Experimental Methods in Economics (ii). Historically, the method and subject matter of economics have presupposed that it was a non-experimental (or ‘field observational’) science more like astronomy or meteorology than physics or chemistry. Based on general, introspectively ‘plausible’, assumptions about human preferences, and about the cost and technology based supply response of producers, economists have sought to understand the functioning of economies, using observations generated by economic outcomes realized over time. The data of the astronomer is of this same type, but it would be wrong to conclude that astronomy and economics are methodologically equivalent. There are two important differences between astronomy and economics which help to illuminate some of the methodological problems of economics. First, based upon parallelism (the maintained hypothesis that the same physical laws hold everywhere), astronomy draws on all the relevant theory from classical mechanics and particle physics -- theory which has evolved under rigorous laboratory tests. Traditionally, economists have not had an analogous body of tested behavioural principles that have survived controlled experimental tests, and which can be assumed to apply with insignificant error to the microeconomic behaviour that underpins the observable operations of the economy. Analogously, one might have supposed that there would have arisen an important area of common interest between economics and, say, experimental psychology, similar to that between astronomy and physics, but this has only started to develop in recent years.

Second, the data of astronomy are painstakingly gathered by professional observational astronomers for scientific purposes, and these data are taken seriously (if not always non-controversially) by astrophysicists and cosmologists. Most of the data of economics has been collected by government or private agencies for non-scientific purposes. Hence astronomers are directly responsible for the scientific credibility of their data in a way that economists have not been. In economics, when things appear not to turn out as expected the quality of the data is more likely to be questioned than the relevance and quality of the abstract reasoning. Old theories fade away, not from the weight of falsifying evidence that catalyses theoretical creativity into developing better theory, but from lack of interest, as intellectual energy is attracted to the development of new techniques and to the solution of new puzzles that remain untested.

At approximately the mid-20th century, professional economics began to change with the introduction of the laboratory experiment into economic method. In this embryonic research programmer economists (and a psychologist, Sidney Siegel) became directly involved in the design and conduct of experiments to examine propositions implied by economic theories of markets. For the first time this made it possible to introduce *demonstrable* knowledge into the economist’s attempt to understand markets.

This laboratory approach to economics also brought to the economist direct responsibility for an important source of scientific data generated by controlled processes that can be replicated by other experimentalists. This development invited economic theorists to submit to a new discipline, but also brought an important new discipline and new standards of rigour to the data gathering process itself.

An untested theory is simply a hypothesis. As such it is part of our *self*-knowledge. Science seeks to expand our knowledge of *things* by a process of testing this type of self-knowledge. Much of economic theory can be called, appropriately, ‘ecclesiastical theory’; it is accepted (or rejected) on the basis of authority, tradition, or opinion about assumptions, rather than on the basis of having survived a rigorous falsification process that can be replicated.

Interest in the replicability of scientific research stems from a desire to answer the question ‘Do you see what I see?’. Replication and control are the two primary means by which we attempt to reduce the error in our common knowledge of economic processes. However, the question ‘Do you see what I see?’ contains three component questions, recognition of which helps to identify three different senses in which a research study may fail to be replicable:

(1) *Do you observe what I observe?* Since economics has traditionally been confined to the analysis of non-experimental data, the answer to this question has been trivially, ‘yes’. We observe the same thing because we use the same data. This non-replicability of our traditional data sources has helped to motivate some to turn increasingly to experimental methods. We can say that you have replicated my experiments if you are unable to reject the hypothesis that your experimental data came from the same population as mine. This means that the experimenter, his/her subjects, and/or procedures are not significant treatment variables.

(2) *Do you interpret what we observe as I interpret it?* Given that we both observe the same, or replicable data, do we put the same interpretation on these data? The interpretation of observations requires theory (either formal or informal), or at least an empirical interpretation of the theory in the context that generated the data. Theory usually requires empirical interpretation either because (i) the theory is not

developed directly in terms of what can be observed (e.g. the theory may assume risk aversion which is not directly observable), or (ii) the data were not collected for the purpose of testing, or estimating the parameters of a theory. Consequently, failure to replicate may be due to differences in interpretation which result from different meanings being ascribed to the theory. Thus two researchers may apply different transformations to raw field data (e.g. different adjustments for the effect of taxes), so that the results are not replicable because their theory interpretations differ.

(3) *Do you conclude what I conclude from our interpretation?* The conclusions reached in two different research studies may be different even though the data and their interpretation are the same. In economics this is most often due to different model specifications. This problem is inherent in non-experimental methodologies in which, at best, one usually can estimate only the parameters of a prespecified model and cannot credibly test one model or theory against another. An example is the question of whether the Phillips' curve constitutes a behavioural trade-off between the rates of inflation and unemployment, or represents an equilibrium association without causal significance.

I. MARKETS AND MARKET EXPERIMENTS. Markets and how they function constitute the core of any economic system, whether it is highly decentralized -- popularly, a 'capitalistic' system, or highly centralized -- popularly, a 'planned' system. This is true for the decentralized economy because markets are the spontaneous institutions of exchange that use prices to guide resource allocation and human economic action. It is true for the centralized economy because in such economies markets always exist or arise in legal form (private agriculture in Russia) and clandestine or illegal form (barter, bribery, the trading of favours, and underground exchange in Russia, Poland and elsewhere). Markets arise spontaneously in all cultures in response to the human desire for betterment (to 'profit') through exchange. Where the commodity or service is illegal (prostitution, gambling, the sale of liquor under Prohibition or of marijuana, cocaine, etc.) the result is not to prevent exchange, but to raise the risk and therefore the costs of exchange. This is because enforcement is itself costly, and it is never economical for the authorities (whether Soviet or American) even to approximate perfect enforcement. The spontaneity with which markets arise is perhaps no better illustrated than when (1979-80) US airlines for promotional purposes issued travel vouchers to their passengers. One of these vouchers could be redeemed by the bearer as a cash substitute in the purchase of new airline tickets. Consequently vouchers were of value to future passengers. Furthermore, since (as Hayek would say) the 'circumstances of time and place' for the potential redemption of vouchers were different for different individuals, there existed the preconditions for the active voucher market that was soon observed in all busy airports. Current passengers with vouchers who were unlikely to be traveling again soon held an asset worth less to themselves than to others who were more certain of their future or impending travel plans. The resulting market established prices that were discounts from the redemption or 'face' value of vouchers. Sellers who were unlikely to be able to redeem their vouchers preferred to sell them at a discount for cash. Buyers who were reasonably sure of their travel plans could save money by purchasing vouchers at a discount. Thus the welfare of every active buyer and seller increased via this market. Without a market, many -- perhaps most -- vouchers would not have been exercised and would thus have been 'wasted'.

