

Theory, Experiment and Economics

Vernon L. Smith

It is now over thirty years since research was initiated in the laboratory experimental study of market behavior and performance.¹ This essay provides my interpretation of what the implications of this type of work are for the study of economics. The essay is not intended as a systematic survey of the field, although examples will be cited where appropriate and necessary. The reader can find the associated references in more general surveys (E. Hoffman and M. Spitzer, 1985; C. Plott, 1979, 1982, 1986a, 1986b; V. Smith, 1976, 1980, 1982a, 1982b, 1986).

Experimentation and Economics

Economics as currently learned and taught in graduate school and practiced afterward is more theory-intensive and less observation-intensive than perhaps any other science. I think the statement that “no mere fact ever was a match in economics for a consistent theory” accurately describes the prevailing attitude in the profession (Milgrom and Roberts, 1987, p. 185). This is because the training of economists conditions us to think of economics as an *a priori* science, and not as an observational science in which the interplay between theory and observation is paramount. Conse-

¹My first supply and demand experiment was done in January 1956, but others were involved at about the same time or earlier. Among the pioneering contributors to experimental economics were a number of scholars in the United States and Germany, many of whom were working quite independently without knowledge of each other's almost simultaneous work: E. Chamberlin (Harvard), A. Hoggatt (Berkeley), H. Sauermann and R. Selton (Germany), M. Shubik (Yale), S. Siegle and L. Fouraker (Pennsylvania State) and J. Friedman (Yale).

■ *Vernon L. Smith is Professor of Economics and Research Director, Economic Science Laboratory, University of Arizona, Tucson, Arizona.*

quently, we come to believe that economic problems can be understood fully just by thinking about them. After the thinking has produced sufficient technical rigor, internal coherence and interpersonal agreement, economists can then apply the results to the world of data.

But experimentation changes the way you think about economics. If you do experiments you soon find that a number of important experimental results can be replicated by yourself and by others. As a consequence, economics begins to represent concepts and propositions capable of being or failing to be demonstrated. Observation starts to loom large as the centerpiece of economics. Now the purpose of theory must be to track, but also predict new observations, not just “explain” facts, *ex post hoc*, as in traditional economic practice, where mere facts may be little more than stylized stories. The professional problem is for the theorist to recognize and respond to this purpose, and to undertake the arduous and challenging task of theory development disciplined by ongoing empirical studies. As Einstein put it, “[T]his theory is not speculative in origin; it owes its invention entirely to the desire to make physical theory fit observed fact as well as possible . . . the justification for a physical concept lies exclusively in its clear and unambiguous relation to facts that can be experienced.” But this process is not tautological so long as every time new data motivates an extension in theory, the new theory can be confronted with new field or laboratory observations, and this confrontation yields at least some victories some of the time.

In any confrontation between theory and observation the theory may work or fail to work. When the theory works it becomes believable in proportion to its predictive “miracle,” instead of only respectable in proportion to its internal elegance or its association with authority. But when it works, you lean mightily upon the theory with more challenging “boundary” experiments designed to uncover the edges of validity of the theory where certainty gives way to uncertainty and thereby lays the basis for extensions in the theory that increase its empirical content. When the theory performs well you also think, “Are there parallel results in naturally occurring field data?” You look for coherence across different data sets because theories are not specific to particular data sources. Such extensions are important because theories often make specific assumptions about information and institutions which can be controlled in the laboratory, but which may not accurately represent field data generating situations. Testing theories on the domain of their assumptions is sterile unless it is part of a research program concerned with extending the domain of applications of theory to field environments.

When the theory fails to work in initial tests, the research program is essentially the same. This is because all theories can be expected to be more or less improvable, and statistical tests of theories, whether the results are initially “falsifying” or not, are simply the means to motivate extensions in theory. Better theory that narrows the distance between theory and observation is always welcome.

From the perspective of experimental methodology, this scenario is what the profession of economics is all about. But it is not always what we economists do very well as a profession, because our publishing incentives are not always compatible with this research paradigm.

What Is There in a Theory to Test?

As is well known, when economists test a theory we make direct comparisons between observations and the predictions of the theory. But what precisely among the elements of a theory do we test when we make these comparisons? To answer this underlying question, it is instructive to distinguish among the following three ingredients of a theory: environment, institution and behavior.

The *environment* consists of the collection of all agents' characteristics; that is, tastes and technology, which in traditional economics are represented by utility or preference functions, resource endowments and production or cost functions. In reduced form these characteristics are the individual demand (willingness-to-pay) and supply (willingness-to-accept) schedules. The *institution* defines the language (messages or actions) of communication; examples include bids by buyers, offers by sellers, acceptances by either, and the characteristics of the commodity. The institution also specifies, either formally as on an organized exchange or informally by tradition, the order in which economics agents move, or that there is no order (moves are free form), and the rules under which messages become contracts and thus allocations. For example, in most retail markets the sellers first post their offer prices, then buyers scan, search and perhaps accept offers for stated quantities. The organized commodity and stock markets use variations on the oral double auction; buyers and sellers freely announce price-quantity bids and offers. A contract occurs when a buyer accepts a seller's offer or a seller accepts a buyer's bid. Consequently, the institution specifies the rules, terms or conditions under which components of market demand make contact with components of market supply to produce binding allocations.

Finally, *behavior* is concerned with agent choices of messages or actions given the agent's characteristics (environment) and the practices (institutional rules) relating such choices to allocations. Theories introduce assumptions about agent behavior: that agents maximize utility, profit or expected utility, that common information yields common expectations, that agents make choices as if they are risk averse, that expectations adjust using Bayes rule, that transactions costs (the cost of thinking, deciding, acting) are negligible.² Theories of behavior make predictions about messages—the bid(s) that an agent will submit at a sealed bid auction, the price that will be posted by an oligopolist, the reservation price below which a price searching agent will buy, and so on. Messages are not outcomes; they translate into outcomes depending upon the allocation and cost imputation rules of the institution.

²The distinction between that which we label "behavior" and that which is called an "agent's characteristic" (environment) will not, nor need it be, *a priori*. Part of the function of experiments is to increase our understanding of the issues involved in being or not being able to make this distinction. For example, is risk aversion an agent characteristic or an element of behavior embedded in his choices? It is both, but can we separate them in the context of experiments? It is yet to be shown empirically that we can operationalize this separation. See Cedric Smith's proposal (discussed below) to risk neutralize subjects by paying them in lottery tickets.

