is not necessarily incorrect or unrealistic. Some maps of a city show local attractions, minor streets, expressways, bus routes, branded gas stations, food outlets, parks, museums, and so on, but no map is all inclusive because such a map would be useless for all objectives. There are wrong maps that show incorrect locations. However, there are numerous correct maps. Only the modeler's objectives dictate superiority (Shugan 2002).

Supposedly Realistic Assumptions Is Also a Bogus Benefit

Given that, criticisms of unrealistic assumptions are bogus, conversely, praise for realistic assumptions is unwarranted. We should be equally skeptical of claims of more realism. The compelling contrapositive argument is simply that if more realistic or detailed assumptions produce better models, then it would be best to avoid models and accept reality without abstraction. However, assumptions create the constancy that allows prediction; otherwise, every situation would be unique because some variables or relationships would differ. Fortunately, many models provide adequate predictions demonstrating the value of assumptions that abstract reality.

Publication bias favors models that are more complex because the misguided quest for realistic assumptions causes rejection of models advocating the

removal of previously published relationships and analyzed variables. Nevertheless, third-party research studies that compare many models often find virtue in simplicity. For example, Meade and Islam (1995) compare 17 forecasting models using 25 time series in 15 countries for the development of telecommunications markets and find that simple logistic and the Gompertz models significantly outperform more complex models. Dror and Steinberg (2006) find that simple clustering for constructing experimental designs for multivariate generalized linear models outperforms more complex and sophisticated methods. Schnaars (1984) finds that more data can diminish forecasting accuracy when applying extrapolation techniques. This is not an argument that simplicity produces better predictions; it is only an argument that the so-called realism of the assumptions is a poor measure of the ability of the model to predict, provide insights, yield implications, and make useful recommendations. The approach advocated by van Heerde et al. (2002) might be best, i.e., adding new variables and new relationships to simple models after we observe output shortcomings. Hence, predictions also iteratively drive assumptions. Only after predictions fail do we attempt to make supposedly more realistic assumptions resulting in more variables, relationships, and complexity.

Making supposedly more realistic assumptions often results in more variables, relationships, indeterminacy, and complexity. Rather than claiming realistic assumptions as a benefit, it is far better to demonstrate better output (e.g., better decisions, better actions, more accurate forecasts) whether unconditional or conditional on specific decisions. We already have reality. Research is only more realistic when it can predict reality in new settings, in future times, or given different actions. If the assumptions produce compelling evidence that an unconventional potentially ingenious strategy would save the firm, the unrealistic assumption suddenly becomes realistic. If the assumptions provide an obviously flawed strategy, one or more assumptions might be flawed or measurement might be an issue. Everything is suspect.

Moreover, quantifying assumptions alone is also a phony benefit. Although mathematics is the language of science and the "indispensable medium by which and within which science expresses…itself (Bochner 1966, p. 256)," mathematical quantification only adds value by communicating the assumptions.

Finally, rather than a benefit, models that require more onerous input (e.g., more observed variables, more data over time, more extensive observation) are inferior to models that can produce the same output with less onerous input requirements. When standard methods (e.g., controlled or natural experiments) for evaluating model performance fail to reveal which model is superior, Occam's (or Ockham's) razor tells us to prefer the simpler model.

Properly Evaluating Research

Agency theory recommends focusing on output for the expert agent evaluation, rather than on input, because agents have the expertise, experience, time, and information to create more and better output from the same input. Focusing on input involves almost futile micromanagement involving myriad unobserved variables, unobserved factors, and unknown processes. Similarly, evaluating research requires focusing on output (e.g., predictions, findings, implications, recommendations), rather than on the input creatively chosen by the researcher.

To properly evaluate research, we should use traditional scientific methods for evaluating model output on criteria such as predictive accuracy relative to extant models, reliability, validity, robustness, interobserver agreement, and generality. Moreover, we should employ traditional scientific criteria for evaluating the research process, including logical validity, objectivity, clarity, meticulousness, ability to replicate, verifiability, and compatibility with scientific procedures. No similar objective scientific criteria exist for evaluating the realism of assumptions, which are mainly a matter of taste.

