PART I

SETTING THE SCENE
I

The Homo Economicus Model

Political economy ... does not treat the whole of man’s nature as modified by the social state, nor of the whole conduct of man in society. It is concerned with him solely as a being who desires to possess wealth, and who is capable of judging the comparative efficacy of means for obtaining that end.

—Mill (1836)

Traditionally, economic theory has been based on an assumption that behavior is “rational,” ... whereas most psychological and sociological theory insists that behavior is, at least largely, “irrational”.

—Rose (1957)

The starting point in economic theory is that the individual or the firm is maximizing something, usually utility or profit. Economists, almost without exception, make constrained maximization the basic building block of any theory.... Few economists are willing to concede that individuals simply do not know what they are doing.

—Lazear (2000)

Summary: Economists have been very successful using a simplified model of human behaviour, the homo economicus model. This model is typically understood as explaining behaviour by the ability and desire to maximise material personal well-being.

In many aspects, economics has been one of the most successful disciplines in the social sciences. Analyses of citation flows show that economists export their results to other social sciences more than they import from them (Fourcade et al. 2015). Economics has greatly influenced other disciplines, such as sociology, political science, law, psychology and even biology, with its tools,
results and ways of thinking. This influence, sometimes described as “economic imperialism”, is a reflection of the success of its “standard model” of human behaviour, the *homo economicus* (Lazear 2000).

But in spite of its success, this model has been the target of much scorn and rejection in social sciences. Inside the economic discipline, a new field of research, behavioural economics, has grown primarily as a criticism of the *homo economicus* model. Behavioural economics has enriched economics’ understanding of human behaviour by bringing new ideas from psychology into the discipline. However, we should not merely dismiss the standard economic approach, and we should understand its strength, before we look into the insights from behavioural economics.

Imagine yourself as a social scientist trying to understand how and why people do what they do. How easy would it be for you to make sense of the behaviour you observe in society? This task quickly appears daunting. People do a wide range of things. Some things may seem trivial, like deciding what to buy for dinner. But many things look quite complicated, such as political demonstrations or religious rituals. The drivers of human behaviours seem innumerable, from basic desires (e.g., hunger, thirst) to more complex ones (e.g., ambitions, moral feelings). Among the complex drivers of behaviour are those social in nature: people engage in many actions influenced by norms, cultures and traditions. You may be tempted to catalogue the wide range of possible drivers of behaviour to explain what you observe. It would be studying social behaviour like early biologists studied wildlife: simply observing it and labelling it. Your study of behaviour could conclude that: “People follow a norm of honesty in situations A, they express national pride in situations B, they care about their social status in situations C, they have friendship norms in situation D, etc.” Such an approach is not useless, but we do not gain an understanding of why people act in the way they act. The strength and appeal of the *homo economicus* model are that, instead, it relies on minimal behavioural assumptions.

In their quest to understand the world, scientists build models that are simplified versions of reality. For example, objects in space are three-dimensional, but physicists often represented them as points (with zero dimension) to study their motions. In the same way, economists and other social scientists build simplified versions of people and societies to study them. How simplified these models should be is an important and non-trivial question. Make a model too simple and it doesn’t look enough like reality to help us understand it. Make a model too complex and it is hard to get insights from it. An excellent way to think about models is to compare them to maps: a map should not be too simple in order to help you find your route, but it should also not be overly crammed with details cluttering it.

A key scientific principle to determine how simple a model should be is Occam’s razor. This principle states that scientific models should be as simple as possible to explain the phenomenon they purport to represent. Isaac
Newton rephrased this principle, saying, “We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances” (Newton 1687/1999). In that regard, one of the strengths of the homo economicus model is its simplicity: it relies on only a small number of assumptions. And still, in spite of this simplicity, it has proven useful to understand a wide range of behaviour.

To start with, economists make two assumptions about behaviour that are so simple they almost seem innocuous. Think about an agent facing different options to choose from. The first assumption is that the agent knows what he wants (completeness of preferences). The second one is that the agent does not have incoherent/conflicting preferences (transitivity). For economists, these two assumptions are what characterise “rational” preferences in consumer theory (Mas-Colell et al. 1995). The notion of rationality is an important one in economics. In everyday use, it typically means that somebody is sensible/reasonable. Economists define rationality according to a few principles of behaviour, like completeness and transitivity.