The previous paragraph illustrates a fundamental hypothesis (theorem) of economics: the ('competitive') market process yields welfare improving (and, under certain limiting ideal conditions, welfare maximizing) outcomes. But is the hypothesis 'true', or at least very probably true? (Lakatos (1978) would correctly ask 'Has it led to an empirically progressive research programmed?') I think it is 'true', but how do I know this? Do you see what I see? A Marxist does not see what I see in the above interpretation of a market. The young student studying economics does not see what I see, although if they continue to study economics eventually they (predictably) come to see what I see (or, at least, they say they do). Is this because we have inadvertently brainwashed them? The gasoline consumer does not see what I see. They see themselves in a *zero* sum game with an oil company: any increase in price merely redistributes wealth from the consumer to the company, which is not 'fair' since the company is richer. What I see in a market is a *positive* sum game yielding gains from exchange, which constitutes the fundamental mechanism for creating, not merely redistributing wealth. The traditional method by which the economist gets others to see this 'true' function of markets is by logical arguments (suppose it were not true, then ...), examples, and 'observations', such as are contained in my description of the voucher market, in which what is 'observed' is hortatively described and interpreted in terms of the hypothesis itself. But if this knowledge of the function of markets is 'true', can it be demonstrated? Experimentalists claim that laboratory experiments can provide a uniquely important technique of demonstration for supplementing the theoretical interpretation of field observations.

I conducted my first experiment in the spring of 1956. Since then hundreds of similar, as well as environmentally richer experiments have been conducted by myself and by others. In 1956, my introductory economics class consisted of 22 science and engineering students, and although this might not have been the 'large number' traditionally thought to have been necessary to yield a competitive market, I though it was large enough for a practice run to initiate a research programmer capable of falsifying the standard theory. I conducted the experiment before lecturing on the theory and 'behaviour' of markets in class so as not to 'contaminate' the sample. The 22 subjects were each assigned one card from a well-shuffled deck of 11 white and 11 yellow cards. The white cards identified the sellers, and the yellow cards identified the buyers. Each white card carried a price, known only to that seller, which represented that seller's minimum selling price for one unit, and each yellow card identified a price, known only to that buyer, representing that buyer's maximum buying price for one unit. On the left of Figure 1 is listed these so-called 'limit' prices, identified by buyer, Bl, B2 etc. (in descending order, D) and by seller, Sl, S2 etc. (in ascending order, S). To keep things simple and well controlled each buyer (seller) was informed that he/she was a buyer (seller) of at most one unit of the item in each of several trading periods. Thus demand, D(supply, S) was 'renewed' in each trading period as a steady state flow, with no carry-over in unsatisfied demand (or unsold stock), from one period to the next. In the airline voucher example, imagine the vouchers being issued, followed by trading; the vouchers then expire, new vouchers are issued, traded and so on. In the experiment, suppose real motivation is provided by promising to pay (in cash) to each buyer the difference between that buyer's assigned limit buying price and the price actually paid in each period that a unit is purchased in the market. Thus suppose seller 5 sells their unit to buyer 2 at the price 2.25. Then buyer 2 earns a 'profit' of $0.75 from this exchange. In this way we induce on each buyer a value (or hypothesized willingness-to-pay) equal to the assigned limit buy price. Similarly, suppose each seller is paid the difference between that seller's actual sales price and assigned limit price ('cost', or willingness-to-sell)in each trading period that a unit is sold. Thus in the previous exchange example, seller 5 earns $0.50 from the transaction.

This experimental procedure operationalizes the market preconditions that (1) 'the circumstances of time and place' for each economic agent are dispersed and known only to that agent (as in the above voucher market) and (2) agents have a secure property right in the objects of trade and the private gains ('profits') from trade (an airline travel voucher was transferable and redeemable by any bearer). The reader should note that 'profit' is identified as much with the act of buying as with that of selling. This is because 'profit' is the surplus earned by a buyer who buys for less than his willingness-to-pay, just as a seller's 'profit' is the surplus earned when an item is sold for more than the amount for which they are willing to sell. Willingness-to-sell need not have, and usually does not have anything to do with accounting 'cost', or production 'cost', from which one computes accounting profit. Willingness-to-sell, like willingness-to-buy, is determined by the immediate circumstances of each agent. Hence, a passenger might be prepared to pay the regular full fare premium on a first-class ticket for an emergency trip to visit a sick relative. The accountant's concept of profit cannot be applied to the passenger's decision any more than it can be applied to that of a passenger willing to sell a voucher at a deep discount. In what follows I will use the term 'buyer's surplus' or 'seller's surplus' instead of 'profit' to refer to the gains from exchange enjoyed by buyers or sellers because the term 'profit' is so strongly, exclusively and misleadingly associated with selling activities.