We must also evaluate research as a whole and not as its parts. Assumptions are the analogous ingredients of a chef's gourmet recipe. Individual ingredients are meaningless, compared with the desired end. Should the desired end prove disappointing, only then does diagnosis look to the ingredients and other parts of the recipe, including the blending of the ingredients. Of course, the task of diagnosis belongs to the chef and not to the epicurean. Similarly diagnosing the source of disappointing findings or implications is the researcher's task. Evaluators of research should focus on the quality of the predictions, findings, implications, recommendation, and so on. The evaluator can also ensure that the recipe, when loyally followed, always produces the same result.

Shugan (1986) estimates brand perceptual maps based on scanner data using assumptions from Hauser and Shugan (1983) that rely on firms forming an inefficient frontier in two-dimensional spaces. It is futile to debate whether two-dimensional spaces are realistic. Certainly, some product categories have hundreds of candidate dimensions. Best is to test the consistency of the model's predictions with observation (e.g., market shares).

Consequently, we can only evaluate models, theories, technologies, and conceptualizations as a whole and not as separate isolated a priori parts. Assumptions are as good as what they produce. Evaluating

assumptions requires evaluation of what they produce. Although formulating the assumptions might be an art requiring idiosyncratic creativity and discovery (Morris 1967), testing the subsequent model predictions is more of a science.

The Burden of Proof Still Lies with the Researcher

Although researchers need not show their assumptions are realistic, researchers must establish the worthiness of their findings, whether those findings are unconditional predictions, conditional predictions, strategic implications, or recommended actions. Each type of claim might require a different type of compelling evidence for its validity and accuracy.

For example, when research claims to produce better explanations, that research should provide compelling evidence against all reasonable competing explanations (particularly, simpler explanations). The research should demonstrate sufficient consistency between its explanation and multiple observed outcomes. Moreover, proposed explanations should be completely internally consistent with no logical flaws. Finally, when models claim causal relationships, they are no longer merely mathematical models. The models become full-fledged descriptive theories subject to standard scientific testing methods and potential falsification. The model must provide testable falsifiable implications and confirm some of the theory's predictions. Theoretical predictions must be consistent with outcomes. Of course, scholarly journals do tolerate some inconsistencies, interpreted as measurement error.

When a theory makes no falsifiable predictions, the theory is usually not a scientific theory and merely represents a conceptualization or framework. Conceptualizations must provide very concrete advantages, such as revealing previously unknown inconsistencies, paradoxes, relationships, hidden implicit assumptions, or creative new assumptions that might be useful for future research.

When proposed models claim better conditional or unconditional predictions than extant models, we should employ standard methods for demonstrating superiority. Beyond various measures of accuracy and fit, other criteria for superiority might consider the ease of making predictions (e.g., less onerous data requirement, less complexity, easier to replicate).

When the research claims to produce more appropriate actions, evaluation should focus on the recommended actions. An actual implementation or simulation might help provide compelling evidence that the recommended actions outperform alternative actions (Shugan 2004b). The logic underlying the recommended actions might be sufficiently novel and compelling to warrant publication. Sometimes, it is sufficient to inspire new strategies or produce situations requiring surprising actions. Remember that playing with models is safer than playing with reality.

When research claims to develop a better technology, e.g., better measures, computation speed, better estimates, then again performance trumps assumptions. However, it is worthwhile delineating the boundary conditions when one technology outperforms competing technologies. Note that these conditions differ from assumptions because we discover these boundary conditions from extensive application or sensitivity analysis rather than creative insight or personal taste. Boundary conditions should not involve unobserved or unobservable information. More importantly, the research should provide evidence that the technology does produce superior results compared with the best competing technologies when the conditions are met.

When research claims to produce a model that cleverly combines common assumptions to produce more general surprising implications, the research must establish that the modeling implications are new and interesting. For example, a research study might demonstrate that the implications are correct, surprising, contradict current wisdom, and could change practice. Although judging this type of contribution involves an element of taste, the judgment still relies on the specific evaluation of whether the logic is correct, the findings generalize, the implications are falsifiable, and the implications can improve practice. For example, the Pythagorean theorem is worthy of publication, either as such a deductive finding or as an observation that is easily empirically verified (at least for Euclidean spaces).

