Economics and Philosophy, 2, 1986, 245–273. Printed in the United States of America.

LABORATORY EXPERIMENTATION IN ECONOMICS

ALVIN E. ROTH

University of Pittsburgh

1 INTRODUCTION

Controlled experiments conducted by economists under laboratory conditions have a relatively short history. Only in the last ten years has laboratory experimentation in economics completed the transition from being a seldom encountered curiosity to a well-established part of the economic literature. (The *Journal of Economic Literature* has this year initiated a separate bibliographic category for "Experimental Economic Methods.") How this came to pass makes for an interesting episode in the history and sociology of science. What I will try to describe here, however, are the different *uses* to which laboratory experimentation is being put in economics.

I think that, loosely speaking, many of the experiments that have been conducted to date fall on an imaginary continuum somewhere between experiments associated with testing and modifying formal economic theories (which I will call "Speaking to Theorists"), and those associated with having a direct input into the policy-making pro-

I have had helpful correspondence and discussion on this subject with Jack Ochs, Paul Milgrom, and Charles Plott. This work has been supported by grants from the National Science Foundation and the Office of Naval Research, and by a Fellowship from the Alfred P. Sloan Foundation. Charles Plott and Charles Holt graciously gave me permission on behalf of themselves and their colleagues to reproduce figures from their experiments. This paper was prepared for delivery at the Fifth World Congress of the Econometric Society, August 1985. It will also be published in *Advances in Economic Theory*, edited by Truman Bewley, which is forthcoming from Cambridge University Press.

cess (which I will call "Whispering into the Ears of Princes").[1] Somewhere in between lie experiments designed to collect data on interesting phenomena and important institutions, in the hope of detecting unanticipated regularities ("Searching for Facts"). Most experimental investigations will contain elements from more than one of these categories.

In the following sections, I will briefly describe examples of each of these activities and how they interact with and contribute to other parts of economic research. I like to think of economists who do experiments as being involved in three kinds of (overlapping) dialogues, and the examples I have chosen are designed to illustrate these. The material I will discuss under the heading "Speaking to Theorists" illustrates the kind of dialogue that can exist between experimenters and theorists, while the material in the section "Searching for Facts" illustrates the kind of dialogue that experimenters can engage in with one another. "Whispering in the Ears of Princes" deals, of course, with the kind of dialogue that experimenters can have with policy makers.

One sign of how experimental economics has grown in recent years is that it would be impossible, in the space available here, even to attempt to survey comprehensively the great variety of experimental investigations that have been undertaken. My aim here will only be to describe briefly some work that I think exemplifies each of the activities and dialogues discussed above. There are many other examples I could have chosen.[2]

2 SPEAKING TO THEORISTS

The experiments I will discuss next, in which I have personally been involved, are concerned with two-person bargaining under the following rules. (For a fuller discussion, see Roth, forthcoming.) The two agents may allocate some valuable resource between themselves in any way they like, provided they both agree. If they fail to agree on how to allocate the resource, they each receive nothing. This is an example of what is sometimes called a "pure" bargaining game, which can be thought of as the opposite of the idealized case of "perfect" competition: whereas individual agents have negligible influence under perfect competition, in a pure bargaining game each agent has an absolute veto over every division of the resource between the bargainers.

At least since the time of Edgeworth (1881) it has been argued by some economists that pure bargaining games are fundamentally inde-

1. This latter kind of laboratory experiment is in some respects close kin to field experiments. See., e.g., Ferber and Hirsch (1982) or Hausman and Wise (1985).
2. One important topic I will not cover is experimental studies of individual-choice behavior, since it was the subject of another paper in the Congress for which this paper was originally prepared. Provocative discussions on the implications of these studies for economics appear in Thaler (forthcoming) and Knez & Smith (forthcoming), which are included in the collection of papers (Roth, forthcoming) presenting the various viewpoints on economic experimenters.

terminate. In the language of cooperative game theory, the problem is that the core of such a game is the large set of outcomes corresponding to the entire set of agreements that leave no part of the resource unallocated.[3] The prediction that observed outcomes will be in the core is therefore a fairly weak (although not an empty) prediction. There has thus been considerable sustained interest in developing theories that attempt to predict specific outcomes in the core. Such theories attempt to make use of information concerning some measure of the *intensity* of agents' preferences, and most often do so by using the kind of information contained in a von Neumann–Morgenstern expected utility function representing those preferences.[4] That is, these theories attempt to predict the outcome of bargaining on the basis of information about each bargainer's willingness to tolerate risk.

In the theoretical literature concerned with these matters, these theories were primarily regarded as offering predictions about bargaining conducted under conditions of "complete information," in which bargainers were assumed to know the information about one another's preferences contained in their expected utility functions. The formal requirements of the complete information condition were often regarded as an idealization, incapable of a concrete realization. Furthermore, it was sometimes argued that basing a theoretical model of bargaining on the "complete information" contained in the agents' utility functions was somehow overdetermining the problem; i.e., the actual phenomena being modeled would be sensitive only to a subset of this information. In any event, since an agent's willingness to tolerate risk cannot be readily observed in uncontrolled, naturally occurring environments, theories of this kind proved unusually difficult to test with field data.

Laboratory experimentation offers (at least) two promising approaches for testing theories whose predictions are sensitive to difficult-to-observe attributes of the agents such as their risk posture and information. The first of these is to design experiments that *control* for individual differences on the relevant dimensions, using the extensive possibilities for control inherent in setting up economic environments in the laboratory (in contrast to [nonexperimental] field studies, in which the environment must be taken as given). The second is to *measure* directly the relevant attributes of agents, taking advantage of the relatively unrestricted

3. And this same set of outcomes can be achieved as the equilibria of the game in strategic form.
4. The most influential single theory of this sort is due to John Nash (1950), whose work led to the development of a family of related theories in the tradition of cooperative game theory; see Roth (1979) for a full account. More recently, there has been a growing body of work on theories of bargaining in the tradition of noncooperative game theory; see Roth (1985a) for a collection of representative work. Some of this work uses information about time preference rather than risk preference in the preferences of the bargainers, in a manner exemplified by Rubinstein (1982).

access to agents that is available in the laboratory (again, in contrast to in the field). My colleagues and I found it convenient to use the first of these approaches to test theoretical predictions concerning what information possessed by the agents would influence the outcome of bargaining, and to use the second approach to test predictions concerning the effects of differences in the bargainers' risk postures.

For the experiments designed to test predictions about the distribution of information on the outcome of bargaining, an experimental design was developed (see Roth and Malouf, 1979) that controlled for the risk preferences of the bargainers. Bargainers play *binary lottery games*, in which each agent i can eventually win only one of two monetary prizes, a large prize λ_i or a small prize σ_i ($\lambda_i > \sigma_i$). The players bargain over the distribution of "lottery tickets" that determine the probability of receiving the large prize: e.g., an agent i who receives 70% of the lottery tickets has a 70% chance of receiving the amount λ_i and a 30% chance of receiving the amount σ_i. Players who do not reach agreement in the allotted time each receive σ_i. Since the information about preferences conveyed by an expected utility function is meaningfully represented only up to the arbitrary choice of origin and scale, there is no loss of generality in normalizing each agent's utility so that $u_i(\lambda_i)=1$ and $u_i(\sigma_i)=0$. The utility of agent i for any agreement is then precisely equal to his probability of receiving the amount λ_i, i.e., equal to the percentage of lottery tickets he receives.[5]

Note that restricting each agent to an environment in which all choices involve lotteries over only two ultimate payoffs (receiving his large prize or his small prize) controls out the effects of risk aversion or risk preference, which are phenomena that arise in making tradeoffs among three or more outcomes. That is, on the restricted choice domain of a binary lottery game, the differences in risk posture that individuals bring to the bargaining table are controlled away; i.e., they have no scope for expression. Consequently, theories such as Nash's (and a broad class of related axiomatic and strategic theories) that distinguish between individual bargainers only on the basis of their expected utility payoffs, treat all bargainers in binary lottery games symmetrically. One prediction of such theories, therefore, is that the resulting agreements will give each

5. Of course, we have not addressed here the (empirical) question of whether a given individual's choice behavior can indeed be summarized by a preference relation exhibiting the regularity conditions needed for it to be accurately represented by an expected utility function. In fact, a substantial body of empirical work has recorded a number of systematic ways in which individual preferences often fail to exhibit these regularities. However, virtually all of these departures from expected utility maximization involve the way tradeoffs are made among three or more riskless alternatives; see Machina (1983). Consequently these kinds of departures cannot occur when only two such alternatives are feasible, as in a binary lottery game. Some methodological innovations that make clear how to extend the use of binary lottery games to explore experimentally a wide variety of economic environments have been described by Berg et al. (forthcoming).

bargainer 50% of the lottery tickets. This is so independently of the value of the monetary prizes, and, specifically, independently of whether the prizes of one bargainer are equal to those of the other.

