COMMENTARY:

The Use of Experimental Economics in Strategy Research
(K. Weigelt, C.F. Camerer, and M. Hanna)

Erhard Bruderer and Will Mitchell

EXPERIMENTAL ECONOMICS AS A "USEFUL LINK" IN STRATEGY RESEARCH

In this paper, Weigelt, Camerer, and Hanna describe and review the experimental economic approach to strategy research. A laboratory experiment attempts to eliminate nonessential individual and environmental effects on strategic behavior. By doing so, the researcher gains control over variables and causal relations, and so derives insights on decision making without confusing cause and effect of extraneous influences.

Experimental economics makes two key contributions to strategic research. Most theories of strategic behavior are either simply asserted, as intuitive logic or the results of a game theoretic model, or are tested with flawed data. Both approaches have problems. The assumptions underlying logical and game theoretic assertions are often questionable, so that conclusions may be wrong. On the other hand, the results of empirical analyses using case studies, surveys, or secondary economic data sources are often unreliable or barely significant, because strategic decisions by companies are heavily moderated by the psychology of the individual decision maker, team effects, standard operating

Advances in Strategic Management, Volume 8, pages 201-205.
Copyright © 1992 by JAI Press Inc.
All rights of reproduction in any form reserved.
procedures, and general economic factors that are outside the scope of the predictions being tested. The first contribution of a well-constructed laboratory experiment is that it controls for such extraneous factors, allowing us to focus on a few plausible and carefully defined theoretical influences on strategic actions. The second value of experimental economics is that it is inexpensive, when compared to the time and cost often required to gather concrete data. For both reasons—control and economy—the experimental economic approach may improve theory development and testing, and so further the process of integrating economic and behavioral theories of individual action in strategy.

Experimental economics may provide a means of improving the strategy research process. The standard pure research approach in strategy is divided into two steps: first, theory development; then, theory testing and application. The first major contribution of the strategy field occurs in the theory development stages, where strategy researchers create applied theories by jointly analyzing economic and behavioral influences on firm actions. This joint analysis of influences is what makes strategy different from disciplines such as economics and psychology. The second contribution of strategy occurs in testing and application, in building generalizable understanding of what may appear to be idiosyncratic decisions made within individual firms. The attempt to generalize gives the field its distinction from managerial practice.

Frequently, though, we do not make a successful transition from theory development to theory testing, often because we lack the time, money, or data to test what seem to be sensible predictions. Experimental economics can be seen as a "useful link" in the strategy research process (see Figure 1), with laboratory experiments becoming an intermediate step. Interposing experimental studies in the strategy research process combines the two advantages of control and economy. The value of an experimental intermediate step becomes even more apparent once we remind ourselves that the pure research approach is a figment of a text writer's imagination. In practice, theory development and testing requires many repetitions of development and analysis. Interposing a controlled and inexpensive step within such iterations is a worthy effort. The control and economy may allow theory building efforts that are never truly completed, or leap directly to prescription, to proceed further before the work is abandoned in favor of a newer popular theory.

APPLICATIONS OF EXPERIMENTAL ECONOMICS

Weigelt, Camerer, and Hanna describe three principal applications for economic experiments: strategy concerning markets, games, and decisions. Within market research, for instance, experimental studies of simultaneous competition in many markets offers particular promise. This issue is underexplored because it is difficult to track the complexity, either analytically or empirically, of strategic interactions between many players in many markets.

Most of the applications described by the authors concern relatively stable situations. That is, the studies involve either decisions made at one point in time, or decisions that change over time but in response to a fixed environment. An example of the second type of stable case is that of repeated auctions, in which bid and offer prices tend to move toward equilibrium levels while costs and demand conditions remain unchanged.

Experimental economics is well suited to studying such issues, and may provide answers to questions that are difficult to analyze with game theory or empirical studies. For example, questions like "How do cooperative strategies emerge?" and "How do strategies evolve over time?" have received only limited attention in game theory and empirical analyses. Experiments can add an empirical component to existing game theory and a theoretical component to empirical studies of evolutionary processes.

Yet, many of the most interesting strategic questions involve situations in which strategic behavior and routines evolve in response to changing environments. Strategic adaptation in the face of changing tastes, product or process innovation, competitor entry, and shifting social environments are issues which are critically important to researchers and practitioners. Experimental economics will make its greatest contribution if it can move in this direction.

NEW DIRECTIONS

Let us offer some suggestions for new experimental applications, including quasi-experimental computer simulations, strategic learning, and aggregation processes. Simulations with artificial adaptive agents, which are experiments with abstracted economic actors performed on a computer, provide a possible means of analyzing strategic change (e.g., Arthur, 1991; Holland & Miller, 1991). Simulations are less expensive than most laboratory experiments, because artificial agents do not need to receive even the nominal payment made
to human subjects. More importantly, simulations can model cases in which changes occur frequently and over long periods of time, and so move beyond the scope of most laboratory experiments. An example of this is Axelrod's (1987) experiments with the iterated repeated prisoner's dilemma. Another application would be to make a computer model of multipoint competition in the airline industry. Such simulations may fit into the theory development and testing process of Figure 1 as an intermediate step between theory development and laboratory experiments.

Strategic learning and change offers another possible venue for the experimental methodology. We often believe that we know what would be optimal and rational behavior. But we know little about how managers learn such behavior (or even if they learn what we define as rational behavior) especially in complex environments. Similar to this question is the issue of whether expert systems and other decision support systems help to improve strategy-making (e.g., Davis & Kotteman, 1991). Experimental economics will help us study such questions.

Finally, aggregation phenomena are difficult to study in practice because the phenomena are complex and take place over long periods. We have little understanding of important questions, such as "How do firms form alliances to create standards?" (e.g., Axelrod, Mitchell, Thomas, Bennett, & Bruderer, 1991). Experimental methods can offer means of testing theories of aggregation when concrete data are difficult to obtain.

CONCLUSION

There are limitations to the experimental approach. Much of the virtue of an apple resides in its peel and flesh. If we pare away the peel and the flesh, and study only the seeds in an attempt to understand the genetic structure of past apples and predict how future apples will grow, we may ignore how the apple looks and tastes. Put in less figurative language, laboratory experiments do not duplicate the complexity of decisions in the practical business world and find it difficult to replicate the pressure under which managers have to make these decisions. In part, the parsing out of complexity is a value of the experimental approach, because it paves away the issue to the critical core. But eliminating the context of the decisions may change the decisions themselves, and do so qualitatively.

The organizational and social-level contexts that are often eliminated from laboratory experiments are major determinants of firm strategy and performance, so that experimental economics must be seen as a complement of other research streams rather than a replacement. At best, current experimentation approaches hold such effects constant, leaving their influences in the residual. At worst, the organizational and social-level influences intersect with the experiment, producing biased results.

With its controlled environment, laboratory experimentation is like testing a small version of an airplane in a wind-tunnel. The conclusions from the wind-tunnel experiments must be validated in practical experience. Part of the research task in experimental economics is to show the validity of such comparisons in the strategic world. For example, experiments need to be performed where one test group consist of executives and another of MBA students. Similarly, as the authors note, the validity of comparing people, places, and periods must be tested. In general, principles of comparison need to be developed, such as those applied in laboratory and applied studies of fluid dynamics. Nonetheless, the experimental approach has already provided significant insights to strategic decisions, and promises even greater understanding in the future.

REFERENCES

REFERENCES


THE USE OF EXPERIMENTAL ECONOMICS IN STRATEGY RESEARCH

Keith Weigelt, Colin F. Camerer, and Mark Hanna

ABSTRACT

Experimental methods in economics are well-suited for strategy research. The control of the laboratory allows researchers to decouple confounding influences of decision and environmental variables. We review the basic tenets of the methodology, and illustrate how economic experiments have examined strategic behavior in three different settings: Markets, games, and decisions. In markets we discuss contestability, multiple point competition, and information-based fads (or "mirages"). In games we discuss coordination, reputation-building, and committee decision making. In decisions we discuss compensation contracts and tournaments.

INTRODUCTION

In the last thirty years, economists have begun to study behavior in experiments that are designed to mimic natural economic settings. Our purpose in this paper is to describe the philosophy and design of economics experiments. We will illustrate design issues through a detailed discussion of studies that address several strategy issues, including entry decisions, managerial compensation, coordination of productive activity, and strategic behavior in markets.

Economics experiments might be useful in strategy research for several reasons. In an experiment, control over variables and causal relation is gained at the expense of realism (or "external validity"). Added control could help answer important strategy questions that are difficult to answer using natural data, like how do managers respond to changes in their compensation contracts. Good experimental designs also clarify theory and provide a common language in which research questions can be debated.