In laboratory market experiments, we test the theory's assumptions about agent behavior. How? Laboratory market experiments begin with an experimental design which seeks to control the environment using the techniques of induced valuation, and to control the institution by defining the language and the rules under which experimental subjects will be allowed to trade.³ With these controls we narrow the interpretation of inconsistency between predictions and observations so that the burden of inconsistency is borne by the behavioral assumptions of the theory. When the experimental observations are consistent with a theory we have our first evidence that the theory—as implemented by the particular environment and institution that was used—has predictive power. If the theory was explicit about the institution (for example, specified a sealed-bid discriminative auction), but made very general assumptions about the environment (for example, Q units offered to $N > Q$ unit bidders), then it is natural to direct the research exercise to variations on the original experimental environment. If the theory was not explicit about the institution (the strong interpretation is that the theory claims to be institution-free, but more likely the demands of tractability led the theorist to make simplifying institutional assumptions), then a reasonable research objective is to explore experimental designs that vary the institution. In this way one lays a foundation of empirical results that can motivate accommodating extensions in the original theory.

The above interpretation of experimental tests of a theory can be contrasted with tests based only on field data. In the latter case the economist has no independent control over the environment and the institution; as a result, the process is a *composite test* of the theory's assumptions about the environment, the institution and agent behavior. If the theory passes the test, it may be because all elements of the theory are "correct," or because "incorrect" elements of the theory had offsetting effects that could not be identified by the test. If the theory fails, the economist cannot know which of its elements accounted for the falsifying outcome.

When various operational forms of a theory are not falsified by laboratory data, we can say that the theory's assumptions about behavior are supported given the environment and institution posited by the theory and the experiment. But we are not finished. Often the theory will specify a particularly simple artificial institution which may fail to coincide with any that we observe in the field: for example, that firms choose price and/or quantities, and that buyers fully reveal demand. These institu-

³Value is induced on buyer i by assigning him/her values $V_i(1) \geq V_i(2) \geq \dots \geq V_i(Q_i)$ for successive units $1, 2, \dots, Q_i$, and guaranteeing to i that he/she will be paid in cash the difference between the assigned value, and the corresponding realized price paid in the market, $V_i(q_i) - Pq_i$, for each unit purchased, q_i . If $U_i(\cdot)$ is i 's unobserved monotone increasing utility of money, then i is motivated to buy an additional unit at any price below its assigned value. The valuation schedule thus becomes the individual's maximum willingness-to-pay for the item. Similarly, a supply schedule is induced upon sellers by assigning them individual "costs" for successive units and paying in cash the difference between the realized prices received and the assigned costs. This procedure easily generalizes to induce utility $U_i[V^i(X_i, Y_i)]$ on two "commodities" (X_i, Y_i) with the dollar valuation function $V^i(X_i, Y_i)$. The environment is defined and controlled by the collection of all assigned value (cost) schedules. The institution is controlled by defining the manner in which individuals interact to yield exchange prices and contracts.

tional assumptions can be reproduced in the laboratory, with real reward-motivated people as firms and simulated demand-revealing buyers. But it is a one-sided partial equilibrium theory of the firm and its market. This modelling tradition has carried over into contemporary game-theoretic theories of behavior in industrial organization.⁴ In recent extensions of search theory it is assumed that firms quote prices, knowing search behavior, and this generates an equilibrium-predicted price distribution. These models can be tested using artificial institutions that impose the conditions postulated by the theories. But the observed institution is free-form so that firms may quote prices while learning about the search behavior of buyers, and simultaneously the latter may choose their search behavior while learning about the prices set by firms. In these examples the experiments are constrained by the (message space) *limitations of the theories* that are tested, *not by limitations in the experimental methodology*. In fact, the easiest experiment is to put no restrictions on the price setting-searching process.

Much of the experimental literature is guided by nothing more sophisticated than the static theory that markets will clear. However, this body of literature has sought to break through the boundaries created by the current limitations of theory, and to establish a less restrictive empirical foundation for theory improvement. One such example is the large literature on the double “oral” auction trading institution, various forms of which are used in the organized stock, commodity, currency and interest rate futures markets. In this institution the messages are bids by buyers, offers by sellers and acceptances by either. There are some rules, such as that all-or-none bids are prohibited, and that a new bid must provide better terms than a standing bid. But beyond these the institution is free-form and is similar to modelling a *two-sided* search equilibrium market. Sellers announce and modify offer prices, while learning about the acceptance behavior of buyers. Simultaneously, the latter choose their acceptance behavior while learning about the prices announced by sellers. Buyers also announce bid prices, and sellers are free to accept such prices. Until recently (in the work of Easley and Ledyard, Wilson and Friedman) this institution has been beyond our analytical efforts. Yet to date no other trading institution studied in the laboratory has exceeded its capacity to exhaust the gains from exchange, or exceeded its speed of convergence to competitive equilibria.

Although this discussion has emphasized laboratory experiments, what I say applies also to field empirical research. Natural experiments occur all the time, and it would be desirable to develop a professional readiness to seize upon these occasions. When Mt. St. Helens started to quiver, it was quickly peppered with geologists and instruments collecting the data that can only be generated during the reactivation of a volcano. By comparison, the tradition of direct observation in economics appears weak; our training does not seem to include the techniques, nor develop the alertness, to respond to contemporary or historical empirical opportunities. One need not romanticize the techniques of other sciences, or exaggerate their applicability to

⁴Several experimental studies suggest that market behavior may be different when there are real buyers than when one simulates a revealed demand schedule in response to seller decisions.

economics, to recognize that economists could benefit from a stronger passion and curiosity for the microeconomics of how things work.⁵

Experiments, Institutions and Economic Theory

Since it is impossible to test a theory using experimental market data without specifying an institution, the experimental study of allocation processes forces an institutional and informational mode of thinking into every research design. People have to know what the message space is, who can move when, (as in posted price institutions), or that moves are unrestricted (as in oral double auction institutions), who knows what, when, and how their message decisions generate allocations, cost imputations and net returns. Experimentalists were therefore primed to welcome, and their research has been much influenced by, the *institution-specific theory* that began to develop about 1960.

The important thing we have learned from these theories, and from many of the experiments testing them, is that *institutions matter*.⁶ This is because agent incentives in the choice of messages (like bids) are affected by the institutional rules that convert messages into outcomes (like whether the high bidder wins an auction and pays the amount of the high bid or the amount of the second highest bid). In pre-1960 theory, by contrast, allocations were derived directly from the environment using ad hoc assumptions about demand revelation, or “price taking” behavior, by agents. Figure 1 illustrates the different ways of thinking. Experiments now address the question of how different institutions affect the incentive to reveal demand and supply. Thus the double oral auction elicits effective full revelation of demand and supply, but everyone in the market is a price maker as well as a price taker. Pre-1960 economic theory was totally unprepared for this kind of result.

