Vernon Smith’s Insomnia and the Dawn of Economics as Experimental Science

Theodore C. Bergstrom*

December, 2002

*Theodore C. Bergstrom is the Aaron and Cherie Raznick Professor of Economics, University of California at Santa Barbara, Santa Barbara, California. Thanks to Mark Isaac and to Gary Charness for helpful discussions. The task of exploring the large body of Smith’s work was eased by the existence of two fine collections of papers by and his coauthors (1991e), (2000a). A full bibliography of Smith’s work can be found at www.gmu.edu/departments/economics/facultybios/smith.html.
1 Introduction

If asked to produce a list of current economists whose work will most influence the development of economics over the next 50 years, I would put Vernon Smith’s name close to the top. As Smith (1987) once remarked, economics has been traditionally considered an “observational” science like astronomy or meteorology, rather than an experimental science like physics or chemistry. The great accomplishment of Smith and his fellow experimentalists has been to convince the economics profession that economics can be an experimental science. A wide range of previously “untestable” propositions in economics become subject to empirical investigation once we realize that controlled laboratory experiments are possible. This has drastic implications for the attitude that we bring to our discipline. Smith (1989) describes this change:

“... 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. Consequently, we come to believe that economic problems can be understood fully by just thinking about them ... But experimentation changes the way you think about economics ... 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.”

A large and growing community of economic researchers now conducts laboratory experiments in economics. The Economic Science Association, an organization of experimental economists that was founded by Smith and a handful of others, currently has about 200 members. In 1995, the Handbook of Experimental Economics (1995), edited by John Kagel and Alvin Roth, appeared with excellent surveys of an impressive array of experimental work. The pace of experimental work has since accelerated and a new Handbook of Experimental Economics Results (2003), edited by Smith and Charles Plott, is about to appear. But perhaps the most significant measure of the impact of experimental economics is the way that experimental results have reshaped the thinking of those who work in game theory, in the theory of consumer choice, and in the applied areas such of public economics, industrial organization, resource economics, labor economics, and finance.

Experimental economists have initiated another new branch of economics that promises to assume great importance. This is the science of experimentally tested
economic design. At least since Jeremy Bentham, (1876), economists have attempted to use economic theory to guide them in proposing institutional arrangements that lead to efficient social outcomes. But only in recent years, have economists thought to use the experimental laboratory as a “wind tunnel” in testing new economic designs.¹

2 Competitive Market Experiments

2.1 Simple Supply-Demand Markets

Smith’s experimental career began with the study of experimental competitive markets. The use of controlled market experiments in economics seems to have been initiated by Edward Chamberlin, who conducted a series of market experiments in his Harvard classroom and reported the results in the *Journal of Political Economy* (1948).

Chamberlin “induced” market demand and supply in his Harvard classroom by distributing cards that assigned each participating student to be either a supplier or a demander. Each supplier was assigned a seller cost at which she could supply a single unit and each demander a buyer value for single unit of the good. In any sale, the seller’s profit is the difference between the price and her seller cost, while the buyer’s profit is the difference between his assigned buyer value and the price. Students were asked to move about the room trying to make the best deal they could with a person of the other type. When a buyer and seller agreed on a price, the transaction was recorded on the blackboard for all to see. Trading continued until no more supplier-demander pairs were willing to make trades.

The assigned distribution of seller costs and buyer values in Chamberlin’s experiment implicitly determine supply and demand curves. If competitive equilibrium theory applied to this environment, then all trades would take place at the same price and the price and quantity observed would be predicted by the point at which the supply and demand curves cross.

Chamberlin argued that his experimental results showed that competitive theory is inadequate to deal with realistic market situations, such as those in his experiments, where individual buyers and sellers don’t know in advance what the equilibrium prices and quantities will be. Chamberlin found that different buyers

¹Al Roth’s game theory web page has a section called “The emerging (consulting) business of economic design” (http://www.economics.harvard.edu/~aroth/aroth.html#design) with many examples of the use of experiments to help devise economic mechanisms for practical problems.
and sellers traded at substantially different prices. He found that in 39 of 46 separate classroom experiments, the number of transactions exceeded the competitive prediction and was never smaller. He also found that usually the observed mean price was lower than the predicted competitive price. Chamberlin concluded that

“... economists may have been led unconsciously to share their unique knowledge of the equilibrium point with their theoretical creatures, the buyers and sellers who, of course, in real life have no knowledge of it whatsoever. ((1948), p 102)

Looking at Chamberlin’s article today, it is remarkable that such a fascinating and provocative line of research could have been ignored by almost the entire economics profession. In fact, it appears that this jewel was almost lost. According to the Social Science Citation Index, Chamberlin’s paper was cited by other authors only four times between its publication in 1948 and its revival in 1962 by Vernon Smith.

Smith recognized the merits of Chamberlin’s experimental method, but brought fresh ideas to the problem of price formation in markets. In so doing, Smith began a series of market experiments and developed a methodology and collection of results that stands today as a cornerstone of a new discipline. Smith describes the genesis of his experimental work in a delightful passage from his essay “Experimental Economics at Purdue,” which can be found in Smith’s Papers in Experimental Economics (1991c). We are treated to a glimpse of the nocturnal churnings of a creative mind as it formed a revolutionary idea.