Of course, assumptions should always be explicit, unambiguous, mutually consistent, and sufficient to produce observable implications. In the rare event when evaluators criticize the assumptions rather than the findings, it is the burden of the critics to demonstrate that different assumptions would produce superior findings.

Conclusions and Implications

Putting aside the philosophical discussions of whether true models exist, whether models represent reality, or whether models approximate reality, we are invariably unable to evaluate models by focusing on their assumptions. Assumptions are analogous to the ingredients that create a gourmet recipe. Without knowing the product of the recipe, we are unable to judge the adequacy of the ingredients. Moreover, the apparent specificity of the assumptions is often unrelated to the generality of the findings. For example, the prisoner's dilemma remains relevant despite the fact that we seldom study actual prisoners. The adage "A bird in the

hand is worth two in the bush" is relevant for more than birds.

If we hope to employ mathematical models to create insights, advance theory, improve technology, create new methods, make better decisions, or merely make models relevant, we must dramatically shift our attention far away from the hopeless pursuit and sophistry of realistic assumptions to the contribution those assumptions purportedly produce. There are scientific methods for evaluating model output (i.e., predictions, findings, implications, recommendations) on criteria such as predictive accuracy relative to extant models, reliability, validity, robustness, interobserver agreement, and generality. Moreover, there are criteria for evaluating the scientific research process itself such as logical validity, objectivity, clarity, meticulousness, replicability, verifiability, and compatibility with scientific procedures. No corresponding objective scientific methods exist for evaluating the realism of assumptions. We must evaluate output despite the relative ease and fun associated with criticizing input. Only after evaluating output (i.e., predictions, findings, implications, recommendations) can we diagnosis possible problems from the input (assumptions, observations, numerical data, etc.).

One special case of this argument is the criticism or rejection of a model, theory, or technology because the assumptions hold for only a small range of parameter values. This criticism is illogical. The parameters describing most observed phenomena (lifetimes of businesses, product prices, the size of markets, the conditions for planetary life, the size of mammals) span only a small sliver of all possible numerical values. Again, we must focus on whether the model unambiguously predicts what we observe and not observe, rather than assumed ranges of parameters.

Good theory requires unrealistic assumptions. Most breakthrough research makes some assumptions that directly contradict current beliefs. Moreover, shifting to empirical analyses, although commendable, fails to resolve this problem. Only by making assumptions that dismiss most of reality can research reduce many rich data points to a few scant numerical values. Only by vastly simplifying the numerical data can empirical research claim to add value to the original data by reporting only a meager number of summary statistics or estimated parameters.

Moreover, authors should never claim advantage by merely asserting that their assumptions are more realistic or their models include more variables. Authors must argue for their contribution based on the output of their research, rather than on the input.

Assumptions create value in almost all research (except research that merely reports raw numerical data or detailed observations). The assumptions, as a whole, allow us to create the theory, predictions, findings, implications, and recommendations that are the useful product of research.

References

Aris, R. 1994. *Mathematical Modelling Techniques*. Dover Publications Inc., New York.

Bachmann, J. W. 2002. Competitive strategy: It's O.K. to be different. *Acad. Management Executive* **16**(2) 61–65.

Barr, S. M. 2006. *A Student's Guide to Natural Science*. Intercollegiate Studies Institute (ISI) Books, Wilmington, DE.

Bass, F. M. 1969. A new product growth model for model consumer durables. *Management Sci.* **15**(5) 215–227.

Bergen, M., S. Dutta, S. Shugan. 1996. Branded variants: A retail perspective. *J. Marketing Res.* **33**(1) 9–19.

Blaug, M. 1980. *The Methodology of Economics: Or How Economists Explain.* Cambridge University Press, Cambridge, UK.

Bochner, S. 1966. *The Role of Mathematics in the Rise of Science*. Princeton University Press, Princeton, NJ.

Boland, L. A. 1979. A critique of Friedman's critics. *J. Econom. Literature* **17**(2) 503–522.

