Contents

1

Introduction

Like other scientists, economists observe naturally occurring data patterns and then try to construct explanations. The resulting theories are then evaluated in terms of factors like plausibility, generality, and predictive success. As is the case in any science, it is often difficult to sort out cause and effect when many factors are changing at the same time. Thus, there may be several reasonable theories that are roughly consistent with the same observations. Without a laboratory to control for extraneous factors, economists often "test" their theories by gauging reactions of colleagues (Keynes, 1936). In such an environment, theories may gain support on the basis of mathematical elegance, persuasion, and focal events in economic history like the Great Depression. Theories may fall from fashion, but the absence of sharp empirical tests leaves an unsettling clutter of plausible alternatives. For example, economists are fond of using the word equilibrium preceded by a juicy adjective (e.g., proper, perfect, divine, or universally divine). This clutter is often not apparent in refined textbook presentations.

The development of sophisticated econometric methods has added an important discipline to the process of devising and evaluating theoretical models. Nevertheless, any statistical analysis of naturally occurring economic data is typically based on a host of auxiliary assumptions. Economics has only recently moved in the direction of becoming an experimental science in the sense that key theories and policy recommendations are suspect if they cannot provide intended results in controlled laboratory and field experiments. This book provides an introduction to the experimental study of economic behavior, organized around games and markets that can be implemented in class.

Notes for the Instructor and Students: The chapter will describe a market simulation that can be done in class with playing cards and a decision sheet from the Pit Market Instructions in appendix 2 at the end of the book. The tone will differ from that of other chapters in that it will describe how *you* (yes, you) could run a class experiment and guide the discussion afterward. This advice is particularly relevant for classes with team presentations based on class experiments and related material collected from Internet searches. If your class is not adapted for this format, just think of the teaching advice as being relevant for when you need to explain some economics concept to work colleagues, students from other majors, or in your younger brother's or sister's high school economics class.

1.1 Smith

It is always useful to start with a sense of historical perspective. Economics did not exist as an academic discipline in the eighteenth century, but the issues like tariffs and trade were as important then as they are today. Thomas Jefferson wrote that he lamented the lack of understanding of "political economy" among the colonists. He recognized and resented the effects of monopoly (as implemented by royal grants of exclusive marketing power) and even wrote to James Madison suggesting inclusion of "freedom of monopoly" in the Bill of Rights. Jefferson was aware of the Scottish philosopher, Adam Smith, and *Wealth of Nations*, which was written in the same year as the Declaration of Independence. Jefferson considered this book to be tedious and wordy, or as he put it, "prolix." This reaction reflects Smith's methodology of thinking based on detailed observations. In fact, Adam Smith wrote *Wealth of Nations* after a European trip, taken while tutoring one of his wealthy students. This trip provided Smith with the opportunity to observe and compare different economies. In France, he encountered the view that the source of wealth was vast fertile farmland, and that the wealth flowed toward Paris, the heart. He was also exposed to the Spanish view that wealth originated with gold and silver. Adam Smith considered these ideas and wondered about what could be the source of wealth in a country like England. His explanation was that England was a commercial nation of shopkeepers and merchants, with a tendency to "truck, barter, and exchange one thing for another." Smith explicitly recognized the importance of fairness in the *language* of trade, which separates humans from animals: "Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog." Each voluntary exchange creates wealth in the sense that both people benefit, so extensive markets with flourishing trade create considerable wealth. In *Theory of Moral Sentiments*, he extended the discussion to include social exchanges with family, friends, and neighbors, exchanges based on pro-social attitudes like altruism, reciprocity, etc.

Smith's detailed accounts of specific markets, prices, and exogenous events like a "public mourning" show a clear understanding of the forces of supply and demand that move prices temporarily or keep them down to cost levels in the long run. (Graphical representations would come later.) But Smith's deepest insights about market systems are about how traders, following their own self-interest, are led to promote the common good, even though that was not their intention. His belief in the power of the "invisible hand" was balanced by a healthy skepticism of those who "affected to trade for the common good." Smith clearly recognized the dangerous effects of the pursuit of self-interest in the political sphere. Although a philosopher, Smith lobbied Parliament for the lifting of tariffs, a development that provided a major boost to the English economy.

Adam Smith casts a long shadow. When the author was a college sophomore many decades ago, his Economics professor, John Gunn, noted that he had studied under Jacob Viner at Princeton, who had studied under . . . and so forth, back to Alfred Marshall and eventually Adam Smith! Even today, introductory books echo Smith's distinctions between land, labor and capital, specialization of labor, the extent of the market, etc. The textbooks from 50 years ago painted a picture of a highly idealized market. Students had to memorize a set of assumptions for perfect competition, which included "an infinity of buyers and sellers," "perfect information about market conditions," and the like. *At that time, Vernon Smith had begun running experiments at Purdue that did not satisfy any of the perfectness assumptions.* There were small numbers of buyers and sellers, and they had no prior knowledge of each others' values or costs, although there was good information about current bids, asks, and recent transactions prices.