Beyond these two core principles, more demanding ones have also been used to characterise a rational agent. Economists, for instance, commonly assume that a rational agent would form beliefs that reflect the strength of the available information using Bayesian updating, which specifies how prior beliefs should be updated after the observation of new evidence. When agents use Bayesian updating, their subjective beliefs about the probability of different possible events respect the rules of consistency imposed on measures of probability. In practice, this extended notion of rationality corresponds to the assumption that people know what they want, have internally consistent preferences and beliefs, and adequately use the means available to reach the best option available among the set of possible options.

Note that, defined in that way, “rationality” is not about the kind of things that people want. For instance, this definition of rationality does not say that some things should be more desirable for people and some other things should not. It only means that people should not be inconsistent in their preferences, whatever these preferences are. In other words, rationality is not about the content of preferences; it is about their coherence. If John decides to go on a cruise with spa and massages while Jane chooses to go on a retreat in the jungle with minimal living standards, neither of them is more or less rational. They just have different preferences.

This definition of rationality can, however, be a bit problematic for the applied economist trying to understand human behaviour in the real world. Any
difference in behaviour observed between people could easily be “explained” as reflecting differences in their inner preferences. For example, if John sacrifices his private life to climb the social ladder and make a lot of money, we could explain his behaviour as being driven by “his preferences”. If Jane works on a low salary in an NGO to reduce world poverty, we could also explain her behaviour as being driven by “her preferences”. Furthermore, any change in behaviour over time could also be explained by changes in preferences: if Jane quits her NGO job to work in a hedge fund, we could simply say that it is because her preferences have changed. Talking of “preferences” in that way, we have not explained anything. Preferences are, in that sense, a tautology: people’s preferences are what they choose to do, and they choose to do it because it is their preference.

For the notion of preferences to have some bite when studying behaviour, economists usually impose some restrictions on the likely content of these preferences. Classical economists like John Stuart Mill had assumed that, when studying economic activities, we could consider people’s goals as being restricted to increasing their wealth and material well-being. In their famous article “De gustibus non est disputandum” Nobel Prize–winners George Stigler and Gary Becker rephrased this principle with specific types of restrictions on people’s preferences. Their first restriction is that preferences are stable over time: “one does not argue over tastes for the same reason that one does not argue over the Rocky Mountains – both are there, will be there next year, too, and are the same to all men”. Their second restriction is that people follow their material interest, and behaviour can for this reason mostly be explained by prices and income: “the economist continues to search for differences in prices or incomes to explain any differences or changes in behavior” (Stigler and Becker 1977). These restrictions consist in looking for explanations of human behaviour that rely on the pursuit of material self-interest.

Stigler and Becker acknowledged in their initial article that these restrictions are not necessarily the only valid way to investigate human behaviour. They likely proposed to restrict the possible explanations of human behaviour primarily for methodological reasons. Even if humans do not have stable preferences and even if they do not care only about money, it may be useful for economists to make these simplifying assumptions. The reason is that a scientific approach is stronger when it can explain a lot of facts with only a few clear assumptions, rather than with a lot of different assumptions. In the case of human behaviour, it is more appealing to explain a wide range of observed behaviour with one single motive, material self-interest, rather than to use different types of preferences for the different types of behaviour observed.

The distinction between a methodological assumption (about how the world should be studied) and a substantive assumption (about how the world is) can, however, be blurry. Economists have often taken the selfish assumption as true in practice. This “self-interest theory” has been one of the tenets of
1 The Homo Economicus Model

the homo economicus model. Behaviours that violate self-interest have, as a consequence, frequently been labelled “irrational”.

One of the noticeable benefits of the economists’ assumptions about rationality is that they make it possible to model human behaviour formally (i.e., using mathematical models). Under these assumptions, we can characterise how individuals will choose between different options: they will choose the option bringing the highest material benefits among all the possible options available. Modelling behaviour that way makes it possible to make predictions. These predictions can then be tested using observed behaviour.

The selfishness hypothesis is convenient as a methodological approach because it reduces the dimensions the agent cares about when making choices (e.g., agents just care about money). Other secondary hypotheses about the homo economicus have also been very convenient to model people’s decisions formally. One of the most important of such hypotheses is that the homo economicus is really good at maths.

