

1 The Emergence of Experiments in Economics

There is a property common to almost all the moral sciences, and by which they are distinguished from many of the physical; that is, that it is seldom in our power to make experiments in them.

Mill (1836), cited in Guala (2005, p. 2).

This statement by John Stuart Mill, or similar remarks, introduces virtually all texts on the methodology of experiments in economics. At the time, and for a long time after that, controlled experiments in the social sciences, and especially in economics, were considered impossible to conduct; it appeared that experiments were reserved to the natural sciences, and that the testing of social and human behaviour in the framework of a controlled experiment would prove completely unworkable. Nowadays, experiments are a widely accepted means of generating knowledge in economics. Among many examples, it is shown by the fact that experimental or behavioural economics is part of the graduate programme of most universities, there are many books, handbooks and textbooks focusing on the field, and even a well-recognised academic journal (*Experimental Economics*) is specialised on research using this method.

Before moving on to a detailed discussion of why and how laboratory experiments are performed in economics, we will explore this intriguing trend. What happened between the time experimental economics first came into existence and when it finally became an established member of the community? We will start by highlighting the progress of experimental methods in economics, from an area that was thought impracticable, meaningless or uninteresting, to an accepted and widely used process in economic research. In describing the reasons why there was such a sudden change of interest in and attitude towards experiments, we will examine some of the very first examples of experiments in economics. These examples are interesting not only from a historical point of view, but also because they underscore the main reasons for the change and how experimental economics has grown since – both in terms of the research questions that are addressed and in the type of answers it provides. These will be followed by three more recent examples which illustrate what the research programme has become today – a unified and also very diverse area of study.

The most obvious and powerful unifying factor of all works using laboratory experiments is, in fact, the methodology applied: a controlled environment allowing use of the observed behaviour of human beings to produce knowledge about economics. As the last section will show, a thorough study and presentation of this methodology requires

a wide-ranging knowledge of economic theory as a whole, and its relation to different application fields, analytical tools and approaches. It will soon become clear that no single textbook can possibly cover all these aspects: this chapter will offer a road map of everything this book is unable to cover, or can only cover in part. Perhaps more importantly, this chapter will try to convince you that in order to fully understand the rationale, contribution and practical lessons of the results generated by experiments in economics, the first step is to be aware of the choices of methodology and the reasoning behind them: this is what this book is all about.

1.1 The End of a Long-Standing Regretful Impossibility

Even if experiments in economics were considered impossible for a long time, they were nonetheless the object of considerable wishful thinking. If experiments could be implemented, they could be designed and put in place in order to provide empirical evidence and serve as a basis to enhance theory. This is implicitly acknowledged in a celebrated remark made by Friedman, ‘We can seldom test particular predictions in the social sciences by experiments explicitly designed to eliminate what are judged to be the most important disturbing influences’ (Friedman, 1953, p. 10). Experiments in the social science are seen as a very attractive, though impossible, way of testing theories. If feasible, experiments would allow researchers to neutralise all forces driving behaviour that are outside the scope of the theory. In that case, experiments would help elicit the empirical content of theory, and therefore identify the main driving forces of behaviour. This opinion was shared by many eminent economists long after 1953. In their groundbreaking principles textbook, Samuelson and Nordhaus noted that ‘economists cannot perform the controlled experiments of chemists or biologists because they cannot easily control other important factors’ (Samuelson and Nordhaus, 1985, p. 8). All of the remarks cited above show quite clearly how recent the appearance of experimental economics as a *bona fide* field of study is and also underline how desirable experiments are for research. Fortunately, the long-standing and powerful belief in the impossibility of experiments in the social sciences, however regretful, is now a thing of the past.

As a matter of fact, in a later edition of their textbook (which appeared less than ten years later) Samuelson and Nordhaus had already adopted a new and different mindset: ‘Experimental economics is an exciting new development’ (Samuelson and Nordhaus, 1992, p. 5). Between these two editions, economists had managed to set up experiments similar to the ones conducted in the natural sciences. But, even more importantly, the results generated by these experiments began to be considered by an increasing number of specialists to be sound empirical evidence.

From then on, the pace and scope of the changes taking place increased so rapidly that today the situation stands in sharp contrast with the earlier views expressed above. This phenomenon is illustrated, for instance, by the rise in the rate of academic publications related to experimental economics over the years. Figure 1.1 shows the results of a survey carried out by Noussair (2011) concerning the percentage of articles including experiments that have appeared in major academic economic journals. The survey

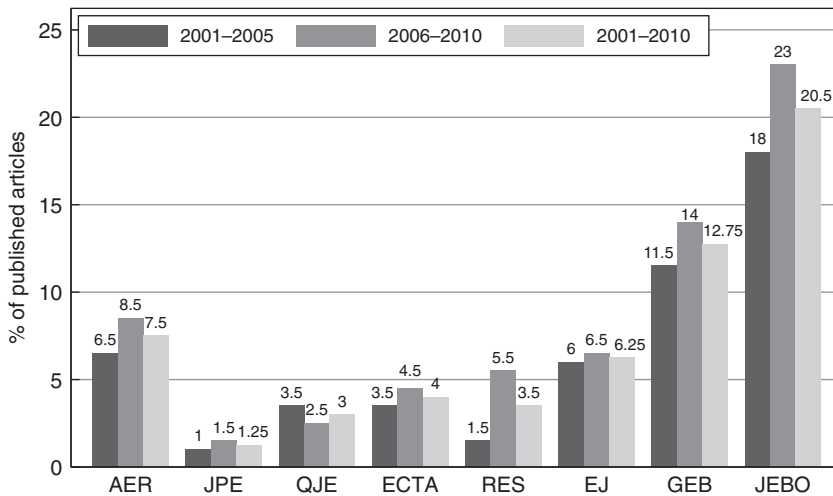


Figure 1.1 Trends in academic publishing in experimental economics

Note. Percentages of experimental articles from those appearing in the journals: *American Economic Review* (AER), *Journal of Political Economy* (JPE), *Quarterly Journal of Economics* (QJE), *Econometrica* (Ecta), *Review of Economic Studies* (RES), *Economic Journal* (EJ), *Games and Economic Behaviour* (GEB), *Journal of Economics, Behavior and Organization* (JEBO).

Source: Noussair (2011, p. 8).

covers the top five journals (AER, JPE, QJE, ECTA, RES) which experts acknowledge as the leading supports in the field; three other journals were added to the list: EJ, GEB and JEBO. These are more specialised and/or lower-ranked journals, but which are, nonetheless, highly influential and open to experimental works. The chart shows the change in the rates from 2001–2005 to 2006–2010. The first ten years of the new millennium saw a slight increase in the percentage of articles in the sample. More importantly, the share of experimental papers is very significant in most of these leading journals: from 2% to 7% in the top five journals, and from 5% to 20% in the more specialised ones. This a clear indication of the growing acceptance and recognition of this type of work by the academic community.