Now let us interpret the previously cited fundamental theorem of economics in the context of the experimental design contained in Figure 1. We note first that the ordered set of seller (buyer) limit prices defines a supply (demand) function (Figure 1). A supply (demand) function provides a list of the total quantities that sellers (buyers) would be willing to sell (buy) at corresponding hypothetical fixed prices. Neither of these functions is capable of being observed, scientifically, in the field. This is because the postulated limit prices are inherently private and not publicly observable. We could poll every potential seller (buyer) of vouchers in Chicago's O'Hare airport on 20 December 1979 to get each person's reported limit price, but we would have no way of validating the 'observations' thus obtained. Referring to Figure 1, we see that in my 1956 experiment, sellers (hypothetically) were just willing to sell three units at price 1.25, nine units at 2.75 and so . on. Similarly buyers (hypothetically) were just willing to buy four units at 2.50, seven units at 1.75 and so on. If seller 3 is indifferent between selling and not selling at 1.25, and if every seller (buyer) is likewise indifferent at his/her limit price, then any particular unit may not be sold (purchased) at this limit price. One means of dealing with this problem in laboratory markets is to promise to pay a small ‘commission’, say 5 cents, to each buyer and seller for each unit bought or sold. Thus seller 3 has a small inducement to sell at 1.25 if he can do no better, and buyer 6 has a small inducement to buy at 2.00 if she can do no better.

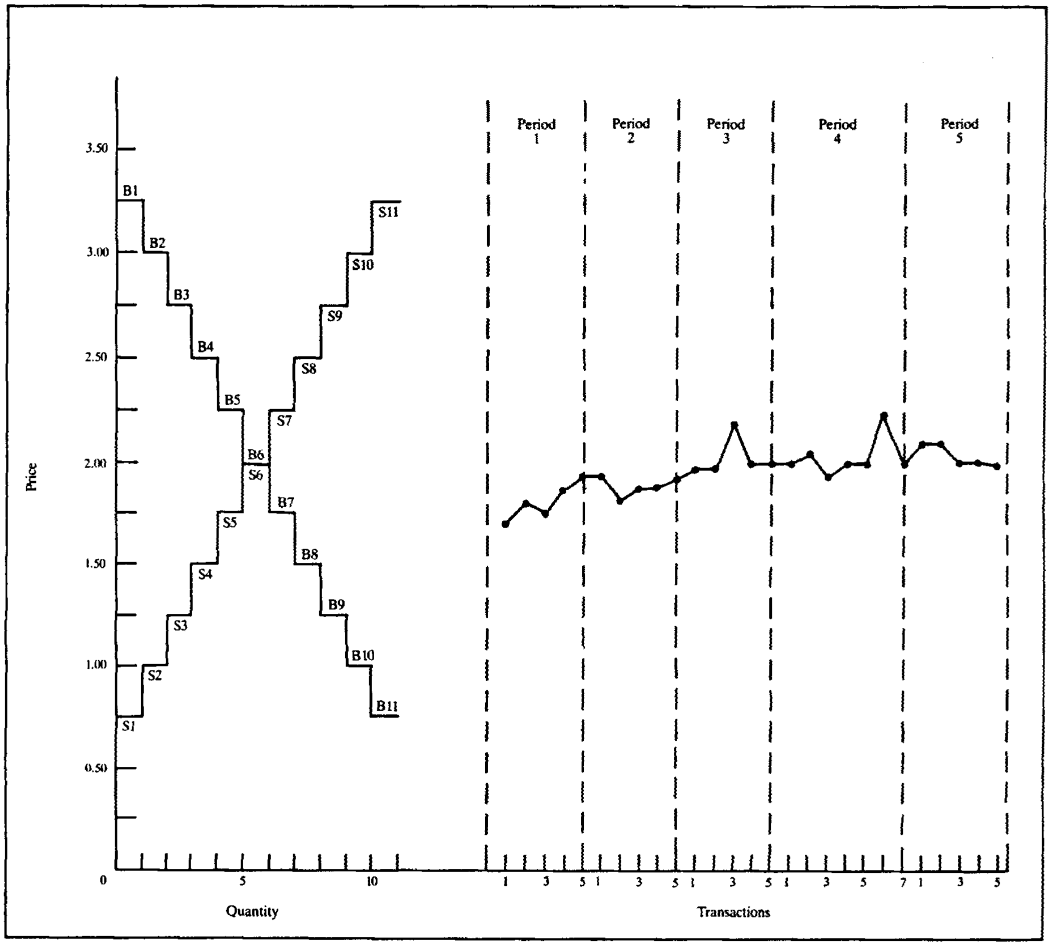

Figure 1

Economic theory defines the competitive equilibrium as the price and corresponding quantity that clears the market; that is, it sets the quantity that sellers are willing to sell equal to the quantity that buyers are willing to buy. This assumes that the subjective cost of transacting is zero; otherwise any units with limit prices equal to the competitive equilibrium price will not exchange. In Figure 1 this competitive equilibrium price is 2.00. If the 5 cent ‘commission’ paid to each trading buyer and seller is sufficient to compensate for any subjective cost of transacting, then buyer 6 and seller 6 will each trade and the competitive equilibrium quantity exchanged will be 6 units. At the competitive equilibrium price, buyer I earns a surplus of 3.25-2.00=1.25 (plus commission) per period and so on. Total surplus, which measures the maximum possible gains from exchange, or maximum wealth *created* by the existence of the market institution, is 7.50 per period, at the competitive equilibrium.

If by some miracle the competitive equilibrium price and exchange quantity were to prevail in this market, sellers 1-6 would sell, buyers 1-6 would buy, while sellers 7-11 would make no sales and buyers 7-11 would make no purchases. It might be thought that this is unfair -- the market should permit some or all of the ‘submarginal’ buyers (sellers) 7-11 to trade -- or that more wealth would be created if there were more than six exchanges. But these interpretations are wrong. By definition, buyer 10 is not willing to pay more than 1.00. Consequently, it is a peculiar notion of fairness to argue that buyer 10 should have as much priority as buyer 1 in obtaining a unit. In the airline voucher example, this would mean that a buyer who is unlikely to redeem a voucher should have the same priority as a buyer who is likely to redeem a voucher. One can imagine a market in which, say, buyer I is paired with seller 9 at price 3.00, buyer 2 with seller 8 at price 2.75, and so on with nine units traded. If this were to occur it would mean buyers 7-9, who are less likely to use vouchers, have purchased them, and sellers 7-9, who initially held vouchers, and were more likely to use them than buyers 7-9, have sold their vouchers. Furthermore, this allocation yields additional possible gains from exchange, and is thus *not sustainable*, even if it were thought to be desirable. That is buyer 9, who bought from seller 1 at price 1.00, could resell the unit to seller 9 (who sold her unit to buyer 2), at price (say) 2.00. Why? Because, by definition a voucher is worth 2.75 to seller 9 and only 1.25 to buyer 9. Similar additional trades can be made by buyers (sellers) 7 and 8. The end result would be that buyers 1-6 and sellers 7-11 would be the terminal holders of vouchers, just as if the competitive equilibrium had been reached initially.