Economics experiments also could help bridge the long-standing gap between rational and behavioral approaches to strategy. The rational view sees decision making as a logical, conscious process; the behavioral view is that strategies emerge from complicated social processes, often driven by unconscious motives. Despite obvious areas of common interest, few bridges exist between the two groups. Experiments can test rational theories of strategy formulation and make it easy to observe behavioral processes. In a similar way, experiments have been useful in "behavioral economics," the interface of economics and psychology (Green & Kagel, 1990).

Principles of Experimental Economics

Experimental economists study behavior in tightly controlled laboratory settings in which the decision-making environment closely mimics those found in the field. In contrast to field studies, the experimenter has more control over environmental parameters, which allows her to disengage the confounding properties of parameters and individual differences. This control is essential for testing models since predicted behavior depends on properties of individuals that are difficult to observe. For example, in bargaining models, the predicted actions of individuals depend on properties of utility functions and information states. Experimenteres can control utility functions and information states so predictions about actions can be tested.

Experimental economics began with the bargaining experiments of Siegal and Fouraker (1960) in the late 1950s. They studied how changes in environmental parameters affect behavior in bargaining situations. Vernon Smith (1962) was the first to look carefully at strategic behavior in a market setting (Chamberlin, 1948, took a more casual look). His experiments were designed to answer a simple, important question about competitive markets: Is the market price determined by the intersection of the supply and demand curve? Many models presume that market participants are fully informed of where market demand and supply intersect, but in natural markets people usually know only their own demands or costs of supply. The full-information model was widely accepted as an explanation of how naturally occurring markets worked, but it appeared to be based on a false assumption (full information) and had never been tested.

Smith constructed a simple laboratory goods market that paralleled natural markets. Subjects’ private valuations or costs determined the prices they were willing to buy or sell at. Subjects knew only their own valuations or costs. The difference between a buyer’s private valuation and the price at which she bought the good was her profit. A seller’s profit was the difference between the price she sold the good for and her cost. Exchanges were organized in a "double oral auction"—buyers and sellers announced bids or offers they were willing to buy or sell at. A trade took place when a buyer accepted an offer, or a seller accepted a bid. Smith discovered that the bidding process drove prices toward the competitive equilibrium "as if" subjects knew where the demand and supply curve intersected (though they did not).

Generally, in economics experiments, subjects make choices and send "messages" in an environment controlled by the experimenter (Smith, 1982). Prescribed rules determine how outcomes are derived from messages. The main features of the environment are the commodity being traded (or the choice being made) and its value or cost (or consequences) to subjects. The value of a commodity or decision is typically "induced" by paying subjects money which depends on outcomes.

Smith’s (1962) markets illustrate the general principles of experimental designs in economics. He induced value for the good by paying buyers a prescribed amount of money for each unit of the commodity they bought (e.g., $2.10 per unit). Sellers were charged a cost for each unit they sold (e.g., $1.05). Messages were oral bids to buy units ("I’ll buy for $1.95"), offers to sell units ("Sell for $1.35"), or acceptances of bids and offers. The rules, or "exchange institution," stated that bids had to top the current bid and offers had to undercut the current offer. (When someone accepted a current bid or offer, subsequent bids and offers could begin at any price.) The experimenter acts as a legal authority enforcing contracts between subjects who made bids and other subjects who accepted the bids.

How Economics and Psychology Experiments Differ

Economics experiments like Smith’s differ from psychological experiments in several ways. We will describe a few of the differences in detail.

Formal economic theory circumscribes the design of most economics experiments. Experimental economists strive for "theoretical validity"—(were all the assumptions of the economic theory met?)—as well as "internal validity" (did the treatment cause the effect?) and "external validity" (do conclusions from the experiment generalize to natural settings?). Since economic theory usually assumes rational, self-interested agents and focuses on equilibrium (when people have no incentive to change their behavior), the most important elements of theoretical validity deal with incentives and learning.
Incentives

Smith (1982) argued that a subject's payoffs must be "salient" (connected to her actions) and "dominant" (greater than the subjective costs of thinking and acting) for an experiment to legitimately test economic theory. Therefore, subjects are paid substantial sums according to the outcomes of their decisions. (On average, subjects make a per-hour wage roughly equivalent to the wage they could earn in regular jobs, around $10/hour in 1990.) Paying subjects a fixed sum for participating lacks salience and dominance, so experimental economists avoid fixed payments. Of course, intrinsic motivation provides a psychic payoff which does not connect payoffs to actions, even when subjects are paid a fixed sum or paid nothing at all. Experimental economists do not trust intrinsic payoffs because their power is hard to judge, and may wear off during a long experiment.

Learning

Repeating the entire process of giving subjects goods and money, letting them choose actions, and announcing their outcomes—called "stationary replication"—is usually necessary for convergence to equilibrium. If the theory being tested describes equilibrium behavior, stationary replication with feedback is required to test the theory fairly (i.e., for theoretical validity).

Data and Their Analysis

Economic presumptions largely dictate the data experimental economists gather and the way they analyze it. Economists are kin to "behavioralists" in psychology: The data they take most seriously are behaviors that are connected to payoffs (or "salient"). They have little faith in surveys or polls. They are skeptical about subjects' explanations of their behavior, and reports of intended behavior, so they usually do not gather such data (cf. Knez, Smith, & Williams, 1985). However, many experimenters talk informally with subjects after experiments; the subjects' comments are helpful in checking for internal validity, and suggesting new hypotheses.

Labelling

Experimental economists dislike using realistic labels for elements of the experimental environment, for fear of inducing a nonmonetary utility for labelled actions (which ruins their control over a subject's incentives). Shares of stock will thus be called "certificates", two random states of nature representing good and bad economic conditions will be called "A and B", levels of effort will be called "decision numbers."
MARKETS

In a competitive market, the number of buyers and sellers is large, entry and exit are free, and products are homogeneous. Microeconomic theory states that price will be equal to marginal cost in equilibrium, at which point there is little room for strategic behavior (firms are “price-takers”). In an imperfectly competitive market, there are fewer firms, entry and exit are difficult, and firms are proactive price-setters rather than reactive price-takers. Then strategic behavior matters, and the optimal strategies are not easily derived. Market structure and imperfect competition are therefore the focus of much strategy research. Many economics experiments have examined strategic behavior in imperfect markets (see Plott, 1989 for reviews).

Behavior in Imperfect Markets

Smith and Williams (1990) tested the range and application of economic laws in imperfect markets, with small numbers of buyers and sellers and unequal distributions of gains from exchange. They used a PLATO computerized double-auction (see Williams, 1980). Subjects entered bids to buy or offers to sell by typing a number on a computer keyboard and touching a display screen. Buyers (or sellers) accepted a seller’s (buyer’s) price by touching a box labeled “Accept,” then touching another box labeled “Confirm” within five seconds. The computer then recorded the transaction.

Supply and demand were induced by paying buyers a “resale value” for each unit bought, and charging sellers a “production cost” for each unit sold. Resale values thus represented the theoretical maximum a buyer was willing to pay; production costs represented the theoretical minimum a seller was willing to accept. An induced demand curve was constructed by aggregating the buyers’ resale values from high to low. The supply curve was constructed by aggregating sellers’ costs from low to high. The intersection of the demand and supply curves defined the price and quantity of trade in competitive equilibrium.

Using this setup, Smith and Williams ran a sequence of trading periods (usually fifteen). Each trading period lasted 5-6 minutes. Subjects were undergraduate and graduate students, and were paid $3 plus their earnings from the trading sessions.

Smith and Williams wanted to know how small the number of sellers could be before market behavior differed from the competitive equilibrium. They fixed the number of buyers at four, and reduced the number of sellers to two (duopoly) or one (monopoly). From four experimental duopoly sessions, they concluded that even with only two sellers, market behavior was close to competitive equilibrium. (Contract prices were closer to the competitive price, though not exactly equal to it, than to the monopoly price.) Competitive equilibrium broke down when they reduced the number of sellers from two to one because monopolists could sometimes charge more than the competitive price.

Figure 1 shows the results of a typical monopoly experiment. The supply and demand curves are drawn at left. (The dotted line, $MR$, is the seller’s marginal revenue curve). $P_M$ and $Q_M$ are the monopolist’s profit-maximizing price ($\$6.60$) and quantity (5 units). $C.E.$ is the competitive equilibrium price ($\$6.30$) and quantity (8 units). The points in each trading period represent prices at which trades occurred, arranged in chronological order. In the session shown in Figure 1, prices tend to stick near the monopoly price in most trading periods. By selling about five units (and sometimes only one or two), the monopolist can keep prices above the competitive price of $\$6.30$. However, prices typically fall toward the end of a trading period. The monopolist appears to be using a “price discrimination” strategy, selling to buyers with high reservation prices early in the trading period and then offering “sale prices” to buyers with lower reservation prices later on.”