The four experimental economists who have been awarded the Nobel Prize in Economics in the first decades of the new millennium, who we will come across a number of times in this book, are another example of this recognition. In 2002, Vernon L. Smith and Daniel F. Kahneman were the joint recipients of the Nobel Prize in Economics. Smith was thus acknowledged as one of the founders of experimental economics and as someone who contributed to establishing it as a conclusive method. The main justification for the award was the introduction of the methodology per se (they received the prize ‘for having established laboratory experiments as a tool in empirical economic analysis, especially in the study of alternative market mechanisms’). In terms of contributions, the field is seen as interdisciplinary in nature, with Kahneman receiving the prize ‘for having integrated insights from psychological research into economic science, especially concerning human judgement and decision-making under uncertainty’. Ten years later, another renowned experimentalist, Alvin Roth, was also granted the

Nobel Prize. But this time, the co-winner was Lloyd Shapley, a pure theorist. Together they were recognised ‘for the theory of stable allocations and the *practice* of market design’. It goes without saying that the Smith and Kahneman contributions are of major importance to the discipline, and that these three Nobel Prizes in themselves are convincing proof that experiments have been widely accepted as part of the field. But there is an interesting change in nature between the two prizes: while the first Nobel Prize was awarded for the methodological advance itself, the acknowledgement of Roth’s contribution was based on actual laboratory results using the toolbox of experimental economics and applied to research issues that are at the core of economic theory. This is further evidence of the wide acceptance of experimental economics by the academic community. Last, Richard Thaler was awarded in 2017 for having incorporated ‘psychologically realistic assumptions into analyses of economic decision-making’. Richard Thaler showed how experimental methods are particularly meaningful for uncovering deep psychological phenomena such as mental processes, self-control behaviour and social preferences. The award also underlines his contribution to public policies based on nudges (see Chapter 9). This is further evidence of the wide acceptance of experimental economics by the academic community, with results from the laboratory now being seen as useful in order to better design choice architectures.

In contrast with the quotes that opened this section, in which experiments were regarded with substantial scepticism, there is now substantial evidence that experimental economics has become a well-established and widely accepted empirical method. One may wonder how an entire new field has managed to surface in such a short period of time. As a first step towards a better understanding of how this change came about, we will show in the next section that this, in fact, was not the case at all: experiments in economics have existed for a long time, producing results that are much in line with the works that appear nowadays in leading publications. It appears that the reason for the lack of experiments in economics comes not so much from their practical impossibility, but rather from the main focus of academic research at the time. Since then, a change in focus occurred towards questions that are closer and closer to the kind of issue that experiments are well suited to investigate.

1.2 Why Such a Change: Two Early Examples

The two examples below are among the best known of the early experiments. They illustrate the state of infancy of experimental economics at the time, although they are now regarded as important and insightful contributions to economic knowledge.

1.2.1 How Do Competitive Markets Work?

In 1948, Harvard Professor Edward Chamberlin organised a game with his students. The aim was to replicate the functioning of a market in perfect competition with rational agents as closely as possible. Students were randomly assigned a card, which made each student either a seller or a buyer. In addition, the card displayed a price for a hypothetical

good to be sold or bought. For students playing as sellers, this price referred to the minimum price at which they were ready to sell. For the buyers, this price indicated the maximum price they were willing to pay to obtain the (hypothetical) good. Afterwards, the students walked freely in the classroom and bargained with their colleagues to either buy or sell the good. Once a deal had been made, the students came to Chamberlin's desk to report the price at which the good had been sold.

In this framework, economic theory predicts outcomes according to the two curves depicted in Figure 1.2, where the supply and demand curves were drawn based on the prices distributed to students – i.e. how many students were willing to buy or sell at each possible price that appears on their card: a 'induced values' design. The game is a textbook example of a market: the demand curve is decreasing in price, whereas the supply curve is increasing. The market equilibrium determines the actual price that should arise from strategic interactions, as well as the resulting quantities exchanged on the market; the unique stable price is the one that clears the market, in such a way that demand meets supply. This point is an equilibrium not only because the two sides happen to be equal, but more importantly because it is the only state of the market in which everyone agrees to stay – there is no possibility of doing better at the individual level by moving out of this situation. For any other price, there is either excess supply or excess demand, in which case either suppliers (sellers) or consumers (buyers) can be in a better situation by moving to another price level. There are thus strong reasons to believe that the equilibrium should result from real interactions in this particular environment.

Surprisingly enough, Chamberlin obtained the results reported in Figure 1.3 based on the actual behaviour of his students. The dashed line depicts the average price at which students traded their goods during the experiment: it is far below the straight line, or the competitive equilibrium price. There was also a huge variation in the actual prices,

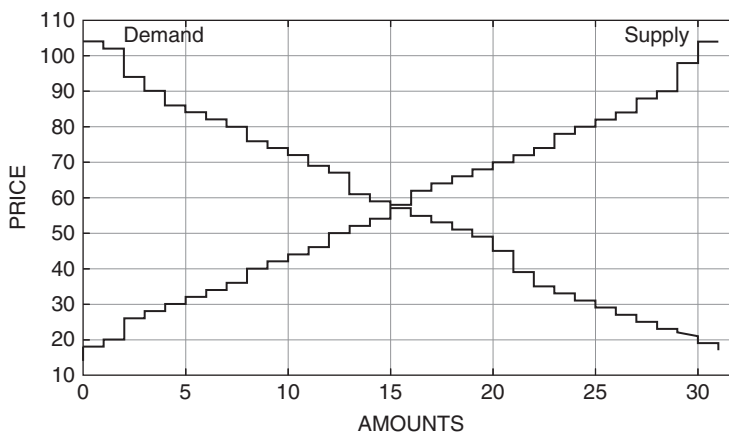


Figure 1.2 Market equilibrium in the Chamberlin (1948) experiment

Note. The figure shows the theoretical equilibrium of the market implemented in the laboratory – at the intersection of the (increasing) supply function and the (decreasing) demand function.

Source: Chamberlin (1948, p. 97, Figure 1).

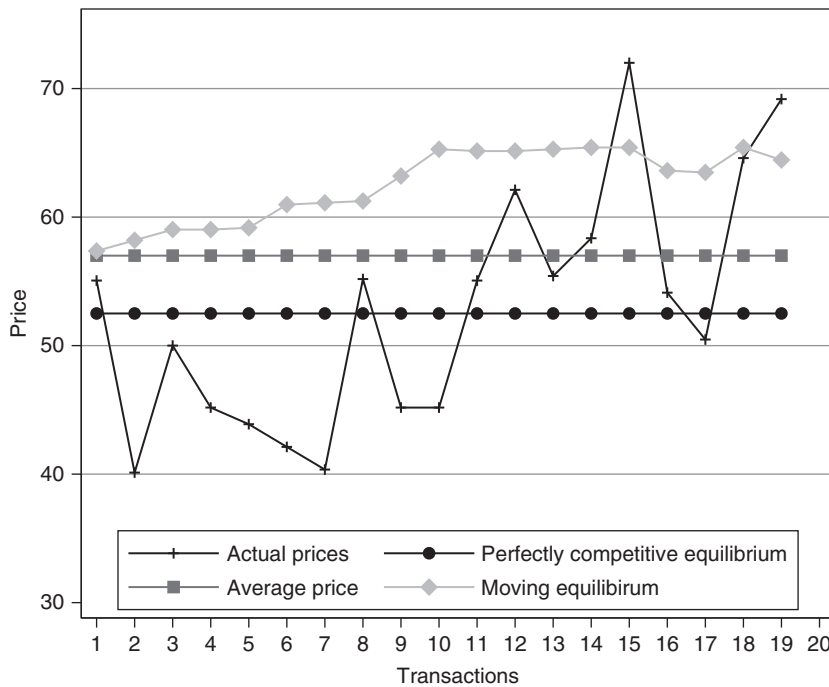


Figure 1.3 Observed behaviour in the Chamberlin (1948) experiment

Note. For each transaction in abscissa, the figure shows the actual price observed in the experiment as well as a recall of the theoretical equilibrium described in Figure 1.2.