To see how an economist would use the good-at-math homo economicus model to represent a human decision process, let’s take a simple life situation. Suppose John is looking for a restaurant in an unknown city. He stumbles upon a first restaurant and observes its menu and prices. Should he stop there, or look for another restaurant?

The homo economicus model gives us a well-defined conceptual framework to think about this problem. First, it would depend on John’s existing preferences: how hungry he is, how tired of walking he is, what his preferences are for the type of food in this restaurant versus other types of food likely available in the neighbourhood, and how much he is willing to pay for better food. The assumption of completeness of preferences means that, when considering other potential restaurants characterised by distance, price, and type of food, John would be able to determine whether he prefers another restaurant to the one he is facing now. In other words, John is able to make trade-offs between all the different restaurants’ characteristics. For example, he knows how much he is willing to walk to save a few dollars off his dinner price and he knows how much he is willing to pay for better quality food. To choose whether

---

3 It is the case, even though theoretical economists working on the foundations of the discipline in the twentieth century were usually agnostic about the content of preferences (see Chapter 10).

4 For instance, in early economic experiments, participants were observed acting in ways not always maximising the monetary gains at stake in small games. Instead, many participants seemed to care about other participants’ payoffs (see Chapter 10). These findings, replicated in many cultural contexts, have been branded as invalidating the “selfishness axiom” at the heart of the homo economicus model (Henrich et al. 2005).

5 I use here the term “secondary hypothesis” to emphasise that these are not canonical hypotheses laid out in the core of the homo economicus model. The core hypotheses are easy to find in textbooks and often labelled “axioms”. Instead, secondary hypotheses are additional hypotheses that are de facto made by economists without being necessarily laid out explicitly in the textbook descriptions of the homo economicus.
to stop or not John would then use his knowledge about the likely presence and characteristics of other restaurants in the neighbourhood and his decision would simply be the one with the highest expected subjective satisfaction: if he is more likely to be satisfied with the present restaurant, he’ll go in; if he is more likely to be satisfied with another one, he’ll try his luck elsewhere. In the latter case, John may end up being unsuccessful. Having failed to find another restaurant, he may have to come back to the first restaurant. But if he is likely enough to find a good restaurant, the risk will be worth it.

The advantage of this economic model is that it will allow the economist to predict how John would make decisions given his utility for food, time and money. Once this solution is determined the economist can then investigate how this choice is likely to change if some of these variables change themselves: if the restaurant’s food quality is lower, if the price is higher, if the nearest alternative option is closer. Being able to predict how people would make choices given some assumptions on their utility is very appealing compared to the inability to make specific predictions. These predictions can then be investigated with empirical observations. With this approach relying on making predictions and investigating them empirically, economics seems to adopt the scientific approach from natural sciences.

Another convenient hypothesis is to assume that the decision maker’s beliefs are correct. Here again, this assumption has some methodological benefits, like the assumptions about stable and selfish preferences. Accurate beliefs constrain the economists’ ability to explain any behaviour with just-so stories. If any belief could be assumed, any behaviour could be explained as maximising self interest for some peculiar belief. For instance, suppose you observe that John prefers butter to margarine; you could assume that it is because he believes it is better for his health. If Jane prefers margarine to butter, you may just assume that it is because she has the opposite belief and think that margarine is better for her health.

In that light, assuming that agents have accurate beliefs is a useful methodological constraint. But it can also become an implausible hypothesis when it leads economists to assume that people never have wrong prior beliefs and that they can perfectly aggregate complex and diffuse elements of information to form accurate beliefs. Going back to our restaurant search example, the assumption of correct beliefs would imply that if John believes there is a 50% chance of finding a better restaurant, in 50% of the cases where he looks for a restaurant elsewhere, he will indeed find a better one.

These different hypotheses do not form a strict and unified definition of rationality, and the word “rationality” has had different meanings in economics, depending on the hypotheses it represented. In some cases, it refers only to an individual having complete and coherent preferences. In other cases, it stands for much stronger assumptions about the content of preferences, the accuracy of beliefs and the ability to solve problems. The
1 The Homo Economicus Model

precise meaning of rationality varies depending on the research area in economics.\(^6\) In its stronger version, the homo economicus is a decision maker who has consistent preferences, is self-interested, is good at maths and has accurate beliefs.