“Experimental economics started at Purdue in the late fall of 1955 ... I had insomnia one night, and ... I found myself thinking about the classroom experiment that Ed Chamberlin used to perform with the Harvard graduate students to prove the impossibility of perfect competition. I didn’t take Chamberlin’s course, ... but I did observe and participate in Chamberlin’s little ‘experiment’. The scuttlebutt among the Harvard graduate students was that the whole exercise was sort of silly ...

So there I was, wide-awake at 3 am, thinking about Chamberlin’s silly experiment. He gave each buyer a card with a maximum buying price for a single unit, and each seller a card with a minimum selling price for one unit. All of us were instructed just to circulate in the room, engage a buyer (or seller), negotiate a contract, or go out to find another buyer (or seller) and so on. If a buyer and a seller
made a contract, they were to come to Chamberlin, reveal the price of the exchange, turn in their cards, and he would post the price on the blackboard for all to see. When it was all over, he would reveal the implicit demand and supply schedules, and we would learn the important lesson that supply and demand theory was worthless in explaining what had happened; namely that prices were not near the equilibrium and neither was the quantity exchanged.

The thought occurred to me that the idea of doing an experiment was right, but what was wrong was that if you were going to show that competitive equilibrium was not realizable . . . you should choose an institution of exchange that might be more favorable to yielding competitive equilibrium. Then when such an equilibrium failed to be approached, you would have a powerful result. This led to two ideas: (1) why not use the double oral auction procedure, used on the stock and commodity exchanges? (2) why not conduct the experiment in a sequence of trading ‘days’ in which supply and demand were renewed to yield functions that were daily flows?”

Smith’s first published discussion of the results of his classroom experiments appeared in the *Journal of Political Economy* (1962). The market environment for this experiment was similar to that of Chamberlin’s experiments, except that it added the two features that he had concocted in his “insomniacal plan” (1) Whereas Chamberlin used an unstructured “trading pit” design, Smith used a double oral auction scheme in which both buyers and sellers call out bids or offers while an auctioneer recognizes and records transactions resulting from accepted bids and offers. This continues until there are no more acceptable bids or offers. (2) At the conclusion of trading, Smith reopened the market for a new “trading day”. In the new day, everyone has the same buyer value and seller cost as in the previous day. No goods are carried over from one day to the next. The only thing that has changed is that market participants have now observed the outcomes of the previous day’s trading and may adjust their expectations accordingly. Typically, the subjects would iterate through four or five trading days.

Smith shares his thrill of astonishment as he observed the results of his newly-designed experiment.

“The following January, I carried through my insomniacal plan . . . I am still recovering from the shock of the experimental results. The outcome was unbelievably consistent with competitive price theory.
If these results are to believed, what was being knocked down was Chamberlin’s hypothesis of the unattainability of supply and demand theory. But the results can’t be believed, I thought. It must be an accident, so I must take another class and do a new experiment with different supply and demand schedules.”

Over the next several years, Smith performed a great many variations of this experiment. He discovered that the convergence of double oral auction results toward competitive equilibrium is robust to variations in the shape of demand and supply curves, to asymmetries in the distribution of profits between demanders and suppliers, and to various permutations in research design. Smith (1965), Smith (1991a), Smith and Williams (1991a), Smith and Williams (2000). He also learned that convergence to competitive equilibrium occurs with as few as six to eight agents. As Smith points out, these results appear to extend the applicability of supply and demand theory far beyond the frictionless, price-taking environment assumed by conventional economic theory.

Smith’s experimental methods evolved as he pursued his research program, and his methodological innovations have profoundly influenced the way that most experimental economics is done today. In his 1962 experiments, as in Chamberlin’s earlier experiments, the “profits” earned by subjects were only hypothetical and no cash changed hands. In his next published experiment (1964), Smith’s hypothetical profits were backed by actual monetary payoffs. Smith appeared in a Purdue colleague’s classroom (with no advance notice to the students) carrying experimental materials and payoff money. Successful subjects in this experiment earned the equivalent of about $35 in today’s currency. In subsequent work (2000d), Smith investigated the effects of monetary rewards on subject performance. His evidence shows that inexperienced subjects converge toward “rational” behavior more rapidly as the size of rewards are increase. Smith emphasizes the importance of using monetary rewards to achieve “salience;” so that subjects’ objectives coincide with what the investigator thinks they are. Thus he maintains that monetary incentives should be sufficient that subjects will try to maximize the payoffs stated by the experimenter, even where this involves mental effort or tedium.

Through the 1960’s and well into the 1970’s, Smith’s experiments were “low tech”, using ordinary classrooms with no special equipment other than paper and pencil. In 1975, Smith moved to the University of Arizona, where he developed a computer laboratory for experiments. Smith and Arlington Williams designed and programmed implementations of the double auction in which subjects would
interact through computer terminals in a carefully monitored laboratory environment.

2.2 Alternative Market Institutions

Having established striking results in the competitive double-auction environment, Smith went on to investigate the performance of a wide variety of alternative market institutions, operating in markets with similar fundamentals of buyer values and seller costs. A very readable summary of this body of work is found in a recent paper by Smith. (2000f).

In (1991b), Smith and Williams explored some of the many double auction designs that are possible when the computer manages the queueing of bids, offers, and transactions. Of the systems they investigate in the laboratory, the one that works best in terms of price stability and market efficiency maintains a “rank queue” in which offers are accepted and placed in a queue, with the lowest not-yet-accepted offer and highest not-yet-accepted bid automatically entered as the current standing bids and offers. These features are similar to procedures of the New York Stock Exchange which feature a “specialists book” with bid-ask reduction rules. This suggests that evolutionary forces in actual stock markets have led them to market designs that promote efficiency and stability.