Source: Chamberlin (1948, p. 101, Figure 3).

which are represented by the curving line. In addition, the equilibrium volume of trade is higher than what the theory would have predicted. Actual behaviour in this environment thus strongly departs from what economic theory expects, leading Chamberlin to conclude, ‘Perhaps it is the assumption of a perfect market which is “strange” in the first place’ (and interpret this as a support for his monopolistic competition model). This result is not, however, the end of the experimental story of markets.

Vernon Smith (who, as mentioned above, was subsequently awarded a Nobel Prize) was one of Chamberlin’s students and participated in his classroom experiment. Around fifteen years later, in 1962, he decided to replicate Chamberlin’s experiment, but with various changes in the environment – aimed at replicating what Smith thought were important actual driving forces of a competitive market. As in Chamberlin’s experiment, each student received a card, making him either a buyer or a seller. This card also gave the student a reservation price: the price above which a buyer would not buy, and below which a seller would not sell. The changes implemented as compared to the seminal experiment are as follows. First of all, instead of having bilateral bargaining (or, at most, discussions in small groups) between students, the announcements of offers and demands become public, meaning that buyers and sellers could call out their offers in the room so that everybody could hear. This is aimed to make the information on prices public, so as to mimic what is achieved by an auctioneer receiving and distributing all

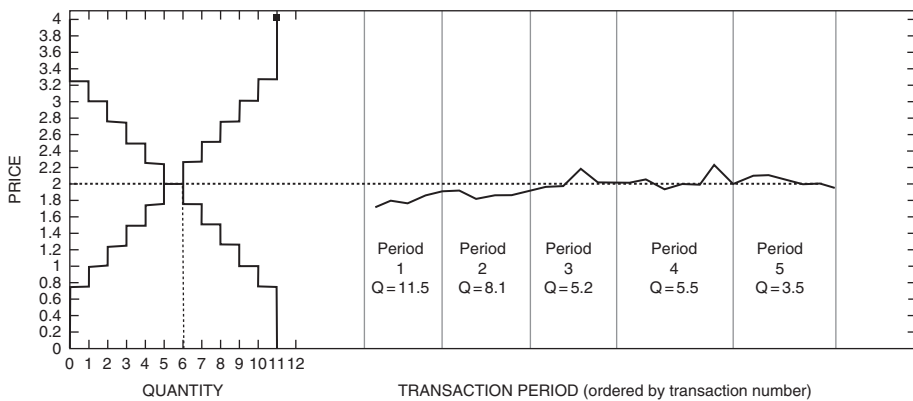


Figure 1.4 Predicted and observed behaviour in the Smith (1962) replication

Note. The left-hand side shows the theoretical market equilibrium – at the intersection of the (increasing) supply function and the (decreasing) demand function. The right-hand side shows the price and number of transactions in each market period.

Source: Smith (1962, p. 113, Figure 1).

offers. Second, the market experiment was repeated over several periods, and allowed the students to better understand the functioning of the market, hence getting closer to market behaviour of professional market traders.

Figure 1.4 reports the observed behaviour and theoretical predictions of the Smith experiment. The theoretical market plotted on the left-hand side shares the same features as the one implemented by Chamberlin. The curve on the right-hand side shows the prices at which market clears for five market periods. The contrast with the previous results is drastic: the observed prices smoothly converge towards the equilibrium price, and the number of transactions (reported on the bottom part of the graph) converges to the equilibrium quantity equal to 6.

Beyond the seminal insights about how the market works, these series of experiments help to describe the methodological issues behind experimental results. Both experiments aim to replicate competitive markets, but with different implementation choices. The best environment to describe markets is a matter of judgement, and the theoretical conclusion drawn will be entirely different whether one or the other experiment is believed to best capture the important features of the economic phenomenon. At the same time, the implementation differences between the two experiments also inform about the key features that explain behaviour in a market situation: the extent of information buyers and sellers receive, for instance, seems to be a critical driving force. Beyond rejection/support of the prediction, the experiment thus informs theory by highlighting the salient dimension to be taken into account. Lastly, as the Smith experiment clearly shows, it is not always the case that the theory is necessarily wrong or that experiments are designed expressly to reject the behavioural assumptions behind the theoretical results (as is sometimes taught, mainly by some academics who view experiment results with scepticism): in this case, experiments serve more to identify the circumstances under which these assumptions are actually accurate.

Table 1.1 The choice sequence of the Allais paradox

A or B?	Option A	Option B
	100% chance of winning 1 million	10% chance of winning 5 million 89% chance of winning 1 million 1% chance of winning nothing
C or D?	Option C	Option D
	11% chance of winning 1 million 89% chance of winning nothing	10% chance of winning 5 million 90% chance of winning nothing

Note. Each respondent was asked to make both choices in turn.
Source: Allais (1953, implemented in 1952).

1.2.2 **Choice Consistency in Risky Decisions**

The second example focuses on individual decision-making, rather than on strategic situations. During the annual conference of the American Economic Society held in New York City in 1953, Maurice Allais presented the economics professors attending the conference – especially those specialised in game theory and decision theory – with two binary choices. Respondents were shown Table 1.1 and asked to choose either A or B, and then either C or D.

Based on the axiomatic framework of decision theory, the first choice and the second choices are strongly related – although the choice between the two options per se is a matter of preferences that nobody can predict. To understand the link between the two decisions, let us first put aside the 89% probability of winning one million – in situations A and B – or nothing – in situations C and D. Apart from this 89% probability, both situations A and C have the same probability (11%) of winning one million. Similarly, situations B and D offer the same expected outcome: nothing with a probability equal to 1%, and five million with a probability of 10%. As a result, still disregarding this 89% probability, an individual who prefers A over B (B over A) should also prefer C over D (D over C). You can note that the outcome that results from the 89% probability is exactly the same for A and B on the one hand, and C and D on the other. Consequently, it only comes down to the addition of an identical outcome for each pair of situations: one million for A and B, nothing for C and D. It sounds reasonable to assume that this should not affect the preference ordering of consistent decision-makers.¹ Because of this very clever feature in the way situations are built, elicited choices provide a test of consistency: depending on individuals’ unknown preferences, either A and C, or B and D, should be picked together; no other combination can be rationalised with classical decision theory. Using these choice situations, Allais was successful at tricking the economists at the conference. As he expected, 45% of the leading theorists (including Savage, one of the leading researchers in the field) to whom Allais submitted the choice

¹ This property of preferences is named the “independence axiom” in decision theory, which implies that if there are two different gambles and one is preferred to the other, then mixing them with another identical gamble should not alter the order of the preferences. This axiom is the one violated by the results of this experiment, which is now known as the common consequence or Allais paradox.

opted for A against B, but D against C. Almost half of the respondents, who were all well versed in economic and decision theory, and some specialised in decision theory, failed to pass the consistency test associated with the two successive choices. A key feature of this experiment is that it is designed in such a way that there is a unique relationship between one, clearly identified, theoretical assumption driving the predictions and the choices available. Therefore, observed behaviour challenges not only theory, but, more importantly, the specific feature of theory that fails to describe behaviour. Beyond simple rejection (which is unambiguous given the magnitude of the result and the sample pool from which it was obtained) it provides a guide to the particular assumptions that have to be reworked so that they correspond to the real driving forces behind behaviour. In the Allais paradox, two features of the available options are of particular interest. On the one hand, certainty generates a strong attractiveness for option A. On the other, the change in probabilities appears to be quite small between options C and D. These two features of behaviour under uncertainty are central in theories that rationalise behaviour in the Allais paradox (Quiggin, 1982; Kahneman and Tversky, 1979).

1.2.3 Why Was There Such a Fast and Sudden Change?

These early experiments marked the beginning of a new field, which has made rapid gains in terms of both acceptance and popularity over the last decade. But many years went by between the time of those first experiments and the time when the economic community truly started paying attention to them. Until recently, experimental economics was thought of as unworkable or of no meaningful importance. What was it, then, that suddenly made the experimental method so widely accepted?