Several questions can be raised concerning this dichotomy between institution-free and institution-specific theory. For example: doesn't the above definition of an institution preclude the possibility that any theory be institution-free? The answer is that a theory can be considered institution-free if it can be shown that the allocations it predicts are the same for all members of some class of institutions. An example is the theory of the four standard auctions: the English ascending price auction, in which the prices increase until only one bidder is left; the Dutch descending price auction, in

⁵An example of what I mean is the opportunistic response of Deacon and Sonstelie (1985) to an unusual natural experiment in which federal regulations temporarily constrained a few California stations to sell gasoline at prices less than those at other stations. Of course queues formed at the lower price stations. The authors put together a survey research program which was applied to both the high-price and low-price outlets which enabled them to measure the characteristics of respondents in the two situations, estimate the welfare cost of a market-wide ceiling on gasoline prices, and estimate the value of time spent in queues.

⁶The influential early contributions to this new conception of theory include Vickrey's (Nash-Harsanyi) models of the four standard auctions, Hurwicz's more abstract “mechanism” theory, and Shubik's emphasis on the extensive form game representation of microeconomics. Since that time there has been an increased development (particularly in bidding, information, and price search theory) of models that show how prices and allocations can or might be generated out of the internal processes of an institutionally mediated information exchange system. These theories have allowed institutions to slip unannounced back into economics, but now as an integral part of theorizing.

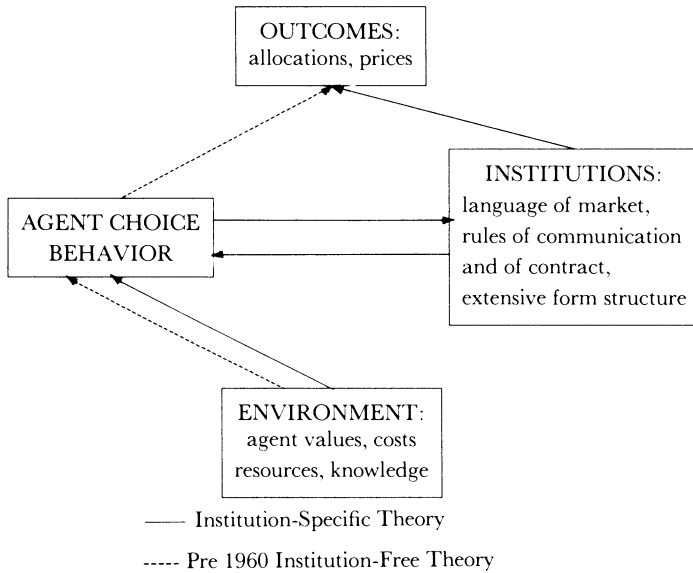


Fig. 1. Institutions in economic theory

which prices fall until some bidder bids; the first price sealed bid auction, in which the high bidder wins and pays his bid; and the second price sealed-bid auction, in which the high bidder wins but pays the amount of the second highest bid. These institutions are all equivalent if the environment is one in which all agents are risk neutral. The English and second price auctions are equivalent and the outcomes are all the same regardless of the risk attitude of agents. The Dutch and first price auctions are equivalent whatever the risk attitude of agents, but the outcomes (except in the risk neutral case) are distinct from those in the English and second price auctions. The experimental data tend to support the equivalence between English and second price auctions; do not support the equivalence of all four auctions (people do not behave as if risk neutral), and do not support the equivalence of Dutch and first price auctions.

Experimental methods can be used to test hypotheses about why certain institutions survive in the economy. Laboratory experiments enable economists to evaluate the performance characteristics of alternative institutions in controlled value-cost environments. These performance measures include: efficiency, speed of convergence, price stability, extent of price discrimination, responsiveness to changes in the environment, and so on. Dutch auctions are less efficient but faster (as implemented in the field), than English, first price, or second price auctions. This difference may account for the tendency to use Dutch auctions for selling perishables such as cut flowers, produce and fish. Posted price institutions are less efficient and yield higher selling prices than oral double auction institutions; but in the former price policy can be centralized, negotiation (a transactions cost) avoided and the products vended do not have to be standardized. These data can provide the basis for more complete theories of markets in which the institution is a variable and whose predictions can be tested with field observations across markets organized under different institutions of exchange.

There now exist many examples of continuing research programs in which theory has been sufficiently institution-rich to allow direct tests of the behavioral assumptions of the theory. I will discuss the first of two such examples in some detail to illustrate how a dialogue between testing and theory development can proceed using the trichotomy: environment, institution and behavior.

The first example involves Vickrey's risk neutral Nash equilibrium model of the first price sealed bid auction, which has stimulated an extensive study of bidding behavior in private value auctions by Cox, Smith and Walker (1988). (In a private value auction, each agent has a distinct value for the item.) Vickrey's theory assumes that each of N bidders, i , derives utility $u = v_i - b_i$ if i is the high bidder with bid b_i and i 's value for the auction objective is v_i . Note that all bidders are identical except for their values, v_i . Vickrey also assumed that the v_i are distributed as a rectangular distribution. Each bidder knows N and his/her own value, but only the distribution of all others values. Given this environment Vickrey deduced the equilibrium bid function $b_i = (N - 1)v_i/N$; that is, all individuals are predicted to bid the same fraction, $(N - 1)/N$, of their respective values.

We first found that about 92 percent of the subjects bid too high to be consistent with the Vickrey model; that is, they bid on the risk averse side of Vickrey's linear bid function. I often encounter the argument that the amounts of money used in laboratory experiments are "too small" for subjects to show risk averse behavior. But there is no theorem stating how small is "locally" in the phrase "locally linear." A lot of data from quite different experimental markets shows systematic deviations from risk neutral predictions. These data are given coherence by the hypothesis that subjects are risk averse.