“Efficiency” is the percentage of total trading gains (producer and consumer surplus) actually earned by subjects. Efficiencies are often below 100% because fewer units are sold than is optimal.
Smith and Williams (1990) concluded that markets are generally competitive with two sellers, and are sometimes competitive with one.

**Contestable Markets**

Contestability is a feature of market structure that captures elements of both perfect and imperfect competition. Demsetz (1968) and others laid the foundations for the theory, and Baumol, Panzar, and Willig formalized it in 1982. The main idea is that the ease of entry and exit, or contestability, of a market determines its competitiveness. Even a monopolist must behave competitively if other firms could enter, compete, and leave without incurring any nonrecoverable costs (other than product costs). For example, suppose Air Monopoly serves a regional market that is too small to support two airlines. If Air Monopoly charges too high a price, entrants can lease planes, undercut the high prices, and reap profits. (If Air Monopoly then cuts its price to retaliate, the entrants can costlessly exit by cancelling their leases.) Therefore, Air Monopoly must charge competitive (marginal cost) prices. In a perfectly contestable market, prices will equal marginal costs regardless of the number of sellers (Sinha, 1986, p. 406).

There have been many articles in economics, public policy, and strategy on contestability. Researchers have applied the theory to specific industries, including airlines (Bailey, 1981; Beesley, 1986; Fawcett & Farris, 1989; Morrison & Winston, 1987; Reed & Waldman, 1988; Sinha, 1986), banking and financial services (Davies & Davies, 1984; Kane & Kidwell, 1984; Martin, 1989; Nelson, 1988), electric utilities (Costello & Hemphill, 1990), and railroads (Meyer & Tye, 1985; Tye, 1985). Some authors are extending contestable market theory with game-theoretic modeling (e.g., Agliardi, 1988, 1990). Despite all the theory on contestability, many empirical questions remain. Does contestability actually discipline monopolists? When are actual markets contestable? How does contestability affect strategic behavior?

The theory of contestable markets states that in markets with infinitely high barriers to entry and strong demand a monopolist should charge monopoly prices. Conversely, when the nonrecoverable cost of entry is zero, and entrants have the same costs, the monopolist should price at the competitive level.

To test these strategic implications of the theory, Coursey, Isaac, and Smith (1984) conducted four single-seller (monopoly) experiments in which the effective cost of entry for a second firm was infinite. They also ran six experiments with two sellers where the cost of entry was zero. Subjects represented firms with decreasing marginal costs (up to their capacity); the capacity of any given firm could satisfy the entire market. The authors used a PLATO posted offer market over eighteen trading periods. Sellers posted public offers and buyers privately chose the seller they wished to buy from. Buyers chose between the two posted prices, thus determining each seller's market share.

### Table 1

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Uncontested monopoly (total number of experiments = 4)</th>
<th>Contested market with entry cost = 0 (total number of experiments = 6)</th>
<th>Contested market with entry cost = $2 (total number of experiments = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monopoly:</td>
<td>( P &gt; \frac{P_e + P_m}{2} )</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Weak contestable:</td>
<td>( P \leq \frac{P_e + P_m}{2} )</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Strong contestable:</td>
<td>( P \leq P_e )</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>


They tested three hypotheses. The monopoly hypothesis predicted that final trading prices would be closer to the monopoly price \( P_m \) than the competitive price \( P_c \). The weak contestability hypothesis said duopoly prices would be closer to \( P_c \). The strong contestability hypothesis said duopoly prices would equal \( P_c \).

Table 1 shows experimental results. All four monopoly experiments supported the monopoly hypothesis (though two of the four did not quite converge to \( P_m \)). In the six duopoly experiments, four of the six sessions converged to \( P_c \), confirming strong contestability. All six sessions satisfied weak contestability. Contestability does discipline monopolists when entry costs are zero. These results are impressive because they occurred with only two sellers. Adding more sellers would presumably increase competition.

Coursey, Isaac, Luke, and Smith (1984) studied contestability when entrants had to pay $2 in sunk costs to enter a market. (The $2 payment corresponds to $2 in fixed costs entrants must pay but cannot recover, like licensing fees or investments in specialized equipment for which there is no market). They ran twelve sessions, which each lasted between 23 and 25 trading periods. All sessions had five buyers and two sellers. Seller A (the incumbent) had to purchase seller permits for periods 1-5 and 6-10 before period 1 began. For the first five periods Seller A was a protected monopolist; Seller B simply
observed market prices. After period five, seller $B$ could contest the market by paying $2 and entering. After period ten, both sellers had to buy permits to be in the market.

The last column of Table 1 shows the results. All sessions satisfied the weak contestability hypothesis and five satisfied the strong contestability. In the seven sessions that did not satisfy strong contestability, either the mere threat of entry kept prices low (but no entry occurred), or sellers cycled between high prices which induced entry and low prices after entry.

Thus, strong contestability appears to be likely when there are no sunk entry costs, and less likely when there are some sunk costs. Weak contestability holds in either case. The data suggest that the sunk costs entrants must incur is a good measure of the height of entry barriers.

These results may be sensitive to changes in experimental design. For example, Harrison (1987) introduced a Stackelberg leader-followers structure into the contestable market experiments—the incumbent chose prices before the entrant did—and found more support for strong contestability.

Milner, Pratt, and Reilly (1990) examined contestable markets in which firms sell a continuous flow of goods or services. Their purpose was to create a laboratory situation closer to the one originally modelled by Baumol, Panzar, and Willig.

The marginal costs of sellers decreased with output. Sellers posted offers to sell flows of output in real time. For example, a seller might offer to supply five units per second for a price of $3.00 (per unit per second). Buyers would then decide whether to buy units at the offered price.

Milner, Pratt, and Reilly conducted 60 experimental sessions. In 35 of the sessions, 2 firms contested the market. In 25 sessions only 1 firm posted offers (natural monopoly). Contestability improved efficiency (compared to efficiency in the uncontested natural monopoly markets), but prices did not converge to a stable competitive level. Instead prices often cycled between low and high levels in the last fifteen minutes of the trading session.

Much work remains. The contestability of natural markets is largely unknown (Schwartz, 1986, p. 52). Furthermore, the connection between contestability and both traditional and game-theoretic models of oligopoly is mostly unexplored (Baumol & Willig, 1986, p. 14). The cycles of low and high prices observed in experiments by Coursey et al. (1984) and Milner et al. (1990) present a theoretical puzzle.

Multiple Point Competition

An interesting area for strategy research is multiple point competition, which occurs when firms compete with each other in two or more markets (e.g., Karnani & Wernerfelt, 1985). For example, major airlines like American and United compete in several city-to-city routes. Bernheim and Whinston (1990) showed that in theory, multiple point competition between two firms will cause some markets to become more competitive (i.e., lower prices), and other markets to become less competitive, relative to single-market competition.

Research in multiple point competition may help answer an important strategy research question about conglomerate firms. Why do some conglomerates have portfolios of unrelated businesses, but no apparent scope economies in managing the different businesses? Multiple point competition suggests an answer since multiple point competition can increase overall firm profits, it creates a kind of strategic scope economy.

Phillips and Mason (in press) studied multiple point competition experimentally. In their experiments, subjects chose quantities to produce. Each pair of subjects represented a selling duopoly. The two subjects in a duopoly picked production levels (quantities) at the same time, and their profits were determined by the two levels. To focus attention on the strategic quantity decision, these experiments collapsed the exchange institution between sellers (the duopolists) and buyers into a simple formula. Subjects read the profits associated with each production level, for each level their duopoly partner might choose, from a chart.

In control groups, each subject pair competed in only one market. Their production levels tended to lie somewhere between two extremes: The Cournot (or Nash equilibrium) point at which each player's production level is the best response to the other's optimal choice; and the collusive level at which joint profits are maximized. In the treatment groups, a pair of subjects competed in two markets simultaneously. Their choices tended to be more collusive in one market and less collusive in the other, precisely as Bernheim and Whinston's theory predicted.

The experiments on contestability and multiple point competition are good examples of how experiments can operationalize and test remarkable theoretical predictions about strategic market behavior that are hard to test with natural data. Tests with natural data would require knowledge of profit margins and historical production levels, and how much of fixed entry costs are sunk. The experiments hardly "prove" the theories, but they suggest the theories should be taken seriously.

Mirages in Markets with Asymmetric Information

Several studies have examined behavior in experimental asset markets where traders have asymmetric information. Corporate investment is analogous to behavior in asset markets because a diversified firm is a type of internal capital market. Managers must choose among investment alternatives (e.g., purchasing equipment, entering new markets, acquiring firms) which yield future rents. Reversing the metaphor, a capital market for equity shares is simply a secondary market for bundles of corporate assets.
Figure 2. Illustrates the transaction number and mean price per experimental period.