As shown by the two previous examples, this was not a matter of feasibility. Both experiments were published in very good journals and existed when some of the quotes opening this chapter were written. Experimentation was thus already a possibility. In fact, it has always been quite straightforward to test results from decision theory or game theory in an experimental setting. It simply amounts to having people make choices within a simple set of rules describing the decision-making environment. The most drastic change was in fact the change in the kind of questions, which in the 1970s and 1980s economics began to focus on, with a growing importance put on these two theoretical tools.² In the middle of the twentieth century, economics was set in the context of a beautiful model of how the entire economy worked and how all the agents in the economy, as a group, made decisions in the present and for the future. This environment was so complex and all-encompassing that the empirical relevance of behavioural assumptions was obviously not a primary concern. But as economics moved away from this representation, more and more attention began to be given to the forces behind individual and strategic decision-making. Microeconomics became one core focus of economic analysis, making an intensive use of game and decision theory. What were considered revolutionary issues at the time have now become orthodox, and the rise of experimental economics was concurrent with the fall of general equilibrium theory. The reinforcing of

² See Fontaine and Leonard (2005), in particular Chapter 3, for an insightful review of these trends.

the economic representation of human behaviour, along with clear-cut definitions of the environment, has now made the long-held dream of testing economics in the laboratory an achievable goal.

The role of experiments in the history of economics helps better understand what experimental economics is all about. First, experiments and economic theory go hand in hand: experiments are about assessing the empirical relevance of the behavioural content of economic models. They are not in contradiction with economic theory, but rather serve as a complement to it. Economic theory provides a deep and subtle understanding of how the economy works when decisions are taken by the *homo æconomicus*. Experiments rather involve a *Homer æconomicus*: the ordinary Joe, endowed with an average level of cognitive and social skills – rather than unlimited computational abilities – under the influence of psychological and environmental factors – rather than driven by a well-defined preference functional.³ They thus allow us to measure whether *homo æconomicus* and *Homer æconomicus* lead to similar or substantially different outcomes in a given economic situation. Second, as the two examples cited above show, these two kinds of people, the *homo æconomicus* and ‘real’ human beings, are not strangers to one another: they sometimes behave differently, calling for a different theory (rather than different people), but there are also many important situations in which the two behave as if they were one and the same. Why and when they do is one of the key questions that remains to be answered.

1.3 The Research Programme: Three Examples

We conclude this quick overview of the recent history of experiments in economics with three examples drawn from a more recent literature. Although chosen at random (with bias) among many other similar studies, these examples clearly illustrate the current state of the art in the experimental field, and the way it helps elucidate what human beings and economic theory – *Homer æconomicus* and *Homo æconomicus* – have in common, and how they differ. To a large extent, the current answer is similar to the main lesson we learned from the early examples described above.

We first present the prisoners’ dilemma (PD), a well-known example of the discrepancy between game-theoretic results in a simple environment, and the behavioural patterns actually observed. This example also shows that, while game theory alone has trouble explaining behaviour – typically, without reference to more general factors related to economic agents’ social environment – it is in fact quite effective in predicting changes in behaviour. The second example shows that experimental economics can help significantly in this aspect as it can easily address difficult questions about the basics of economic rationality. The centipede game is a typical example of a simple experiment that calls into question some common principles of rationality. Lastly, we proceed to a more complicated game, a zero-sum game with incomplete information, in which one would expect the gap between economic theory and observed behaviour to be larger

³ The terminology is due to Hall (2005); see e.g. Beggs (2013); Hall (2014) for a full statement of the parallel between economics and the Simpsons.

than in any other context. The example shows that this foregone conclusion is definitively not applicable in this case. These three examples serve as a tour of the type of research question addressed thanks to experiments in economics and of the variety of answers it offers.

1.3.1 Nash Equilibrium and Pareto Efficiency

The first example is that of a *non-cooperative game*: a game in which outcomes are determined by the decentralised and independent actions of players. Figure 1.5 presents the payoffs each of the two players gets according to the actions they choose. It is a simultaneous-move game, as each player decides without knowing what the other one is doing. This type of representation of a game will be used often in this book. For readers who might not be fully familiar with it, we will take the opportunity here to describe it step by step.

According to the normal form representation in Figure 1.5, each player can choose between two actions. Player 1 is the row decision-maker, and Player 2 is the column decision-maker. Player 1 chooses either Top or Bottom, Player 2 either Left or Right. Together, both players' actions determine the outcome of the game: the state of the world resulting from all the players' actions. The numbers in the matrix show the payoffs linked to each of the four possible outcomes for each player. In each cell, the number on the left is the payoff Player 1 gets in this particular outcome, and the one on the right the payoff Player 2 gets. For example, if Player 1 plays Top and Player 2 chooses Right, then Player 1 loses 10 and Player 2 earns 10.

A quick inspection of Figure 1.5 shows that one outcome seems intuitively preferable: if the players choose Top of Left, they reach an outcome that maximises what they collectively get. It is a Pareto-dominant outcome: that particular outcome makes it impossible for one agent to improve his lot by unilaterally modifying his action without making the other player worse off. However, this outcome is not sustainable when the actions are decentralised and non-coordinated. This is so because, given the Pareto-dominant situation, both agents have an incentive to deviate: given the action of Player 2, Player 1 can earn more by playing Bottom than Top against Left, and similarly Player 2 can earn more by playing Right against Top rather than Left. Because of these individual incentives to move away from the Pareto-dominant outcome, the equilibrium coincides with the worst outcome of the game: that which occurs when Player 1 chooses Bottom and Player 2 chooses Right. This is a Nash equilibrium because there is no longer any

	Left	Right
Top	5; 5	-10; 10
Bottom	10; -10	-5; -5

Figure 1.5 Table of payoffs in a non-cooperative game

individual incentive to deviate – none of the players can be better off by moving away from the equilibrium strategy when the others are playing it.

The Nash equilibrium of the game, Bottom-Right, is the outcome that is predicted to occur from uncoordinated simultaneous decisions. Because it does not coincide with the Pareto-dominant situation, this game is a textbook example of the failure to reach an efficient outcome via non-cooperative decisions. It is often called the prisoners' dilemma game, in which case the moves are 'to denounce' or 'not to denounce' for two prisoners who are separately offered leniency if they provide information about the crime they committed together. This strategic framework can be applied to a great many economic situations. Collusion between firms on markets is a typical example of the dilemma of cooperation and defection (which will be studied in length in Chapter 4, Section 4.4.2). Collusion occurs when firms agree to set the market price to a level higher than its competitive value. All firms prefer the collusive outcome, as profits are higher. But each firm has a strong temptation to slightly decrease its price so as to make even higher profits, at the expense of others. This incentive to deviate from the collusion agreement is a natural force against the ability to sustain a non-competitive equilibrium. Another example of non-cooperation when cooperation would be optimal is the Kyoto Protocol, an international agreement which aims to commit countries to reducing their greenhouse gases. A Pareto-optimal outcome would be that all countries sign the agreement. Nonetheless, countries have an incentive to let the other countries sign and to free-ride, thus benefiting from the reduction in greenhouse gases without having to pay the price of the treaty.