Note further that bargaining under this design meets the requirement of being a game of "complete information," since the information determining each agent's utility for any outcome is his percentage of lottery tickets at that outcome, which is common knowledge. The design thus provides the opportunity to test the predictions of various economic theories of bargaining precisely on the domain to which they are intended to apply. Since the bargainers have "complete" information whether or not they know the value of one another's prizes, a second prediction made by theories of bargaining that use as data only this information is that the outcome of bargaining will be the same when bargainers know one another's prizes as when they do not.

The experiment of Roth and Malouf (1979) was designed in part to test these predictions, and to determine whether or not changes in the size of the prizes, and whether the bargainers knew one another's prizes, influenced the outcome. Each bargainer played games with different prizes[6] against different opponents in one of two information conditions. In the "full information" condition, each bargainer knew both his own prize and that of the other bargainer; while bargainers who were assigned to the "partial information" condition each knew only their own prize value.

The results of the experiment were that, in the partial information condition, and also in those games of the full information condition in which the two bargainers had equal prizes, agreements were observed to cluster, with very low variance, around the "equal probability" agreement that gives each bargainer 50% of the lottery tickets. In the full information condition, in those games in which the bargainers' prizes were not equal, agreements tended to be distributed bimodally between two "focal points": the equal probability agreement and the "equal expected value" agreement that gives each bargainer the same expected value. That is, in these games the bargainer with the lower prize tended to receive a higher share of the lottery tickets. Contrary to the prediction of the theory, both the monetary value of the bargainers' prizes and whether each bargainer knew his opponent's prize were clearly observed to influence the outcome.

My first inclination on seeing these results was that they would be explainable within the framework of traditional noncooperative (strategic) game-theoretic models, by modeling the game in greater detail so as to take into account the larger strategy sets that more information gives to the bargainers. However, an experiment utilizing a strategically equivalent bargaining environment in which bargainers knew the value of one another's prizes in terms of an artificial intermediate commodity (rather

6. In these games the small prize was always equal to zero for both bargainers.

than in terms of their monetary value), showed that this was not the case (Roth, Malouf, and Murnighan, 1981). Information about the artificial commodity did not affect the outcomes in the same way as did strategically equivalent information about money. The results of this experiment supported the hypothesis that there is a *social* aspect to the bimodal, "focal point" phenomenon, that depends on something like the players' shared perceptions of the credibility of any bargaining position.

While it is still not clear what theoretical tools can fruitfully be brought to bear on the task of explaining this class of unpredicted phenomena, it was not too difficult to design experiments to explore it further. The experiment of Roth and Murnighan (1982) was designed to separate the observed effect of information into components that could be attributed to the possession of specific information by specific individuals.

Each game of that experiment was a binary lottery game in which one player had a $20 prize and the other a $5 prize. In all eight conditions of the experiment, each player knew at least his own prize. The experiment used a 4 (information) × 2 (common knowledge) factorial design. The information conditions were: 1) *neither knows* his opponent's prize; 2) the *$20 player knows* both prizes, but the $5 player knows only his own prize; 3) the *$5 player knows* both prizes, but the $20 player knows only his own prize; and 4) *both players know* both prizes. The second factor made this information common knowledge for half the bargaining pairs, but not common knowledge for the other half.[7] Note that the two conditions that made it common knowledge that neither player knew both prizes or that both players knew both prizes provide a replication of the experiment of Roth and Malouf (1979).

The results of this experiment permitted three principal conclusions. First, the equal expected value agreement becomes a focal point if and only if the player with the smaller prize knows both prizes. When the $5 player knew that the other player's prize was $20, this was reflected not only in his messages and proposals, but also in the mean agreements (when agreement was reached), and in the shape of the distribution of agreements. (See Figure 1.) Note that, in the four conditions in which the $5 player does *not* know his opponent's prize, the distribution of agreements has a single mode, corresponding to the 50/50 equal probability agreement. However, in the four conditions in which the $5 player *does* know that the other player has a $20 prize, the distribution of agreements is *bi*modal, with a second mode corresponding to the 20/80 equal expected

7. For example, when the $20 player is the only one who knows both prizes, then the (common) instructions to both players in the common knowledge condition reveal that both players are reading the same instructions, and that after the instructions are presented, one player will be informed of only his own prize, and the other will be informed of both prizes. In the noncommon knowledge condition, the instructions simply state that each player will be informed of his own prize, and may or may not be informed of the other prize.

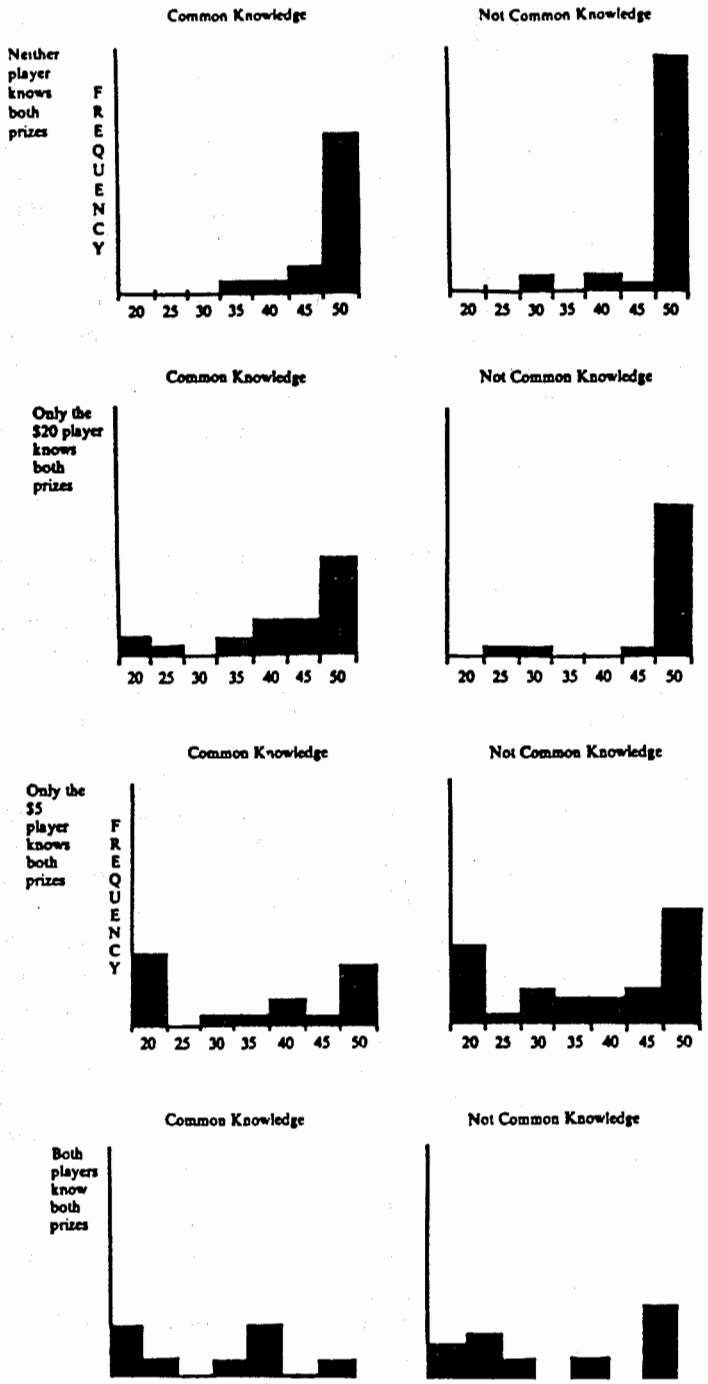

FIGURE 1 Frequency of Agreements in Terms of the Percentage of Lottery Tickets Obtained by the $20 Player

value agreement. The mean agreements reached when neither player knows both prizes and when both players know both prizes replicate the results of Roth and Malouf (1979), both in direction and magnitude.

Second, whether it is common knowledge what information the bargainers possess influences the frequency of disagreement. The frequency of disagreement in the two noncommon-knowledge conditions in which the $5 player knew both prizes is significantly higher than in the other conditions. The highest frequency of disagreement (33%) occurs when the $5 player knows both prizes, the $20 player does not, but the $5 player doesn't *know* that the $20 player doesn't know both prizes. (In this situation the $5 player cannot accurately assess whether or not the $20 player's [honest] skepticism that his opponent's prize is only $5 is just a bargaining ploy.)