- Experiment 1 (Mirror in Period 1):
  - Transaction Number:
    - Period 1: 100
    - Period 2: 50
  - Mean Price (Francs): 300

- Experiment 2 (Mirror in Period 2):
  - Transaction Number:
    - Period 1: 50
    - Period 2: 100
  - Mean Price (Francs): 300

- Experiment 3 (Mirror in Period 3):
  - Transaction Number:
    - Period 1: 100
    - Period 2: 50
  - Mean Price (Francs): 300

- Experiment 4 (Mirror in Period 4):
  - Transaction Number:
    - Period 1: 50
    - Period 2: 100
  - Mean Price (Francs): 300

- Experiment 5 (Mirror in Period 5):
  - Transaction Number:
    - Period 1: 50
    - Period 2: 100
  - Mean Price (Francs): 300

- Experiment 6 (Mirror in Period 6):
  - Transaction Number:
    - Period 1: 100
    - Period 2: 50
  - Mean Price (Francs): 300
Asymmetric information is present in the market for corporate investments because managers may have different amounts of knowledge about markets they compete in. (Presumably, incumbent firms probably know more about their markets than entrants do.) The analogy between financial investment and corporate investment implies that experiments on asset pricing with asymmetric information may suggest interesting experiments on corporate investment.

Camerer and Weigelt (1991) studied experimental markets in which assets lived one period, then paid either a high or a low dividend. Before each trading period a coin was flipped to determine if there would be any insiders in that period, and then all subjects received a “clue card,” saying whether the dividend was high or low (inside information) or saying nothing. If there were no insiders, then six traders got clue cards with the actual dividend value (high or low) and six got blank cards. If there were no insiders, everyone got a blank card.

A trader with a blank card could not tell whether six others knew the dividend or whether everyone had blank cards. Therefore, a few random price changes could trigger an avalanche of mistakes in periods with no insiders. One trader could mistakenly use price movements to infer that others knew the dividend was high, and begin bidding prices up. Others would observe her and conclude that she knew the dividend was high, “learn” from her mistake, and bid prices up further. Once prices rise to the high dividend level subjects are unlikely to question their inference that others knew the dividend was high, so prices will not fall.

Camerer and Weigelt tested whether traders overreacted to uninformative trades, mistakenly inferring information from them. They observed only a few “information mirages” which lasted through an entire trading period. Figure 2 shows them. Each of the four pairs of trading periods comes from a different experimental session. In each pair, the first period shows trading with insiders and the second period shows trading with no insiders. The solid line represents a time series of prices at which trades took place; the dotted line is the rational price.

For example, the top two graphs show periods 5 and 6 in one session. Prices in period 5, in which there were insiders, rose to the rational level (375). Prices in period 6 began at the high level where prices in period 5 had left off, falsely signaling that some insiders knew the dividend was high (because there were no insiders). Prices fell because everyone thought someone else knew the dividend was high.

The mirage experiments suggest a way to study fads or bandwagon effects in corporate behavior. Suppose firms are imperfectly informed about the returns to an investment, and rely on the actions of other firms to signal what the other firms know. Then fads could occur in which several firms enter a new market, creating overcapacity, because they think the others know a lot about the market’s potential growth (cf. Bikhchandani, Hirshleifer, & Welch, 1991). Or firms might neglect a market with high growth potential because they observe other firms neglecting it. The same logic could be applied to fads in adoption of corporate form (e.g., leveraged buyouts), diversification patterns, and managerial practices (e.g., quality circles). We are not arguing that managerial investments or practices are always faddish. But studying the process of price formation in experiments may help us understand how managers form expectations, and whether fads occur.

GAMES

Many economics experiments study decisions made in strategic situations. Most of these experiments test predictions of game theory. (In fact, some game-theoretic predictions are difficult to test in any other way.) A debate is brewing over the usefulness of game-theoretic analysis in strategy research (e.g., Camerer, 1991; Saloner, 1991; cf. Zajac & Bazerman, 1991). Experiments can contribute to this important debate by testing whether people are as rational as the theories assume, or can learn to behave in ways consistent with rationality.

The methodology of game-theoretic experiments and their potential application to strategy questions can be illustrated by coordination games. See Rasmussen (1988), Kreps (1990a), and Tirole (1989) for useful semitechnical introductions to game theory and its application.

Coordination Games

Defined formally, a game occurs when a set of players makes choices, in a specified sequence, from a set of strategies. The strategy choices by each player determine an outcome, which has a utility payoff for each player; a player's payoff therefore depends on the choices by others. (With no such dependence, the situation is a decision rather than a game.) In noncooperative game theory, players are assumed to be unable to make enforceable agreements about what strategies they will choose.

In a coordination game, players prefer to do what other players do. Table 2 shows an example (often called “the battle of the sexes”). The two numbers in each cell show the payoffs to the row player and the column player. Much of game theory tries to predict what players will do, or should do, in a game. There are many alternative predictions (or “equilibrium concepts”); none are universally agreed upon.

The most familiar solution concept is “Nash equilibrium” (an “equilibrium,” from here on). An equilibrium is a set of strategies, one for each player, which
are mutual best responses. For example, in the battle of the sexes the strategies $(T, R)$ form an equilibrium, because $T$ is a best response to $R$ (it yields 1, while $B$ yields 0) and $R$ is a best response to $T$ (it yields 3, while $L$ yields 0).

In the Table 2 game, there are two Nash equilibria: $(T, R)$ and $(B, L)$.

<table>
<thead>
<tr>
<th>Column Player</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>0.0</td>
<td>1.3</td>
</tr>
<tr>
<td>Bottom</td>
<td>3.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The game requires coordination because if players mismatch the equilibria they pick, the result is not an equilibrium and their payoffs are low. For example, suppose the row player chooses $T$ (thinking the $(T, R)$ equilibrium is the right one) and the column player chooses $L$ (thinking the $(B, L)$ equilibrium is the right one). Then they will play $(T, L)$, and both get nothing.

Many strategic decisions within and between firms are coordination problems. Within firms, agents would like to coordinate productive activity. For example, divisions want to receive the right materials promptly from other divisions, salespeople would like to make promises to customers that can be fulfilled by the shipping department, and so forth. Between firms, strategic decisions about pricing and the choice of technology standards can create coordination games.

Schelling (1960) was the first to suggest that certain equilibria in coordination games might be “focal points,” psychologically prominent points that would be chosen by tacit agreement. Consider the game of “matching pennies.” Two players say either “heads” or “tails” at the same time. If they match they both get $1. (Written in a matrix, the game is like the one shown in Table 2, except the payoffs (3,1) and (1,3) are replaced by ($1,$1).) Schelling thought most people would say “heads,” because the unwritten cultural rule which makes heads more prominent than tails causes expectations to converge. Mehta, Starmer, & Sugden (1990) found that 67% of their subjects did choose heads.

Focal points result from a kind of tacit cultural communication, similar to verbal communication between players. Kreps (1990b) suggested corporate culture might usefully be viewed as a set of focal principles which dictate how agents in a corporation should behave in coordination games with other agents (see also Camerer & Vepsalainen, 1988).

There have been several experimental studies of coordination without focal points or communication. Cooper, DeJong, Forsythe, and Ross (1989) studied the game in Table 2. They took a large group of subjects (usually 20) and paired subjects with each other randomly, and anonymously, across a computer network. To test predictions about equilibrium behavior, subjects made a series of choices with a different partner each time. After each choice, they got feedback about their partner’s choice. Players chose the strategy that could give them the high payoff of 3 (i.e., $L$ for the column player, $B$ for the row player) 63% of the time. Their choices matched (i.e., they chose $(T, R)$ or $(B, L)$) 48% of the time, slightly less than if they were choosing randomly. They did not coordinate very well.

Van Huyck, Battalio, and Beil (1990, 1991) studied two interesting coordination games. In their games, each player chooses an “effort level” from 1 to 7. A player’s payoff depends on the group’s effort, and on the deviation between her effort and the group effort. Table 3 shows a payoff matrix.