Hundreds, if not thousands, of experiments have been run to assess the empirical relevance of this analysis. As an example, Figure 1.6 presents the results of one of the earliest experiments of this type, conducted by Cooper et al. (1996). The x-axis represents each of the ten different periods of the game, while the y-axis depicts the frequency of the cooperative play (i.e. when the collectively optimal, but not individually rational, actions are chosen) when the action leading to the efficient outcome is chosen. The upper curve represents the frequency of cooperative play in the case of a prisoners' dilemma game with repeated interactions, where the same two players play together ten times. The lower curve represents the outcome with different partners for ten periods, each game being a one-shot game. Both curves show a departure from theoretical predictions. Theory predicts a 0% rate of cooperation in the game. It is far from the observed patterns not only in the repeated games – which do not, in the strict sense, implement the model – but also in the one-shot games. For example, in the first period, about 60% of the subjects decided to cooperate in the case of finitely repeated games, but around 35% of the people did so in one-shot games. At the same time, it is not true that these results fit with a view of human behaviour only driven by the well-being of everybody and disregarding self-interest. Free-riding behaviour, based on the temptation to increase one's payoffs at the expense of the other players, accurately describes the results in 70–50% of observed outcomes. Because these two kinds of behaviour (cooperation and deviation) are widespread, both should be accounted for by any accurate theoretical representation. As a result, neither the Nash equilibrium, nor alternative motives leading to full cooperation, are enough to account alone for the observed behaviour in the prisoners' dilemma game.

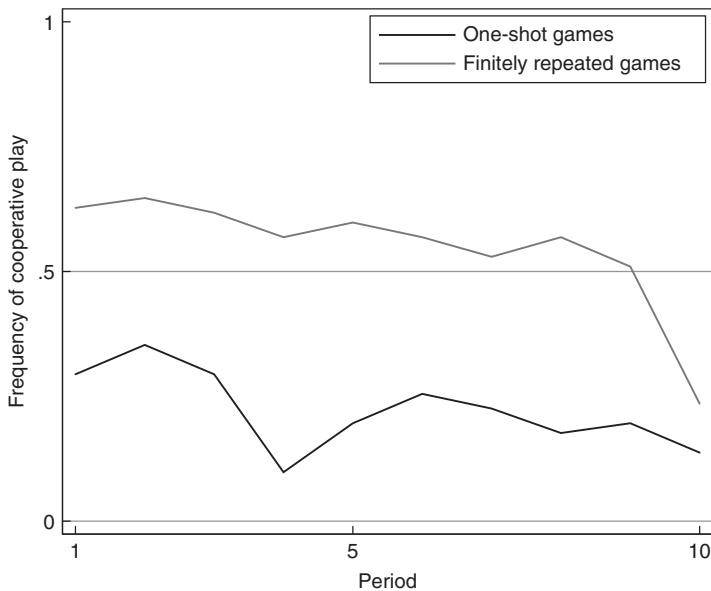


Figure 1.6 Empirical behaviour in prisoners' dilemma games

Note. The figure reports the share of participants who decide to cooperate in each of the ten periods of the game.

Source: Cooper et al. (1996, p. 199, Figure 1).

1.3.2 A Simple Two-Player Sequential Game

The previous example focused on a simultaneous-move game, in which the players decide without knowing what the others will be doing. Another branch of game theory studies behaviour in sequential-move games, in which the players decide one after the other. The big change in terms of strategic interaction is that each player now observes what the other did before choosing an action. Figure 1.7 provides a well-known example of such a game, introduced by Rosenthal (1981) as the *centipede* game. The structure of the game is quite simple. Two players alternately get a chance to take the larger portion of a continually increasing pile of money – the number on each node indicates which of the two players has to decide, with the two payoffs being those experienced by each of the two players respectively if the game ends at this point. For Player 1 payoffs are given in the first row, for Player 2 payoffs are given on the second row. The amount keeps increasing as long as the players continue to play (denoted P in Figure 1.7 and Figure 1.8). But as soon as one of the two players decides to take (denoted T in Figure 1.7 and Figure 1.8), they get a larger portion of the pile while the other gets the smaller part. The trade-off is not easy to resolve from an intuitive point of view: conditional on the game continuing, it is always better to go as far as possible along the tree (the original form had 100 nodes, hence the name *centipede*), but at the same time each player wants to be the one who stops the game. The question that remains open, then, is when the players will stop and at which stage.

The way game theory resolves this trade-off is, in a sense, even less intuitive than this simple explanation suggests. The key point to note is that the number of steps in the

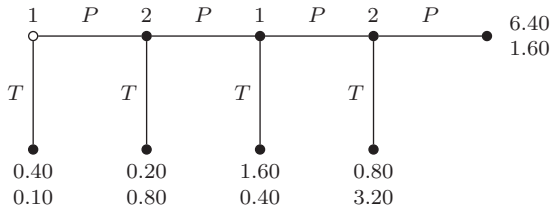


Figure 1.7 A simple four-moves sequential game

Note. Each of the two players (1 or 2) decides in turn at each node to either Pass or Take. For each state, the payoffs of Player 1 appear on the first row, the payoffs of Player 2 on the second row.

Source: McKelvey and Palfrey (1992, p. 806, Figure 1).

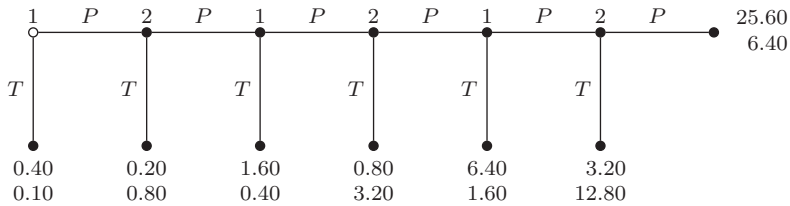


Figure 1.8 A six-moves centipede game

Note. Each of the two players (1 or 2) decides in turn at each node to either Pass or Take. For each state, the payoffs of Player 1 appear on the first row, the payoffs of Player 2 on the second row.

Source: McKelvey and Palfrey (1992, p. 806, Figure 2).

game (four in the example) is known for sure from the beginning. The usual approach to this type of situation is to predict that the players will play in such a way that actions in each sub-game (i.e. the sub-tree that extends from any node to the end) is a Nash equilibrium. Because of this property, the equilibrium can be elicited through backward induction. Starting from the terminal node of the game, the equilibrium behaviour is relatively straightforward: the last player will decide to take, because it earns more than to pass and there is no point in waiting. For Player 1, in the node just before, it means that the decision actually faced is between taking now or having Player 2 take at the last stage. But then the best thing to do is to take at this node so as to avoid letting the other player take at the following one. And this reasoning applies to all the steps leading backward taken one after the other. The result of this reasoning would be the *sub-game perfect equilibrium*, where the first mover takes at the very first node and the game stops. What is startling in this result is that the outcome is not predicted to depend on either the rate at which the pie grows from one step to the other, or on the number of steps – as long as the number is known right from the beginning.⁴

This striking prediction was first tested against actual behaviour by McKelvey and Palfrey (1992). When the participants in their experiment were asked to play the four-move game described in Figure 1.7, only 7% of them actually played according to the sub-game perfect equilibrium, stopping at the very first node. The top part of Table 1.2

⁴ See Reny (1993); Aumann (1995); Ben-Porath (1997); Aumann (1998) for theoretical attempts to weaken this paradoxical result, and Chapter 4, Section 4.4.2, for a more detailed discussion of finite- and infinite-horizon games.