Hence, either the competitive equilibrium prevails, or if inefficient trades occur at dispersed prices, then further 'speculative' gains can be made by some buyers and sellers. If these gains are fully captured the end result is the same allocation as would occur at the competitive equilibrium price and quantity.

Having specified the environment (individual private values) of our experimental market, what remains is to specify an exchange institution. In my 1956 experiment I elected to use trading rules similar to those that characterize trading on the organized stock and commodity exchanges. These markets use the 'double oral auction' procedure. In this institution as soon as the market 'opens' any buyer is free to announce a bid to buy and any seller is free to announce an offer to sell. In the experimental version each bid (offer) is for a single unit. Thus a buyer might say 'buy, 1.00', while a seller might say 'sell, 5.00', and it is understood that the buyer bids 1.00 for a unit and the seller offers to sell one unit for 5.00. Bids and offers are freely announced and can be modified. A contract occurs if any seller accepts the bid of any buyer, or any buyer accepts the offer of any seller. In the simple experimental market, since each participant is a buyer or seller of at most one unit per trading period, the contracting buyer and seller drop out of the market for the remainder of the trading period, but return to the market when a new trading 'day' begins. The experimenter announces the close of each trading period and the opening of the subsequent period, with each trading period timed to extend, say, five minutes. Each contract price is plotted on the right of Figure 1 for the five trading periods of the experiment. This result was not as expected. The conventional view among economists was that a competitive equilibrium was like a frictionless ideal state which could not be conceived as actually occurring, even approximately. It could be conceived of occurring only in the presence of an abstract 'institution' such as a Walrasian *tâtonnement* or an Edgeworth recontracting procedure. It was for teaching, not believing.

From Figure 1 it is evident that in the strict sense the competitive equilibrium was not attained in any period, but the accuracy of the competitive equilibrium theory is easily comparable to that of countless physical processes. Certainly, the data clearly do not support the monopoly, or seller collusion model. The total return to sellers is maximized when four units are sold at price 2.50. Similarly, the monopsony, or buyer collusion model requires four units to exchange at price 1.50.

Since 1956, several hundred experiments using different supply and demand conditions, experienced as well as inexperienced subjects, buyers and sellers with multiple unit trading capacity, a great variation in the numbers of buyers and sellers, and different trading institutions, have established the replicability and robustness of these results. For many years at the University of Arizona and Indiana University we have been using various computerized (the PLATO system) versions of the double 'oral' auction, developed by Arlington Williams, in which participating subjects trade with each other through computer terminals. These experiments establish that the 1956 results are robust with respect to substantial reductions in the number of buyers and sellers. Most such experiments use only four buyers and four sellers, each capable of trading several units. Some have used only two sellers, yet the competitive equilibrium model performs very well under double auction rules. Figure 2 shows the supply and demand design and the market results for a typical experiment in which subjects trade through PLATO computer terminals under computer-monitored double auction rules.