<table>
<thead>
<tr>
<th>Your</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1.30</td>
<td>1.15</td>
<td>0.90</td>
<td>0.55</td>
<td>0.10</td>
<td>-0.45</td>
<td>-1.10</td>
</tr>
<tr>
<td>6</td>
<td>1.25</td>
<td>1.20</td>
<td>1.05</td>
<td>0.80</td>
<td>0.45</td>
<td>0.00</td>
<td>-0.55</td>
</tr>
<tr>
<td>5</td>
<td>1.10</td>
<td>1.15</td>
<td>1.10</td>
<td>0.95</td>
<td>0.70</td>
<td>0.35</td>
<td>-0.10</td>
</tr>
<tr>
<td>4</td>
<td>0.85</td>
<td>1.00</td>
<td>1.05</td>
<td>1.00</td>
<td>0.85</td>
<td>0.60</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>0.50</td>
<td>0.75</td>
<td>0.90</td>
<td>0.95</td>
<td>0.90</td>
<td>0.75</td>
<td>0.50</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>0.40</td>
<td>0.65</td>
<td>0.80</td>
<td>0.80</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.50</td>
<td>-0.05</td>
<td>0.30</td>
<td>0.55</td>
<td>0.70</td>
<td>0.75</td>
<td>0.70</td>
</tr>
</tbody>
</table>

Van Huyck, Battalio, and Beil (1990, 1991) studied two interesting coordination games. In their games, each player chooses an “effort level” from 1 to 7. A player’s payoff depends on the group’s effort, and on the deviation between her effort and the group effort. Table 3 shows a payoff matrix.

In one version of the game, the group effort is the median of the effort levels chosen by all the players. In that case, the game mimics a social production situation in which the median effort determines total output. Rewards increase if total output is higher, and players are penalized for exerting too little effort or too much effort. In organizations, most production situations are of this type or close to it.

In another version of the game, the group effort is the minimum of all the effort levels. Keeping a secret is an example: The secret is out, its damage done, if any one player tells it. Another example is performance in high-reliability organizations, in which one error is enough to cause catastrophe and threaten survival of the organization (e.g., Roberts, 1990).

Both versions of the game have several Nash equilibria. If players think others will choose effort levels of 1, they will all choose 1 too since 1 is a best response (i.e., in Table 2 choosing row 1 is a best response if the group effort is column 1). If players think others will choose 7, they will choose 7 too. In fact, any effort level is an equilibrium, but players would like to coordinate and choose 7, if they can.
Table 2. A Battle of the Sexes Game

<table>
<thead>
<tr>
<th>Row Player</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>0.0</td>
<td>1.3</td>
</tr>
<tr>
<td>Bottom</td>
<td>3.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

are mutual best responses. For example, in the battle of the sexes the strategies $(T,R)$ form an equilibrium, because $T$ is a best response to $R$ (it yields 1, while $B$ yields 0) and $R$ is a best response to $T$ (it yields 3, while $L$ yields 0).

In the Table 2 game, there are two Nash equilibria: $(T,R)$ and $(B,L).$ The game requires coordination because if players mismatch the equilibria they pick, the result is not an equilibrium and their payoffs are low. For example, suppose the row player chooses $T$ (thinking the $(T,R)$ equilibrium is the right one) and the column player chooses $L$ (thinking the $(B,L)$ equilibrium is the right one). Then they will play $(T,L)$, and both get nothing.

Many strategic decisions within and between firms are coordination problems. Within firms, agents would like to coordinate productive activity. For example, divisions want to receive the right materials promptly from other divisions; salespeople would like to make promises to customers that can be fulfilled by the shipping department, and so forth. Between firms, strategic decisions about pricing and the choice of technology standards can create coordination games.

Schelling (1960) was the first to suggest that certain equilibria in coordination games might be “focal points,” psychologically prominent points that would be chosen by tacit agreement. Consider the game of “matching pennies”: Two players say either “heads” or “tails” at the same time. If they match they both get $1. (Written in a matrix, the game is like the one shown in Table 2, except the payoffs (3,1) and (1,3) are replaced by ($1,$1).) Schelling thought most people would say “heads,” because the unwritten cultural rule which makes heads more prominent than tails causes expectations to converge. Mehta, Starmer, & Sugden (1990) found that 67% of their subjects did choose heads. Focal points result from a kind of tacit cultural communication, similar to verbal communication between players. Kreps (1990b) suggested corporate culture might usefully be viewed as a set of focal principles which dictate how agents in a corporation should behave in coordination games with other agents (see also Camerer & Vepsalainen, 1988).

There have been several experimental studies of coordination without focal points or communication. Cooper, DeJong, Forsythe, and Ross (1989) studied the game in Table 2. They told a large group of subjects (usually 20) and paired subjects with each other randomly, and anonymously, across a computer network. To test predictions about equilibrium behavior, subjects made a series of choices with a different partner each time. After each choice, they got feedback about their partner's choice. Players chose the strategy that would give them the high payoff of 3 (i.e., $L$ for the column player, $B$ for the row player) 63% of the time. Their choices matched (i.e., they chose $(T,R)$ or $(B,L)$ 48% of the time, slightly less than if they were choosing randomly. They did not coordinate very well.

Van Huyck, Battalio, and Beil (1990, 1991) studied two interesting coordination games. In their games, each player chooses an "effort level" from 1 to 7. A player's payoff depends on the group's effort, and on the deviation between her effort and the group effort. Table 3 shows a payoff matrix.

In one version of the game, the group effort is the median of the effort levels chosen by all the players. In that case, the game mimics a social production situation in which the median effort determines total output. Rewards increase if total output is higher, and players are penalized for exerting too little effort or too much effort. In organizations, most production situations are of this type or close to it.

In another version of the game, the group effort is the minimum of all the effort levels. Keeping a secret is an example: The secret is out, its damage done, if any one player tells it. Another example is performance in high-reliability organizations, in which one error is enough to cause catastrophe and threaten survival of the organization (e.g., Roberts, 1990).

Both versions of the game have several Nash equilibria. If players think others will choose effort levels of 1, they will all choose 1 too since 1 is a best response (i.e., in Table 2 choosing row 1 is a best response if the group effort is column 1). If players think others will choose 7, they will choose 7 too. In fact, any effort level is an equilibrium, but players would like to coordinate and choose 7, if they can.
In the experiments, strategic uncertainty about what others will do keeps players from choosing 7 (since choosing 7 when other do not, is a mistake). A median of 4 or 5 usually results in the median-effort game. A minimum of 1 invariably results in the minimum-effort game. In both cases coordination is poor.

One way to enhance coordination is through "cheaptalk," preplay announcements of intended strategies that are not binding (Farrell, 1987). Cheaptalk expands the set of strategies available to players, because announcements of future strategies are themselves strategies. The study of cheaptalk is a first step in bringing communication formally into game theory, and has natural application to bargaining (Farrell & Gibbons, 1989a,b), diplomacy, strategy implementation, and strategizing through product preannouncements.

Cooper, DeJong, Forsythe, and Ross (1990) studied cheaptalk in the battle of the sexes game shown in Table 2. They found that a preplay announcement by one player improved coordination dramatically (95% then chose T,R) (or B,L)). Simultaneous cheaptalk by both players did not help: Players either said nothing or both said they would play the strategy that gave them their highest payoffs. Since those strategies were not best responses to each other, their announcements only caused confusion when they played the game.

Cheaptalk suggests a game-theoretic way to study the role of authority in organizations (with application to strategy implementation, for example). One can imagine a model in which authorities get things done by announcing plans (e.g., profit goals or deadlines). If an agent believes that other agents will believe the announcement and carry out the plan, then he will help carry out the plan, too. Authoritative announcements can work because they coordinate productive activity (without punishments for those who disobey, or because other enforcement mechanisms). Leadership corresponds to making announcements credible (so people believe them, and believe others believe them). The fact that one-way cheaptalk works well in the battle of the sexes experiments, but two-way cheaptalk does not, suggests that a unified voice of authority is important for some kinds of coordination.

Reputation in Games

A firm's reputation is an intangible asset that can generate a competitive advantage, but the reputation-building process itself is unexplained. Recently, game theorists have formally described the reputation-building process (see Kreps & Wilson, 1982; Milgrom & Roberts, 1982). The models assume individuals use, or adapt toward, complex equilibria in determining their actions and beliefs. Testing the descriptive validity of reputation models is important because they have described many strategic behaviors, including: signalling product quality, entry deterrence, auditing, behavior of corporate directors, and corporate culture (see Weigelt & Camerer, 1988 for a review). The requirements for reputation-building behavior are common in corporate settings. Two or more players, information asymmetries, and multiperiod interaction. If reputation models are to provide insights into strategic behavior, there should be some evidence that individuals act as the models predict they will.

Reputation models use a subtle refinement of Nash equilibrium. This idea, called sequential equilibrium, simply requires that players have beliefs which are consistent with their choices. (Nash equilibrium has no such requirement.) Beliefs are important because in reputation games players do not know the privately known characteristics of others. These characteristics are summarized in the player's "type." A person or firm's characteristics, or type, could be anything others care about that is better known internally than externally, including product costs, strategic direction, etc.) A player's reputation is simply the beliefs other players have about her type, based on the history of past actions. Players are assumed to update their beliefs after observing behavior.