Table 1.2 Observed continuation decisions in centipede games

		Session	N	f_1	f_2	f_3	f_4	f_5	f_6	f_7
Four Move	1	(PCC)	100	.06	.26	.44	.20	.04		
	2	(PCC)	81	.10	.38	.40	.11	.01		
	3	(CIT)	100	.06	.43	.28	.14	.09		
	Total	1-3	281	.07	.36	.37	.15	.05		
High Payoff	4	(High-CIT)	100	.15	.37	.32	.11	.05		
	5	(CIT)	100	.02	.09	.39	.28	.20	.01	.01
Six Move	6	(PCC)	81	.00	.02	.04	.46	.35	.11	.02
	7	(PCC)	100	.00	.07	.14	.43	.23	.12	.01
	Total	5-7	281	.01	.06	.20	.38	.25	.08	.01

Note. Actual behaviour in the four-move (upper part) and six-move (bottom part) centipede game. N denotes the number of subjects, each column f_i provides the share of subjects who decide to take at the i^{th} node.

Source: McKelvey and Palfrey (1992, p. 808, Table IIA).

shows the full distribution of the share of subjects who stopped at each node of the game (denoted f_i for the i^{th} decision stage). While it is true that the subjects were not anyway near to playing the sub-game perfect equilibrium, at the same time few of them reached the last stage of the game – less than 5% did. From this evidence, the question remains open: what is it that makes subjects decide to stop or go on?

To help answer this question, McKelvey and Palfrey (1992) consider a second experiment that implements the six-move centipede game displayed in Figure 1.8. The results are displayed in the bottom part of Table 1.7. From 7% in the four-moves game, the share of subjects who play the sub-game perfect equilibrium is now almost 0 – only two subjects out of 281 do so. But again, the distribution of subjects according to the node at which they decide to stop is not concentrated at the end. Half the subjects rather decide instead to stop at node 5 or 6, two steps before the last stage.

From these two examples, it appears that sub-game perfectness clearly fails to predict behaviour in the extremes. But at the same time, this theory accurately mirrors the trade-off people face in this type of situation: as subjects reach a node closer and closer to the end, it becomes more and more difficult for them to maintain a decision to pass, and more and more likely that the decision they will take as the game proceeds is to stop a few rounds (two to four) before the end.

1.3.3 The Use of Private Information

The first two examples were simple games whose results challenge theory in one way or another. As a third example, we will move on to another quite different environment in which both the rules and the strategies are far more complicated. Its full name is a zero-sum repeated game with incomplete information – each part is explained in turn below. Figure 1.9 shows the stage games of two different versions of the game. We first focus on the non-revealing (NR) version of the game – the difference with the fully revealing (FR) version will be described later.

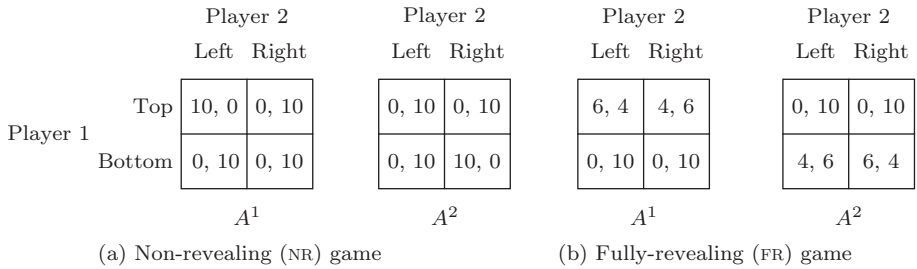


Figure 1.9 Payoff matrices of two zero-sum games

It is a zero-sum game because the payoffs are such that everything that is won by one player is lost by the other – as opposed, e.g., to the prisoners' dilemma game – so that concerns about the situation of other players have no influence over the results. All stage games shown in Figure 1.9 involve two players and two actions. Player 1 chooses either Top or Bottom, and Player 2 chooses either Left or Right – both players decide simultaneously. The numbers in the matrix represent the payoffs of both players after they have chosen their move. In matrix A^1 of Figure 1.9.a, for instance, '10, 0' indicates that Player 1 gets 10 and Player 2 receives 0 if Top/Bottom is played.⁵ Inspecting the payoff tables, the game is straightforward to play for both players, because they both have a dominant strategy, i.e. an action which is preferable whatever the action chosen by the other player. For instance, in this matrix A^1 , Player 1 does better by playing Top rather than Bottom, and Player 2 by playing Right rather than Left – whatever the decision of the other player. Similarly, in matrix A^2 , choosing Bottom is a dominant action for Player 1 and Left is a dominant action for Player 2.

The information structure of the game makes it more interesting than a complete-information zero-sum game. In fact, a random draw (with equal probability) decides on the 'state of the world' before any decision is made. This state of the world is the payoff matrix, either A^1 or A^2 , that players are facing. There is incomplete information (on one side) because players are asymmetrically informed about the result of this draw: only Player 1 is given this information. First consider the situation in which the stage game is played only once. Player 1 is privately informed of the consequence of each action and can thus pick up the dominant action of the matrix that has been drawn. Player 2, by contrast, needs to decide without being aware of the state of the world, and will thus randomise between Left and Right. But this information structure in fact becomes interesting when the game is repeated – players face the stage game together several times, and the state of the world is drawn once for all at the beginning. In this context, Player 2 can infer some information about the state of the world from the observed decisions of Player 1.

To see it more clearly, suppose you are Player 1 facing the stage games of Figure 1.9.a and knowing which state of the world you, and the other player, are in. You have to

⁵ The sum of players' payoff is positive rather than equal to 0, but since the sum is constant across decisions, it is conceptually equivalent to a zero-sum game.

Table 1.3 Theoretical predictions in the non-revealing and fully revealing games

	Value of the game, v_t if t is						Optimal use of information
	1	2	3	4	5	∞	
FR	5.00	4.50	4.33	4.25	4.20	4	Fully revealing
NR	5.00	3.75	3.33	3.21	3.07	2.50	Non-revealing

Note. Theoretical predictions on behaviour in the NR and FR games.

Source: Jacquemet and Koessler (2013, p. 110, Table 1).

choose between Top and Bottom and you know the game will be repeated. You also know that Player 2 would like to play Right in A_1 and Left in A_2 . First imagine that you decide to use your dominant action: you play Top if A_1 is drawn and Bottom if it is A_2 . If this is an equilibrium strategy, then Player 2 knows this is how you react to the draw: observing Top delivers perfect information to Player 2 that A_1 has been drawn. At the next stage, Player 2 will thus play Right. But the combination Top, Right is clearly not in your interest, since you get 0: by revealing your information you no longer benefit from it. The other options for you are either not to use your information at all (deciding with equal probability between the two decisions as if you did not receive the information about the draw) or to use it only slightly, by playing the dominant action a bit more often than the other. The game thus features a trade-off in the way private information is used by Player 1, and how beneficial it is to hold such private information (the only exception is the last stage of the game, when the dominant action will always be chosen, because there is no longer any possibility to exploit the signal contained in your choices). This kind of game thus allows one to study the extent of the use of information, and the value of private information, i.e. how much more the informed player is able to earn.

The equilibrium strategies depend on two crucial features of the game: the length, denoted T , which is the number of stages during which both players play in the same matrix, and the structure of payoffs. The payoff structure we just described (the one shown in Figure 1.9.a) is called a non-revealing game, because the optimal strategy for Player 1 is to not reveal their private information in all stages but the last one: at equilibrium, it is best for Player 1 to behave as if the information were not available and the randomly drawn matrix were unknown. In the payoff structure shown in Figure 1.9.b, the prediction is exactly the opposite: the optimal strategy is for Player 1 to actually reveal private information about the true state of the world, by going straight for the stage game dominant action despite the loss incurred through sharing this information with Player 2. These theoretical predictions are summarised for different lengths of the game in Table 1.3. It is worth noting that this change in the predictions is entirely due to the change in the payoffs. Before turning to empirical evidence on this game, you should try to think of each of the two matrix pairs, and ask yourself whether the way you will play the game will change so dramatically with the payoff structure.