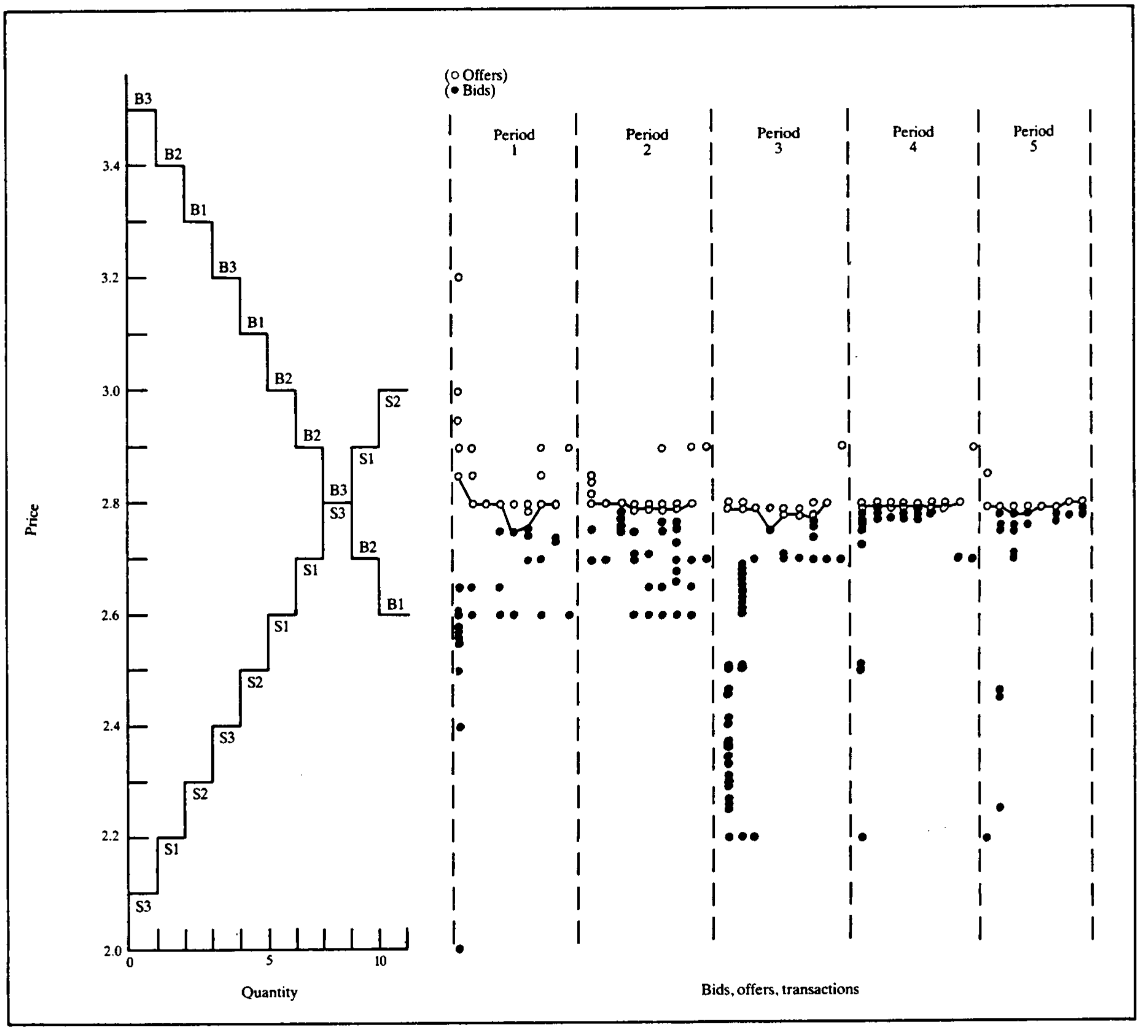

Figure 2

In addition to its antiquarian value, Figure 1 illustrates the problem of monitoring the rules of a ‘manual’ experiment. Observe that in period 4 there were seven contracts which are recorded as occurring in the price range between $1.90 and $2.25. This is not possible since there are only six buyers with limit buy prices above $1.90. Either a buyer violated his budget constraint, or the experimenter erred in recording a price in his first experiment. In Figure 2 there is plotted each contract (an accepted bid if the contract line passes through a ‘dot’; an accepted offer if the line passes through a ‘circle’) and the bids (‘dots’ and offers (‘circles’) that preceded each accepted bid or offer. One of the several advantages of computerized experimental markets is that the complete data of the market (all bids, offers, and contracts at their time of execution) are recorded accurately and non-invasively, and all experimental rules are enforced perfectly. In particular the violation of a budget constraint revealed in Figure 1, which is a perpetual problem with manually executed experiments, is not a problem when trading is perfectly computer monitored.

The rapid convergence shown in Figures 1 and 2 has not always extended to trading institutions other than the double auction. For example, the ‘posted offer’ pricing mechanism (associated with most retail markets), in which sellers post take it or leave it non-negotiable prices at the beginning of each period, yields higher prices and less efficient allocations than the double auction. This difference in performance becomes smaller with experienced subjects and with longer trading sequences in a given experiment (Ketcham et at., 1984). Similarly, a comparison of double auction with a sealed bid-offer auction finds the latter to be less efficient and to deviate more from the competitive equilibrium predictions (Smith et at., 1982). Thus, institutions have been demonstrated to make a difference in what we observe. The data and analysis strongly suggest that institutions make a difference because the rules (legal environment) make a difference, and the rules make a difference because they affect individual incentives.

II. BRIEF INTERPRETIVE HISTORY OF THE DEVELOPMENT OF EXPERIMENTAL ECONOMICS. The two most influential early experimental studies represent the two primary poles of experimental economics: the study of individual preference (choice) under uncertainty (Mosteller and Nogee, 1951) and of market behaviour (Chamberlin, 1948). The investigation of uncertainty and preference has focused on the testing of non Neumann-Morgenstern-Savage subjective expected utility theory. Battalio, Kagel and others have pioneered in the testing of the Slutsky-Hicks commodity demand and labour supply preferences using humans (1973) and animals (1975). A series of large-scale field experiments in the 1970s extended the experimental study of individual preference to the measurement of the effect of the negative income tax and other factors on labour supply and to the measurement of the demand for electricity, housing and medical services.

Since the human species has been observed to participate in market exchange for thousands of years, the experimental study of market behaviour is central to economics. Preferences are not directly observable, but preference theory, as an abstract construct, has been *postulated* by economists to be fundamental to the explanation and understanding of market behaviour. In this sense the experimental study of group market behaviour depends upon the study of individual preference behaviour. But this intellectual history should not obscure the fact that the study of markets and the study of preferences need not be construed as inseparable. Adam Smith clearly viewed the human propensity to truck, barter and exchange' (and *not* the existence of human preferences) as axiomatic to the scientific study of economic behaviour. Obversely, the work of Battalio and Kagel showing that animals behave as if they had Slutsky-Hicks preferences makes it plain that substitution behaviour is an important cross species characteristic, but that such phenomena need not be associated with market exchange.

A significant feature of Chamberlin's (1948) original work is that it concerned the study of behaviorally complete markets; that is all trades, including purchases as well as sales, were executed by active subject agents This feature has continued m the subsequent bilateral bargaining experiments of Siegel and Fouraker (1960) and in market experiments (Smith, 1962, 1982; Williams and Smith, 1984) such as those discussed in section I. This feature was not present in the early and subsequent experimental oligopoly literature (Hoggatt, 1959; Sauermann and Selten, 1959; Shubik, 1962; Friedman, 1963), in which the demand behaviour of buyers was simulated, that !5s programmed from a specified demand function conditional n the prices selected in each 'trading' period by the sellers. This simulation of demand behaviour is justified as an intermediate step in testing models of seller price behaviour that assume passive, simple maximizing, demand-revelation behaviour by buyers. But the conclusions of such experimental studies should not be assumed to be applicable, even provisionally, to any observed complete market without first howled that the experimental results are robust with respect to the substitution of subject buyers for simulated buyers.

III. THE FUNCTIONS OF MARKET EXPERIMENTS IN MICROECONOMIC ANALYSIS. A conceptual framework for clarifying some uses and functions of experiments in macroeconomics can be articulated by suitable modification and adaptation (Smith, 1982) of the concepts underlying the adjustment process, welfare economics literature (see the references to Hurwicz and to Reiter in Smith, 1982). In this literature a microeconomic environment consists of a list of agents {1, . . ., N}, a list of commodities and resources {1, . . . K}, and certain characteristics of each agent *i*, such as the agent's preferences (utility) u_i, technological (knowledge) endowment T^i, and commodity endowment w_i. Thus agent *i* is defined by the triplet of characteristics $E^i = (u^i, T^i, w^i)$ defined on the K-dimensional commodity space. A microeconomic *environment* is defined by the collection $E = (E^1, \ldots E^N)$ of these characteristics. This collection represents a set of primitive circumstances that condition agents' interaction through institutions. The superscript *i*, besides identifying a particular agent, also means that these primitive circumstances are in their *nature* private: it is the individual who likes, works, knows and makes.