Third, in the noncommon-knowledge conditions, the relationship among the outcomes is consistent with the hypothesis that the bargainers are rational utility maximizers who correctly assess the tradeoffs involved in the negotiations. That is, in the noncommon knowledge conditions there is a tradeoff between the higher payoffs demanded by the $5 player when he knows both prizes (as reflected in the mean agreements reached), and the number of agreements actually reached (as reflected in the frequency of disagreement). One could imagine that, when $5 players knew both prizes, they might have tended, as a group, to persist in unrealistic ambitions about how high a percentage of lottery tickets they could expect to get. The mean overall payoffs (which include both agreements and disagreements) indicate that this is not the case. The increase in the number of disagreements just offsets the improvement in the terms of agreement when the $5 players know both prizes, so that the overall expected payoff to the $5 players does not change. This means that the behavior that $5 players were observed to employ in any one of these conditions could not profitably have been substituted for the behavior observed in any other condition. The same is true of the $20 players: in particular, the expected payoff of $20 players who knew both prizes does not differ from that of those who knew only their own prize (although it is significantly affected by what the $5 player knows), so that a $20 player who knew both prizes, for example, could not have profited from behaving as he would have if he knew only his own prize.

Thus, although the effect of information and the bimodal distribution of agreements between "focal points" observed in all of these experiments is unpredicted by existing theory, the data nevertheless exhibits observable regularities that resemble other phenomena for which theoretical models exist, such as equilibrium behavior in certain games of incomplete information. It is common in incomplete information models to observe a "screening" effect, in which a bargainer whose (private) attributes put him in an unusually strong bargaining position must distinguish himself from weaker bargainers by demonstrating a willingness to

accept some higher cost, such as higher probability of disagreement (see, e.g., Chatterjee, 1985). While these models will not be directly applicable (since the experimental observations discussed here occur under conditions of complete information), the pattern of observable regularities suggests that theories capable of describing these phenomena may not necessarily have to be radical departures from existing theories.

However, the evidence also suggests that descriptive theories of bargaining will have to take account of some information beyond the traditional game-theoretic specification of what constitutes "complete" information. (Recall that the experiment of Roth, Malouf, and Murnighan, 1981, identified a social component to the observed effect of information about prize values.) A subsequent experiment (Roth and Schoumaker, 1983) lends support to the hypothesis that the effect of information about the cash value of prizes is attributable to its effect on the *expectations* of the bargainers about what constitutes a credible bargaining position. Such information may help bargainers (and theorists) to select from among the multiplicity of equilibria that are found in bargaining games (cf. Rubinstein, 1985).

So far, the bargaining experiments discussed have focused on observable effects due to experimental manipulations that existing theory predicts should have no effect. Such manipulations are a powerful experimental device for exploring the underlying assumptions of a given theory or class of theories.[8] However, the significance of these unpredicted effects for our evaluation of the overall state of the theory cannot be evaluated in isolation. We must also consider the descriptive power of the theory's *predicted* effects, and conduct experiments that will permit us to evaluate how well the theory predicts the effects of manipulating those factors that it designates as important.

Risk posture is obviously an important factor in the predictions of theories that are stated in terms of the von Neumann–Morgenstern utilities of the bargainers. It turned out, however, that the predictions of these theories about the specific effects of risk aversion had not been developed in a way that yielded experimentally testable hypotheses. (One indirect virtue of experimentation is that it can provide a discipline to theoretical work, and suggest directions in which theory ought to be explored.) To this end, therefore, a systematic study of the predictions of these models relating to the risk posture of the bargainers was carried out in Roth (1979); Kihlstrom, Roth, and Schmeidler (1981); Roth and Rothblum (1982); and Roth (1985b). Rather surprisingly, a very broad class of apparently quite different models, including all the standard

8. Compare this to the research strategy followed by, e.g., Kahneman and Tversky on utility theory and "framing effects" in individual choice. See Kahneman and Tversky (1984); Tversky and Kahneman (1981); and Thaler (forthcoming).

axiomatic models[9] and the strategic model of Rubinstein (1982),[10] yield a common prediction regarding risk aversion. Loosely speaking, these models all predict that risk aversion is disadvantageous in bargaining, except when the bargaining concerns potential agreements that have a positive probability of yielding an outcome that is worse than disagreement. That is, these models all predict that a personal attribute of the bargainers—specifically, their risk aversion—will have a decisive influence on the outcome of bargaining. This prediction concerning risk aversion is important to test because it connects the theory of bargaining with what has proved to be one of the most powerful explanatory hypotheses in a number of other areas of economics—e.g., in explanations of investment behavior, futures markets, and insurance.

Three studies exploring the predicted effects of risk aversion on the outcome of bargaining are reported in Murnighan, Roth, and Schoumaker (1985). Whereas binary lottery games were employed in the earlier experiments precisely in order to control out the individual variation due to differences in risk posture, these studies employed *ternary* lottery games having *three* possible payoffs for each bargainer i. These are large and small prizes λ_i and σ_i obtained by lottery when agreement is reached, and a disagreement prize δ_i obtained when no agreement is reached in the allotted time. (In binary lottery games, $\sigma_i=\delta_i$).

The bargainers' risk postures were first measured by having them make a set of risky choices. Significant differences in risk aversion were found among the population of participants, even on the relatively modest range of prizes available in these studies (in which typical gambles involved choosing between receiving \$5 for certain or participating in a lottery with prizes of λ_i=\$16 and σ_i=\$4).

Those bargainers with relatively high risk aversion bargained against those with relatively low risk aversion in pairs of games such that $\delta_i>\sigma_i$ in one game and $\delta_i<\sigma_i$ in the other. The prediction of game-theoretic models such as Nash's is that agreements reached in the first game should be more favorable to the more risk averse of the two bargainers than agreements reached in the second game.

The results of these experiments support the predictions of the game theoretic models, but suggest that, in the (relatively modest) range of payoffs studied here, the effects due to risk aversion are smaller than some of the "focal point" effects observed in previous experiments. The significance of this is that the classical models of bargaining attribute the entire variability of bargaining outcomes to the differences in bargainers' risk posture, and predict no influence for the

9. Including those of Nash (1950); Kalai and Smorodinsky (1975); and Perles and Maschler (1981).

10. But see Binmore, Rubinstein, and Wollinsky (1985) for an alternative theoretical interpretation.

information about the prizes that allows a second focal point to come into play. Our observations here are that the predicted cause of variability has less effect in this range of prizes than the unpredicted cause. Even if subsequent experiments should show an increased effect due to risk aversion when the prizes are larger, these results give ample reason for devoting additional effort to theoretical and experimental study of the focal point phenomenon.

Taken together, these experiments do several things. First, they disconfirm some important aspects of the received theory of bargaining, chief among which is what constitutes a "complete" specification of the information available to bargainers. This is particularly notable in light of the fact that much of the theoretical criticism of the complete information assumption is founded on the assumption that bargaining is better modeled by assuming that the bargainers have strictly *less* than complete information. The experimental evidence suggests that, while bargainers may certainly be expected in general to have less of *some kinds* of information, the outcome of bargaining is also highly sensitive to information the bargainers may have about each other *in addition* to what is included in "complete" information about utility functions. These experiments also provide preliminary support for some of the subtle but robust predictions of these bargaining theories about the effect of risk aversion. They also serve to illustrate what I suspect will prove to be fairly typical interactions between formal theories and experiments designed to explore them. But this last point will be easier to discuss once we have looked at some other uses of experimentation.

3 SEARCHING FOR FACTS

The set of experiments discussed next involves markets in which buyers and sellers seek to trade with one another under one of a variety of rules of market organization.[11] In each market, sellers were told how many units of a good they could "produce," and given a schedule of unit costs—i.e., a seller might be told that he may sell up to two units, the first of which will cost him \$1.00 to produce and the second \$1.25. At the conclusion of the experiment, each seller is paid the sale price of any unit he has sold minus its cost: i.e., if the above seller sells only one unit, at a price of \$1.10, he will receive (a profit of) \$0.10 for the transaction. Similarly, buyers are told how many units they can consume, and what the value of each unit is to them. That is, a buyer might be told he can only consume one unit, and it is worth

11. See Smith (1964), (1976), (1982b); Plott and Smith (1978); and Plott (1982). Here I will be concerned with the general design and results of these experiments, and will not discuss details of procedure.

$1.00 to him. At the conclusion of the experiment, each buyer is paid the value to him of any units he has purchased, minus the purchase price. Thus if the above buyer makes a purchase at a price of $0.75, he will receive $0.25 for the transaction.[12] Each agent knows only his own costs or values. All goods are indivisible and homogeneous.