The homo economicus is not the type of human you and I encounter in our daily lives. And many critiques have simply rejected this model as ill-conceived and deeply flawed. They consistently pointed out that economists made unrealistic assumptions about the cognitive ability of individuals. To appreciate this critique, let’s consider again the example of John having to choose whether to stop at a given restaurant or to look for another one. Economists using the homo economicus model would study human behaviour assuming that people can find the best approach to solve this problem. Specifically, a homo economicus would find the optimal stopping strategy that sets a rule about when to stop looking for another restaurant and enter into the one they currently face. To find this rule, an economist would spend hours solving complex equations where the best decision depends on the beliefs of the agent about the possibility of finding a better restaurant elsewhere (using some probability distribution), on the cost of searching for another option, and on a utility function that converts menu quality, search time and food price into one single dimension of satisfaction for the agent. Having found the solution for these equations (a stopping rule that maximises the utility of the agent), the economist would write a computer program that runs calculations to find what the agent would do in practice, given the parameters of the situation considered (e.g., search costs, utility for food quality). Depending on the complexity of the problem, the program could take minutes or hours to find the solution. The economist would then use this solution to describe how John would quickly choose whether to stop at a restaurant or not. It is not unreasonable to question the validity of his approach.

Nonetheless, we should not reject the old homo economicus approach offhandedly. This model has been a very successful research program that has led to many great insights about human behaviour and society. The reductionism of the homo economicus model proposes to look for a unifying explanation behind all these different behaviours: people are reasonable and follow their self-interest. Reductionism, explaining many different things with a few principles, is one of the most effective scientific principles. A demonstration of the strength of the homo economicus model is the insights it can give from rejecting the conventional explanations provided by people for their behaviour. People may have interests to put forward some of their motives and to hide others. Politicians are more likely to be elected if they present their motives as “defending the national interest”, rather than enjoying the prestige of high

\(^6\) I delve into an in-depth discussion of the economic notion of rationality at the end of this book.
office. Pop stars are more likely to be likeable if they claim to sing because they enjoy making their public happy, rather than to make money.\textsuperscript{7}

Beyond individual agents, social organisations also build narratives that buttress justifications for their existence. The government “represents the interest of citizens”, the police force “protects citizens and decrease crime”, and the Church “helps guide believers through their lives”. These narratives purposely ignore the possible existence of other motives from the agents within these organisations: politicians can favour making decisions against the national interest to increase their chance of being elected; the police force may sometimes prefer covering up a mistake, even if a crime is left unresolved, or even if it is uncertain whether the person put in prison is actually guilty; some religious representatives may engage in criminal behaviour and be protected by their church.

In their famous book \textit{Freakonomics}, Levitt and Dubner (2005) describe the two Japanese words that reflect this dual reality in society: \textit{honne} (people’s true feelings and desires) and \textit{tatame} (the behaviours and opinions people display in public). By rejecting the official narratives given by society about human behaviour (\textit{tatame}), the economic approach has been subversive. It has allowed in many instances to get a glance at the real motives of observed behaviour (\textit{honne}). And this has been a major source of insights that has made economics very successful.

Levitt and Dubner give a striking illustration of this fact through the investigation of corruption in the world of sumo wrestling. Sumo is a discipline firmly embedded in a culture of honour and the respect of old traditions. If there is a domain where you would expect economic incentives to be left at the door, it is this one. But winning fights provide prestige and material rewards. Equipped only with his homo economicus model, Levitt asked a simple question, a question that would normally be dismissed in the world of sumo wrestling: Do wrestlers react to material incentives even when these conflict with the traditional norms of the game?

The answer was a resounding “yes”. The statistical evidence pointed out the fact that, once sumo wrestlers were already qualified to move up in rank, they sometimes threw away matches they didn’t need to win. These matches were frequently lost against opponents who needed to win the match in order to move up in ranks. Noticeably, the result tended to be reversed the next time the same two wrestlers met. It suggested that the wrestler who dearly needed

\textsuperscript{7} The best illustration of such incentives was given by Paul McCartney recalling sitting down with John Lennon in 1964 to write a new song. “John would be getting an extension on his house or something, and the joke used to be, ‘Okay! Today let’s write a swimming pool.’ It was great motivation. Then in the next three hours, ‘Help!’ appears from nowhere, you’d suddenly get the idea, this’ll be a hit, this is a good one. You become aware what you were doing was making money. Making good money” (Roller 2012).