Our experiments also found that the relationship between bids and values tends to be systematic for each individual, but to differ across individuals. This result is inconsistent with the hypothesis that individuals have identical risk aversion.⁷ Given these empirical results, we derived a new theoretical model based on the assumption that each bidder has constant relative risk averse utility from winning the auction with a particular bid given a particular private value of the prize. This model was consistent with the highly linear distinct bidding behavior of subjects. It can be shown that this is the *only* utility model which predicts that bids in relation to value will not be affected if the payoffs are increased by any multiple. We reran some of the

⁷A comprehensive survey of bidding theory (McAfee and McMillan, 1987) contains only one paper which admits of an environment in which agents may have differing nonlinear utility functions. Yet here is a critique of the above experimental results: "Of course, we don't have to go to an experimental situation to refute the hypothesis that individuals have the same degree of risk aversion!" Two comments are in order. First, if there is a widespread consensus on this "stylized fact" how come so many bidding models assume it away? I would suggest that the requirements of tractability, and the incentive to publish, loom large in charting this "low-apple-picking" course of least resistance. Otherwise, why do theory that assumes to be true that which we "know" to be false? Second, our auction experiments reject the hypothesis that people have the same degree of risk aversion, but a valuable field study by Binswanger (1980, p. 395), framed within the context of choice among uncertain prospects, found that "at high payoff levels, virtually all individuals are moderately risk-averse with little variation according to personal characteristics. Wealth tends to reduce risk aversion slightly, but its effect is not statistically significant."

experiments but tripled the payoffs, and observed no significant change in the relationship between an individual's bid and his/her value.⁸

Our model, however, implies linear homogeneous bid functions, and some 22 percent of the fitted linear individual bid functions had intercepts significantly different from zero. We hypothesized that the positive intercepts were due to a utility of winning in addition to the utility of the surplus won, while the negative intercepts were due to a threshold income necessary to provide positive utility. Does our *ex post facto* attempt to explain these nonzero intercepts increase empirical content? Yes; the new model is testable in that it implies that paying winning subjects a cash lump sum in addition to the auction surplus will increase their estimated linear bid function intercepts. Similarly, charging winning subjects a lump sum is predicted to decrease these intercepts. These new implications are not falsified by the indicated new experiments.

Despite this deepening chain of empirical successes, like *all* theory, this class of models is not without some unresolved empirical anomalies. For example, many years ago it was pointed out (C. Smith, 1961, pp. 13–14) that one could induce risk neutral behavior on risk averse expected utility maximizing agents. Instead of paying them cash for the outcomes of their decisions you give them chances, or lottery tickets, on a fixed reward if they win, and a smaller amount if they lose. This makes expected utility linear in outcomes whatever the utility of money function, and subjects are predicted to bid as if risk neutral in lottery tickets. However, this prediction is clearly falsified by auction experiments using this environment. It seems likely that it is the compound probability axiom of expected utility theory that fails in this application.⁹

⁸I have been asked: "How do you react to criticisms which say that from market data we can reject the assumption of constant relative risk aversion? We can look at how individuals change their portfolio with wealth, and it does not conform even to a much looser specification of the utility function? Why test a theory which has been rejected by market data?" Here are my reactions: (1) We can't reject the theory from this kind of market data. That data tells us how portfolios change with some measure of "wealth," confounded with changes in time, income, expectations, information, unmeasured probability assessments, and so on *ad infinitum*. We can't learn what we want to know from this sort of exercise independently of more rigorous tests, although market evidence and experimental evidence can illuminate each other. (2) Constant relative risk averse utility has been reported to do well in accounting for U.S. Treasury bill dealers' behavior (Wolf and Pohlman, 1983) both as elicited and as observed in bidding, but the constant relative risk averse coefficient shows greater risk aversion for actual bids than for the (Kahneman and Tversky) hypothetical assessments. (3) Non-constant relative risk averse utility cannot account for first price auction bidding behavior because the latter appears not to change when we triple payoffs. (4) Constant relative risk averse utility need not be valid over the entire interval of positive income to yield predictive accuracy over the relevant range of observations. Probably no functional form will be satisfactory everywhere.

⁹This interpretation raises the question, "If the compounding axiom fails, what does it imply about individual behavior? You can't have an important prediction of the theory rejected and still accept the theory." It is correct to say that not all observed behavior is consistent with expected utility maximization, but it is incorrect to say that you cannot therefore accept the theory. You can accept it, indeed you have little choice, until a better theory emerges. Theories are not accepted because all tests have failed to refute them. Lakatos (1978, pp. 4–5) wrote, "...scientists talk about anomalies, recalcitrant instances, not refutations... When Newton published his *Principia* it was common knowledge that it could not properly explain even the motion of the moon; in fact, lunar motion refuted Newton. Kaufman, a distinguished physicist, refuted Einstein's relativity theory in the very year it was published." People accept theories, in spite of anomalies, because they account for some evidence, and in particular are able to predict novel, even

In any case these results should alert scholars to the hazards of applying this procedure and assuming that their subjects are rendered neutral to risk. Any test of a theory which applies this procedure is necessarily a joint test of Smith's hypothesis and the theory; if the hypothesis fails to test out, the test of the theory is open to doubt. But even if C. Smith's hypothesis had been consistent with these tests, the exercise has limited transfer value in natural field and laboratory environments if real people are risk averse in the rewards that mediate their choices. If this is the case one must sooner or later confront the empirical demand for theories based on heterogeneous risk averse agents.

The second example is one in which the demand for new theory requires more than modifications that introduce individual differences in risk aversion. This is the extensive study of common value auctions by J. Kagel and D. Levin (1986) showing that subjects generally do not satisfy the predictions of a particular risk neutral Nash model of bidding. In this environment (see Thaler in the Winter 1988 issue of this journal for a discussion and a different interpretation) subjects do not know the common value of the item when they bid; but they each have an independent unbiased sample ("signal") which is positively related to true value. Unless one's bid is suitably discounted, as in this Nash model, the high bidder tends to be the one with the most optimistic sample estimate of value, and is said to be a victim of the "winner's curse." In small groups (3–4) experienced subjects make consistently positive profits and bid closer to the Nash prediction than the "winner's curse" prediction. (Profits are about 65 percent of the Nash prediction so that even here there appears to be room for theory improvement by appropriately introducing risk aversion). However, bids are found to increase with the number of bidders in larger groups (6–7), and contrary to this theory, experienced subjects suffer losses and bankruptcies. These data suggest that, *ceteris paribus*, the theory needs extensions that make the number of active bidders endogenous, and which predict equilibrium adjustment over time through some exit (entry) survival process. This is implied by the fact that the endogenous forces tending to vary N frustrated the Kagel-Levin attempt to fix N as a treatment variable in the larger groups. But the theory under test assumes that N is given. The data suggest an alternative zero profit market clearing model (it might be formulated as a Nash model) with N as a variable, in which entry would occur for $N = 3$ or 4 and exit when $N = 6$ or 7, yielding an equilibrium industry size, $N^* = 5$, for the parameters used by Kagel and Levin. One would seek a model in which N^* is a function of parameters characterizing the environment and, in

stunning, facts that cannot be accounted for by alternative theories. Newton's theory enabled Halley to make the stunning prediction that a certain comet would return in seventy-two years, as indeed it did, and Einstein's theory made the stunning prediction that a star's light was bent by the gravitational pull of the sun. Vickrey's model, suitably modified for heterogeneous risk averse bidders, rather astoundingly, I would say, accounts for and predicts the behavior of naive reward-motivated subjects in a first price auction. What alternative theory shows a comparable capacity to organize this immense data set?

addition to explaining the Kagel-Levin results, would lead to new testable implications.