1.2 A Class Pit Market

As a graduate student, Vernon Smith had participated in a market simulation in a Harvard class taught by Edward Chamberlin (1948), who argued that his class experiments highlight failures of the standard model of perfect competition. Before discussing such a market simulation, it is useful to actually run one. Experimental economists, including this author, often begin class on the first day with an experiment. The easiest way to proceed and maximize active student involvement is to distribute a numbered playing card to each person, after dividing the deck(s) into two stacks—one for seller's costs (clubs and spades) and another for buyer's values (hearts and diamonds). The pit market instructions in the appendix at the end of the book should be read out loud to ensure that everybody is on the same page and ends at the same time. Those instructions explain the process:

> Buyers and sellers will meet in the center of the room (or other designated area) and negotiate during a 5-minute trading period. When a buyer and a seller agree on a price, they will come together to the front of the room to report the price, which will be announced to all. Then the buyer and the seller will turn in their cards, return to their original seats, and wait for the trading period to end.

For example, if a buyer with a 7 of hearts makes a trade at \$5 with a seller who has a 4 of clubs, then the buyer earns $\$7 - \$5 = \$2$, and the seller earns $\$5 - \$4 = \$1$. Traders typically hold their cards so that they cannot be seen, but

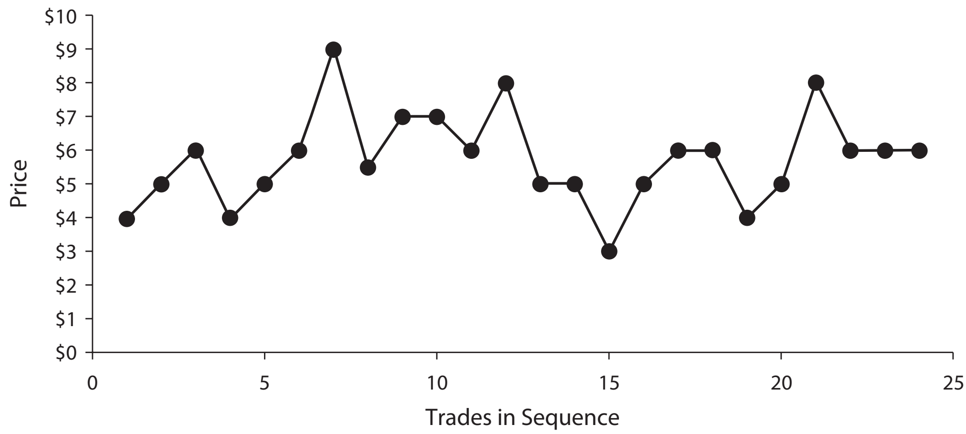

Figure 1.1. Transactions Prices for a Class Pit Market

when they come to the front to make a report, the cards are checked to be sure that the price is no lower than the cost or no higher than the value. The price should be called out, so that others who are still negotiating are aware of the "going" prices. It is useful to pre-select a student assistant or two to help with the checking and announcement process.

Figure 1.1 shows the price sequence for a market with about 60 public policy students at the University of Virginia on the first day of summer "math camp." Prices tend to converge to about $6 (the average was $5.73 last year and $5.57 this year). Sometimes it is necessary to repeat, collecting and re-shuffling buyers' and sellers' cards, and redistributing before the next trading period.

At this point, it is useful to reveal the cards used, which in this case were:

Buyers's Cards: 5, 6, 7, 8, 9, and 10 (five cards of each number)
Seller's Cards: 2, 3, 4, 5, 6, and 7 (five cards of each number)

The discussion can be focused on why the prices are near observed levels. The most important thing is to lead the discussion instead of just announcing the correct economic prediction. Consider the question: at a price of $7, would there be a larger number of willing buyers or of willing sellers? And with more willing sellers at $7, what do you think will happen to the price (they will undercut each others' prices). The parallel question is: at a price of $3, would there be a larger number of willing buyers or willing sellers? And what would tend to happen to the price? Then the question is: at what price would there be no pressures, upward or downward, on price? This is essentially a question about finding a price at which the quantity demanded equals the quantity supplied, but it is better if the students figure this out in the discussion process. The point is that the discussion of an experiment can be structured to maximize the benefits of having students discover the principles for themselves. As students

Table 1.1. Ordered Values and Costs for the Class Pit Market

Buyers' Cards	Quantity Demanded		Sellers' Cards	Quantity Supplied
\$11 (0 cards)	0		\$11 (0 cards)	30
\$10 (5 cards)	5		\$10 (0 cards)	30
\$9 (5 cards)	10		\$9 (0 cards)	30
\$8 (5 cards)	15		\$8 (0 cards)	30
\$7 (5 cards)	20	excess supply	\$7 (5 cards)	30
\$6 (5 cards)	**25**	**equilibrium**	**\$6 (5 cards)**	**25**
\$5 (5 cards)	30	excess demand	\$5 (5 cards)	20
\$4 (0 cards)	30		\$4 (5 cards)	15
\$3 (0 cards)	30		\$3 (5 cards)	10
\$2 (0 cards)	30		\$2 (5 cards)	5
\$1 (0 cards)	30		\$1 (0 cards)	0

of experimental economics, you will have opportunities to use experiments in class, and it helps to plan a structured de-briefing discussion.