Various rules for making transactions were compared in these experiments. In a *double auction,* sellers may make offers and buyers may make bids. Sellers and buyers may make new offers and bids whenever they like; the lowest outstanding offer is the "market asked price," and the highest outstanding bid is the "market bid price." If any seller offers the market bid price, or if any buyer bids the market asked price, a transaction is consummated between the seller and buyer whose prices coincide.

In a *single auction,* only one side of the market is allowed to suggest prices. There are thus two kinds of single auctions, the *bid* auction in which buyers may make bids, and the *offer* auction in which sellers may make offers. Agents on the passive side of the market may merely accept the current market (bid or asked) price, or remain silent. Agents on the active side of the market may revise their prices whenever they wish. A transaction occurs when an agent on the passive side of the market accepts the market price.

In a *posted-price market,* as in the single auction, only one side is allowed to suggest prices. However, each agent on the active side of the market must "post" his price when the market opens, and is thereafter not free to change it. There are, of course, two kinds of posted-price markets, depending on whether sellers are posting offers or buyers are posting bids.

In each of these experiments, some small numbers of buyers and sellers (typically from 4 to 12 of each) engaged in some moderate number of repetitions of such a market (typically 5 to 25 "trading days") under one of the market institutions described above. Each trading day constitutes a new market, with the same agents and parameters. (Thus in the various auctions agents could revise their prices as often as they wished each trading day, but in the posted-price markets, prices could be posted only at the beginning of each trading day.) The final payment that subjects received was equal to their total profit on all transactions in all trading days.

In all of the experiments, transaction prices in the final period were observed to be substantially closer to the competitive price than in the first period. This was most clearly observed in the double-action markets, in which many transactions in the final periods occurred precisely

12. In the experiments discussed here, agents were effectively forbidden to make trades that would bring them a negative return. In some of these experiments, agents were also credited with a small "commission" for each trade they made.

at the competitive price, and in which the quantity traded was virtually always within a single unit of the competitive quantity.

While the other market institutions also exhibited a tendency to converge toward the competitive price and quantity, Plott and Smith (1978) observed that, as compared with the results of the double-auction markets, single auctions and posted-price markets each exhibited a bias toward one or the other side of the market. Curiously, the bias in the single auctions is in favor of the *passive* side of the market, while the bias in the posted-price markets is in favor of the *active* side. That is, compared to the prices observed in the double auction (and to the competitive price), a single auction yields a higher price when the buyers can make bids, and a lower price when the sellers can make offers, while the posted-price market yields a higher price when the sellers can post offers, and a lower price when the buyers can post bids. Thus it is an advantage to be on the active side of a posted-price market rather than on the passive side, and an advantage to be on the passive side of a single auction rather than on the active side.[13]

As far as I know, no presently available theory is able to account for these results, which are a good example of the kind of unanticipated regularity that can arise from using experimental methods to "search for facts." (However, recent theoretical advances concerning games of incomplete information have made important progress toward elucidating some similar phenomena—see, e.g., Wilson, forthcoming.) It is difficult to imagine how empirical support for such a proposition might have been gathered without using experimentation, since, in naturally occurring economic environments, we normally observe different market institutions in operation only in quite different markets. In the laboratory, however, it is possible to conduct an experiment designed so as to hold constant all other market features (such as the number of buyers and sellers, and the distribution of supply and demand) and compare the outcomes of otherwise identical markets under these different institutions.[14]

Turning back to the double-auction institution, the fact that the experimental markets conducted under these rules were often observed to converge precisely to the competitive equilibrium is rather

13. A considerable amount of subsequent data gathered on experimental posted-price markets largely replicate this result. As far as I know, much less additional data has been gathered on single auctions of this kind, so that the conclusions regarding single auctions must be regarded as being of a more tentative nature.
14. To avoid confusion, a note on terminology is in order here. I use the term "experiment" to denote a set of experimental observations, e.g., of auction markets, that permit some comparisons to be made, and the term "experimental design" to denote the relationship among the parameters and procedures used to make the different observations. In some of the literature discussed here, the term "experiment" is used to denote a single observation of a multiperiod market, and "experimental design" is used to denote the parameter settings in each market.

surprising in view of the small numbers of traders involved, and the active role that each plays in determining the price. This is clearly not an environment that approximates "perfect" competition as it has traditionally been formulated. What seems to be happening is that a great deal of information is conveyed to the agents in this kind of market, not merely by the transactions that take place, but also by the bids and offers that are not accepted and do not lead to transactions. For example, if the first few trades that take place on the first trading day happen to occur at a price far above the competitive price, then some potential buyers will be priced out of the market. As the trading day progresses, sellers who have units left to sell lower their asking prices a little, hoping to transact with some of these buyers. Each buyer learns from these lower asking prices that some units are available at a lower price than previous transactions, and this influences the bids made on the next trading day.[15]

There remains a spectrum of opinion on how robust is the convergence to competitive equilibrium in double auctions—i.e., on the extent to which such convergence is due to the double-auction institution itself, as opposed to other features of the way the experiments were conducted. Smith (1982b) states:

> The double auction . . . is a remarkably robust trading institution for yielding outcomes that converge to the C.E. It achieves these results with a small number of agents, under widely different supply and demand conditions, with each individual agent having strict privacy, that is, the agent only knows his/her own value or cost conditions. Several sets of experiments have been conducted to test the boundary of application of these conditions. One set of experiments [Smith, 1981; Smith and Williams, 1981] used only one or two sellers. . . . Only in the one-seller experiments is there a failure to arrive consistently at C.E. outcomes, thus establishing "one" as the limiting number of sellers at which competitive price theory fails under double auction trading.

This is a strong and simply stated conclusion. One of the virtues of the experimental method is that, if a conclusion drawn from experimental data appears to other investigators to be too strong or too simple, further experiments can be designed and data gathered to reopen the question. Two experiments with a bearing on the above conclusion are briefly discussed below.

Holt, Langan, and Villamil (1984) considered whether convergence to competitive equilibrium in double auctions might be influenced by the parameters determining the supply and demand curves. They considered, for example, a double auction examined by Smith, which he

15. Recall that information about supply and demand obtained from offers and transactions in one trading day is highly germane to the next trading day since these experiments were conducted so that the market starts over again each trading day, with supply and demand unchanged by the trades that occurred on the previous trading day.

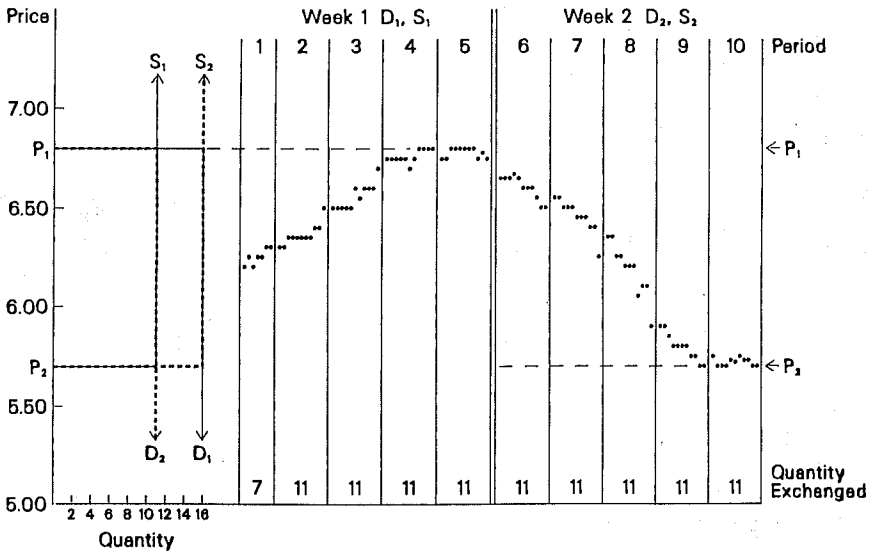

FIGURE 2

reported (Smith, 1982a, p.172), "provides the most stringent of all reported tests of the equilibrating tendency in double auction trading," since the competitive equilibrium gave all of the exchange surplus[16] to one side of the market. They observed that while in this dimension the test is indeed stringent, nevertheless the supply and demand is such that "the lack of market power is so severe in this design that even if one buyer unilaterally withholds demand for all four of his units, he has no effect on market price." They proposed to examine double-auction markets that differed from those previously examined in that agents on one side of the market had market power, in the sense that, by foregoing some trades, they could move the competitive price sufficiently in their favor so that they would earn a larger profit on their remaining transactions than if they had made all the trades that would be profitable for them.