In (1964), Smith investigated whether outcomes were sensitive to who could make offers: sellers only, buyers only, or both. When agents on only one side of the market can make offers, Smith found that “silence is golden” in the sense that the contracts tended to favor the side that did not make offers. While this difference from the double auction results is statistically significant, the magnitudes of the price difference is not large and the outcomes from one-sided markets were as efficient as those with double auctions.

A series of papers, Smith (1991a), Smith and Williams (1991b), Smith et al. (1982) compared the outcomes under double oral auctions with such alternative institutions as posted prices and sealed-bid auctions. In (1981), Smith and Williams follow up on earlier experimental work by Charles Plott and Mark Isaac (1981) to study surprising effects of non-binding price controls on the dynamics of competitive markets. In Williams et al. (1986) and Smith (2000c), Smith and his coauthors discuss experiments in which subjects participate in two simultaneous markets where the demand in each market depends on prices in the other. Remarkably, it turns out that after four or five rounds of trading, prices and quantities converge quite closely to the levels predicted by competitive theory.

An alternative to the double auction that is more familiar to most economic
theorists is Leon Walras’ *tâtonnement* mechanism (1954). The difference between these two mechanisms is that with *tâtonnement*, no trades take place in any period until a price is found at which demand and supply for that period are equal, while with the double auction, actual trading takes place continuously between agents who do not know the equilibrium price, and thus may trade at prices that differ from equilibrium.

According to Vernon Smith (2000b),

> Walras’ experience with the operations of the Paris Bourse and his need for a price mechanism that, in principle, could coordinate general equilibrium price adjustments let him to invent the *tâtonnement* mechanism.”

Even today, the performance of the Walrasian mechanism is of considerable practical as well as theoretical interest. Each morning, for example, the New York Stock Exchange determines the opening prices of securities by a method that is essentially the *tâtonnement* mechanism.

Modern theorists have long known that in general, the Walrasian mechanism can produce unstable dynamics when there are multiple markets. But the actual performance of this mechanism in simple functioning markets has not been well understood. The first laboratory experiments investigating the performance of the *tâtonnement* mechanism were conducted by Patrick Joyce, (1984), who found Walrasian *tâtonnement* to perform very well in a stationary market where each buyer and seller could trade at most one unit.

In (2000b), Smith and coauthors compared the performance of the Walrasian mechanism in a more challenging market environment than that considered by Joyce. Instead of a stationary environment, the mechanisms are confronted with supply and demand conditions that shift between periods. Furthermore, individuals are also allowed to engage in multiple transactions. These two changes make it more likely that with *tâtonnement*, traders could manipulate the Walrasian price by shading their orders. Indeed Smith and his coauthors found in their experiments that the *tâtonnement* mechanism consistently performed less efficiency than the continuous double-auction mechanism. Although the prices that emerged with the Walrasian mechanism were close to those predicted by competitive theory, buyers and sellers appeared to engage in strategic withholding of trade, and consequently there were fewer trades than would be efficient.

The study of alternative mechanisms has led Smith to conclude that “institutions matter” in ways that *a priori* economic theory would not have predicted.
This work has had the healthy effect of inducing economists to pay much greater attention to the influence of institutional structure on economic results.

2.3 Bidding and Auctions

In recent years, the study of auctions has become a prominent showplace for demonstrating the power of economics. Economists have found elegant and surprising theoretical results, which have helped them to provide effective assistance to policy-makers in designing auctions suitable to a variety of institutional and technical settings. This enterprise has been greatly enriched by the interaction of laboratory work with economic theory and policy advising.

“The Biggest Auction Ever: The Sale of the British 3G Telecom Licenses”, (2002) by Ken Binmore and Paul Klemperer is an engaging account of the part played by economists in organizing the British government’s sale of spectrum licenses for the third-generation mobile-phone technology. This auction in March and April, 2000, raised about $34 billion, or about 2.5% of British GNP. According to the authors, “not since the Praetorian Guard knocked down the entire Roman Empire to Didius Julianus in AD 195 had there been an auction quite as large.” To design the auction mechanism used for marketing licences, the authors made extensive use of laboratory tests with subjects playing for controlled stakes. These experiments helped to expose potential flaws in their original designs. In addition, the authors found that letting government officials play roles in the laboratory simulations was highly effective in helping the officials to understand the nature of the problem.

These experiences would come as no surprise to Vernon Smith, who in 1991 explained the affinity between auction theory and the laboratory:

... Much of game theory, as with general equilibrium theory, is stillborn, unable to guide meaningful empirical investigation because of its failure to come to grips with exchange institutions and thus with process. But the modelling of auctions is directly predicated upon the allocation and message rules of alternative market institutions. ... In environments in which alternative auction institutions are equivalent, this institution-free property is derived as a theorem instead of an implicit assumption. Auction theory does more than begin with the extensive form of a game, it begins with various extensive forms we observe in the economy... Consequently it is able to guide empirical
testing programs . . . Where theory fails under experiment, which in-
evitably it will do with sufficiently rigorous “boundary” experiments,
it is easier to see which part of the theory has failed and to see where
the theory needs improvement. (1991e) p 509”

Smith and his coworkers have run a large variety of experiments that test the
theoretical propositions of auction theory Coppinger et al. (1991), Cox et al.
the regularities that they discovered are the following: (i) English auctions and
second-bidder sealed-bid auctions, which are theoretically isomorphic in private
values environments, produce very similar results in the experimental laboratory.
(ii) Dutch auctions and first-bidder sealed-bid auctions, which are also theoreti-
cally isomorphic in private goods environments do not produce the same results
in the laboratory. (iii) In laboratory experiments, Nash equilibrium models that
assume identical attitudes toward risk by all participants are rejected. (iv) In lab-
aboratory experiments, as predicted by auction theory, English and second-bidder
auctions result in more efficient outcomes than Dutch and first-bidder sealed bid
auctions, while first-bidder sealed-bid auctions tended to be more efficient than
Dutch auctions.