Economics textbooks typically show demand and supply as lines, so it is useful to summarize the results of the previous discussion by organizing the values from high to low, and organizing the costs from low to high, as in table 1.1. In the top row of the left side of the table, the price is \$11, and there are 0 units demanded at that price since all buyers' values are \$10 or below. At a price of \$10 (or slightly below) there are 5 units demanded, since there are 5 buyers with values of \$10. Coming down the left side of the table, as the price falls to \$5 or below, all 30 units are demanded. For sellers, the quantity supplied is 0 units at a price of \$1 in the bottom row, right side, since all sellers have costs that are higher. As the price goes up, moving up the right side of the table, more and more units are supplied. Notice that there is excess demand at low prices below \$6 and excess supply at prices above \$6, with equality (equilibrium) at \$6.

The prices and quantities trace out a supply line with a vertical intercept of \$1 at 0 units (bottom row of the right side of the table), and a slope of 1/5, since the line rises by \$1 in subsequent rows for each 5-unit increase in quantity, i.e., $P = 1 + 0.2Q$. Similarly, the prices and quantities for buyers trace out a demand line with a vertical intercept of \$11 at 0 units (top row, left side of the table) and a negative slope of −1/5, i.e., $P = 11 - 0.2Q$. The intersections of these lines, solved by substitution, are at $Q = 25$ and $P = 6$, which are approximately at the observed price and quantity levels. (The familiar looking supply and demand lines mask little "steps" due to the discreteness of units or 5-unit blocks, to be clarified later.)

From Adam Smith's perspective, however, the key point is not the price prediction, but rather that the trades create wealth in the sense that both buyer and seller benefit from a voluntary trade. Each trade adds several dollars to the total benefit, and the transactions price merely determines how that "surplus" between value and cost is divided. Participants in the market focus on their own gains from trade, but the big picture is in terms of the total gains from trade from all transactions combined. In fact, the equilibrium that occurs at the intersection of supply and demand *maximizes* the gains from trade. This realization can be driven home by considering the question of which sellers are excluded if all trades are at a price of $6 (i.e., those with higher costs). The reader should think of an example of a market for a service requiring low skills (lawn care), and thinking about how the going price excluded those with high skills and high opportunity costs for their time. Parallel observations apply to the exclusion of low-value buyers. If the cards for untraded units in a pit market are kept separate, they are generally the high costs and low values, and exceptions typically involve small losses, e.g., if a seller with a $6 cost displaces one with a cost of $4, the loss is $2. This year, the remaining cards were all numbered 5, 7, and 6, so there were no efficiency losses.

The standard measure of performance in a market experiment is the actual earnings (gains from trade), which represents the wealth from all of the voluntary transactions. This total is expressed as a percentage of the maximum, which provides an efficiency measure. The maximum is determined by differences between values and costs of units that are predicted to be traded in equilibrium. These included values and costs are listed in table 1.2. First consider the column on the left side, with values of $10 and costs of $2. If all 5 of the $10 value units trade, and all 5 of the $2 cost units trade, even though those people do not necessarily trade with each other, then the surplus is the difference between value and cost (10–2) times the number of units, 5, for a total of $40 in surplus, as shown in the bottom row. Similar calculations add amounts of $30, $20, and $10, for a total of $100 in surplus. (Note that these surplus calculations do not depend on who trades with whom, as long as those with values above the equilibrium buy from those with costs below.) If some of the low-value and high-cost units in the columns on the right side of table 1.2 do not actually trade, the loss in surplus would be several dollars, which would be small relative to the total surplus. And notice that adding in excluded units with costs of $7 and values of $5 would reduce efficiency. Efficiency is often quite high in these markets. To summarize:

Table 1.2. Included Values and Costs at a Price of $6

Included Values	10	9	8	7	6
Included Costs	2	3	4	5	6
Surplus	(10−2)5=40	(9−3)5=30	(8−4)5=20	(7−5)5=10	(6−6)5=0

Market Trading: *In equilibrium, the price provides a bright line boundary that excludes low-value buyer units and high-cost seller units, and thereby tends to maximize the gains from trade as measured by market efficiency.*

1.3 Early Developments in Experimental Economics

Markets

The value and cost configuration used in the previous section's pit market is symmetric. Vernon Smith (1962, 1964) considered asymmetric designs that might cause prices to start too low (if supply is relatively steep) or too high (if demand is steep). But unlike Chamberlin, he would conduct a series of market trading periods with identical values and costs. In addition, Smith used a "double auction" that collected all buyers' bid prices and sellers' asking prices into a single auction process. Double auction procedures (discussed in the next chapter) provide a price signal that is more organized than the decentralized negotiations in a pit market. A major advantage of the experimental approach is that it let Smith measure the surplus achieved, as a percentage of the theoretical maximum. Values and costs are typically not observed in naturally occurring markets, but are induced in laboratory experiments. Thus, controlled experiments permit better measurements. Smith observed efficient competitive outcomes as prices converged in a series of double auctions, even with as few as 6–10 traders. This result was significant, since the classical "large numbers" assumptions were not realistic approximations for most market settings.

An introductory class today will begin with discussions of the benefits of competitive market allocations, but the narrative switches to various imperfections like asymmetries in quality information that can cause markets to fail, as documented by George Akerlof's (1970) analysis of the market for low-quality "lemons." These observations have motivated experimental studies of market failures in insurance and other settings where the selection of those who make purchases or sales is "adverse" to those on the other side of the transactions.