Holt et al. first replicated the results of the double auction discussed by Smith, confirming the tendency of the market to converge to competitive equilibrium (see Figure 2). They then conducted seven multiperiod double auctions each involving five buyers and five sellers, using the procedures of previously published experiments, but with parameters that gave market power either to some of the buyers or to some of the sellers. Three of their double auctions used experienced subjects. Contrary to the convergence to competitive equilibrium uniformly observed in earlier experiments (Smith, 1982b), four of the seven auctions

16. Except for the "commissions" that were paid for each transaction.

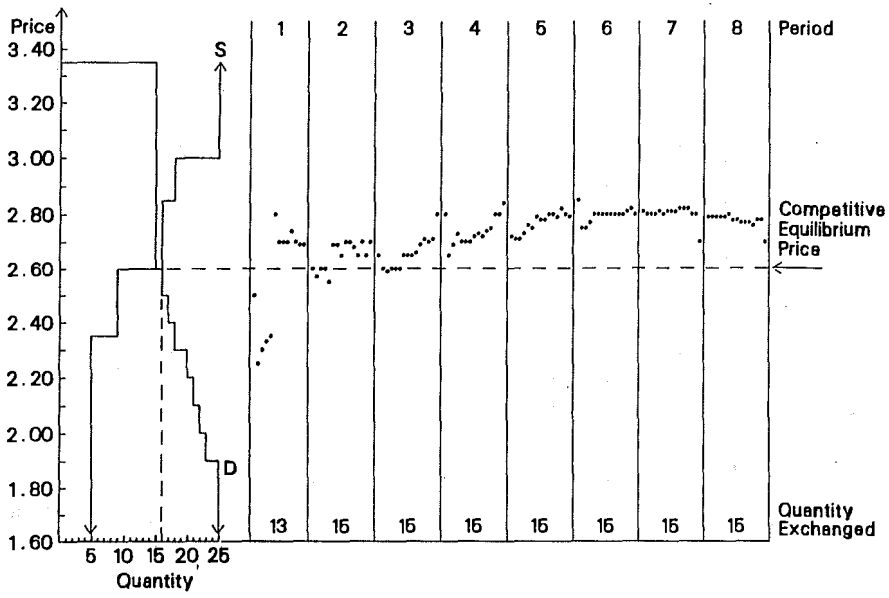

FIGURE 3

observed here, including all of those with experienced subjects, failed to converge to the competitive price, and converged instead to a price reflecting the distribution of market power (see, e.g., Figure 3). The results of this experiment thus support the hypothesis that the parameters of the market may influence the convergence to competitive equilibrium previously observed in double-auction markets, particularly when experienced subjects are involved. (For closely related results, see also the subsequent paper by Holt and Villamil, 1984).

In most of the multiperiod repeated auction experiments conducted to date, the number of periods was determined fairly arbitrarily, typically with an eye toward the time constraints in running the experiment. Alger (1984) noted that this left it impossible to tell whether or not these markets had stabilized; i.e., if they would continue to behave much as in the final periods of observation if they were allowed to continue to operate. He designed an experiment, using posted-price markets, to see whether behavior in the number of periods typically examined in previous experiments resembled the market behavior when it had been in operation enough periods to settle down fully. His operational definition was to terminate each market only after there had been no deviation in seller profits for "at least five periods [usually ten periods] or if an obvious cyclic pattern developed." In order to allow market periods to run very quickly in order to accommodate a potentially large number of periods, Alger simulated the buyer side of the market, by having buyers

behave competitively. Therefore, the only active agents in his markets were sellers. He examined the two-seller case.

His results are thought provoking. First, he observed that markets often took many more periods to stabilize than were allowed in previous experiments, and that while the initial periods of his markets looked a great deal like the same number of periods in the earlier literature[17] the final periods, when the market had stabilized, did not. In particular, a typical pattern that was observed is that prices initially fell to somewhere near the competitive level, but later rose to near the monopolistic price, as the sellers succeeded in signaling each other of their willingness to cooperate, by raising their posted prices. The striking thing about these results is not merely that these markets finally settle down near the monopolistic price, while previous experiments involving posted-price markets had supported the conclusion that they tended to settle down near the competitive price; what is particularly interesting is that if only the same number of (initial) periods of this experiment were examined as were observed in previous posted-price markets, the conclusion would be that, in this experiment, too, prices tended to converge to the competitive price.

Alger quite properly emphasizes that it would be premature to try to draw direct conclusions about previous experiments from the results of this one, because of the many design differences other than the number of periods observed. However, the time pattern of his results raises questions that will likely stimulate further experimental study.

Note that this whole group of experiments has been conducted without reference to a well-articulated body of formal theory. Instead, what we are seeing is a series of experiments inspired by previous experiments, seeking to isolate, identify, and delimit experimentally observed empirical regularities. These experiments have shown that competitive equilibrium outcomes can be observed with very few buyers and sellers. This simple fact seems to provide a powerful challenge and stimulus to developing new theories of competitive behavior. And, since there is evidence supporting a diversity of opinion about how robust this observation is likely to prove, it would be surprising if these results do not stimulate further experimental work as well.

4 WHISPERING IN THE EARS OF PRINCES

The next experiments I will discuss have been motivated by questions of policy, of the kind raised by government regulatory agencies, typically about the effects of changes in the rules by which some market is organized. These investigations offer the possibility of bringing scien-

17. In particular, he compares his data to posted-price markets discussed by Ketcham, Smith, and Williams (1984) and Fouraker and Siegal (1963).

tific methods to bear on one of the traditional nonscientific vocations of economists, which is whispering in the ears of princes who require advice about pressing practical questions whose answers lie beyond the reliable scientific knowledge of the profession.

One of the studies discussed (Hong and Plott, 1982) arose in a matter of concern to the Interstate Commerce Commission; the other (Grether and Plott, 1984) in a case before the Federal Trade Commission. Both cases had to do with complex posted-price markets, and in both cases an attempt was made to mirror as closely as possible in the laboratory the industrial structure of the market in question.[18]

The ICC case concerned whether barge operators should be required by the ICC to post their prices, and announce price changes in advance. The existing market allowed rates to be set by individual negotiations between barge operators and their customers, so that the terms of each contract were private information. Plott (forthcoming) reports that the question arose because railroad companies were lobbying to require such price posting. The reasons offered by the railroads were that

> the public information feature of posted rates would make the industry more competitive, allow the railroads to compete better, and aid the small barge owners who were allegedly secretly being undersold by the large barge companies (Plott, forthcoming).

In their introductory comments, Hong and Plott (1982) say the following about their use of laboratory experimentation to illuminate the issues raised by the proposed change:

> The full consequences of a rate filing policy are unknown. Plausible theoretical arguments can be made on both sides of the policy argument. When existing theory does not yield a definitive answer, one can usually turn to previous experiences with policies, but in this case we are aware of no industrial case study that would provide direct evidence on either side of the controversy.

They go on to note that it would be difficult to draw any compelling policy conclusions regarding the barge industry from previous laboratory experiments concerning posted-price markets, since

> any extrapolation from published experimental results to the barge industry itself is open to two potential criticisms, the reasonableness of which this study was designed to assess. First, the barge industry has several prominent economic features that are not incorporated in existing laboratory market studies. Examples include the relative sizes of buyers

18. See also Plott, 1985, forthcoming.

and sellers, the demand and supply elasticities, and the cyclical nature of demand. Naturally, we can never be certain that all the important features have been included in the present design. If something important has been misspecified or omitted, then the observed behavior of the laboratory market may not extend to the barge industry, and additional appropriately modified experiments can be conducted as checks on our conclusions. The second potential criticism is that the effects of price posting in laboratory studies have only been measured relative to the performance of oral auction markets. Since auction markets differ from the negotiated price markets of the industry, the relevance of the comparison can be questioned.

Hong and Plott proceeded to design their experiment around a laboratory market scaled to resemble, in the features mentioned above, the market for transporting grain along the upper Mississippi River and Illinois Waterway during the fall of 1970. (This market was chosen because it was believed to be representative of a significant portion of the dry-bulk barge traffic in the United States, and because adequate data about the market parameters were available.) Aggregate supply and demand functions for the laboratory market were scaled to estimates available for the target market, as was the distribution of large and small firms on each side of the market. The laboratory market was divided into periods representing two weeks of the target market, and the seasonal aspects of the target market were modeled by having demand in the laboratory market scaled to resemble two months of normal demand, followed by two months of high demand, followed by two months of normal demand. The experimental design involved running the market under both posted-price and negotiated-price policies.

In presenting the data from this experiment, Hong and Plott report that

> The results are easy to summarize. The posted price policy causes higher prices, reduced volume, and efficiency losses. Furthermore, the posted price policy works to the disadvantage of most market participants, especially the small ones, and helps only the large sellers.

They also conclude that the posted-price markets react more slowly to the seasonal change in demand than do the negotiated price markets.[19]

We will consider the relationship of these experimental results to policy conclusions after briefly considering the experiment of Grether and Plott, which was motivated by an FTC complaint that also involved posted prices, among other things.