Klemperer (2002) emphasizes the importance of designing tailor-made auc-
tions to solve the unique institutional problems that arise in particular circum-
stances. He points out several examples where unsuitable “off-the-shelf” have
led to auction “fiascos” and inefficient outcomes. Smith’s earlier papers describe
interesting laboratory tests of auction mechanisms designed for special problems.
Smith, Rassetti, and Bulfin test a mechanism that allows airlines to submit various
contingent bids for flight-compatible combinations of airport landing or takeoff
slots. In “Designing Smart Computer-Assisted Markets: An Experimental Auc-
tion for Gas Networks”, Smith, Kevin McCabe, and Rassetti (1991) study an auc-
tion market for dealing with the complexities of simultaneously pricing natural
gas at each delivery outlet, source, and on all pipelines that connect sources with
delivery points.

2.4 Intertemporal Asset Markets

Working with several collaborators, Smith has used experimental markets to study
intertemporal asset markets. Smith, Plott and Miller (1977) performed an elegant
experiment with a “two-season market.” In each round of this experiment, a good
is demanded and supplied in each of two periods. In the “autarkic” outcome, with no storage, the equilibrium price would be higher in the second period than in the first. However, if the good can be stored at zero cost, it would be worthwhile for some individuals to purchase and store the good either for their own use or for resale. In this environment, the competitive equilibrium has the same price in both periods. Experimental subjects are given no direct information about aggregate demand or supply in either period, but they participate in five rounds of this two-season market, all with the same fundamentals. In the initial round of play, prices are quite close to the autarkic outcome, with a higher price in the second round than in the first, but by the final round, prices are essentially the same in both rounds and both prices and quantities are very close to the levels predicted by competitive theory. The authors emphasize that this convergence to competitive outcomes occurs even with relatively small numbers of traders and even though no individual trader has “perfect knowledge” of market conditions.

Smith and his collaborators have produced a remarkable body of work on laboratory-induced stock market bubbles Smith et al. (1988), Porter and Smith (1995), Knez and Smith (1991). In these experiments, there is typically a series of spot markets for “stocks” that pay random dividends drawn from a known probability distribution for each of a fixed number of periods. Under a wide variety of conditions, they find a robust pattern of outcomes. With inexperienced traders, trading over the lifetime of the stock exhibits dramatic price bubbles, in which stock prices rise far above the fundamental values based on their dividends and then crash back to their fundamental values at some time late in the stock’s lifetime. As subjects become more experienced with trading in this kind of market, the bubbles tend to disappear and spot prices approach those that would be predicted by rational expectations models. In a recent summary of this work, David Porter and Smith (2000) conclude that:

“Financial theorists, and economic theorists in the rational expectations mode will tell you . . . that these bubbles should not happen; that ‘something’ is wrong with the experiments, although all specific ‘somethings’ such as use of student subjects, lack of short selling and margin buying . . . have all been tested, and the predictions were not born out. . . . Psychologists . . . love the bubbles because they see bubbles as violating the rationality of expectations.

Both the theorists and the psychologists, however, are wrong in thinking that rational expectations are falsified by the experiments. . . . A rational expectations equilibrium, if attained, cannot be instan-
there must be a process whereby people go from wherever they start to the ending equilibrium. ... As we view it, the experiments provide the theory with a dynamic learning directive. Rational expectations theory does not define a process whereby agents come to have rational expectations.”

3 Public Goods Experiments

The search for institutions that could result in efficient provision of public goods along with equitable methods of paying for them has a long and interesting history. Knut Wicksell argued in 1896 (1958) that a just and efficient system of public finance should require that any incremental change in public expenditures be voted on simultaneously with a proposed set of taxes to cover its cost, and that such changes should be accepted only if there is nearly unanimous consent among taxpayers. Erik Lindahl in 1919 (1958) proposed a related scheme in which tax rates and the supply of public goods are determined in such a way that given the assigned tax rates, all voters agree on the same quantity of public goods. Although these proposals have intrigued economic theorists, they apparently have not been implemented by real governments.

Modern economists understand that Wicksell’s and Lindahl’s proposals are incomplete in the sense that they do not specify the way in which the necessary information for their implementation would be obtained from self-interested citizens, who are likely to reveal their true preferences only if it is in their interest to do so. Paul Samuelson (1955) expressed strong skepticism that any decentralized market or voting scheme could find or attain Pareto-efficient provision of public goods. Leonid Hurwicz (1972) established that it is in general not possible to find an “incentive-compatible and individually rational” mechanism that implements Pareto optimal allocations of private goods and John Ledyard and J. D. Roberts (1974) showed that this result extends to public goods.