Today, many of the market experiments are focused on the design and testing of new trading institutions, e.g., auctions for broadcast spectrum, water, or emissions permits to limit greenhouse gases. Auction design is an area where game theory and experimentation continue to have a major impact on public policy. In situations where price is not appropriate, e.g., allocation of slots in schools, there has also been theoretical and related work on matching mechanisms based on ranked preference lists submitted by participants. Students will recognize a typical sorority rush procedure as an application of matching mechanisms. Al Roth's 2012 Nobel Prize was given in recognition of his theoretical and experimental work on matching mechanisms. Another common application of market experiments is the study of asset markets and macroeconomics issues with interrelated markets.

Game Theory

A parallel development is based on game-theoretic models of strategic interactions. In a "matching pennies" game, for example, each player chooses heads or tails with the prior knowledge that one will win a sum of money when the coins match, and the other will win when the coins do not match. Similarly, an accountant will want to be especially well prepared when an audit occurs, but the auditor wants to catch cases when the accountant is not prepared. Each person's optimal decision in such situations depends on what the other player is expected to do. The systematic study of strategic interactions began with John von Neumann and Oscar Morgenstern's (1944) *Theory of Games and Economic Behavior*. They asserted that standard economic theory of competitive markets did not apply to the bilateral and small-group interactions that make up a significant part of economic activity. Their "solution" was incomplete, except for the case of "zero-sum" games in which one person's loss is another's gain. While the zero-sum assumption may apply to some extremely competitive situations, like sports contests or matching pennies games, it does not apply to many economic situations where all players might prefer some outcomes to others.

Economists and mathematicians at the RAND Corporation in Santa Monica, California began trying to apply game-theoretic reasoning to military tactics at the dawn of the Cold War. In many strategic scenarios, it is easy to imagine that the "winner" may be much worse off than would be the case in the absence of nuclear war. At about this time, a young graduate student at Princeton entered John von Neumann's office with a notion of equilibrium that applies to a wide class of games, including the special case of those that satisfy the zero-sum property. John Nash's notion of equilibrium (and the half-page proof that it generally exists) were recognized by the Nobel Prize committee about 50 years later. With the Nash equilibrium as its keystone, game theory has recently achieved the central role that von Neumann and Morgenstern envisioned. Indeed, with the exception of supply and demand, the "Nash equilibrium" is probably used as often today as any other construct in economics.

A *Nash equilibrium* is a set of strategies, one for each player, with the property that nobody could increase their payoff by unilaterally deviating from their planned action given the strategies being used by the other players. To illustrate this idea, consider the most famous simple game, known as a prisoner's dilemma, or more generally, a *social dilemma*. Suppose that there are two producers who sell a product that is needed by the other. The delivered product can be of high or low quality, with a high-quality delivery costing more. The value of a high-quality delivery is 3 for the recipient, the value of a low-quality delivery is only 0 for the recipient, and the producer's costs are 1 for high and 0

Table 1.3. A Prisoner's Dilemma (Row's payoff, Column's payoff)

	Column Player	
Row Player	High	Low
High	2, 2	−1, 3
Low	3, − 1	0, 0

for low. The payoff situation can be represented as a matrix in table 1.2, where the row player's decisions, High or Low, are listed on the left, and the column player's decisions, also High or Low, are listed across the top. Therefore, if both deliver high to each other, the payoffs are 3 (from the other's high-quality delivery) minus the cost of 1 (from the person's own delivery to the other), or 2 for each person, as shown in the upper left corner of table 1.3. In the lower-right corner, the payoffs are 0 (for receiving low quality and incurring no cost of delivering it to the other). The asymmetric payoffs on the "off" diagonal are for cases in which one person receives high quality, worth 3, at no cost, and the other incurs the cost of 1 but receives low quality, as indicated by the (3, −1) outcome. For a single play of this game, a Nash equilibrium would be a pair of strategies, one for each person, for which neither would have a unilateral incentive to change. Notice that (Low, Low) is a Nash equilibrium, since if the other is going to deliver Low anyway, one's own payoff only goes down by incurring the cost of high delivery. Moreover, this is the only Nash equilibrium. For example, (High, High) is not a Nash equilibrium, since each would have an incentive to accept the other's generosity, but cut cost on one's own delivery to earn 3 instead of 2. In general, a prisoner's dilemma is a 2 × 2 game with a unique Nash equilibrium, and another non-equilibrium outcome that involves higher payoffs for both players than they receive in the equilibrium. More generally, the term *social dilemma* refers to games with two or more decisions, for which the unique equilibrium provides lower payoffs for each player than can be achieved with another outcome.

John Nash's equilibrium definition and existence proof caught the attention of researchers at the RAND Corporation headquarters in Santa Monica, who knew that the Princeton graduate student was scheduled to visit RAND that summer. Two RAND mathematicians immediately conducted a laboratory experiment designed to stress-test Nash's new theory. Nash's thesis advisor was in the same building when he noticed the payoffs for the experiment written on a blackboard. He found the game interesting and made up a story of two prisoners facing a dilemma of whether or not to make a confession. According to the story, not confessing benefits both, but the prosecutor makes promises and threats that provide each person with a unilateral incentive to confess, which is the Nash equilibrium. This story was used in a presentation to the psychology department at Stanford, and the "prisoner's dilemma" became the most commonly discussed paradigm in the new field of game theory. The actual experiment involved multiple plays (over 100) of the same game with the same two players.