There can be no such thing as a credible institution-free economics. Institutions define the property right rules by which agents communicate and exchange or transform commodities within the limits and opportunities inherent in the environment, *E*. Since markets require communication to effect exchange, property rights in messages are as important as property rights in goods and ideas. An institution specifies a language, $M = (M^1, \ldots, M^N)$, consisting of message elements $m = (m^1, \ldots, m^N)$, where M^i is the set of messages that can be sent by agent i (for example, the range of bids that can be sent by a buyer). An institution also defines a set of allocation rules $h = (h^1(m), ..., h^N(m))$ and a set of cost imputation rules $c = (c^1(m), ..., c^N(m))$, where $h^i(m)$ is the commodity allocation to agent i and $c^i(m)$ is the payment to be made by i, each as a function of the messages sent by all agents. Finally, the institution defines a set of adjustment process rules (assumed to be common to all agents), $g(t_0, t, T)$, consisting of a starting rule, $g(t_0, \cdot, \cdot)$, a transition rule, $g(\cdot, t, \cdot)$, governing the sequencing of messages, and a stopping rule, $g(\cdot, \cdot, T)$, which terminates the exchange of messages and triggers the allocation and cost imputation rules. Each agent's property rights in communication and exchange is thus defined by $I^1 = (M^i, h^i(m), c^i(m), g(t_0, t, T))$. A microeconomic *institution* is defined by the collection of these individual property right characteristics, $I = (I^1, ..., I^N)$.

A microeconomic *system* is defined by the conjunction of an environment and an institution, $S = (E, I)$. To illustrate a microeconomic system, consider an auction for a single indivisible object such as a painting or an antique vase. Let each of N agents place an independent, certain, monetary value on the item $v_1, ..., v_N$, with agent i knowing his own value, v_i, but having only uncertain (probability distribution) information on the values of others. Thus $E^i = (v_i; P(v), N)$. If the exchange institution is the 'first price' sealed-bid auction, the rules are that all N bidders each submit a single bid any time between the announcement of the auction offering at t_0, and the dosing of bids, at T. The item is then awarded to the maker of the highest bid at a price equal to the amount bid. Thus, if the agents are numbered in descending order of the bids, the first price auction institution $I_1 = (I_1^1 = [h^1(m) = 1, c^1(m) = b_1]$ and $I_1^i = [h^i(m) = c^i(m) = 0], I > 1$, where $m = (b_1, ..., b_N)$ consists of all bids tendered. That is, the item is awarded to the high bidder, $i = 1$, who pays b_1, and all others receive and pay nothing. This contrasts with the 'second price' sealed-bid auction $I_2 = (I_2^1, ..., I_2^N)$ in which $I_2^1 = [h^1(m) = 1, c^1(m) = b_2]$ and $I_2^2 = [h^i(m) = 0, c^i(m) = 0], i > 1$; that is, the highest bidder receives the allocation but pays a price equal to the second highest bid submitted.

Another example is the English or progressive oral auction, whose rules are discussed under the entry AUCTIONS. It should be noted that the 'double oral' auction, used extensively in stock and commodity trading and in the two experimental markets discussed in section I, is a two-sided generalization of the English auction.

A microeconomic system is activated by the behavioural choices of agents in the set M. In the static, or final outcome, description of an economy, agent *behaviour* can be defined as a function (or correspondence) $m^i = \beta^i(E^i|I)$ carrying the characteristics E^i of agent i into a message m^i, conditional upon the property right specifications of the operant institution I. If all exchange-relevant agent characteristics are included in E^i, then $\beta \equiv \beta^i$ for all i. Given the message-sending behaviour of each agent, $\beta(E|I)$, the institution determines the outcomes

$$h^i(m) = h^i[\beta(E^1|I), ..., \beta(E^N|I)]$$

and

$$c^i(m) = c^i[\beta(E^1|I), ..., \beta(E^N|I)].$$

Within this framework we see that agents do not choose allocations directly; agents choose messages with institutions determining allocations under the rules that carry messages into allocations. (You cannot choose to 'buy' an auctioned item; you can only choose to raise the standing bid at an English auction or submit a particular bid in a sealed bid auction.) However, the allocation and cost imputation rules may have important incentive effects on behaviour, and therefore messages will in general depend on these rules. Hence, market outcomes will result from the conjunction of institutions' and agents' behaviour.

A proper theory of agents' behaviour allows one to deduce a particular β function based on assumptions about the agent's environment and the institution, and his motivation to act. Auction theory is perhaps the only part of economic theory that is fully institution specific. For example, in the second price sealed bid auction it is a dominant strategy for each agent simply to bid his or her value; that is

$$b_i = \beta(E^i|I_2) = \beta(v_i|I_2) = v_i, \quad I = 1, ..., N.$$

The resulting outcome is that $b_1 = v_1$ is the winning bid and agent 1 pays the price v_2. Similarly, in the English auction, agent 1 will eventually exclude agent 2 by raising the standing bid to v_2 (or somewhat above), and obtain the item at this price. In the first price auction Vickrey proved that if each agent maximizes expected surplus (v_i-b_i) in an environment with $P(v) = v$ (the v_i are drawn from a constant density on [0, 1]), then we can deduce the noncooperative equilibrium bid function, $b_i = \beta\ (E^i|I_1) = \beta\ [v_i; P(v), N|I_1] = (N - 1)v_i|N$ (see the entry on AUCTIONS for a more complete discussion).