A few economic theorists, however, took up the challenge of finding “satisfactory” if not ideal mechanisms for allocating public goods. Edward Clarke (1971) and Theodore Groves (1973) independently devised a scheme that is now known as the Groves-Clarke mechanism. The Groves-Clarke mechanism is an extension of William Vickrey’s (1961) incentive compatible auction mechanism. Groves and John Ledyard (1977) devised another scheme, now known as the Groves-Ledyard mechanism. These mechanisms are similar in spirit to the Wicksell and Lindahl proposals, but have explicit incentive-compatible rules for information
transmission. A drawback of the Groves-Clarke mechanism is that it applies only if willingness to pay for public goods does not depend on income (the case of quasi-linear utility). Moreover, the Groves-Clarke mechanism does not quite yield a Pareto efficient outcome. The Groves-Ledyard mechanism applies under fairly arbitrary preferences, but unlike the Groves-Clarke mechanism, it does not have a dominant strategy equilibrium. On the other hand, Nash equilibria of the game generated by the Groves-Ledyard mechanism are fully Pareto efficient. Bergstrom, Simon, and Titus (1983) show that although Nash equilibrium for the Groves-Ledyard mechanism is unique in the case of quasi-linear preferences, this is not true in general. In the more realistic case where demand for public goods increases with income, there are generally many distinct Nash equilibria and these differ significantly in the distribution of benefits among participants. When the Groves-Ledyard mechanism generates multiple Nash equilibria that are not Pareto ranked, it remains an open question whether there is a way to coordinate the players on a single Nash equilibrium.

Perhaps more disconcerting than their theoretical drawbacks is the fact that these new mechanisms seem to be rarely if ever observed in the real world. Can it be that these mechanisms have not been adopted because nobody had thought of them until recently and that it takes time for new ideas to be adopted? Or is it that they have serious flaws that have prevented their adaptation and spread?

Vernon Smith and his frequent collaborator and bass-fishing companion, Charles Plott, were among the first to see that this situation cries out for laboratory experiments. Smith spent the year 1973-74 as a research fellow at Cal Tech, where he, Plott, and a lively group of economists and political scientists established the new field of experimental public choice. Smith and his coworkers designed a series of experiments, testing alternative mechanisms for determining the amount of public goods and the way in which they are paid for. Admirable surveys of the large and growing literature on experimental tests of public goods mechanisms have been written by John Ledyard (1995) and more recently by Yan Chen (2002).

In (1977) and (1991b), Smith proposed a mechanism, which he calls The Auction Method, and which “can be interpreted as an implementation of Wicksell’s principle of unanimity.” In (1980), Smith pointed out that his mechanism is fairly close to existing practice in some fund-raising drives, where pledges will be collected from donors only if some target amount of funds is collected. In Smith’s Auction Method, each participant proposes a quantity of the public good to be supplied and also a share of the total cost of the public good that he would pay. Proposed quantities and shares are collected by an auctioneer, who then proposes a quantity of public goods that is the mean of the quantities named by participants.
The auctioneer also proposes that each participant pay the share of total cost that is not covered by the bids of other participants. If the auctioneer’s proposal receives unanimous agreement, it will be implemented. It is easy to see that if the auctioneer’s proposal is accepted, the agreed-on cost shares must add to at least one. If they add to more than one, excess revenue is rebated according to a specified rule. Each participant then gets the implied payoff and the experiment ends. If anyone rejects the auctioneer’s proposal, it is scrapped and the participants try again, with new proposed shares and quantities. The cycle is continued until either there is unanimous agreement to some proposal or until 6 rounds of play are completed and no agreement is reached. In this environment, the reason that individuals agree to pay significant amounts of taxes and request reasonable quantities of public goods is that they don’t want the other players to veto the outcome proposed by the auctioneer. One of the Nash equilibria for the game defined by this experimental is the Lindahl equilibrium, which is Pareto optimal, but because of the finite horizon for bargaining, the mechanism has many other Nash equilibria, not all of which are not efficient.

A series of Vernon Smith’s papers (1977), (1991b), (1991d), (1980) report on experiments with his Auction Method for public goods provision. Smith finds that with quasi-linear utility and with groups of 5-8 participants, quantities of public goods selected are fairly close to efficient, but the distribution of costs is not very close to Lindahl equilibrium. In about 10 percent of the sessions of the auction method, the participants fail to reach unanimous agreement, in which case no public goods are provided. In (1980), Smith tried the Auction Method for Cobb-Douglas rather than quasi-linear preferences. The experimental outcomes differed significantly from Lindahl predictions and did not perform well as measured by Pareto efficiency.

In (1991b) and (1991d), Smith reports on experiments with other mechanisms for public goods experiments, including an implementation of the incentive-compatible Groves-Ledyard scheme and a non-incentive compatible implementation of Lindahl equilibrium. For groups of four, five, and eight members, his Groves-Ledyard mechanism usually resulted in near-optimal provision, while his implementation of the Lindahl mechanism usually resulted in outcomes that were quite far from efficiency. In later work, Harstad and Marrese (1981), found that mechanisms similar to Smith’s implementation of the Groves-Ledyard mechanism frequently failed to converge to efficient outcomes.