It is easy to implement a prisoner's dilemma in class by writing the payoffs on the board and giving each person two playing cards: hearts or diamonds correspond to the cooperative decision (deliver high quality), and clubs or spades correspond to the uncooperative decision ("defect" with low quality). Then pairs of people can be asked to show a card simultaneously. A Nash equilibrium would have the property that neither person would have an incentive to choose a different card after seeing the other's card decision, a kind of "announcement test." The point would be to highlight the tension between the socially optimal, cooperative outcome, and the privately optimal defect outcome, and to discuss features of ongoing business relationships that would help solve this problem.

One problem with the notion of a Nash equilibrium, which arises in multi-stage games, is that it might involve a threat that is not credible. Consider an *ultimatum bargaining game* in which one player makes a take-it-or-leave-it proposal, that the other must either accept or reject. In particular, suppose that the proposer can offer a fair split (2 for each) or an unfair split (3 for the proposer, 1 for the responder). The responder observes the initial proposal and must decide whether to accept the proposal or reject, in which case both receive 0. This game has two stages, with the proposer moving first and the responder moving second after seeing the proposer's decision. A *strategy* specifies a decision for each possible contingency. In other words, a strategy is a plan of action that could be given to an assistant to play the game, with no need for the assistant to check back about what to do. The proposer has only two strategies, which are labeled fair and unfair. The responder's strategy, in contrast, must specify decisions in two contingencies, so the responder has four strategies: (accept fair, accept unfair), (accept fair, reject unfair), (reject fair, accept unfair), and (reject fair, reject unfair). If the responder's strategy is (accept fair, accept unfair), the proposer's best response is unfair, which would be accepted and is a Nash equilibrium. If the responder's strategy is (accept fair, reject unfair), the proposer's best response is to choose fair and avoid the 0 payoff from rejection. This is also a Nash equilibrium, but it requires a responder, who receives an unfair proposal (an offer of only 1), to reject it and end up with 0. The problem is that the responder might have trouble making this rejection threat credible. Another way to think about the situation is to consider the second stage as a game (a "subgame") in which only the responder has a decision, to choose between a payoff of 1 or 0, given the proposal that was made earlier. A rejection is not an equilibrium if attention is restricted to that subgame.

One of the major advances in game theory was Reinhard Selten's work on *subgame perfection*, which involves ruling out Nash equilibria that are not also equilibria in the subgames. The point here is not to fully develop the precise theoretical definitions or illustrate them with experiments, but to point out an important connection with laboratory experiments. The author always imagined that Selten, who also started doing experiments in the 1950s, was wearing his "theory hat" while working on subgame perfection, and forgetting about experiments. The realization that Selten's insight was actually motivated by experiments was communicated by him in a "witness seminar" discussion of a group of early contributors to the field, held in Amsterdam several years ago:

> I wanted to tell you a story of how the sub-game perfect equilibrium came about. What we did first, in my early experiments, I also did experiments with situations where I didn't know or where there was no theory for it. There were many oligopoly situations, which didn't involve clear theory. For example, I looked at the situation of oligopoly with demand inertia. Demand inertia means that future demand or future sales depend on prior sales. . . . It was quite complex when we explored it. My associate *Otwin Becker* and I tried to make a theory for this experiment. What should be the theoretical solution? Then, I think I simplified it completely. I simplified it to a high degree and kept demand in it. And I computed the equilibrium. [But] I suddenly found that this was not the only one. There were many other equilibria. And then I invented the idea of sub-game perfectness in order to single out the one equilibrium. (Svorenčík and Mass, 2016, p. 155)

This passage highlights an essential advantage of experimentation: the setup can be controlled enough to allow the application of relevant theory, and looking at the resulting data motivated changes in theory, making it more behaviorally relevant.

Social Preferences

After his initial exposure to the RAND experiment results in 1950, John Nash briefly considered the implications for bargaining behavior. Bargaining had been one of Nash's interests ever since taking an international trade course in college and realizing that economists did not have a good way to model it. He soon gave up on bargaining experiments, presumably because there was no well-developed theory of fairness at that time, at least among economists. Initial results of subsequent experiments with take-it-or-leave-it "ultimatum" offers were sharply at odds with the predictions of subgame perfection, as summarized in Selten's account in the Amsterdam Witness seminar:

> This is a psychologist-led experiment where the subject played against a computer, but didn't know it was a computer. The computer was

> programmed to have a fixed concession rate. And they played for 20 periods. They made alternated offers and the subject was the last one to accept or reject the offer. We saw that the subject often left their $3 on the table or so. And it all resulted in a conflict. I was completely surprised about this. The psychologists were not surprised at all. They did not think anything about this, but I was surprised about it. And I discussed it with *Werner Güth* and later he made then this experiment. *Werner* simplified the whole thing to just one period to the ultimatum game, which happens at the end of this game. And of course, he got the result that very low offers are not accepted. But it was foreshadowed in these psychological experiments. It was not even remarked by these people that there was something extraordinary happening. (Svorenčík and Mass, 2016, p. 156)

Today, there is a large literature on ultimatum bargaining and social dilemma games that focus on fairness, reciprocity, altruism, and other factors that affect behavior. In particular, Elinor Ostrom's work combined laboratory and field studies of how small groups solve resource management issues. Trained as a political scientist, she was the first woman to win a Nobel Prize in economics, in 2009. Today there is a Social Dilemmas Workshop that meets regularly to discuss this line of research that Ostrom initiated.