19. Plott (forthcoming) reports that this experimental evidence helped to deter the lobbying of the railroads on this matter, and that the price-posting policy they had advocated was not pursued.

The FTC case involved a complaint by the FTC against the pricing practices of the Ethyl Corporation, E.I. du Pont de Nemours and Company, PPG Industries, Inc., and Nalco Chemical Corporation, the four domestic producers of tetraethyl and tetramethyl lead, the additives in "leaded" gasoline that raise its octane level. The FTC sought to have the producers cease and desist from a number of unusual pricing practices used for these additives, that, according to the FTC theory, had the effect of reducing price competition.

One of these pricing practices was that suppliers agreed to give at least a thirty-day notice of all proposed price increases, and usually such announcements were made with even more than the contractual thirty-day warning. Another practice was that "most favored nation" clauses were commonly included in contracts, by which a buyer was assured that he would receive the best terms being offered by the seller to any customer. (Apparently "meet or release" clauses were sometimes also used, which assured a buyer that the seller would meet the price offered by any other seller, or else release the buyer from any contractual obligation to buy from that seller.) In addition, all prices were quoted in terms of "delivered prices," for goods delivered to the purchaser regardless of his location.

Some of these practices might appear to favor the buyers, but the FTC theory was that, together, they worked to allow the producers to cooperate in raising prices. One way to explain this idea is as follows. If a producer thought that a price increase was desirable, he could announce, with something more than the required thirty-day warning, his intention to raise his price. This would not cause a customer with a "meet or release" clause to start searching for a supplier with a better price, because such a customer is assured that, in any event, he will only be charged the lowest price. (The lowest price is known, since prices are announced, and it is unambiguously defined, since only delivered prices are quoted, so there can be no hidden discounts in transportation costs.) If the other producers agree that this price increase is desirable, they can also announce it; otherwise, it will be rescinded by the initial producer. So a producer faces little cost in exploring the possibility of a price rise, while at the same time has little incentive to explore a price cut, since he will not be able to increase his market share (again, because of advance announcements, and "meet or release" clauses).

In its defense, the industry advanced the competing theory that price levels were determined entirely by the concentrated structure of the industry, and that in such a concentrated industry, prices were unaffected by the pricing practices described.

Expert testimony by economists was available in support of both positions. The experiment of Grether and Plott was intended to be a

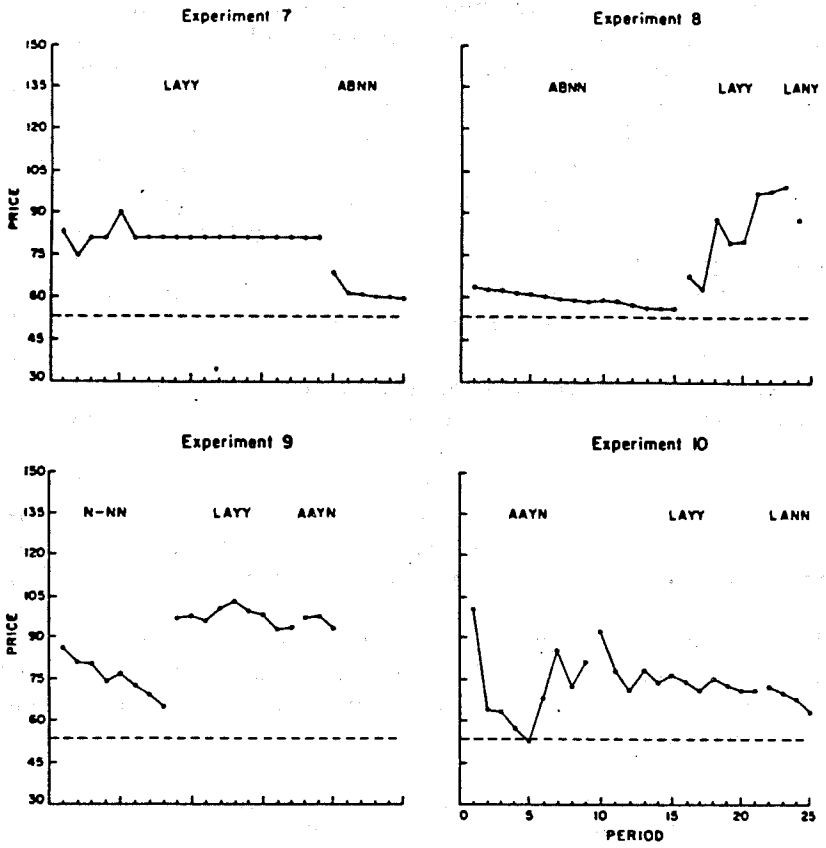

FIGURE 4

possible source of evidence for rebuttal of the industry theory that the indicated pricing policies could not be affecting the price in such a concentrated industry. A scaled-down model of the industry was implemented in the laboratory, with careful attention paid to preserving the relative costs, capacities, and numbers of participants. The experimental design involved a number of multiperiod repeated markets, each one of which would be examined both with and without (some or all of) the pricing practices in question. The results are fairly clear—when all the practices are in force, the observed prices are above those that are observed when none of the practices are employed (see Figure 4).[20]

20. Plott (forthcoming) reports that this experimental evidence was ultimately not used in court testimony. The government won the case, but was reversed on appeal. For some subsequent theoretical work that supports the general conclusions of this experiment, see Holt and Scheffman (1985).

Looking at these two experiments together, it is apparent that one of the differences between these experiments and those described in the previous sections has to do with the complexity of the economic environment being studied. There is some tension between the goal of designing an experimental market to resemble a particular naturally occurring market, and the goal of designing an experiment whose results will be likely to support some fairly general conclusion.

On the other hand, it should also be apparent that these "policy-oriented" experiments have something in common with the "theory-testing" experiments described in Section 2, since both involve the testing of hypotheses, whether those hypotheses arise from formal economic theories, or from the arguments of lawyers, lobbyists, and expert witnesses. However, in contrast to hypotheses drawn from general economic theories, which are presumably applicable to any market in which the conditions of the theory are met, the hypotheses of interest in this case are explicitly concerned with the target market, and not with the experimental market. Therefore, the bearing that the experimental evidence has on the hypotheses is different in these policy-oriented experiments than in the theory-testing experiments discussed in Section 2.

Plott (forthcoming) aptly describes the role of experimental evidence in policy debates of this kind as serving to *shift the burden of proof.* Hong and Plott, speaking of the experiments modeled on the barge industry, put it this way:

> From a scientific point of view, we have solid evidence only that price posting markets do not necessarily operate better than negotiated price markets under the parametric conditions we considered. From a policy point of view, this evidence presumably shifts the burden of proof to the price posting advocates, who must now identify the specific features of the barge industry which, if incorporated in the experiment, would reverse the conclusions.

5 SOME IMPORTANT OMISSIONS

While any survey of this sort must inevitably have more omissions than inclusions, there are two bodies of work that I will mention here in order to highlight briefly some different aspects of experimental research.

Recall that the section on theory-oriented experiments discussed a research program that started with a body of formal theory, and proceeded to develop a set of experiments that allowed some conclusions to be drawn about the theory. This is by no means the only direction that the dialogue between experimenters and theorists can take. A good example of a complementary approach is given by the work of

Selten on coalition formation. (See particularly Selten, forthcoming; but also Selten, 1972, 1982, and Selten and Krischker, 1982.) Selten considers the formation of coalitions in three-person games presented in what game theorists call "characteristic-function form." He starts, not with a theory of coalition formation that he wishes to test in the laboratory, but rather with a body of data from a number of experiments involving such games, conducted under varying experimental conditions by a variety of experimenters, motivated by different theoretical considerations.

Selten identifies some empirical regularities in this diverse set of experimental data, and constructs a formal theory of "equal division payoff bounds" to describe and explain them. He proposes statistical tests that allow the descriptive accuracy of this theory to be compared to that of previously existing theories applicable to this data, and shows that these measures support the hypothesis that this theory may have superior predictive power for experiments of this kind. There are two points that I would like to make concerning how this work exemplifies different aspects of the dialogue that can go on between experimenters and theorists.

The first is that, whereas the next step of the work described in Section 2 is to construct new theory that is able to describe the observed data, the logical next step of Selten's work is to conduct new experiments specifically designed to test the predictive value of the proposed theory. In general, this back-and-forth process between conducting experiments and constructing theories can be expected to keep iterating, with experimental data motivating new theory, which can in turn motivate new experiments, and so forth.