A last theoretical prediction about this kind of game is that the expected payoff of Player 1 (known as the ‘value of the game’) is bounded above by the value of the infinitely repeated game (shown in the last column of the left-hand side of Table 1.3), and bounded below by the value of the average game. These theoretical predictions have

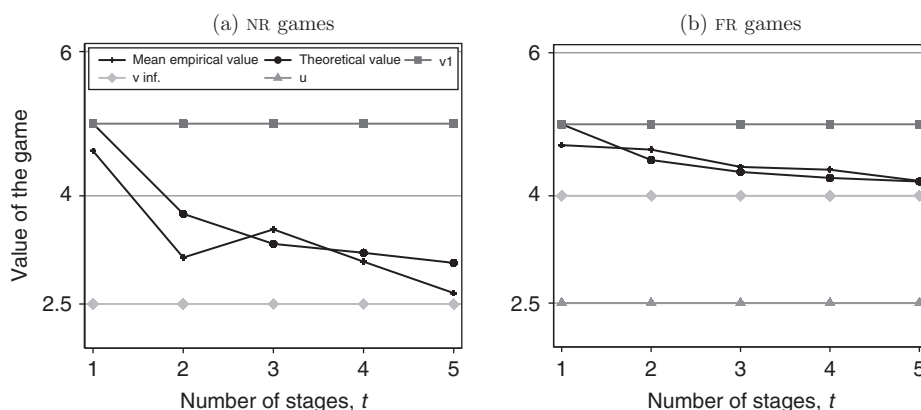


Figure 1.10 Empirical value functions

Note. Observed average payoff in the NR and FR games, along with the theoretical upper (v_1) and lower (∞) bounds.

Source: Jacquemet and Koessler (2013, p. 112, Figure 8).

been tested by Jacquemet and Koessler (2013) in an experiment in which participants play either the NR game or the FR game.

Figure 1.10 provides an overview of a comparison between the average observed values of the games in the experiment (measured as the average payoff earned by Player 1), and the predicted values as presented in Table 1.3. The empirical value functions confirm the theoretical bounds discussed above: the empirical value in both games lies between the value of the infinitely repeated game and the value of the average game. The empirical value is decreasing and smoothly converges towards its lower bound. This provides support for the theoretical analysis of the game. But the most challenging prediction is about the individual strategies, and their change according to the payoff structure.

Figure 1.11 provides information in that regard, through a summary of how information is used in each treatment. Remember that all the treatments have one prediction in common: Player 1, who knows which matrix has been drawn, has nothing to lose by using their private information (i.e. playing the stage-dominant action) at the last stage of the game. The figures are thus separated according to the stage within each game: the last stage of all games is reported on the left-hand side and the intermediate stages of all repeated games (in stages $t = 1$ to $t = T - 1$ for all $T > 1$) are reported on the right-hand side. From both the left-hand figure and the frequency of the stage-dominant action observed in the FR and NR games, experimental subjects unambiguously use information whenever it is worthwhile to do so. The relative frequency of the dominant action in the FR games is always higher than 90% and is much the same as in the last stage of the NR games. This frequency is much lower during intermediate stages of NR games, and is lower and lower as the overall duration of the game increases – when the revelation of information becomes more and more costly. Thus experimental subjects adjust their use of information not only as a reaction to experimental treatments, but also according to the decisions taken during the different stages of a given game.

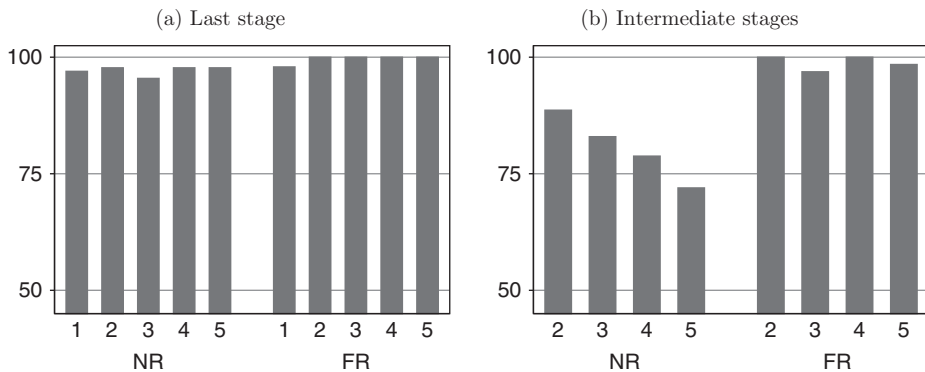


Figure 1.11 The actual use of information: informed players' behaviour

Note. For each treatment and each length, the figures display the mean share of the informed player's decisions that are the current stage-dominant action, in the final stage (*left-hand side*) and in intermediate stages (*right-hand side*).

Source: Jacquemet and Koessler (2013, p. 116, Figure 10).

Overall, empirical behaviour is relatively consistent with theoretical predictions in this environment, in sharp contrast with the two previous examples. This shows that complexity – in the game structure, but also in the theoretical predictions it induces – does not necessarily induce a larger gap between theory and empirical behaviour. The reasons for this consistency, in sharp contrast with the previous examples, is still a largely open question.

1.3.4 Beyond the Examples: Experimental Economics and Behaviour

These examples are not meant to provide a complete picture of the state of the art. But they do offer several important insights as to how experiments can help us better understand decision-makers. First, they show that experiments and economic theory are closely related. Empirical questions and the way data can be most usefully analysed are all based on a theoretical understanding of the situation. Second, and perhaps more importantly, the results described above shows a wide range of conclusions regarding the empirical relevance of theoretical results. Theory seems to accurately predict the outcomes in some games, and fails to do so in others based on similar behavioural assumptions. But the empirical relevance of theory goes beyond predicting outcomes. In particular, it accurately identifies the trade-offs and incentives people face, and how they are likely to resolve these issues. The above examples show that theory is often empirically influential in achieving this goal.

At the same time, it is also true that many behaviours and observed outcomes differ radically from theoretical expectations. Over the years, observations of this type have led specialists to enlarge the scope of the driving forces behind behaviour, to include psychological and sociological motives (this is the aim of behavioural economics). As the examples illustrate, the behaviour observed in economic experiments is related to theory in a complicated way: at times the *homo æconomicus* and human beings act as if

they were perfect strangers, and at other times they are surprisingly close to one another. How, why and under what circumstances do behavioural economics and economic theory converge or diverge? These are the core matters now being taken into consideration in the field (see, e.g. McFadden, 1999, for a survey).

1.4 Experimental Economics Today: What Every Newcomer Must Know

Since its tentative first steps, described at the beginning of the chapter, the use of experiments in economics has grown rapidly and dramatically. A very large number of contributions in economics nowadays rely on assumptions on individual decision-making, about which experiments definitely have something to say. What every newcomer must know in order to become familiar with experimental economics is so vast that no single work could possibly cover the whole field. This book is no exception. Instead, the following section offers an overview of the must-knows of experimental economics. Each of the items listed below corresponds to an index entry (see p. 441) that will refer the reader to sections of the book that discuss or illustrate this particular aspect. The section concludes the outline of the book, describing the must-knows this book will focus on.

1.4.1 Must-know 1: Microeconomic Theory and Decision Sciences

As explained above, experimental economics has grown together with game theory and decision theory. As a consequence, an important part of experimental economics focuses on assessing the empirical content of theories of behaviour. This requires familiarity with a vast number of topics from microeconomic theory. The most important of them are as follows.