With the above framework it is possible to explicate the roles of theory and experiment, and their relationship, in a progressive research programme (Lakatos, 1978) of economic analysis. But to do this we must first ask two questions:

(1) 'Which of the elements of a microeconomic system are not observable?' The nonobservable elements are (i) preferences, (ii) knowledge endowments, and (iii) agent message behaviour, $\beta\ (E^i|I)$. Even if messages are available and recorded, we still cannot observe message behaviour functions because we cannot observe, or vary, preferences. The best we can do with field observations of outcomes is to interpret them in terms of models based on assumptions (Cobb-Douglas, constant elasticity of substitution, homothetic), knowledge (complete, incomplete, common), and behaviour (cooperative, noncooperative). Any 'tests' of such models must necessarily be joint tests of all of these unobservable elements. More often the econometric exercise is parameter estimation, which is conditional upon these same elements.

(2) 'What would we like to know?' We would like to know enough about how agents' behaviour is affected by alternative environments and institutions so that we can classify them according to the mapping they provide into outcomes. Do some institutions yield Pareto optimal outcomes, and/or stable prices, and, if so, are the results robust with respect to alternative environments?

These two questions together tell us that what we want to know is inaccessible in natural experiments (field data) because key elements of the equation are unobservable and/or cannot be controlled. If laboratory experiments are to help us learn what we want to know, certain precepts that constitute proposed sufficient conditions for a valid controlled microeconomic experiment must be satisfied:

(1) *Non-satiation* (or monotonicity of reward). Subject agents strictly prefer any increase in the reward medium, π; that is $U_i(\pi_i)$ is monotone increasing for all I.

(2) *Saliency.* Agents have the unqualified right to claim rewards that increase (decrease) in the good (bad) outcomes, x_i, in an experiment; the institution of an experiment renders these rewards salient by defining outcomes in terms of the message choices of agents.

In both the field and the laboratory it is the institution that induces value on messages, given each agent's (subjective) value of commodity outcomes. In the laboratory we use a monetary reward function to induce utility value on the abstract accounting outcomes ('commodities') of an experiment. Thus, agent i is given a concave schedule, $V_i(x_i)$, defining the 'redemption value' in dollars for x_i units purchased in an experimental market, and is assured of receiving a net payment equal to $V_i(x_i)$ less the purchase prices of the x_i units in the market. If the x_i units are all purchased at price p (which is the assumption used to derive a hypothetical demand schedule) the agent is paid $\pi_I = V_i(x_i) - px_i$, with utility $u^i(x_i) = U_i(\pi_i(x_i))$. In defining demand it is assumed that the agent directly chooses x_i (that is $x_i = m_i$). Therefore, if i maximizes $u^i(x_i) = U_i[V_i(x_i) - px_i]$, then at a maximum we have $U'_i \cdot [V'_i(x_i) - p] = 0$ giving the demand function $x_i = V'^{(-1)}_i(p)$ if $U'_i > 0$, where $V'^{(-1)}_i$ is the inverse of i's marginal redemption value of x_i units. (The same procedure for a seller using a cost function $C_j(x_j)$ and paying $px_j - C_j(x_j)$ allows one to induce a marginal cost supply of j.) This illustration generalizes easily: if the joint redemption value is $V_i(x_i, y_i)$ for two abstract commodities (x_i, y_i), $u^i = U_i[V^i(x_i, y_i)]$ induces an indifference map given by the level curves of $V^i(x_i, y_i)$, on (x_i, y_i), with marginal rate of substitution $U'_i\ V^i_x|U'_iV^i_y = V^i_x|V^i_y$, if $U'_i > 0$. If $V^i(x_i, X)$ is the reward function, with x_i, a private and X a common (public) outcome good, we are able to control preferences in the study of public good allocation mechanisms, or if

$$X = \sum_{i=1}^{N} x_i$$

we are poised to study allocation with an 'atmospheric' externality (Coursey and Smith, 1985).

The first two precepts are sufficient to allow us to assert that we have created a microeconomic system $S = (E. I)$ in the laboratory. But to assure that we have created a controlled microeconomy, we need two additional precepts:

(3) *Dominance.* Own rewards dominate any subjective costs of transacting (or other motivation) in the experimental market.

As with any person, subject agents may have variables other than money in their utility functions. In particular, if there is cognitive and kinesthetic (observe the traders on a Stock Exchange floor) disutility associated with the message-transaction process of the institution, then utility might be better written $U_i(\pi_i, m^i)$. To the extent that this is so we induce a smaller demand on i with the payoff $V_i(x_i)$ than was computed above, and we lose control over preferences. As a practical matter experimentalists think the problem can usually be finessed by using rewards that are large relative to the complexity of the task, and by adopting experimental procedures that reduce complexity (e.g. using the computer to record decisions, perform needed calculations, provide perfect recall, etc.). Another approach, as noted in section I, is to pay a small commission for each trade to compensate for the subjective transaction costs.

(4) *Privacy.* The subjects in an experiment each receive information only on his/her own reward schedule.

This precept is used to provide control over interpersonal utilities (payoff externalities). Real people may experience negative or positive utilities from the rewards of others, and to the extent that this occurs we lose control over induced demand, supply and preference functions. Remember that the reward functions have the same role in an experiment that preference functions have in the economy, and the latter preferences are private and non-observable.

If our interest is confined to testing hypotheses from theory, we are done. Precepts (1) (4) are sufficient to provide rigorous tests of the theorist's ability to model individual and market behaviour. But one naturally asks if replicable results from the laboratory are transferable to field environments. This requires,

(5) *Parallelism.* Propositions about behaviour and/or the performance of institutions that have been tested in one microeconomy (laboratory or field) apply also to other microeconomies (laboratory or field) where similar *ceteris paribus* conditions hold.

Astronomy, meteorology, biology and other sciences use the maintained hypothesis that the same physical laws hold everywhere. Economics postulates that when the environment and institution are the same, behaviour will be the same, that is, behaviour is determined by a relatively austere subset of life's parameters. Whether this is 'true' is an empirical question. Hence, when one experimentalist studies variations on the treatment variables of another it is customary to replicate the earlier work to check parallelism. Similarly, one must design field experiments, or devise econometric models using non-experimental field data, that provide tests of the transferability of experimental results to any particular market in the field. Only in this way can questions of parallelism be answered. They are not answered with speculations about alleged differences between the experimental subject's behaviour and (undefined) 'real world' behaviour. The experimental laboratory *is* a real world, with real people, real institutions, real payoffs and commodities just as real as stock certificates and airline travel vouchers, both of which have utility because of the claim rights they legally bestow on the bearer.