In a study funded by the National Aeronautics and Space Administration, Jeffrey Banks, Charles Plott, and David Porter (1988) conducted a series of laboratory tests of the Smith auction mechanism, which was being considered as a pos-
sible device for allocating resources in the development and operation of a space station for the mid 1990’s. Their experiment was conducted with groups of 10 subjects. When agreement was reached with the Smith Auction Method, the outcomes were close to efficient, but unfortunately, agreement was reached in only 14 of 28 trials. Since the amount of public good supplied is zero when agreement is not reached, this suggests that Smith’s mechanism is not very efficient. For comparison, Banks, Plott and Porter also ran laboratory experiments in which public goods were funded by simple direct contributions. With this treatment, contributions fell far short of efficiency and far below the average contributions realized by Smith’s method. Thus, on average, the Smith auction method outperforms a simple direct contribution method, but falls far short of full efficiency.

In (1991e) Smith suggests that, given our current knowledge, economists need to take a somewhat cautious and humble approach to the problem of wholesale economic design.

“Like languages, economic institutions . . . are not the product of one mind or someone’s logical experimental design, but are the product of thousands of minds over many generations of trial-and-error filtering, combined with a societal memory for those arrangements that are in some sense best, or good enough . . . Can we consciously design new and better property-right exchange systems? There is good reason to be skeptical about whether any of us professionals knows or understands enough about the elements of institutional success to allow an affirmative answer to this question. But it is also true that we have made significant progress in the last quarter of a century in our abstract and empirical understanding of incentives in institutions.

Although some interesting experimental work has been done with mechanisms for public goods provision, this body of work is much smaller and its results are much less conclusive than the research on private goods markets. As Chen (2002) maintains, this seems to be an area where much interesting research is yet to be done.
4 Bargaining, Psychology, and Evolution

4.1 Rationality and Context

Most of economic theory is motivated by the study of highly rational agents, able to solve problems of arbitrary complexity. Experimental psychologists and behavioral economists maintain that real people are not nearly so rational. They have accumulated a large stock of experiments in which people make decisions that are inconsistent with rational decision theory. Vernon Smith (2000e) proposes that neither side has the story quite right and that experimental economics offers “a third view, which documents a growing body of evidence that is consistent with rational models, although there are many exceptions.” He argues that much of the tension between the views of the rationalist economists and those of the psychologists’ can be attributed to a fundamental misconception shared by both camps. According to Smith:

“the numerous areas of claimed disagreement (stem from) two unstated premises on which there is implicit agreement between psychology and mainstream theory: (1) rationality in the economy emanates from and derives from the rationality of individual decision-makers in the economy, and (2) individual rationality is a cognitively intensive, calculating process of maximization in the self-interest. A third shared tenant, which is a correlate of points 1 and 2 is that (3) an acceptable and fundamental way to test economic theory is to test directly the economic rationality of individuals isolated from actual experience in social and economic institutions. (2000e)

Smith believes that evidence from experimental economics suggests that economists and psychologists need to abandon or at least revise implicit premises (1) and (2) quoted above, along with the methodological implication (3) that economic theory can be satisfactorily tested by examining subjects’ choices in abstract situations devoid of institutional context. In Smith’s view, “institutions serve as social tools that reinforce, even induce economic rationality.”

Smith’s most compelling exhibit in support of this view is the success of laboratory market experiments. In (2000e), Smith argues that:

“What has emerged from 30 years of experimental research is that preceding premisses 1-3 are false. . . . What these and many hundreds of other experiments have shown is that (1) prices and allocations
converge quickly to the neighborhood of the predicted rational expectations competitive equilibrium, and (2) these results generalize to a wide variety of posted-price, sealed-bid, and other institutions of exchange, although convergence rates tend to vary and can be influenced by extreme parameter distributions.”

Smith points out that convergence to the outcomes predicted by rational behavior occurs despite the fact that subjects have little understanding of the economic processes at work, and in post-experiment interviews often describe the market situation as confused and disorderly.

Smith (2000e) also suggests that many of the behavioral anomalies noted by economists and psychologists in laboratory experiments will either disappear or be much less pronounced when framed in the context of a market. He cites evidence that this is the case for “preference reversals,” differences between “willingness-to-pay,” and “willingness to accept,” confusion about opportunity cost and sunk costs, and unwillingness to accept a small share in ultimatum games.

### 4.2 A Propensity to Barter and Truck?

Smith conjectures that the human ability to operate successfully in market environments may be an evolved capacity, similar to the capability for learning languages. Evolutionary psychologists (see for example Leda Cosmides and John Tooby (1992)) propose that evolution has endowed humans with mental modules for solving social problems. These modules are as much a part of the adapted mind as our ability to hear and see. Among these modules might be an aptitude for rational trading and for maintaining cooperative reciprocal relations. Vernon was not the first Smith to propose such a propensity. In 1776, Adam Smith stated a strikingly similar view.

“This division of labour, from which so many advantages are derived, is not originally the effect of any human wisdom, which foresees and intends that general opulence to which it gives occasion. It is the necessary, though very slow and gradual consequence of a certain propensity in human nature which has in view no such extensive utility; the propensity to truck, barter, and exchange one thing for another. Whether this propensity be one of those original principles in human nature of which no further account can be given; or whether, as seems more probable, it be the necessary consequence of the faculties of reason and speech, it belongs not to our present subject to
inquire. It is common to all men, and to be found in no other race of animals, which seem to know neither this nor any other species of contracts. . . .