Noncooperative Equilibrium Theory and Experiment

Two prevailing criticisms about noncooperative (Nash) equilibrium theory require modification in the light of experimental evidence.

First, it is widely believed that the concept of a noncooperative Nash equilibrium is “inherently” unsatisfactory because of its strong (or unrealistic) information requirements that each agent must know the preferences of all other agents.

Second, noncooperative equilibrium concepts are of questionable value because there are so many different such concepts leading to distinct theoretical implications; even with any given concept, such as Nash-Harsanyi, there are sometimes multiple solutions that leave open the question of a criterion for choosing among them.

The first criticism does not reflect the experimental results from many different environments which support noncooperative theory. The second criticism reflects a mode of thinking that excludes the prospect that a theory might be taken seriously enough to test it.

There are now numerous experimental studies based on noncooperative equilibrium concepts in which the results support the proposition that such equilibria are *more* likely to obtain under the minimal information requirement that no agent knows the preferences (or in an experiment, the payoffs used to induce preferences) of any other agent, and are *less* likely to obtain, or take longer to obtain, under the complete information conditions that critics argue are needed for equilibrium.

Fouraker and Siegel provided the first evidence on this point a quarter of a century ago in their classic work on bargaining behavior. Their bilateral bargaining, repeat transaction experiments were run under both complete and incomplete information. In these experiments, the seller chooses a price message, followed by the buyer choosing a quantity message, so that the noncooperative message equilibrium corresponds to the monopoly price and quantity. With incomplete information, eight of nine observations support a Nash outcome, and one supports the competitive equilibrium. With complete information 11 of 21 observations support a Nash noncooperative equilibrium, while 10 support competitive. Similarly, in their Cournot quantity message triopoly markets, support for Nash was 15 of 33 under complete information, 20 of 33 under incomplete. Finally, in their Bertrand price message markets, Nash was supported in 17 of 17 duopoly pairs under incomplete information, but only 11 of 17 under complete information. Another example based on large group double auction markets is discussed in the next section.

Two rudimentary fallacies underlie the criticism that the Nash equilibrium is not useful because it requires full information. First, *theorists* have to assume complete information in order to calculate a noncooperative equilibrium. But it does not follow that agents either require such information, or would know how to make the calculations if they had the information. An equilibrium is a state, and analysts can

ask whether a market tends to settle upon that state independently of the process used to calculate the state.

The second fallacy is that if agents have complete information, why should they use it to identify a noncooperative equilibrium? With complete information one can also identify more lucrative cooperative outcomes, and this is precisely the way real people deviate from Nash in most of the above experiments. Subjects are not so irrational as to satisfy “rational” models of behavior when it is contrary to their self-interest!

As theorists, we have been lax in the assumptions we make as to what follows from the state of common knowledge. First, what can objectively exist—say, in an experiment—is a state of common information, which is not the same thing as common knowledge or *expectations*. Real people have their own diverse ideas as to the knowledge implications of common information. Second, people have their own agenda as to what it is rational to expect and do given common information, and their presuppositions need not accord with economists’ presumptions about rational behavior.

The second major criticism of the concept of noncooperative equilibrium, that there are so many such concepts all with different theoretical implications, is symptomatic of a research program which attempts to answer scientific questions independently of observations. The obvious criterion in most sciences for selecting among alternative theoretical predictions is empirical, not *a priori*. One can design experiments capable of yielding results that may support any of the theoretical predictions, then see if the data “select” (are closer to) one among the competitors. Multiple theoretical predictions are exactly what the experimentalist likes to see in any science.

Concepts of noncooperative equilibrium have performed well in a large number of experimental markets—better than they have had a right to perform given that they were never intended to be tested, and that their acceptability is judged on internal rather than empirical criteria. Furthermore, the nonuniqueness of Nash equilibrium concepts (and even of equilibria) is a strength, not a weakness. When one concept fails empirically (as in common value auctions with N fixed and certain) there may be other formulations that do not contradict observation.

Laboratory Experiments, Psychology and Economics

The considerable literature that has developed using experimental methods has examined many of the mainstream issues in microeconomic theory and is helping to illuminate an increasing number of applied economic questions. If one seeks common general themes in this literature, particularly concerning behavior, I see three such empirical propositions, with far-reaching implications for how economists think about economics and how we interpret, and perhaps do, theory. The first proposition is that economic agents do not solve decision problems by thinking about them and calculating in the same way as economists. Second, from this first proposition it should not be presumed that economic agents will fail to get the “right” answers in the context of

markets: namely, the answers that are predicted by market theories. The experimental evidence is often consistent with the predictions of market theories. Third, economists have little in the way of formal understanding of how people are able to get the “right” answers without consciously performing our logic and calculations.

Experimentalists in both psychology and economics have provided overwhelming evidence in market experiments, in informal subject debriefing, and in choice surveys, that supports the first proposition. The second receives support from hundreds of supply and demand, oligopoly, bargaining and other experiments stretching back three decades in the work of many experimental economists and the psychologist S. Siegel. The third proposition calls attention to the lack of a satisfactory analytical-empirical integration of two bodies of evidence: one relating to the introspective and sometimes actual cash motivated responses of people in surveys soliciting choices among alternatives; the other to cash motivated choices in the context of repetitive market exchange.

Several psychologists (Edwards, Kahneman, Lichtenstein, Slovic, Tversky) and economists (Allais, Ellsberg) have pioneered the development of experimental designs for collecting evidence on hypothetical and actual individual choice behavior. This evidence generally tends to be inconsistent with expected utility theory, and with some of the fundamental behavior hypotheses in the theory of demand (Kahneman, Knetsch and Thaler, 1986; Knetsch and Sinden, 1984), such as the opportunity cost principle, and the implication of demand theory that there should exist only “small” differences between willingness-to-pay and willingness-to-accept. Some of this work has been replicated using cash payoffs and the conclusions (preference or demand theory is violated) are not changed. Although replication using cash payoffs (where this has not been done) is certainly needed, I think it is a mistake to assume that the economist’s paradigm will somehow be rescued in the context of these particular experimental designs, if experimenters would just pay money.