IV. CLASSIFYING THE APPLICATION OF EXPERIMENTAL METHODS. There are many types of experiments and many fields of economic study to which experimental methods have been applied.

The experimental study of auctions makes the most extensive use of models of individual behaviour based explicitly on the message requirements of the different institutions. This literature provides test comparisons of predicted behaviour, $m^i = \beta\ (E^i|I)$, with observations on individual choice, $\hat{m}_i = \beta(\hat{E}_i|I)$ for given realizations, $\hat{E}_i$ (such as values, $\hat{v}_i$, where they are assigned at random). The large literature on experimental double auctions makes no such individual comparisons, because the theoretical literature had not yielded tractable models of individual bid-offer behaviour (but recent contributions by Friedman (1984), and Wilson (1984) are providing such models). Here as in most other areas of experimental research the comparisons are between the predicted price-quantity outcomes of static theory (such as competitive, monopoly, and-Cournot models), and observed outcomes. But double auctions have been studied (see references in Smith, 1982) in a variety of environments; for example, the effect of price floors and ceilings have been examined (see references in Plott, 1982). In all cases these studies are making comparisons. In *nomotheoretical* experiments one compares theory and observation, whereas in *nomoempirical* experiments one compares the effect of different institutions and/or environments as a means of documenting replicable empirical 'laws' that may stimulate modelling energy in new directions. The idea that formal theory must precede meaningful observation does not account for most of the historical development of science. *Heuristic* or exploratory experiments that provide empirical probes of new topics and new experimental methods should not be discouraged.

In industrial organization, and antitrust economics, experimental methods have been applied to examine the effects of monopoly, conspiracy, and alleged anticompetitive practices, and to study the concept of natural monopoly and its relation to scale economics, entry cost and the contestable markets hypothesis (see references in Plott, 1982; Smith, 1982; Coursey et at., 1984).

An important development in the experimental study of allocation processes has been the extension of experimental market methods to majority rule (and other) committee processes, and to market-like group processes for the provision of goods which have public or common outcome characteristics (loosely, public goods). These studies have examined public good allocation under majority (and Roberts') rules for committee including the effect of the agenda (see the references to Fiorina and Plott, and Levine and Plott in Smith, 1982), and under compensated unanimity processes suggested by theorists (see the references in Coursey and Smith, 1985). Generally, this literature reports substantial experimental support for the theory of majority rule outcomes, the theory of agenda processes (the sequencing of issues for voting decisions), and for incentive compatible models of the provision of public goods.

VERNON L. SMITH

See also ALLAIS PARADOX; EFFICIENT ALLOCATION OF RESOURCES; PREFERENCE REVERSALS; PSYCHOLOGY AND ECONOMICS.

BIBLIOGRAPHY

Battalio, R., Kagel, J., Winkler, R., Fisher, E., Basmann, R and Krasner, L. 1973. A test of consumer demand theory using observations of individual consumer purchases. *Western Economic Journal* 11(4), December, 411-28

Chamberlin, E. 1948. An experimental imperfect market. *Journal of Political Economy* 56, April, 95-108.

Coursey, D., Isaac, M., Luke, M. and Smith, V. 1984. Market contestability in the presence of sunk (entry) costs. *Rand Journal of Economics* 15(1), Spring, 69-84.

Coursey, D. and Smith, V. 1985. Experimental tests of an allocation mechanism for private, public or externality goods. *Scandinavian Journal of Economies* 86(4), 468-84.

Friedman, D. 1984. On the efficiency of experimental double auction markets. *American Economic Review* 74(1), March, 60-72.

Friedman, J. 1963. Individual behavior in oligopolistic markets: an experimental study. *Yale Economic Essays* 3(2), 359-417.

Hoggatt, A. 1959. An experimental business game. *Behavioral Science* 4(3), July, 192-203.

Kagel, J., Battalio, R., Rachlin, H., Green, L., Basmann, R. and Klemm, W. 1975. Experimental studies of consumer behavior using laboratory animals. *Economic Inquiry* 13(1), March, 22-38.

Ketcham, J., Smith, V. and Williams, A. 1984. A comparison of posted-offer and double-auction pricing institutions. *Review of Economic Studies* 51(4), October, 595-614.

Lakatos, I. 1978. *The Methodology of Scientific Research Programmer*, *Philosophical Papers*, Vol. 1, Ed. J. Worrall and G. Carrie, Cambridge: Cambridge University Press.

Mosteller, F. and Nogee, P. 1951. An experimental measurement of utility. *Journal of Political Economy* 59, October, 371-404.

Plott, C. 1982. Industrial organization theory and experimental economics. *Journal of Economic Literature* 20(4), December, 1485-527.

Sauermann, H. and Selten, R. 1959. Ein Oligopolexperiment. *Zeitschrift für die Gesamte*

Staatswissenschaft 115(3), 427-71.
Shubik, M. 1962. Some experimental non zero sum games with lack of information about the rules. *Management Science* 81(2), January, 215-34.
Siegel, S. and Fouraker, L. 1960. *Bargaining and Group Decision Making*. New York: McGraw-Hill.
Smith, V. 1962. An experimental study of competitive market behavior. *Journal of Political Economy* 70, April, 111-37.
Smith, V. 1982. Macroeconomic systems as experimental science. *American Economic Review* 72(5), December, 923-55.
Smith, V., Williams A., Bratton, K. and Vannoni, M. 1982. Competitive market institutions: double auctions versus sealed bid-offer auctions. *American Economic Review* 72(1), March, 58-77.
Williams, A. and Smith, V. 1984. Cyclical double-auction markets with and without speculators. *Journal of Business* 57(1) Pt 1, January, 1-33.
Wilson, R. 1984. Multilateral exchange. Working Paper No. 7, Stanford University, August.