Whoever offers to another a bargain of any kind, proposes to do this. Give me that which I want, and you shall have this which you want. . . .” (1937), Wealth of Nations, Book I Chapter II

4.3 Structure-Induced Rationality and Zero-Intelligence Traders

A competing view to the notion that humans have evolved to act rationally in market contexts, is the idea that the structure of markets themselves may produce rational outcomes, regardless of the rationality of the participants. In a paper titled “ Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Rationality” (1993), D.K. Gode and Shyam Sunder suggest that the structure of a competitive market may enforce “rational” outcomes, regardless of the rationality of decision-makers.

Gode and Sunder conducted a simulated double auction with robotic players. These players were assigned buyer values and seller costs exactly as in the experiments of Chamberlin and Smith. Demanders and suppliers made bids and offers randomly drawn from a uniform distribution. However, the program prohibited market participants from making bids and offers that would cause them to lose money. Gode and Sunder discovered that a market made up of “zero-intelligence” traders achieved outcomes that were almost as efficient as the outcomes achieved by human subjects. They also found that took place near the end of any “trading day” usually were at prices close to the competitive equilibrium prices.

Gode and Sunder argue that:

“Allocative efficiency of a double auction derives largely from its structure, independent of traders’ motivation, intelligence, or learning. Adam Smith’s invisible hand . . . can generate aggregate rationality not only from individual rationality but from individual irrationality.”

As Gode and Sunder demonstrate, their robotic market performs efficiently only if buyers and sellers are prohibited from losing money on any trade. The reason that the zero-intelligence traders achieve nearly full efficiency is that given the no-loss constraint, the only way that inefficiency can arise is if a mutually profitable, but socially inefficient deal is struck between two traders. This can happen
if, for example, a supplier with costs higher than the competitive equilibrium price happens to meet a demander with a buyer value higher than his costs and if the random numbers chosen by the two agents happen to permit a mutually profitable transaction. With the demand and supply curves imposed by Gode and Sunder, these inefficient transactions turn out to be quite improbable and hence there are typically very few such transactions.

The fact that a market with zero-intelligence traders, constrained only by the need to make non-negative profits, performs quite efficiently in simple Chamberlin-Smith markets should not be taken as compelling evidence that market structure in general eliminates the need for learning and rationality. For example, in the slightly more complex intertemporal models that Smith studied, there are more possible ways for inefficient outcomes to arise although traders make a profit on every trade. In these environments, markets made up of zero-intelligence traders would not perform very efficiently, yet the experimental evidence shows that after a few rounds of trading, human subjects reach efficient outcomes close to those predicted by competitive theory.

Even in the simple environments where zero-intelligence traders achieve nearly efficient outcomes, the time path of prices differs significantly from that observed among human traders and there is greater variation in prices than is seen with human traders. Moreover, in experiments with human traders, the variance in transaction prices diminishes drastically in later trading days. In contrast, since the robots never learn anything, the variance of transaction prices is as large on later days of trading as on the first day.

### 4.4 Ultimatum and Dictator Games

In *ultimatum game* experiments, subjects are matched anonymously. One agent, the first mover, is asked to propose a division of a fixed sum of money between himself and the other agent. The other, the second mover, can either accept or reject the proposed division. If the proposal is accepted, then each agent gets the agreed amount. If the proposal is rejected, both get zero. A simple implication of rational, selfish behavior in this one-shot game is that the second mover would accept any small amount of money rather than zero and hence the first mover would offer a tiny amount to the second player and keep the rest for himself.

The ultimatum game experiment has been run hundreds of times and in dozens of countries and societies Roth (1995), Henrich et al. (2001). In almost all of these experiments, the first mover’s proposals tend to be more generous than they would be if both players were rational self-interested agents. Moreover, if the first mover
makes an ungenerous offer, the second mover is likely to reject it. These results constitute a serious challenge to rational-choice game theory. If people do not behave as rational, self-interested agents in a laboratory environment as simple as the ultimatum game, how can we expect game theory to guide our thinking about real-world economic interactions? Some have suggested that the results of ultimatum game experiments indicate that consumers have a degree of altruism or perhaps a “taste for fairness” that traditional economists have failed to recognize.

Vernon Smith (2000a) suggests that the problem is one of context. The abstract, one-shot game that subjects are asked to play in one-shot ultimatum games is alien to their ordinary economic experience. In Smith’s words,

“Subjects come to the lab in a social context, a world of repeat interaction in which single transactions are not isolated but part of an ongoing sequence... what is needed... is to reevaluate the experimenter/theorist’s premise that subjects will view such an experiment as a single-trial game... What may be wrong is the very idea that instances of human decision interaction can be construed as without a history or a future. (pp 80-82)

In the ultimatum game experiments, it appears that instead of dispassionately maximizing the money payoffs offered by the experimenter, many subjects apply decision rules borrowed from similar, but more familiar contexts that involve repeated social interaction. In the ultimatum game, a second mover who is offered a very small share of the money to be divided may be likely to feel a visceral urge to punish the first mover’s greed. Where the second mover is offered only a small share, it costs very little to punish the first mover by refusing the offer and sending both players’ payoffs to zero. In common real-world interactions, those who acquire a reputation for willingness to punish exploiters (especially when punishment is cheap) are likely to prosper relative to those who are found to be pushovers.