Given the high replicability of most of these studies, I think further such replications are of marginal value. What would be of much greater value is research directed at closing two gaps: the gap between decision theory and decision behavior, and the gap between evidence concerning how people think about economic questions and evidence concerning how people behave in experimental markets. Closing these gaps is crucial if we are not to get stuck on a research plateau. This is because it is clear from the work of Lakatos and other philosopher-historians of science that “there is no refutation without a better theory” (Lakatos, 1978, p. 6). Scientists in all disciplines simply ignore falsifying evidence until new and better theory emerges. Theory selection is based on opportunity cost, not absolute falsificationist standards.

There are two distinct paths whereby the distance between decision theory and behavior can be narrowed. The first lies in the empirical interpretation of expected utility theory: What are the prizes as perceived by the decision maker?

Almost uniformly we economists have assumed *a priori* that the objects to which the axioms of utility theory “should” apply are different amounts of wealth. Psychologists have followed this tradition in their empirical interpretation of data on choice behavior. For example, Kahneman and Tversky have a class of decision problems in

which the observations are consistent with expected utility theory if the utility function is S-shaped and is applied not to wealth, but to changes in wealth (income) from a reference point representing the individual's current wealth state. This curve describes risk seeking behavior below the reference point, and risk averse behavior above. But this result does not violate the theory. The axioms of the theory do not tell us what the prizes are. The theory simply postulates objects that can be preference ordered. It is an extra-theoretical subsidiary hypothesis to assert that these objects are amounts of final wealth measured in some particular way. Empirical evidence going back 35 years to H. Markowitz has suggested that the theory does much better if the prizes are changes in wealth, not absolute wealth. Selecting among hypotheses that are subsidiary to the axioms of a theory, but essential to its empirical interpretation, ought to be one of the more important functions of experimental methods.

In our application of utility theory to a risk averse model of the first price auction discussed earlier, we explicitly apply it to income because subjects participate in a sequence of auctions, with each auction representing a potential increment of wealth for each bidder. It is this form of the theory that organizes the data. This interpretation is consistent with other experimental evidence, and with the observed tendency of gamblers to make repetitive small stakes bets, as against a single bet wagering one's entire gambling budget.

Also in the auction example, notice that when we encountered linear empirical bid functions with nonzero intercepts, instead of dumping all utility theory forthwith, we asked whether it was possible to redefine the prizes so as to account for this contrary evidence, but in a way that was testable. The modified utility function has the Kahneman-Tversky S-shape, but with a "reference point" whose position varies with individual characteristics. The point to be emphasized is not that we have shown that prize reinterpreting extensions of expected utility theory will always work, but that they work in the context of auctions. Furthermore, such extensions in one context may be applicable in others, increasing coherence.

The second path in bringing together decision theory and behavior is to modify the axioms. Chew and Machina have done this expertly with various modifications of independence. These modifications account for some but not all of the violations. The Chew-Machina modifications accommodate Allais-type violations, and under certain conditions, preference reversals¹⁰; they provide new testable propositions not deducible from expected utility theory, and thereby yield an increase in empirical content. For a comprehensive discussion of these developments, see Machina (1987).

The fact that expected utility theory is consistent with some of the empirical evidence, especially when reinterpreted in terms of the prizes, and when extended along more fundamental Chew-Machina lines, argues strongly against any serious

¹⁰A preference reversal occurs when a subject says that he prefers A to B (or B to A) and that his willingness-to-pay (or willingness-to-accept) for A is less (greater) than that for B . A great many subjects' choices exhibit such reversals when A and B are different gambles (or different dividend paying assets). It should be added that the Chew-Machina modifications can accommodate preference reversals if they are the result of violations of the independence axiom. This is implied by the contributions of Holt and Karni and Safra. However the experimental tests by Cox and Epstein (1987) support the proposition that preference reversals are not due to violations of independence.

proposal to abandon it. At this juncture there is no alternative so that to abandon it is to substitute the void for a foundation of sand. The predictions of the theory and its extensions are noncontroversial, and it provides a powerful problem solving heuristic (Lakatos, 1978). Without this theory we have several empirical regularities whose predictive power in new situations is a function of the experience and opinion of the observer. There is a world of difference between having a theory whose predictions for all parameterizations are noncontroversial and having names to assign to certain empirical regularities.

How do we close the second gap, between the psychology of choice and agents' economic behavior in experimental exchange markets? The introspective empirical results make it clear that people's choices violate such basic tenets of demand theory as opportunity cost considerations and the approximate equality between willingness to accept and willingness to pay.¹¹ I think we economists need to accept these replicable empirical results as providing meaningful measures of *how people think about economic questions*. For their part, psychologists need to accept the dominating message in experimental research on the performance of a wide variety of bidding, auctioning and customer (posted price) markets: markets quite often "work" in the sense that over time they converge to the predictions of the economist's paradigm.¹² The few studies that simultaneously measure what people *say* (as questionnaire respondents), and what people subsequently *do* in experimental markets, confirm both kinds of evidence (Coursey, Hovis and Schulze, 1987; Knez and Smith, 1987). That is, the introspective responses of those subjects replicate the similar results of psychologists, but the responses are not necessarily consistent with the subjects' own subsequent market convergence behavior over time.¹³ I suggest that we assume that these two bodies of evidence correctly measure the thoughtful choices and revealed actions of

¹¹It should be emphasized that what is at stake here is the traditional utility-based theory of excess demand functions whose validity or falsity is separable from the theory of markets. The observation that willingness-to-accept exceeds willingness-to-pay, and that this is due to "loss aversion," implies that excess demand functions are discontinuous at the initial endowment. This means that gains from exchange are predicted to be lower than the prediction of the standard utility analysis. But market theory begins with postulated demand functions, which need not come from a utilitarian theory of demand. The Swedish economist Gustav Cassel argued that market theory should only begin with market demand and supply because the utilitarian derivation of market behavior was doomed to fail. Markets can be doing their thing whether or not demand theory is of any relevance to expressed demand.

¹²A new example of the discrepancy between behavior in one-shot choice experiments and behavior in markets is in the study of experimental markets for insurance by C. Camerer and H. Kunreuther (1987). Choice experiments often show evidence of violating expected utility theory when low probability significant loss events are involved. Camerer and Kunreuther study double auction markets for insurance contracts with these characteristics, and conclude that there is nothing special about risky losses in the range of parameters they study using this particular trading institution.