Bounded Rationality

Reinhard Selten was inspired by earlier work on *bounded rationality*, a term that originated with Herb Simon and was the basis for his Nobel Prize. Although trained as a political scientist, Simon spent his career at Carnegie Mellon in the business school, and later in the Psychology and Computer Science departments. Simon stressed that decision makers often rely on rules of thumb or *heuristics*. He favored studies of actual behavior, with a focus on adaptive responses to events and situations. The Carnegie School in the 1950s had a behavioral focus, with some experimentation and a year-long business game used for teaching in the MBA program, which was one of the first computerized market simulations.

When the author arrived at Carnegie Mellon as a graduate student in 1970, this behavioral/experimental focus was dismissed by graduate students as being dated (relative to the exciting work on rational expectations being done by Carnegie Mellon economists Bob Lucas and Ed Prescott). In fact, Lucas once credited his own work on rational expectations to feeling exasperated with notions of adaptive learning and process, with little focus on steady states and final outcomes. Although the author wrote a theoretical thesis on auctions under Prescott with a rational-expectations assumption that "closed" the model, he also worked on behavioral economics projects with Richard Cyert (a Simon coauthor who was then president of the university) and statistician

Morris Degroot. One of those projects was motivated by Cyert's observation that investment decisions made by CEOs and boards of directors were, at the time, typically framed by the amounts of retained earnings, which seemed to be more important than classical interest rate considerations. This is an example of a behavioral bias later identified as *mental accounting*, where sources of funds are used to constrain uses. Experiments played an important role in the documentation of mental accounting and other biases discussed in Richard Thaler's (1992) book, *The Winner's Curse*, and in his "Anomalies" column in the *Journal of Economic Perspectives* (Thaler, 1988, 1989; Tversky and Thaler, 1990). The insights summarized in this book and the research it cites provide the basis for Thaler's 2017 Nobel Prize.

Much of the experimental economics research being done today builds on Simon's notions of bounded rationality and the behavioral insights of psychologists and others who studied actual business behavior. Anyone who has looked at experimental data will have noticed situations in which people respond to strong incentives, although some randomness is apparent, especially with weaker incentives. Psychologists, who would ask people to identify the brighter light or louder sound, came to model such behavioral response as being probabilistic. Today, probabilistic choice models (logit, probit, etc.) are standard in econometric work when there are discrete choices, e.g., whether or not to enroll in a treatment or rehabilitation program. This work was pioneered by Dan McFadden, who was awarded a Nobel prize in Economics for it. The games used in experiments also typically have discrete choices, and incorporating probabilistic responses to incentive differences can be done in a manner that generalizes the notion of a Nash equilibrium, which is known as a *quantal response equilibrium* (McKelvey and Palfrey, 1995), an idea that will explain seemingly anomalous deviations from Nash predictions. There are several places in this book in which sharp "best response" lines in figures will be replaced by curved "better response" lines.

There have also been important advances in understanding the heuristics that people use when playing a game just once, when past observation on others' decisions is not available. For example, think of a "level 0" person as being totally random, a "level 1" player as someone who makes a best response to a level 0 player, a level 2 player makes a best response to a level 1, etc. Even though many economic interactions are repeated, this is often not the case, especially in politics, law, or military conflict, and *level-k thinking* models have been used in the analysis of experimental data for games played once.

Game theory has been developed and applied in disciplines like law, politics, and sociology. In economics, game theory has had a major impact on public policy, especially in the design of auctions and market mechanisms. Experiments have stimulated the development of a more behaviorally relevant theory for subdisciplines like behavioral finance, behavioral law and economics, and behavioral business operations ("B-Ops"). Indeed, game theory is the closest thing there is today to a unified theory of social science.

Decisions and Risk

A game or market may involve relatively complex interactions between multiple people. Sometimes it is useful to study key aspects of individual behavior in isolation. It is straightforward to set up a simple decision experiment by giving a person a choice between gambles or "lotteries," e.g., between a sure \$10 and a coin flip that yields \$30 in the event of heads and \$0 otherwise. The expected value of the lottery is calculated from products of payoffs and their associated probabilities, e.g., $30(1/2) + 0(1/2) = 15$. Which would you choose in this case? What if, instead, the choice were between a sure \$100,000 and a coin flip that provides a 50-50 chance of \$0 and \$300,000? Risk aversion is indicated by a preference for a sure outcome, even though it has a lower expected payoff. The intuition for risk aversion is analogous to diminishing marginal utility, i.e., the third \$100,000 is not as important as the first. Risk is a feature of many games, and von Neumann and Morgenstern (1944) developed a theory based on the expected value of a nonlinear utility function.

Not long after the formalization of expected utility, Allais (1953) presented anomalous results that show up in choices between pairs of lotteries. The *Allais paradox* subsequently generated an outpouring of experimental work and was the basis for his Nobel Prize. At about the same time, Harry Markowitz (1952) noticed that people seem to have a reference point at the current or normal wealth level, and that they treat risks above and below that point differently, with losses being more salient. He even offered a formal definition of what has come to be known as *loss aversion*. Experimental methods were not well developed at that time, and Markowitz based his conclusions on having approached colleagues and asking them questions like: "Would you rather owe me \$1 or have a 1/10 chance of owing me \$10?" He used both gains and losses, at various scales. Psychologists Danny Kahneman and Amos Tversky (1979) further developed these ideas and others, e.g., over- or underweighting of extreme probabilities. The result, known as *prospect theory*, was shown to explain a wide range of anomalies in experimental data, and was the basis for Kahneman's 2002 Nobel Prize in Economic Science.