Boland, L. A. 1984. Methodology: Reply ("Response to Critics"). *Amer. Econom. Rev.* **74**(4) 795–797.

Caldwell, B. J. 1980. A critique of Friedman's methodological instrumentalism. *Southern Econom. J.* **47**(2) 366–374.

Chintagunta, P. K., R. Desiraju. 2005. Strategic pricing and detailing behavior in international markets. *Marketing Sci.* **24**(1) 67–80.

Crilly, M. J. 1974. A perspective of the functions and criticisms of trip distribution models. *Oper. Res. Quart.* **25**(1) 111–121.

Crilly, M. J. 1975. Functions and criticisms of models—A rejoinder. *Oper. Res. Quart.* **26**(1, Part 2) 219.

Dewey, J. 1938. *Logic: The Theory of Inquiry*. Holt, Rinehart and Winston, New York.

Dror, H. A., D. M. Steinberg. 2006. Robust experimental design for multivariate generalized linear models. *Technometrics* **48**(4) 520–529

Ehrman, C. M., S. M. Shugan. 1995. The forecaster's dilemma. *Marketing Sci.* **14**(2) 123–147.

Eliashberg, J., S. M. Shugan. 1997. Film critics: Influences or predictors? *J. Marketing* **61**(2) 68–78.

Friedman, M. 1953. The methodology of positive economics. *Essays in Positive Economics: The Methodology of Economics and Other Essays*. University of Chicago Press, Chicago, IL, 3–43.

Fudenberg, D., J. Tirole. 1991. *Game Theory*. MIT Press, Cambridge, MA.

Goeree, J. K., C. A. Holt. 2001. Ten little treasures of game theory and ten intuitive contradictions. *Amer. Econom. Rev.* **91**(5) 1402–1422.

Hauser, J. R., S. M. Shugan. 1983. Defensive marketing strategies. *Marketing Sci.* **2**(4) 319–360.

Hauser, J. R., G. J. Tellis, A. Griffin. 2006. Research on innovation: A review and agenda for *Marketing Science. Marketing Sci.* **25**(6) 687–717.

Helm, D. 1984. Predictions and causes: A comparison of Friedman and Hicks on method. *Oxford Econom. Papers* **36**(Nov., Suppl.) 118–134.

Hicks, Sir John Richard. 1979. *Causality in Economics*. Basic Books, New York.

Ho, T.-H., Y.-H. Park, Y.-P. Zhou. 2006. Incorporating satisfaction into customer value analysis: Optimal investment in lifetime value. *Marketing Sci.* **25**(3) 260–277.

Keynes, J. N. 1891. *The Scope and Method of Political Economy*. Macmillan and Company, London, UK, 359.

Lachenbruch, P. A., M. Goldstein. 1979. Discriminant analysis. *Biometrics* **35**(1) 69–85.

Lev, B. 2006. An annotated timeline of operations research: An informal history. *Interfaces* **36**(2) 177–179.

Lewis, M., V. Singh, S. Fay. 2006. An empirical study of the impact of nonlinear shipping and handling fees on purchase incidence and expenditure decisions. *Marketing Sci.* **25**(1) 51–64.

Little, J. D. C. 1966. A model of adaptive control of promotional spending. *Oper. Res.* **14**(6) 1075–1097.

Little, J. D. C. 1970. Models and managers: The concept of a decision calculus. *Management Sci.* **16**(8) B466–B485.

Little, J. D. C. 2004. Comments on "Models and managers: The concept of a decision calculus." *Management Sci.* **50**(12) 1841–1860.

Maddala, G. S. 1991. A perspective on the use of limited-dependent and qualitative variables models in accounting research. *Accounting Rev.* **66**(4) 788–807.

Mayer, T. 1993. Friedman's methodology of positive economics: A soft reading. *Economic Inquiry* **31**(2) 213–223.

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

Meade, N., T. Islam. 1995. Forecasting with growth curves: An empirical comparison. *Internat. J. Forecasting* **11**(2) 199–215.

Mitra, D., P. N. Golder. 2006. How does objective quality affect perceived quality? Short-term effects, long-term effects, and asymmetries. *Marketing Sci.* **25**(3) 230–247.