Camerer and Weigelt (1988) used a laboratory setting to test the descriptive validity of a reputation model in a repeated game. A mnemonic story will make the discussion easier to follow: An entrepreneur E needs to sequentially borrow from several banks B. In each period, a B must decide whether to make a loan at a fixed interest rate. If the loan is given, the E must decide whether to pay it back. There are two types of E. The dishonest E prefers to renege on loans, the honest E prefers to pay them back. E knows his type, but B only knows the probability of E being honest or not.9

Figure 3 shows the game for one period. Note how E's preference for a strategy depends on her type. Honest type E's (Y-types) received a higher payoff if they paid back the loan P. Dishonest E's (X-types) received a higher payoff from reneging R. Hence, monetary payoffs were used to induce nonmonetary preferences (like those associated with honesty in a natural setting).

Each experiment had eleven subjects who were brought to a room and randomly assigned a role: 8 were B players (bankers) and 3 were E players (entrepreneurs). After publicly reading the instructions, the 3 E players were taken to another room. Administrators in each room communicated via walkietalkies. Subjects did not communicate directly but could hear the conversations between the administrators.7

An experimental sequence consisted of 8 periods in which each B player sequentially decided whether to give a loan to the E player. A random preexperiment drawing determined the order of B players for each sequence. Before each sequence, Camerer and Weigelt used predetermined random numbers to choose which of the 3 E players would play that sequence.

The B player assigned to the first period would choose whether to make a loan. If no loan was made, the E player had no decision to make, so the B player assigned to period 2 would then decide whether or not to make a
The Use of Experimental Economics in Strategy Research

If the first period B player made a loan, then the E player would decide whether to pay it back, or renge. After the decision of the E player was publicly announced in the room with the B players, the B player assigned to period 2 would then announce his choice. This process continued for 8 periods. Then a new sequence would begin, with a new order of B players, and a randomly determined E player.

Before the beginning of each sequence, the administrator in the room with the E players conducted a public drawing from a bingo cage. In most experiments the bingo cage held 3 balls numbered 1, 2, and 3. If the ball drawn was 1 or 2, the E player was an X-type (dishonest) for the entire 8-period sequence. If the ball drawn was a 3, the E player was a Y-type (honest). (The bingo cage thus created probabilities of being an X- or Y-type each period, P(X) = 2/3, P(Y) = 1/3.) The E player knew her type before the sequence began, but B players did not. After the sequence was completed, B players were told the type of the E player for that sequence.

Subjects were MBA students. They were paid the sum of the payoffs from all their decisions, in cash, at the end of the experiment (an average of $18). Each experiment had 75-100 8-period sequences and lasted about 2 1/2 hours.

To solve for a sequential equilibrium one must begin in the last period (i.e., the 8th period), and derive optimal actions conditional on the beliefs held when entering the last period. Then one “unfolds” to the second-to-last period, and derives optimal actions conditional on beliefs, considering the effect of second-to-last period actions on last-period beliefs. (The solution technique, “backward induction,” is similar to dynamic programming and other recursive methods.) This unfolding procedure continues back to the initial period. The sequential equilibrium predicts precisely the frequency with which certain actions will be observed. Table 4 summarizes the theoretical predictions.

Recall that X-type (dishonest) E subjects received a higher payoff from reneging on a loan than from paying it back. The theory predicts the frequencies with which X-types will renge in each period. For example, an X-type E should always pay back loans in periods 1-3. Beginning in period 4, the frequency of reneging should monotonically increase, until finally in period 8, X-type E's should always renge. The intuition is that an E's reputation for honesty (being a Y-type) is only valuable because it makes B's give loans. As period 8 approaches, a reputation for honesty becomes less and less valuable, so dishonest players are more likely to renge, earning a high payoff but destroying their reputation (revealing they are actually X-types).

Data from the last two-thirds of each experimental session were pooled because behavior in these sequences was quite similar. Some results are shown in Table 4. The actual frequency of reneging was generally lower than predicted, although the frequency of reneging did rise almost monotonically in later rounds.

The theory also predicts how often B players should give loans to E players in each period (see Table 4). B players should always give loans in periods...
Table 4. Reputations

Predicted vs. Actual Lending by B Players, Experiments 3-5

<table>
<thead>
<tr>
<th>Period</th>
<th>First Third</th>
<th>Second Third</th>
<th>Last Third</th>
<th>Third Thirds</th>
<th>Actual</th>
<th>Predicted z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>.819 (.244/.298)</td>
<td>.930 (.304/.127)</td>
<td>.959 (.328/.142)</td>
<td>.945 (.612/.669)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>5-8</td>
<td>.834 (.151/.181)</td>
<td>.626 (.877/.139)</td>
<td>.608 (.101/.166)</td>
<td>.616 (.188/.305)</td>
<td>.643</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Predicted vs. Actual Reneging by E Players, Experiments 3-5

<table>
<thead>
<tr>
<th>Period</th>
<th>First Actual</th>
<th>Second Actual</th>
<th>Last Actual</th>
<th>Third Pooled</th>
<th>Actual</th>
<th>Predicted z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>.909 (10/11)</td>
<td>.909 (10/11)</td>
<td>.818 (9/11)</td>
<td>.664 (9/12)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.600 (15/25)</td>
<td>.625 (5/8)</td>
<td>.445 (4/9)</td>
<td>.529 (9/17)</td>
<td>.560</td>
<td>2.66</td>
</tr>
<tr>
<td>6</td>
<td>.167 (6/36)</td>
<td>.120 (8/25)</td>
<td>.240 (6/25)</td>
<td>.280 (14/50)</td>
<td>.415</td>
<td>1.91</td>
</tr>
<tr>
<td>5</td>
<td>.051 (2/11)</td>
<td>.156 (7/45)</td>
<td>.119 (5/42)</td>
<td>.138 (12/87)</td>
<td>.347</td>
<td>4.10</td>
</tr>
<tr>
<td>4</td>
<td>.075 (1/40)</td>
<td>.082 (5/61)</td>
<td>.018 (1/56)</td>
<td>.051 (6/117)</td>
<td>.192</td>
<td>3.92</td>
</tr>
<tr>
<td>3</td>
<td>.089 (4/45)</td>
<td>.000 (0/59)</td>
<td>.013 (2/61)</td>
<td>.017 (2/120)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.007 (4/46)</td>
<td>.048 (3/62)</td>
<td>.016 (1/62)</td>
<td>.012 (4/124)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.222 (10/45)</td>
<td>.581 (5/62)</td>
<td>.015 (1/65)</td>
<td>.047 (6/127)</td>
<td>.000</td>
<td></td>
</tr>
</tbody>
</table>

Previous experimental work on games (see Colman, 1982 for a review) suggests that subjects may have beliefs about the preferences and tastes of others that even a laboratory setting fails to control for. For example, in many repeated prisoner's dilemma experiments, subjects cooperate when the theory predicts they won't. In the reputation experiments, Camerer and Weigelt tested a particular homemade prior theory, that subjects believed a percentage Q X-type E subjects will pay back their loans (like Y-types) even though our monetary rewards should induce them to renege. If Q is substantial, then renege will begin later in the sequence than predicted by the initial assumption that Q = 0.

The homemade prior that best explained observed behavior was approximately Q = .17; i.e., subjects believed that 17% of X-type E subjects would pay back their loans (though they could earn more money by reneging). In new experiments with different parameters, predictions based on the homemade prior of Q = .17 were reasonably accurate. This "home-made" probability of altruism has since been observed by McKelvey and Palfrey (in press), Cooper, DeJong, Forsythe, and Ross (in press), and Neral and Ochs (1989).

Cooperative Game Theory and Committee Decision Making

Many political scientists have studied cooperative game-theoretic approaches to committee decision making. In cooperative game theory, it is assumed that players can agree on (and enforce) a collective choice of strategies which maximizes their total payoff. The interesting question is not what players do, but how they divide the collective gains they create.

Most experiments have studied collective choices of points in a two-dimensional spatial representation (see Ordeshook, 1986). Figure 4 gives an example. The two dimensions represent two features, attributes, or issues (in a political context). Each player is assumed to have a clear preference about which point in the two-dimensional space is best. The point a player likes best is called her "bliss point." For example, in Figure 4 player 3's bliss point is point Xi. It is usually assumed that a player's utility decreases as the distance between a point and the bliss point increases.

In committee experiments of this sort, players propose various points for a group to adopt. Given certain rules, cooperative game theories make predictions about which points will be chosen by the group. The crucial treatment variables are the size of the group and the rules by which proposals can be made and adopted. Most experiments use five subjects, and require that a majority (three of five) approves the final decision.