The second point is that the demands on theory made by experimental data can be quite different from those imposed by traditional deductive considerations. In this respect, Selten conducts something of a dialogue with himself. As a theorist, he is well-known for his seminal work on perfect equilibria, which forms the basis for much of the current theoretical work on rational and "hyperrational" behavior.[21] But on the basis of his reading of the experimental data he finds himself constructing a theory of a radically different sort. Indeed, he observes:

> The success of the theory of equal division payoff bounds confirms the methodological point of view that the *limited rationality* of human decision behavior *must be taken seriously* (italics added).

and

> The optimization approach fails to do justice to the structure of human decision processes (Selten, forthcoming).

21. See, e.g., Rubinstein (forthcoming).

It is the mark of a committed scientist to be able to adjust his theoretical ideas in the face of compelling evidence, and a characteristic of experimental evidence is that it will often have the power to compel such adjustments in economic theories.

The second body of experimental work addresses the fact that powerful general theories may often have application on unexpected domains. Specifically, a number of experiments have now been conducted to test the applicability of economic theories of consumption and labor supply to the behavior of laboratory animals, namely pigeons and rats. (See particularly Kagel, forthcoming; and Kagel, Battalio, and Green, 1981.) In these experiments, the laboratory animals are observed in controlled environments where they may perform specific tasks (such as pecking a key for pigeons, or pressing a lever for rats) to obtain food or water. By changing the number of key pecks required to obtain a given amount of food or a given amount of water, the relative prices of food and water can be changed, and observations about changes in the consumption of food and water (and leisure) can be compared to the prediction of economic theories. There is a large body of theory and experimental evidence on animal behavior in the behavioral psychology literature, and it appears that economic theories of behavior compete quite effectively in terms of their predictive and explanatory power.

However, economic theories also suggest a host of questions that it is unlikely would have been raised by psychologists studying animal behavior. Kagel (forthcoming) discusses a number of such questions. For example, the "welfare-trap" hypothesis is that agents who receive unearned income will get "hooked on leisure" and subsequently reduce their labor supply. Kagel reports observing a significant but quantitatively small effect in this direction among pigeons who had in earlier periods been given "unearned income" in the form of free access to food and water.[22] Similar experiments were motivated by the "cycle of poverty hypothesis" that low-income agents will tend to discount the future more heavily than high-income agents. Kagel reports observing the *opposite* effect among liquid-deprived rats who were able to choose[23] between small immediate payoffs (of saccharin water) or larger delayed payoffs. In another experiment, rats were observed to exhibit violations of the independence axiom of expected-utility theory, similar to those observed in human subjects in choice situations like the Allais paradox.[24] Kagel (forthcoming) presents a very clear and provocative discussion of the interplay

22. Needless to say, some of the technical questions of experimental design that arise in experiments with animal subjects (such as controlling for subjects' body weight) are quite different from those that arise in experiments with human subjects.
23. By pressing one of two levers.
24. See Machina (1983).

between economic theories and animal experimentation, both from the point of view of better understanding animal behavior, and from the (somewhat more controversial) point of view of understanding fundamental biological components of human economic behavior.

6 DISCUSSION

In evaluating a new method of research like laboratory experimentation in economics, two different kinds of questions seem important. The first of these is "What new economic phenomena have been elucidated so far? What facts have been established through experimentation? What new avenues of research have been opened?" The second is "How do these new methods relate to the other approaches available to economists?"

I think it is fair to say that there already exists a sufficiently large body of successful experimental economic research so as to make a comprehensive answer to the first kind of question impractical in an essay of this scope. At the same time, the experimental approach is still too new to allow a definitive answer to the second kind of question. But we have considered some examples of what has been learned from experimental studies, and I hope that the organization of this essay gives some perspective on the way experimentation is related to other things that economists do.

One thing that should now be clear is that when we think of what we have learned from a particular experimental study, we need to think not only of what we now *know*, but also of what we now know that we *don't* know. Good experimental research is in this regard like good research of other kinds: it raises new questions even as it provides answers to others. But the *kinds* of questions raised by experiments may be novel in some respects.

This is in part because the process of designing an experiment requires a very detailed specification of all aspects of the economic environment in question. In fact, perhaps the most important conclusion that is supported by most experimental studies of economic issues, and one of the first things one notices when setting out to do economic experiments, is that many of the factors that must be considered in setting up an experiment are not spoken of in theoretical papers, nor specified in conventional descriptive accounts of economic institutions. Furthermore, many of these factors have an important influence on the behavior observed.[25] There is nevertheless no easy

25. In this sense, simply designing an experiment is an exercise in model building akin to specifying a game in extensive form. But whereas factors that are left out of a theoretical model can subsequently have no influence on its behavior, factors overlooked in designing an experiment have a way of showing up in unanticipated ways when the experiment is conducted.

way to summarize how economic theory in general fares when tested experimentally, since the results of some experiments falsify the predictions of the theory they are designed to test, while the results of other experiments support them.[26]

As economists further explore the uses of laboratory experimentation in economics, new uses are sure to emerge. For example, it seems that there is considerable potential for using laboratory experiments in conjunction with more traditional kinds of empirical research. Although it is not yet clear what shape this might take, it is reasonable to expect that, in addition to the dialogue discussed here, a fruitful interaction can develop between experimenters and applied economists concerned with a variety of empirical questions.

I am often asked to explain how (or whether) experimental methods can have a bearing on the explanation and understanding of naturally occurring economic phenomena involving large-scale and highly complex systems. Perhaps a useful parallel can be drawn with the experience of evolutionary biology, which, like economics, is a science that deals largely with historical data. Although experiments cannot be conducted on the fossil record, biologists obtain much of their present understanding of selection and evolution from controlled experiments in molecular biology, genetics, and plant breeding. While it is probably impossible to draw precise analogies between economics and other sciences, I think that experimental methods have the potential to play a roughly similar role in economics. They give us a way of learning the answers to certain kinds of questions that we have no other way of answering.

That being the case, it is a promising sign that experimental economics appears to have secured a solid foothold in economic discourse. Not only is it becoming more commonplace to find experiments reported in the major journals of the profession, but a wider circle of economists are now conducting experiments.[27] Looking into the future, which (if any) of the experimental *results* discussed here will remain important will depend both on how future experimental results refine

26. For those of you for whom a little philosophy of science is a dangerous thing, I should probably not say that evidence ever "supports" a theory: please feel free to understand me as saying that the evidence fails to falsify the theory.

27. In this respect, the development of laboratory experimentation in economics may parallel in some ways the growth of game theory as a major tool of economic theory. Not so long ago, game theory was done almost exclusively by "game theorists," whereas today it provides tools that are used as needed by a broad spectrum of economic theorists. Similarly, we are beginning to see experiments conducted not just by a small group of "experimentalists," but by a broader group of economists who turn to experimental methods when they seem appropriate. (In the bargaining literature, for example, in addition to the experiments discussed in Section 2, see Binmore, Shaked, and Sutton, 1985, 1986.)

our understanding of what has been observed to date, and on what kinds of questions remain or become important in light of new theoretical developments. But if I were going to try my hand at prophecy, I would say that there is a good chance that one of the things that will be remembered about science in the latter part of the twentieth century is that laboratory experimentation entered the portfolio of tools that economists use to study the world.

REFERENCES

Alger, Dan. 1984. "Equilibria in the Laboratory: Experiments with Oligopoly Markets where Goods are Made-to-Order." Working paper no. 121, Bureau of Economics, Federal Trade Commission.

Berg, Joyce E., Lane A. Daley, John W. Dickhaut, and John R. O'Brien. Forthcoming. "Controlling Preferences for Lotteries on Units of Experimental Exchange." *Quarterly Journal of Economics.*

Bewley, Truman. Forthcoming. *Advances in Economic Theory,* Fifth World Congress. New York: Cambridge University Press.

Binmore, Ken, Ariel Rubinstein, and Asher Wollinsky. 1985. "The Nash Bargaining Solution in Economic Modelling." Theoretical Economics Discussion Paper Series, London School of Economics.

Binmore, Ken, Avner Shaked, and John Sutton. 1985. "Testing Noncooperative Bargaining Theory: A Preliminary Study." *American Economic Review* 75:1178–80.

———. 1986. "An Outside Option Experiment." Mimeo, London School of Economics.

Chatterjee, Kalyan. 1985. "Disagreement in Bargaining: Models with Incomplete Information." In *Game-Theoretic Models of Bargaining,* edited by A. E. Roth. Cambridge: Cambridge University Press, pp. 9–26.

Edgeworth, F. Y. 1881. *Mathematical Psychics.* London: Kegan Paul.

Ferber, Robert, and Werner Z. Hirsch. 1982. *Social Experimentation and Economic Policy* (Cambridge Surveys of Economic Literature). Cambridge: Cambridge University Press.