¹³In Knez and Smith (1987) we solicited willingness-to-pay and willingness-to-accept responses from potential buyers and sellers for two assets with known probability distributions of dividends. In subsequent double auction trading of these assets 40 percent of the buyers and 34 percent of the sellers announced bids or offers that violated their earlier stated responses. Furthermore, the stated preferences of subjects between the two assets changed considerably after experiencing market trading in them. From these results one should not conclude that the original hypothetical measures were wrong, but that people adapt their beliefs in the light of market experience. Here the incidence of "irrationality" is fairly common in the measures of what people think, but rare in their actual transactions. Only 3 of 146 transactions violated rational principles.

human subjects in these contexts. Reconciling these two sets of observations might be accomplished along the following lines.

People have their own homegrown beliefs about how markets work, or should work. (This is why economic concepts are difficult to teach to beginners.) Their questionnaire responses reflect these beliefs, which are often couched in terms of “fairness” criteria. Also, their initial behavior in a market may reflect these beliefs. But over time their behavior adapts to the incentive properties of markets as often (but certainly not always) represented in the standard economic analysis or its extensions.

It seems to me that we are confronted with two experimental research programs both of which have weak theoretical underpinnings. The economist’s maximizing paradigm often performs well in predicting the equilibrium reached over time in experimental markets, but this theory is not generally able to account for short-run dynamic behavior, such as the contract price paths from initial states to final steady states. Similarly, the psychologist’s “reference frame” descriptive paradigm performs well in explaining subject introspective responses, and their short-run, or initial, decision behavior, but it provides no predictive theory of reference frame adjustment over time. In fact, the statement (Kahneman, Knetsch and Thaler, 1986, p. 731) “that they (people) adapt their views of fairness to the norms of actual behavior” can be interpreted as a description of what is observed in experimental markets.

Initial choices may reflect all manner of beliefs and expectations, but if these choices are not sustainable in a market clearing or a noncooperative equilibrium, subjects adapt their expectations and behavior until they attain such an equilibrium. For example, it has been demonstrated that in an experimental design in which all the exchange surplus is captured by the buyers, convergence to the competitive equilibrium is slower and more erratic under complete (all values and costs are common knowledge) than under incomplete (values and costs are private) information (C. Smith, 1982, pp. 945–946). The condition of complete information gives the freest play to expectations based on social norms and beliefs. The latter are often inconsistent with equilibrium and retard full convergence until traders learn to adapt their beliefs about what they think “ought” to occur to what is attainable and can be sustained. Real people abandon their *a priori* beliefs when they find that their interest is poorly served by such beliefs. Under incomplete information people have little contextual basis for applying their *a priori* beliefs, and can be presumed to be more accepting of the behavior that sustains equilibrium. For the theorist, perhaps one way to model these phenomena in common information environments is to introduce agent *uncertainty about the behavior* of other agents.¹⁴ Now the theory will no longer predict that agents will come off the blocks straight into full rational expectations equilibrium. But as people adapt, and behavioral uncertainty is reduced, the theory can account for equilibrium convergence under some learning scenarios. This accords with observed price bubbles in experimental asset markets. With increasing subject experience

¹⁴ For example, in the various decision problems presented to subjects (see Kahneman, Knetsch and Thaler (1986) and the literature they cite) about 20 to 40 percent of the subjects respond with answers consistent with standard economic reasoning, while the majority responds contrarily, using “fairness” or other “nonrational” criteria.

the incidence of bubbles decreases and prices tend to converge to intrinsic dividend value.

The proposition that people adapt their beliefs about markets to the incentives of markets may also apply to disparate bodies of evidence on opportunity cost and sunk cost. Survey instruments show that, contrary to standard economic analysis, people do not ignore sunk costs and do not treat opportunity costs as equivalent to out-of-pocket costs. These concepts have not been generally examined in experimental markets. However, in my joint work with A. Williams and J. Ledyard in double auction trading with three commodities and two markets, the results fail to falsify the opportunity cost principle. In this environment each demand function depends on the prices of both commodities, and therefore willingness-to-pay in each market is based on foregoing the opportunity to buy an additional unit in the alternative market. These markets converge to the competitive equilibrium, supporting the effectiveness of opportunity cost in this context.

In general, one might think of changes in the reference frame or norms of behavior over time as being induced by the invisible reality of opportunity costs, entry or exit, and the irrelevance of sunk costs. Adaptation, where it is observed, may therefore be forced and agents need not have a cognitive grasp of the causes that are driving changes. Such a model implies sluggish nonoptimal intertemporal adjustment. If some agents (20 to 40 percent) are aware of the effects of opportunity cost, the effects of entry and exit, and the irrelevance of sunk costs, they may approximate optimal adjustment over time, and expedite the adjustment of the less perceptive agents.

Postscript

Experimentalists in economics frequently encounter an argument that proceeds roughly as follows: (a) If a theory is well articulated with clearly stated assumptions, and if there are no errors in the logic and the mathematics; then, (b) certain correct conclusions follow from the theory. So (c), what is there in a theory to test? The punch-line (c) often comes out in other forms without the conditionals (a) and (b) being stated. For example, when the data are consistent with the predictions of a theory, it is sometimes said that the results are not interesting because they merely confirm what economists already knew (or teach?), which seems to suggest that “truly” authoritative theory cannot be doubted seriously. When the data are inconsistent with the predictions of theory it is not uncommon to assert that there must be “something” wrong with the experiments.

Such objections are not without precedent in the history of any science. They tend to impose a double standard: if your theory says that the world is flat, then the tendency of some travelers to be “lost” (they never return home) is taken as evidence that they fell off the edge, while the fact that other travelers return home is interpreted to imply that they did not travel far enough to fall off the earth’s edge. Similarly, my experience has been that questions about experimental procedure are more likely to be

raised when the results appear to disconfirm accepted theory than when they appear to confirm such theory. However, if one wants to gain a greater understanding of economic phenomena, the most productive knowledge-building attitude is to be skeptical of *both the theory and the evidence*. This is likely to cause you to seek improvements in both the theory and the methods of testing.