The committee decision-making paradigm could be applied to any setting in which strategic decisions are made collectively, such as decisions by corporation boards, finance committees, or R&D teams.
The Use of Experimental Economics in Strategy Research

COMPENSATION AND AGENCY

Firms devote considerable resources to creating internal wage structures and assigning managers to hierarchical positions. Despite this financial commitment, there are many aspects of managerial compensation that are not understood (see Baker, Jensen, & Murphy, 1988). Much of the work is normative theory that derives optimal contracts conditional on managers behaving in specified ways. Experiments in three compensation-related areas—tournaments, agency theory, and compensation contracts—have provided empirical evidence on the behavior of subjects. They have helped determine the descriptive validity of the theories, and suggested some responses not predicted by the theories.

Theory of Tournaments

In a tournament, an agent’s compensation depends on her performance relative to the performance of other agents. A series of papers have derived theoretical properties of tournaments (Green & Stokey, 1983; Lazear & Rosen, 1981; Nalebuff & Stiglitz, 1983; O’Keefe, Viscusi, & Zeckhauser, 1984). Tournament schemes are used in a variety of organizations: The best minor league baseball players play in the majors; the most prolific assistant professors receive tenure; and vice presidents compete for the position of CEO. A recent study showed that corporate pay levels are consistent with the idea that promotions are determined by sequential tournaments: As the organizational pyramid narrows, the top-ranked agents are promoted and become eligible for a smaller set of higher-level promotions. In this pyramid structure, tournament theory implies that the increase in pay will be larger at each promotion, and it is (Lambert, Larcker, & Weigelt, 1992).

Tournaments can be either symmetric or asymmetric. Agents in symmetric tournaments are equal in ability and are treated equally by the rules. In asymmetric tournaments, agents differ in ability (tournaments are “uneven”) or agents are equal but the rules favor some agents (tournaments are “unfair”).

Bull, Schotter, and Weigelt (1987), Weigelt, Dukerich, and Schotter (1989), Dukerich, Weigelt, and Schotter (1990) and Schotter and Weigelt (in press) used laboratory experiments to test behavioral implications of the theory of rank-order tournaments. We will sketch the theory first, then describe the experiments.

In rank-order tournaments an agent’s compensation depends only on the rank of her performance relative to other agents. The size of performance differences does not matter. A tournament is a game, because the optimal strategy of an agent (a level of effort) is a function of the actions of others. Agents must form beliefs about the actions others will take, then choose strategies based on these beliefs.
Consider a symmetric tournament with two agents. Denote the reward to the highest-producing agent by $M$ and the reward to the lowest-producing by $m$. Suppose productivity depends on a random component, and an agent's effort, $e$, which has a cost $e^2$. The random component of productivity may arise from measurement error in a monitoring system. Let $A$ represent the precision of the monitoring system (e.g., $A$ might be the range of possible random errors). Deriving the equilibrium is straightforward. The optimal amount of effort is $e^* = (M-m)e/4A$. Optimal effort reflects the tradeoff between working harder (at a cost of $c$), earning an incremental reward $M-m$ from winning the tournament, and losing the tournament even after working hard (which depends on the amount of random error through the precision parameter $A$). Note that all agents are predicted to choose the same effort level.

In the experiments, large groups of subjects came into a room. Each subject took twenty envelopes, each containing a random number, from a large pile. Subjects were paired with an anonymous opponent for twenty trials. In each trial, both subjects in a pair picked a "decision number" from 0 to 100. In most of the experiments, picking the number $e$ cost the subject $e^2/15,000$ dollars (e.g., $e = 50$ cost $\.17$). After picking a decision number, subjects opened one of their envelopes and saw a random number $r$. Their output in that trial was $e + r$. In each pair, the subject with the higher output earned a payment of $M$ ($\$2.04$); the subject with the lower output earned $m$ ($\$0.86$). Usually subjects were told only whether they won the tournament or not at the end of each trial. (More detailed feedback made little difference in behavior.)

The experiment operationalized the constructs in the theory as closely as possible. Decision numbers represent effort levels. The cost $e^2/15,000$ represents the monetary equivalent of the disutility of exerting $e$ units of effort. The random numbers $r$ represent random components in perceived performance, due to measurement error. $M$ and $m$ correspond to rewards for winning and losing a tournament. The instructions to subjects used abstract language (to keep nonmonetary utilities or predispositions created by natural language from determining behavior). Repeating the entire process through stationary replication of twenty separate tournaments (and announcing outcomes after each tournament) gave subjects time to learn. Subjects were actually paid according to whether they won or lost the twenty tournaments, making their decisions salient.

In 25 experimental sessions (with variations in rules and parameters), the results were very robust. The individual-level prediction that subjects in each pair will choose the same effort is easily rejected. However, aggregate behavior is remarkably supportive of the theory. Figures 5 and 6 show effort levels, averaged across 15 pairs, in two symmetric tournaments. The averages are quite close to the predicted levels, 73 and 37, (which were chosen to be distant from round numbers like 50 or 25), and very little across the twenty trials. Further work showed that choices in unfair and uneven tournaments were reasonably well described by tournament theory, too.

![Figure 5. Tournament (Equilibrium = 73)](image)

The tournament experiments illustrate how a simple theory can be operationalized and tested experimentally. The parameters and rules can be varied to "stress test" the theory, and determine its boundaries of application. Experiments are also useful in comparison to other tests, since it is hard to imagine direct tests which definitively reject (or support) the theory using natural data. For example, Ehrenberg and Bognanno (1990) found that golf scores in professional tournaments improved when stakes were higher, but the test did not measure effort directly. O'Reilly, Main, an Crystal (1988) tested whether firms use tournament schemes in setting the compensation level of CEOs. They reject the use of tournament theory because their results suggest that the number of vice presidents is negatively correlated with the compensation levels of CEOs. (The theory predicts that if more agents compete
models of agency relationships, there are few empirical studies testing the theory's basic propositions. Experimental tests include Plott and Wilde (1982), DeJong, Forsythe, and Lundholm (1985), DeJong, Forsythe, and Uecker (1985), and Berg, Daley, Dickhaut, and O'Brien (in press).

The crucial agency problem is that agents can be motivated to work hard by tying their compensation more closely to their output, but closer ties increase the agents' risk because their output is affected by random elements. A weak link between output and compensation creates "moral hazard"—agents need not work hard, because their actions are hidden from principals (and low output is not penalized).

Earlier field studies (Wolfson, 1985) suggested that moral hazard does not influence behavior, contrary to the theory. By setting up a contracting situation in the laboratory, Berg et al. (in press) could study how well principals trade off incentives and risk-sharing, the costs of moral hazard, and the use of institutions and information systems to reduce the costs.

Berg et al. treated contracting as a Stackelberg game (i.e., the principal moves first). An agent and principal must agree on a contract. The principal chooses an incentive scheme \((I_1, \ldots, I_n)\) which specifies a wage \((I_i)\) for each possible output \((Q)\). (In another condition, principals choose wages for each possible action \((a_i)\).) After seeing the contract offered, agents choose whether to accept it then select an action \((a_1, \ldots, a_n)\). Principals are risk neutral, while agents are risk averse. The utility function of agents is assumed to be \(U(x) - G(a)\), where \(x\) is wealth and \(G(a)\) is the disutility of effort.

Principals could pick one of three contracts. One contract was worse for both the principal and the agent than a second contract (it was "Paredominated"), because it did not motivate the agent sufficiently and forced her to bear more risk as well. A third contract was simply a flat wage. Subjects generally picked the second, Pared-optimal contract. Moral hazard—less productive actions by agents—did occur when principals could only pay agents according to their output, rather than according to their actions.

**Compensation**

Theorists have studied how compensation contracts can align the divergent interests of managers and stakeholders (Dukes, Dyckman, & Elliott, 1980; Horwitz & Kolodny, 1980; Lambert & Larcker, 1985). For instance, firms have tried to align managers' discount rates with those of shareholders by using long-term bonus schemes. Such schemes were used by 99% of the top 200 industrial firms and 95% of the top 20 banks in 1987 (Sibson & Company, 1988). The use of long-term bonuses is relatively new (73% were instituted after 1980), and managers think their use will increase (O'Dell & McAdams, 1987).

Implicit in the normative models and the long-term bonus schemes firms use is the assumption that managerial behavior is affected by incentive
structures. But there is little evidence on the association between compensation schemes and managerial behavior. Establishing this association is difficult using natural data because key parameters are hard to measure and control (e.g., utility functions, effort levels, payoffs).