1.4 Advantages of the Experimental Methods

It is important to emphasize that the overall design of an experiment should address an important policy or theoretical issue. The setting should be simple enough so that the results can be interpreted without the need to explore alternative explanations. A clear focus is often achieved by having *treatments* of primary interest that can be compared with a baseline or *control* condition. These themes will be further developed in the subsequent chapters, which include a chapter on methodology and statistical testing.

A *laboratory experiment* is done in a controlled setting, e.g., a closed room with visual separation and minimal outside distractions. Subjects are typically exposed to different treatment conditions, e.g., a sealed bid or an ascending bid auction. A *within-subjects* design has the same person (or group of people in a market or game) being exposed to two or more treatments, so that each group is its own control. A *between-subjects* design exposes each group or person to a single treatment. In this case, a typical "between" design would be to recruit 10 separate groups for one treatment and another 10 groups for the other, with no interaction across groups. Within-subjects designs are attractive when there is considerable heterogeneity between individuals, so that it is important for each person or group to serve as its own control. But between-subjects designs should be considered if there are "sequence effects" that cause outcomes for one treatment to bias those of a second treatment with the same subjects that follows.

A *field experiment* takes place in a natural setting, in which treatment conditions are implemented without the subjects being aware that they are in an experiment. For example, potential voters might be approached with either a phone call or a knock on the door, with the same message about the citizen obligation to vote. For both laboratory and field experiments, the exogenous assignment of treatments is essential to making inferences about causality. An ex post study of get-out-the-vote methods used in actual elections could be biased, for example, if political operatives target the knock-on-the-door resources to districts that are expected to be close races. Practical considerations generally dictate between-subjects designs for field experiments. Field experiments obviously provide more realistic context and subject selection, although there may be some loss of control due to extraneous events in the field environment. Moreover, many key variables, like values and costs, cannot typically be measured in the field. And the lab can be used to create a "perfect storm" that stresses the performance of proposed policies or auction procedures. Roughly speaking, lab experiments are generally better for evaluating theoretical predictions and stress tests of policy proposals, whereas field experiments may be better for evaluating policies under "normal" conditions. A mix of lab and field treatments is sometimes quite effective if it can be accomplished.

With this terminology, it is useful to highlight some of the advantages of experimental methodology:

Motivation: Monetary or other payoffs can be used to induce preferences that place people in environments that correspond to theoretical models or proposed policies.

Control: Experimental trials can be conducted under identical conditions, without the need to adjust for systematic or unexpected changes in weather or economic conditions that complicate the interpretation of results.

Replication: Laboratory trials are typically repeated with new groups of participants, which smooths the effects of random variations due to personalities, etc. Replication for field experiments is more difficult, but sometimes can be done by going to different locations or target populations.

Economy: Obviously, laboratory tests are considerably less costly than large-scale trials with otherwise untested conditions.

Measurement: In the laboratory, it is possible to "look inside the box" and measure things like efficiency using specified values and costs, which are generally more difficult to observe or estimate with data from ongoing markets. Secondary measurement is sometimes an option with field experiments, e.g., using follow-up surveys, or using data provided by charitable foundations about their donor base for a field experiment exposing different people to different donation match treatments.

Discovery: Everyone has heard the phrase that "correlation does not imply causation," e.g., if there is a third factor that is driving both the supposed "cause" and "effect." By holding other factors constant, the effects of a treatment change can be isolated.

Exploration: It is possible to "think outside the box" and use experiments to test new markets or political institutions that have not been used previously. This advantage is amplified by advances in information processing and social media technology that permit novel types of political and economic interactions, e.g., emissions permit sales that are responsive to changes in market conditions.

Stress Testing: It is important to evaluate the performance of alternative types of markets or mechanisms under adverse "perfect storm" conditions.

Demonstration: Experiments can be used to show that a procedure is feasible and to give policy makers the confidence to try it. Experiments can also be used for teaching; there is often no better way to understand a process than to experience it. Experiments that implement simple situations provide a counterbalance to more abstract economic models and graphs.

1.5 *Experimental Economics* and the Economic Science Association (ESA)

In the 1980s, Vernon Smith and his colleagues and students at Arizona established the first large experimental economics laboratory and began the process of developing computerized interfaces for experiments. A community formed around a series of conferences in Tucson. The Economic Science Association (ESA) was founded at one of those conferences in 1986, and the subsequent presidents constitute a partial list of key contributors: Vernon Smith, Charlie Plott, Ray Battalio, Elizabeth Hoffman, Charlie Holt, Bob Forsythe, Tom Palfrey, Jim Cox, Andy Schotter, Colin Camerer, Ernst Fehr, John Kagel, Jim Andreoni, Tim Cason, Al Roth, Jacob Goeree, Yan Chen, and Cathy Eckel (president-elect in 2017!).