Smith and other experimental economists have explored this question by varying the context and details of the experiment in several ingenious ways. In Hoffman et al. (2000b) and Hoffman et al. (2000a), Smith and his coauthors investigate whether the results of the ultimatum game are robust to increases in the amount of money at stake and to variations in the way that the game is framed. They found that whether the amount of money to be divided was $100 or $10, about half of the first movers offered an equal split and only rarely did first movers offer less
than 30 per cent of the sum to the second mover. In another treatment, Smith *et al* used scores on a current events quiz to determine who would be first and second movers. They also framed the ultimatum game as a payoff equivalent market situation, where the first mover makes an all-or-nothing offer to sell an object to the buyer. With this treatment, the first movers were significantly less generous, though most offers were higher than the game-theoretic predictions for selfish, rational players.

If the generosity of first movers and the rancorous behavior of ill-treated second movers in ultimatum games stems from a “taste” for equal division of winnings or from an expectation of altruism, then we might expect the effects of these tastes to extend to markets with more than one trader on each side of the market. But the evidence from a large body of experiments suggests that does not happen. In (1991a), Smith and Williams report that in double-auction market experiments, prices closely approached competitive equilibrium after three rounds of trading despite strong asymmetries in the division of profits between buyers and sellers. In (2000), they take asymmetry to the limit by considering markets with four suppliers and four demanders where the competitive price either gives all profits to buyers and none to sellers or *vice versa*. In the initial round of trading, most transactions divide profits approximately equally between buyers and sellers, but by the end of the third round of trading, most prices are close to the competitive equilibrium price with very unequal division of profits and in all subsequent rounds, prices remain close to competitive levels. Another striking illustration of subjects’ willingness to accept unequal division in a market environment comes from one of the experiments in Ted Bergstrom and John Miller’s introductory textbook, *Experiments with Economic Principles* ((2000), Experiment 2). In this experiment, sellers have zero marginal cost and a $10 fixed cost. Some buyers value the good at $25, some at $20 and some at $5, but the demand curve crosses the supply curve at an equilibrium price of zero. The market is conducted in the same way as Chamberlin’s (1948) early experiments. Students move around the classroom seeking the best deal they can and if they reach agreement, they report the price to the instructor who writes it on the blackboard. Trading takes place until no more buyer-seller pairs choose to make deals. In these classroom experiments, nothing seems to deter demanders from “taking advantage of” the sellers. Typically after three or four rounds of trading, almost all sales are made at prices of $1 or less. Buyers are making large profits while the poor sellers, unable to cover their fixed costs, are all losing money.

In a one-shot ultimatum game, there is no question that a second mover will maximize his payoff by accepting any positive share. Therefore if he is rational,
selfish, and truly believes that the game is a one-shot game, he will not refuse a contract that gives him a positive payoff, however small. On the other hand, the best strategy for the first mover depends on his beliefs about how the second mover will respond. Robert Forsythe and his collaborators (1994) devised a simpler game intended to help untangle the motives of first movers. In this game, which is called the “dictator game”, the first mover proposes a division of a fixed sum of money between himself and an anonymous second mover who has no choice but to accept the division. In the dictator game, rational first movers will not be motivated by concern that a low offer will be rejected. Thus positive offers in the dictator game might be seen as direct evidence of altruism or of concern for fairness. Forsythe et al find that first movers are on average much less generous in the dictator game than in the ultimatum game, but a substantial proportion of the “dictators” continue to offer 30-50 percent to the other player.

In Hoffman et al. (2000a), Smith and his coauthors replicate the results of Forsythe et al. They also try a treatment in which the dictator game framed as a market exchange. With this treatment, about 40 percent of the dictators offer zero about 60 percent offer 10 percent or less. In addition they run the dictator game with a double blind treatment, in which the experimenters use a device designed to convince the dictators that not only are they anonymous to the recipients, but also to the experimenters. Remarkably, with this treatment, the generosity of dictators almost disappears. Now more than 60 percent of the subjects offer zero and about 80 percent offer 10 percent or less.

The dictator game experiments with and without double-blind treatment suggest likely sources of the generous play in experimental environments. In the ultimatum game, equal division is the most common offer of first movers. In the dictator game without the double-blind treatment, most offers are less generous, but still only 20 percent offer zero and the modal offer is 30 percent. One might take this as evidence of altruism or a taste for equal division. But the results of the double-blind experiment suggest otherwise. The selfish play under double-blind conditions suggests that many of the subjects in the dictator experiments were not convinced that they were playing a one shot game, but instead were motivated, as Smith suggests, by "a social concern for what others may think and for being held in high regard by others."
5 Conclusion

A recurring theme in Smith’s research is that competitive theory works well to predict outcomes in laboratory markets over a much broader class of environments than those usually assumed by economic theorists. Neither complete information nor a large number of traders is required. While this news is a comfort for those who are fond of familiar tools, it is also a challenge to theorists and applied economists to produce better theories and more sharply observant empirical work.

It seems appropriate to let Vernon Smith have the last word.

“At the heart of economics is a scientific mystery: How is it that the pricing system accomplishes the world’s work without anyone being in charge . . . Smash it in the command economy and it rises as a Phoenix with a thousand heads, as the command system becomes shot through with bribery, favors, barter and underground exchange. . . . No law and no police force can stop it, for the police become as large a part of the problem as of the solution. . . . The pricing system . . . is a scientific mystery as deep as that of the expanding universe or the forces that bind matter. For to understand it is to understand something about how the human species got from hunter-gathering through the agricultural and industrial revolutions to a state of affluence.” (1982)

References


26