- **Decision theory.** This strand of literature tries to better understand how individuals make decisions under risk and uncertainty, what role time-preferences and discounting of the future play and what leads to choice inconsistencies.
- **Game theory.** Agents in an economy interact with one another; their behaviour is directly influenced by the decisions of other agents and, in particular, by the beliefs they may hold about future behaviour of these other agents. Such considerations lead to strategic decision-making, which is a major topic in experimental economics.
- **Non-standard preferences.** The focus on the driving forces of individual behaviour led to challenging the standard way of looking at preferences. Alternative views of behaviour have been developed and are now part of the economists' toolbox. This includes non-standard decision models, such as prospect theory, where contingent states of the world influence decisions; and social and other-regarding preferences, according to which people's preferences not only are defined by consequences for themselves, but also account for the situations of others.
- **Aggregation.** Society has to make decisions, and thus needs to aggregate in one way or another individual tastes. This is the focus of auction theory, the analysis of markets and studies of collective decision-making, such as voting.

- **Psychology of behaviour.** The focus on individual decision-making makes it natural to borrow much from psychology. A large part of this literature is devoted to exploring the systematic deviations from rational decision-making, associated with several well-documented biases such as anchoring and status quo bias, endowment effect, confirmation bias, conjunction fallacy, framing effect, illusion of control, loss aversion.
- **Neuroeconomics.** The analysis of individual behaviour also borrowed in recent years from decision theory in medical sciences, leading to the field of neuroeconomics, which uses physiological measures to relate behaviour to its physiological driving forces.

1.4.2 Must-Know 2: Games and Decision-Making Frameworks

The implementation in a laboratory of the theoretical frameworks described above often makes use of environments, procedures and rules of particular types. They are tools designed to study different aspects of individual behaviour. They are nowadays considered part of the standard toolbox of anyone working in the field.

- **Elicitation procedures** are mechanisms that force agents to reveal something about themselves, such as risk or intertemporal preferences, or beliefs about what others will do.
- **Experimental games** are games structured with specific theoretical properties that are widely used and studied in experimental economics. These key games include the prisoners' dilemma, the trust game, the stag hunt game, the dictator game, the guessing game, the ultimatum bargaining game, the voluntary-contribution mechanism, the minimum effort game and many others.
- **Psychological questionnaires** can be used to gather data on how people think through their decisions and how they consider different situations. Psychometric questionnaires include, for instance, measures of cognitive and non-cognitive skills, personality traits or emotions.

1.4.3 Must-Know 3: Fields and Applications

The insights from (micro)economics that are implemented in the laboratory can be applied to a wide range of field applications. As a result, there is a growing literature of experiments contributing to a better understanding of issues related to the various fields of interest to economics. Among them, the most important are:

- **Labour economics**, which focuses on the effects of labour market policies, the trade-off between consumption and leisure, the education production function, etc.
- **Personel economics** focuses on how people behave in firms, dealing with questions such as how people choose jobs and the reasons why they choose these jobs, how much people work and how they respond to monetary and non-monetary incentives.
- **Industrial organisation** focuses on how firms interact with one another under decision variables of different kinds, such as volumes, prices or levels of advertisement,

and different market structures, such as auctions, oligopoly or perfectly competitive markets.

- **Environmental economics** studies the policies designed to discipline behaviours that are detrimental to the environment, dealing with problems such as greenhouse gas emissions, air pollution, water quality, toxic waste or global warming. The issue of collective decision-making and the problem of free-riding are of critical importance in this domain.
- **Health economics** is a field concerned with the health of individuals, the health care market, the supply of health services or the public health system in general. Preventive health care is an important behavioural issue, for example, and the supply of health care services by physicians raises intriguing questions about incentive design and payment schemes.
- **Law and economics** tries to understand how individuals react to different sets of legal rules. The focus is on circumstances that make people comply with the law, and how the law changes social norms and equilibria.

1.4.4 Must Know 4: Methodological Issues: Outline of the Book

Lastly, laboratory experiments are a very precise way of gathering data and providing an empirical counterfactual to microeconomic theory. This comes with drawbacks and advantages, with several constraints on how experiments are run, and with questions regarding what they tell us about relevant economic issues. The aim of this book is to provide a review of the current answers to this strand of questions.

Chapter 2 is an introduction to the field, by describing step by step what an experiment looks like from the point of view of a participant, before turning to the analysis of the same experiment. This is a critical starting phase in becoming an experimenter, as running experiments is all about understanding how people behave, and avoiding any misunderstanding they may have about the environment. The best way to deal with this issue is to imagine how you would act if you were a participant in an experiment. The second important lesson from this introductory chapter is that experiments involve many unusual procedures and implementation rules, which may not appear quite appealing at first glance. The last part of the chapter describes the reasons why each of these features is required to make an experiment convincing – and will discuss what *convincing* means for an experiment.

This book is divided into four parts. As is explained above, Part I provides an overview of what experiments are. Part II explains why experiments in economics are needed and to what extent they are useful for empirical research in economic science. Each chapter provides a specific answer to this question. In Chapter 3, we describe how experimental economics is related to other empirical methods in economics. Basically, experiments provide a way to choose the data-generating process, enhancing the ability to measure unknown quantities relevant to economic analysis. Chapter 4 turns to the relationships between experiments and economic theory. We will see that experiments serve three different purposes: testing theory in a controlled environment, searching for facts and

whispering in the ears of princes. Theory and experiments share a dynamic of mutually informing each other in this process.

Part III describes how laboratory experiments can achieve these goals. Each chapter explains how to produce experimental results, one step after the other. Chapter 5 focuses on how to design an experiment such that observed behaviour can be related to the institutions under study – i.e. which is internally valid. Chapter 6 covers all the practical aspects required for running an experiment. These practical aspects include all the phases, from building a laboratory well upstream to the final laboratory session. Last, in Chapter 7, we review the main statistical methods that are commonly used to analyse experimental data.

The focus of Part IV is to assess the relevance of what laboratory experiments tell us. Each of the chapters presents an overview of areas in which experimental results are able to shed additional light on existing knowledge. Chapter 8 begins with a question called the ‘external validity’ of experiments: what do decisions taken in the artificial framework of a laboratory tell us about real life? When an experiment satisfies the conditions so as to be both internally and externally valid, then the experimental results can be used by economic theory and public policy. This opens the way to a more general discussion on the possibility of inductive reasoning in economics, an issue covered in the first section of Chapter 9. This discussion will also show that observed behaviour in the lab has drastically changed the way economists think of institutions and how to organise collective decisions. This point will be the focus of the last sections of this chapter, on the design of public policies thanks to the lessons drawn from the laboratory.

Summary

This introductory chapter presented the field of experimental economics from a general perspective. Originally, experiments in the social sciences and in economics, in particular, were thought to be impossible. The first experiments beginning in the second half of the twentieth century showed otherwise. However, experimental economics did not truly break through until the focus of economics changed with the fall of general equilibrium as the central theory and the questions started to turn more towards issues related to human behaviour. To illustrate the current state of the art, we reviewed three examples from the experimental literature testing behavioural insights from game theory: the prisoners’ dilemma, the centipede game and a repeated zero-sum game with incomplete information. Observed behaviour and theoretical predictions may not match up perfectly but they are not perfect strangers to each other either. This summarises the current state of the art in the field: the core issue at stake in ongoing experimental research is identifying situations where theory goes wrong and where it performs well. Since experimental economics has developed together with the use of decision and game theory in economics, the range of topics to which experiments are applied is now far too wide to be reviewed in a single book. This book focuses on experiments as an empirical methodology to inform economic science.