Schotter and Weigelt (in press) designed a stylized version of the compensation contract of a Fortune 100 firm and tested it experimentally. The firm had hoped the newly designed contract would induce managers to be less myopic. Schotter and Weigelt argued the contract would not change managerial behavior because it did not affect the discount rates of managers. They designed two alternative contracts that did affect discount rates, and predicted the two new contracts would affect managerial behavior.

Their experiments showed that subject behavior did not change under the firm's proposed compensation contract. Behavior did change under the alternative plans. This research shows how experiments can cheaply “pre-test” elements of organizational strategy. A badly designed compensation contract is costly in managerial salaries and in firm performance, but there are few ways to pre-test contracts. A field test is costly in time and implementation. (It is also considered unethical or unfair in many field experiments to choose members of a control group who must forego a treatment that may be desirable.) Testing in the laboratory is cheaper and faster.

CONCLUSION

We have reviewed several examples of how economists conduct experiments. Of course, external validity is always a concern in generalizing findings from experiments. There are three dimensions along which experiments differ from natural settings: People, place, and period (time).

People: The biggest concern in these experimental studies is people. Most of the studies we described (though not all) used undergraduate or MBA students as subjects, and paid them $5-$20 per hour. There are obviously reasons to be skeptical that the students' behavior corresponds to the behavior of experienced executives who make strategic decisions. There are several responses to this skepticism. The first response is that several studies have compared students and professionals in experiments. Burns (1985) compared students and wool traders in auctions like those in wool markets. DeJong, Fosythe, and Uecker (1988) compared accountants and students in a product-quality experiment (formally equivalent to some models of accounting services). Smith, Suchanek, and Williams (1988) compared students and businessmen trading in an asset market. Anderson and Sunder (1989) compared students and professional traders in an asset market. Dyer, Kagel, and Levin (1989) compared construction executives and students in a low-bid common-value auction, like bidding for construction contracts. Weber (1991) compared students and New York Stock Exchange floor clerks (many of whom become floor traders) in a game coordinating order flow between two markets.

In all these studies, only Anderson and Sunder (1989) found a substantial difference between students and professionals. (The professionals were more risk-averse and made fewer errors in judging probabilities.) The skills and experience of professionals build up in their jobs do not seem to affect their behavior in these experimental environments, probably because the environments are artificial and much simpler than the worlds the professionals work in. There is no doubt that students are different in some dimensions—age, income, education, and so forth. But in the laboratory these differences do not seem to matter.

Place: Since students and professionals appear to behave similarly in the experimental environments, the most important dimension of external validity is the correspondence of the experimental environment with natural settings.

The art of experimental design is the art of abstracting the important features of natural settings and building those into the experiment. Other features are necessarily neglected. Empiricists who construct surveys or work with data from secondary sources, and theorists, face the same basic problem of abstraction that experimenters face. Furthermore, the design choices experimenters make are disciplined by scientific procedures—replication, peer review, competition from other experimenters, statistical tests—just as the choices of theorists and empiricists are.

An experiment that tests a specific theory (or several theories), as many of those reviewed above have, is on the firmest philosophical ground because a theory should give clear guidance about which features to include in an experiment. The objection that an experiment is not a fair test of the theory implies that a variable which has been omitted from the theory is likely to matter, which means the theory is incomplete.

Strategy researchers, who discover facts by using data from secondary sources or surveys, are probably concerned about the unreality of laboratory settings. We are too. We would like to see more complex designs incorporating the ambiguity and dynamics of actual strategic decisions, with decisions made by teams or groups of subjects.11 (In all the experiments above, subjects made decisions by themselves.) At the same time, there is a danger in making experiments too complex.12

Experiments which strike the right balance of control and complexity should be informative for strategy research. For example, many experiments that economists have conducted test theories or study settings of interest to strategy researchers—contestability of markets, multiple point competition, strategic coordination, reputation-building, compensation, agency relationships. (We have left out hundreds of experiments less germane to strategy, especially the many studies of price formation under different auctions and trading institutions.)
Economics experiments might also be useful in studying managerial cognition, board behavior, strategic entry of product markets, first-mover advantages, network externalities (which create coordination games), vertical integration and transfer pricing, transaction-cost explanations of governance, and so forth. The resource-based view of the firm raises many fundamental questions that might be attacked experimentally: How are intangible resources (like reputations) created? Does causal ambiguity make imitation of resources difficult? Do managers recognize their firms’ valuable resources, and the extent to which resources can be extended to related and unrelated businesses? How do firms, as bundles of resources, compete? We can also envision highly ambitious experiments studying the complex intersection of competitive strategic decisions made by executives, capital market valuation of those decisions, and compensation those executives get (based on their decisions and capital market valuations of them). Such experiments could help establish causal relations that tests with natural data can only guess at.

ACKNOWLEDGMENT

The authors wish to thank the Reginald H. Jones Center for Management Policy, Strategy, and Organization for its support.

NOTES

1. The economists’ idea of theoretical validity is kin to construct validity in psychology. The idea is that a good economics experiment operationalizes constructs—usually tastes, endowments, and messages—the way theorists intended. The difference in the two fields is that the theory in economics is typically expressed mathematically, so the proper way to operationalize it is more clear.

2. The most common form of price discrimination involves separating buyers by their impatience or flexibility, which is presumably correlated with their reservation prices. Charging people with advanced airline reservations lower fares, or pricing hardcover books much higher than paperback books (and releasing the latter several months after the former) are examples.

3. The idea is that the degree of imperfect competition in a market depends on how much joint profit to the two firms is lost if perfect competition occurs instead. (In game-theoretic jargon, the noncooperative Nash equilibrium is the payoff from a “trigger strategy” that firms play if the other firms cheat on a tacit agreement to collude.) Mathematically, the degree of imperfect competition in each market depends on a profit constraint that must be satisfied to keep firms from competing perfectly. In multiple point competition, each firm must satisfy an overall profit constraint combining profits from multiple markets. The combined constraint enables the firms to oversatisfy the constraint in one market (relative to the case where firms compete in separate markets) and relax it in another.

4. The original experiments of this sort were conducted by Plott and Sunder (1982). Replications and extensions were conducted by Banks (1985); Friedman, Harrison, and Saltman (1984); Plott and Sunder (1988); Forsythe and Lundholm (1990); and Sunder (1992).

5. A third equilibrium involves “mixed strategies” (a probabilistic combination of strategies) in which the row player chooses R 3/4 of the time, and the column player chooses R 3/4 of the time.

6. In our experiments, we used blander labels than “banker” and “entrepreneur.”

7. We were much more careful to control the flow of information than in earlier experiments of others, because recent theory makes it clear how sensitive equilibria might be to small perturbations in beliefs. If players can recognize the other player they are paired with in a period by the other player’s voice, for instance, the equilibrium might change.

8. Note that the cost of effort is increasing at the margin—each hour worked is more costly than the one before—reflecting fatigue or increasing utility for foregone leisure.

9. See Ewalt et al. (1987) and Schotter and Weigelt (in press) for details.

10. A Pareto-optimal contract is one in which the parties cannot both improve their payoffs.

11. One advantage of experiments for strategy researchers is that experimental observation naturally breaks down the academic distinction between strategy formulation and process. One cannot help but observe process in the lab.

12. Many people have suggested that business strategy games used in MBA courses would make for good experiments, but most of these games are so complex that control over the experimental influence on behavior is lost. (These games may fail the test of “internal validity.” Are variables operationalized or manipulated the way the theory requires? Do subjects realize it?) We suspect business games are useful for studying managerial cognition (e.g., Hogarth and Makridakis, 1981), but little else.

REFERENCES


The Use of Experimental Economics in Strategy Research


The Use of Experimental Economics in Strategy Research


COMMENTARY:

The Use of Experimental Economics in Strategy Research
(K. Weigelt, C.F. Camerer, and M. Hanna)

Erhard Bruderer and Will Mitchell

EXPERIMENTAL ECONOMICS AS A “USEFUL LINK” IN STRATEGY RESEARCH

In this paper, Weigelt, Camerer, and Hanna describe and review the experimental economic approach to strategy research. A laboratory experiment attempts to eliminate nonessential individual and environmental effects on strategic behavior. By doing so, the researcher gains control over variables and causal relations, and so derives insights on decision making without confusing cause and effect of extraneous influences.

Experimental economics makes two key contributions to strategic research. Most theories of strategic behavior are either simply asserted, as intuitive logic or the results of a game theoretic model, or are tested with flawed data. Both approaches have problems. The assumptions underlying logical and game theoretic assertions are often questionable, so that conclusions may be wrong. On the other hand, the results of empirical analyses using case studies, surveys, or secondary economic data sources are often unreliable or barely significant, because strategic decisions by companies are heavily moderated by the psychology of the individual decision maker, team effects, standard operating