Fouraker, Lawrence E., and Sidney Siegel. 1963. *Bargaining Behavior.* New York: McGraw Hill.

Grether, David M., and Charles R. Plott. 1984. "The Effects of Market Practices in Oligopolistic Markets: An Experimental Examination of the Ethyl Case." *Economic Inquiry* 22:479–507.

Hausmann, Jerry A., and David A. Wise (editors). 1985. *Social Experimentation.* National Bureau of Economic Research: University of Chicago Press.

Holt, Charles A., Loren Langan, and Anne P. Villamil. 1984. "Market Power in Oral Double Auctions: Convergence to Competitive Equilibrium Prices Reconsidered." Working paper.

Holt, Charles A., and David T. Scheffman. 1985. "The Effects of Advance Notice and Best-Price Policies: Theory, with Applications to *Ethyl.*" Working paper (October).

Holt, Charles A., and Anne P. Villamil. 1984. "The Effect of Market Power on the Direction of Convergence in Oral Double Auctions." Working paper (December), University of Virginia, University of Minnesota.

Hong, James T., and Charles R. Plott. 1982. "Rate Filing Policies for Inland Water Transportation: An Experimental Approach." *Bell Journal of Economics* 13:1–19.

Kagel, John H. Forthcoming. "Economics According to the Rats (and Pigeons Too): What Have We Learned, and What Can We Hope to Learn." In *Laboratory Experimentation*

in Economics: Six Points of View, edited by Alvin E. Roth. New York: Cambridge University Press.

Kagel, John H., Raymond C. Battalio, and Leonard Green. 1981. "Income-Leisure Tradeoffs of Animal Workers." *American Economic Review* 71:621–32.

Kahneman, Daniel, and Amos Tversky. 1984. "Choices, Values, and Frames." *American Psychologist* 39:341–50.

Kalai, Ehud, and Meir Smorodinsky. 1975. "Other Solutions to Nash's Bargaining Problem." *Econometrica* 43:513–18.

Ketcham, J., V. L. Smith, and A. Williams. 1984. "A Comparison of Posted-Offer and Double-Auction Pricing Institutions." *Review of Economic Studies* 51:595–614.

Kihlstrom, R., A. E. Roth, and D. Schmeidler. 1981. "Risk Aversion and Solutions to Nash's Bargaining Problems." In *Game Theory and Mathematical Economics,* edited by O. Moeschlin and D. Pallaschke. Amsterdam: North-Holland, pp. 65–71.

Knez, Mark, and Vernon L. Smith. Forthcoming. "Hypothetical Evaluations and Preference Reversals in the Context of Asset Trading." In *Laboratory Experimentation in Economics: Six Points of View,* edited by Alvin E. Roth. Cambridge: Cambridge University Press.

Machina, Mark J. 1983. "The Economic Theory of Individual Behavior Toward Risk: Theory, Evidence, and New Directions." Technical Report #433, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Murnighan, J. Keith, Alvin E. Roth, and Françoise Schoumaker. 1985. "Risk Aversion in Bargaining: An Experimental Study." Working paper.

Nash, John. 1950. "The Bargaining Problem." *Econometrica* 28:155–62.

Perles, M. A., and M. Maschler. 1981. "The Super-Additive Solution for the Nash Bargaining Game." *International Journal of Game Theory* 10:163–93.

Plott, Charles R. 1982. "Industrial Organization Theory and Experimental Economics." *Journal of Economic Literature* 20:1485–1527.

———. Forthcoming. "Dimensions of Parallelism: Some Policy Applications of Experimental Methods." In *Laboratory Experimentation in Economics: Six Points of View,* edited by Alvin E. Roth. New York: Cambridge University Press.

———. 1985. "Laboratory Experiments in Economics: The Implications of Posted Price Institutions." Mimeo, California Institute of Technology.

Plott, Charles R., and Vernon L. Smith. 1978. "An Experimental Examination of Two Exchange Institutions." *Review of Economic Studies* 45:133–53.

Roth, Alvin E. 1979. *Axiomatic Models of Bargaining.* Berlin: Springer Verlag.

———. 1983. "Toward a Theory of Bargaining: An Experimental Study in Economics." *Science* 220: 687–91.

———, (editor). 1985a. *Game-Theoretic Models of Bargaining.* New York: Cambridge University Press.

———. 1985b. "A Note on Risk Aversion in a Perfect Equilibrium Model of Bargaining." *Econometrica* 53:207–11.

———. Forthcoming. "Bargaining Phenomena and Bargaining Theory." In *Laboratory Experimentation in Economics: Six Points of View,* edited by Alvin E. Roth. New York: Cambridge University Press.

———, (editor). Forthcoming. *Laboratory Experimentation in Economics: Six Points of View.* New York: Cambridge University Press.

Roth, Alvin E., and Michael W. K. Malouf. 1979. "Game-Theoretic Models and the Role of Information in Bargaining." *Psychological Review* 86:574–94.

Roth, Alvin E., Michale W. K. Malouf, and J. Keith Murnighan. 1981. "Sociological Versus Strategic Factors in Bargaining." *Journal of Economic Behavior and Organization* 2:153–77.

Roth, Alvin E., and J. Keith Murnighan. 1982. "The Role of Information in Bargaining: An Experimental Study." *Econometrica* 50:1123–42.

Roth, Alvin E., and Uriel G. Rothblum. 1982. "Risk Aversion and Nash's Solution for Bargaining Games with Risky Outcomes." *Econometrica* 50:639–47.

Roth, Alvin E., and Françoise Schoumaker. 1983. "Expectations and Reputations in Bargaining: An Experimental Study." *American Economic Review* 73:362–72.

Rubinstein, Ariel. 1982. "Perfect Equilibrium in a Bargaining Model." *Econometrica* 50:97–109.

———. 1985. "Choice of Conjectures in a Bargaining Game with Incomplete Information." In *Game-Theoretic Models of Bargaining*, edited by A. E. Roth, pp. 99–114. New York: Cambridge University Press.

———. Forthcoming. "A Sequential Theory of Bargaining." In *Advances in Economic Theory*, edited by Truman Bewley. New York: Cambridge University Press.

Selten, Reinhard. 1972. "Equal Share Analysis of Characteristic Function Experiments." In *Contributions to Experimental Economics III*, edited by H. Sauermann. Tübingen: J. C. B. Mahr, pp. 130–65.

———. 1982. "Equal Division Payoff Bounds for 3-Person Characteristic Function Experiments." In *Aspiration Levels in Bargaining and Economic Decision Making*, edited by R. Tietz. Berlin: Springer Verlag, pp. 265–75.

———. Forthcoming. "Equity and Coalition Bargaining in Experimental 3-Person Games." In *Laboratory Experimentation in Economics: Six Points of View*, edited by Alvin E. Roth. New York: Cambridge University Press.

Selten, Reinhard, and Wilhelm Krischker. 1982. "Comparison of Two Theories for Characteristic Function Experiments." In *Aspiration Levels in Bargaining and Economic Decision Making*, edited by R. Tietz. Berlin: Springer Verlag, pp. 259–64.

Smith, Vernon L. 1964. "Effect of Market Organization on Competitive Equilibrium." *Quarterly Journal of Economics* 78:181–201.

———. 1976. "Bidding and Auctioning Institutions: Experimental Results." In *Bidding and Auctioning for Procurement and Allocation*, edited by Y. Amihud. New York: New York University Press, pp. 43–64.

———. 1981. "An Empirical Study of Decentralized Institutions of Monopoly Restraint." In *Essays in Contemporary Fields of Economics*, edited by G. Horwich and J. Quirk, W. Lafayette: Purdue University, 83–106.

———. 1982a. "Markets as Economizers of Information: Experimental Examination of the 'Hayek Hypothesis.' " *Economic Inquiry* 20:165–79.

———. 1982b. "Microeconomic Systems as an Experimental Science." *American Economic Review* 72:923–55.

Smith, Vernon L., and Arlington W. Williams. 1981. "The Boundaries of Competitive Price Theory: Convergence, Expectations and Transaction Cost." Paper presented at the Public Choice Society Meetings, New Orleans, March 13–15.

Thaler, Richard. Forthcoming. "The Psychology of Choice and the Assumptions of Economics." In *Laboratory Experimentation in Economics: Six Points of View*, edited by Alvin E. Roth. New York: Cambridge University Press.

Tversky, Amos, and Daniel Kahneman. 1981. 'The Framing of Decisions and the Psychology of Choice." *Science* 211:453–58.

Wilson, Robert B. Forthcoming. "Game Theoretic Analysis of Trading Processes." In *Advances in Economic Theory*, edited by Truman Bewley. Fifth World Congress. New York: Cambridge University Press.