One often hears it said that there is “too much theory spinning” or “not enough empirical work” in economics. Neither of these complaints adequately targets our professional weaknesses. Empirical studies would certainly benefit from more theory built directly on observed institutional processes. But not every testable theory may be worth laboratory testing. We need to think ahead to the domain of applicability of such efforts to field environments and institutions. Similarly, we could benefit from an increase in the kind of empirical research, both laboratory and field, that identifies and collects new data sources under the control and responsibility of the scientist; research that seeks to establish, rigorously, those empirical regularities worthy of stimulating deeper theoretical treatment. But we are particularly weak in ongoing research programs in which there is a progressive dialogue between theory development and particular results from laboratory and field tests; that steadily increase the empirical content of theory; and that build usable knowledge and a deeper understanding of things. The process will sometimes yield lags in empirical research, but just as often lags in theory. As the physicist Steven Weinberg described a similar situation in particle physics recently: “[T]here is not one iota of direct experimental evidence for supersymmetry, yet we study it because it looks so much like the sort of theory we would like to believe in. This is symptomatic of the terrible state we are in . . . The salvation of elementary particle physics is, at least for the moment, in the hands of the experimentalists.”

In economics the tendency of theory to lag behind observation seems to be endemic, and, as theorists, few of us consider this to be a “terrible state.” But as noted by Lakatos (1978, p. 6), “where theory lags behind the facts, we are dealing with miserable degenerating research programmes.”

Theory should be ever more demanding of our empirical resources. Simultaneously, data should be ever more demanding of the empirical relevance of theory and of the theorist’s expertise in working imaginatively on problems of the world, rather than on stylized problems of the imagination.

■ *I am grateful for research support from the National Science Foundation and from the Sloan Foundation to the University of Arizona, Economic Science Laboratory. I wish also to express my thanks to J. Cox, D. Kahneman, M. Machina, D. McCloskey, C. Shapiro, J. Stiglitz, T. Taylor and R. Thaler for dozens of pages of commentary on earlier drafts of this essay. I have attempted to incorporate their many valuable comments into the final version. That the final result is an improvement is as certain as is my failure to do as well as I would have liked. Words, pictures and formulas cannot convey a lifetime of experiences under the able tutoring of one’s experimental subjects.*

References

- Binswanger, Hans P.**, "Attitudes Toward Risk: Experimental Measurement in Rural India," *American Journal of Agricultural Economics*, August 1980, 62, 395-407.
- Camerer, Colin, and Howard Kunreuther**, "Experimental Markets for Insurance," Department of Decision Sciences, University of Pennsylvania, July 1987.
- Coursey, Don, John Hovis, and William Schulze**, "On the Supposed Disparity Between Willingness-to-Accept and Willingness-to-Pay Measures of Value," *Quarterly Journal of Economics*, August 1987, 102, 679-690.
- Cox, James, Vernon Smith, and James Walker**, "Theory and Individual Behavior of First Price Auctions," *Journal of Risk and Uncertainty*, March 1988, 1, 61-99.
- Cox, James C., and Seth Epstein**, "Preference Reversals Without the Independence Axiom," Department of Economics, University of Arizona, Discussion Paper No. 87-10, Sept. 1987.
- Deacon, Robert, and Jon Sonstelie**, "Rationing by Waiting and the Value of Time: Results from a Natural Experiment," *Journal of Political Economy*, August 1985, 93, 627-647.
- Hoffman, Elizabeth, and Matthew L. Spitzer**, "Experimental Law and Economics," *Columbia Law Review*, June 1985, 85, 991-1036.
- Kagel, John, and Daniel Levin**, "The Winner's Curse and Public Information in the Common Value Auctions," *American Economic Review*, December 1986, 76, 894-920.
- Kahneman, Daniel, Jack Knetsch, and Richard Thaler**, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, September 1986, 76, 728-741.
- Knetsch, Jack, and John Sinden**, "Willingness to Pay and Compensation Demanded: Experimental Evidence of An Unexpected Disparity in Measures of Value," *Quarterly Journal of Economics*, August 1984, 99, 507-521.
- Knez, Marc, and Vernon Smith**, "Hypothetical Valuations and Preference Reversals in the Context of Asset Trading." In Roth, Alvin, ed., *Laboratory Experiments in Economics: Six Points of View*. Cambridge: Cambridge University Press, 1987, pp. 131-154.
- Lakatos, Imre**, *The Methodology of Scientific Research Programmes*. Vol 1. Worrall, J., and G. Currie, eds. Cambridge: Cambridge University Press, 1978.
- Machina, Mark J.**, "Choice Under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives*, Summer 1987, 1, 121-154.
- McAfee, R. Preston, and John McMillan**, "Auctions and Bidding," *Journal of Economic Literature*, June 1987, 25, 699-738.
- Milgrom, Paul, and John Roberts**, "Information Asymmetries, Strategic Behavior, and Industrial Organization," *American Economic Review*, May 1987, 77, 184-193.
- Plott, Charles R.**, "The Application of Laboratory Experimental Methods to Public Choice." In Russell, C. S., ed., *Collective Decision Making: Applications from Public Choice Theory*. Baltimore: Johns Hopkins University Press, 1979, pp. 137-160.
- Plott, Charles R.**, "Industrial Organization Theory and Experimental Economics," *Journal of Economic Literature*, December 1982, 20, 1485-1527.
- Plott, Charles R.** (a), "Laboratory Experiments in Economics: The Implications of Posted-Price Institutions," *Science*, 9 May 1986, 232, 732-738.
- Plott, Charles R.** (b), "Rational Choice in Experimental Markets," *Journal of Business*, October 1986, 59, S301-S327.
- Smith, Cedric**, "Consistency in Statistical Inference and Decision," *Journal of the Royal Statistical Society*, Ser. B, 1961, 23, 1-25.
- Smith, Vernon L.**, "Bidding and Auctioning Institutions: Experimental Results." In Amihud, Y., ed., *Bidding and Auctioning for Procurement and Allocation*. New York: New York University Press, 1976, pp. 43-64.
- Smith, Vernon L.**, "Relevance of Laboratory Experiments to Testing Resource Allocation Theory." In Kmenta, J., and J. Ramsey, eds., *Evaluation of Econometric Models*. New York: Academic Press, 1980, pp. 345-377.
- Smith, Vernon L.** (a), "Reflections on Some Experimental Market Mechanisms for Classical Environments." In McAlister, L., ed. *Choice Models for Buyer Behavior*. Greenwich: JAI Press, 1982, pp. 13-47.
- Smith, Vernon L.** (b), "Microeconomic Systems as an Experimental Science," *American Economic Review*, December 1982, 72, 923-955.
- Smith, Vernon L.**, "Experimental Methods in the Political Economy of Exchange," *Science*, 10 October 1986, 234, 167-173.
- Thaler, Richard H.**, "Anomalies: The Winner's Curse," *Journal of Economic Perspectives*, Winter 1988, 2, 191-201.
- Wolf, Charles, and Larry Pohlman**, "The Recovery of Risk Preferences from Actual Choices," *Econometrica*, May 1983, 51, 843-850.