Mood, A. M., F. A. Graybill, D. C. Boes. 1974. *Introduction to the Theory of Statistics*, 3rd ed. McGraw-Hill, New York.

Morris, W. T. 1967. On the art of modeling. *Management Sci.* **13**(12) B707–B717.

Naik, P. A., K. Raman, R. S. Winer. 2005. Planning marketing-mix strategies in the presence of interaction effects. *Marketing Sci.* **24**(1) 25–34.

Pasa, M., S. M. Shugan. 1996. The value of marketing expertise. *Management Sci.* **42**(3) 370–388.

Pelloni, G. 1996. De Finetti, Friedman, and the methodology of positive economics. *J. Econometrics* **75**(1) 33–50.

Popper, Sir Karl Raimund. 1985. *Realism and the Aim of Science: From the PostScript to the Logic of Scientific Discovery*, Vol. 1. Hutchinson and Company, London, UK.

Porter, M. E. 1996. What is strategy? *Harvard Bus. Rev.* **74**(6) 61–78.

Press, S. J., S. Wilson. 1978. Choosing between logistic regression and discriminant analysis. *J. Amer. Statist. Assoc.* **73**(364) 699–705.

Radas, S., S. M. Shugan. 1998. Seasonal marketing and timing new product introductions. *J. Marketing Res.* **35**(3) 296–315.

Schnaars, S. P. 1984. Situational factors affecting forecast accuracy. *J. Marketing Res.* **21**(3) 290–297.

Shugan, S. M. 1986. Estimating brand positioning maps using supermarket scanning data. *J. Marketing Res.* **24**(1) 1–18.

Shugan, S. M. 2002. Marketing science, models, monopoly models, and why we need them. *Marketing Sci.* **21**(3) 223–228.

Shugan, S. M. 2004a. Endogeneity in marketing decision models. *Marketing Sci.* **23**(1) 1–3.

Shugan, S. M. 2004b. Consulting, research, and consulting research. *Marketing Sci.* **23**(2) 173–179.

Shugan, S. M. 2006. Errors in the variables, unobserved heterogeneity, and other ways of hiding statistical error. *Marketing Sci.* **25**(3) 203–216.

Silk, A. J., G. L. Urban. 1978. Pretest market evaluation of new packaged goods: A model and measurement methodology. *J. Marketing Res.* **15**(2) 171–191.

Sriram, S., P. K. Chintagunta, R. Neelamegham. 2006. Effects of brand preference, product attributes, and marketing mix variables in technology product markets. *Marketing Sci.* **25**(5) 440–456.

Staelin, R. 1998. Last reflections of the editor. *Marketing Sci.* **17**(4) 297–300.

Stigler, G. J. 1949. Monopolistic competition in retrospect. *Five Lectures on Economic Problems*. Longmans, Green & Co., London, UK, 12–24.

Stigler, G. J., G. S. Becker. 1977. De gustibus non est disputandum. *Amer. Econom. Rev.* **67**(2) 76–90.

Teller, E., W. Teller, W. Talley. 1991. *Conversations on the Dark Secrets of Physics*. Perseus, Cambridge, MA.

Trefil, J. S. 2003. *The Nature of Science: An A-Z Guide to the Laws and Principles Governing Our Universe*. Houghton Mifflin, Boston, MA.

Tucker, A. W. 1950. A two-person dilemma. Unpublished, Stanford University, Stanford, CA.

Urban, G. L., J. S. Hulland, B. D. Weinberg. 1993. Premarket forecasting for new consumer durable goods: Modeling categorization, elimination, and consideration phenomena. *J. Marketing* **57**(2) 47–63.

van Heerde, H. J., P. S. Leeflang, D. R. Wittink. 2002. How promotions work: Scan*Pro-based evolutionary model building. *Schmalenbach Bus. Rev.* **54**(3) 198–220.

Wible, J. R. 1982. Friedman's positive economics and philosophy of science. *Southern Econom. J.* **49**(2) 350–360.

Wible, J. R. 1984. The instrumentalisms of Dewey and Friedman. *J. Econom. Issues* **18**(4) 1049–1070.