The author approached the ESA members in the mid-1990s, with the idea of starting a journal, but the idea was initially rejected. The fear was that a journal would be too specialized and that important ideas would become marginalized and have less impact on the development of economic thinking. At this time, a series of small experimental economics conferences were held at the University of Amsterdam, which also had a lab. After a couple of attempts to start a journal with ESA, the author teamed up with one of the Amsterdam conference organizers, Arthur Schram, and began making plans to edit a journal that would be published with a Dutch publisher, Kluwer. The four advisory editors were Vernon Smith, Reinhard Selden, Al Roth, and Charlie Plott. At this point the ESA officials were asked if they wanted this to be their journal. Tom Palfrey, then president, responded that *yes*, they would be willing for it to be *one* of their journals! Moreover, in the process of gathering support from European ESA members, he negotiated agreements that confirmed regular European ESA meetings that rotated with those in the United States. In addition, the Kluwer representative, Zac Rolnik, agreed to provide the journal free of charge for several years to members of a long-standing German experimental economics association.

Experimental Economics was launched in 1998, and the second ESA journal, the *Journal of the Economic Science Association*, began in 2016. *Experimental Economics*, which received about one submission per week when it started 20 years ago, now receives about one submission per day. The journal is quite selective, the "impact factor" is high, and its size has more than doubled. Membership in the ESA includes access to these publications, as well as a listserv for queries and announcements of upcoming conferences. The ESA welcomes student members and meets in the United States and Europe each year, along with regional meetings in Asia and the Pacific.

Figure 1.2 provides a perspective on the explosive growth in the experimental economics literature. The upper gray line is based on the author's count, beginning with Chamberlin's 1948 paper. The black line begins when the *Journal of Economic Literature* classifications were revised in 1991, and it includes only those papers that list the "design of experiments" code, which excludes books and most papers in collected volumes and in other disciplines. The annual number of publications has more than tripled since the first edition of this book was published in 2007, and many of the new ideas have been incorporated in the chapters that follow.

Many exciting developments are in the works. Economics experiments are being integrated into high school and introductory economics. Theorists

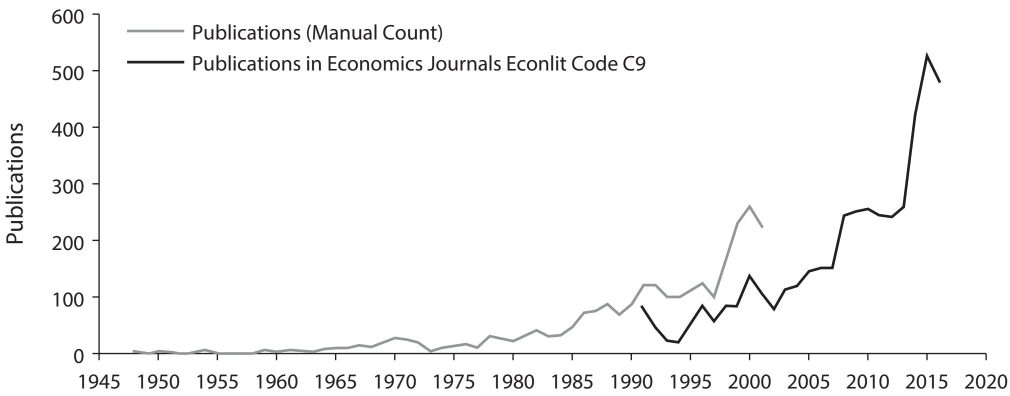

Figure 1.2. Publications in Experimental Economics

are looking at laboratory results for applications and tests of their ideas, and policy makers are increasingly willing to consider how proposed mechanisms perform in controlled tests before risking a full-scale implementation. As indicated in the final chapters of this book, experimental methods have been used to design large auctions (e.g., the FCC spectrum auctions and emissions permit auctions) and systems for matching people with jobs (e.g., medical residents and hospitals). There are also experimental economics subfields in law, business, macroeconomics, finance, and other areas. The ever-expanding list of Nobel laureates with behavioral interests is a reminder of the impact that this work has had. Economics is well on its way to becoming an experimental science!

Chapter 1 Problems

(Hints for all problems are provided in appendix 1 at the end of the book, but you should try to work the problems before turning to the appendix for help.)

1. Use the supply and demand formulas given as approximations for the setup in table 1.1 to solve for price and quantity. These formulas were: $P = 11 - 0.2Q$ for demand and $P = 1 + 0.2Q$ for supply.

2. Consider a market with 8 buyers and 8 sellers. The buyers' values are 10, 10, 10, 10, 4, 4, 4, 4. The sellers' costs are: 2, 2, 2, 2, 8, 8, 8, 8. At a price of 9, would there be excess supply or excess demand? (Explain briefly.)

3. For the market structure in problem 2, at a price of 3, would there be excess supply or excess demand? (Explain briefly.)

4. For the setup in problem 2, at a price of 6, would there be excess supply or excess demand? (Explain briefly.) At this price, how many units would trade, and what would be the total surplus (sum of value-cost differences for traded units)?

5. (non-mechanical) For the setup in problem 2, how would it be possible for 8 units to trade if prices for some trades could be higher than prices for others? Hint: think about how you might put buyers and sellers into separate groups to get more units traded.

6. For the 8 units traded at different prices, as in the answer to problem 5, what would the total surplus (sum of value-cost differences) be? What would the efficiency measure be?

Index