A COURSE IN

BEHAVIORAL ECONOMICS

A Course in Behavioral Economics

A Course in Behavioral Economics

ERIK ANGNER

Third edition

macmillan international HIGHER EDUCATION

RED GLOBE PRESS

This edition published 2021 by
RED GLOBE PRESS

Previous editions published under the imprint PALGRAVE

Red Globe Press in the UK is an imprint of Macmillan Education Limited, registered in England, company number 01755588, of 4 Crinan Street, London, N1 9XW.

Red Globe Press® is a registered trademark in the United States, the United Kingdom, Europe and other countries.

ISBN 978-1-352-01080-0 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

To Iris

The master-piece of philosophy would be to develop the means which Providence employs to attain the ends it proposes over man, and to mark out accordingly a few lines of conduct which might make known to this unhappy biped individual the way in which he must walk within the thorny career of life, that he might guard against the whimsical caprices of this fatality to which they give twenty different names, without having as yet come to understand or define it.

The Marquis de Sade, *Justine*

BRIEF CONTENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

As a Ph.D. student in economics, behavioral economics struck me as the most exciting field of study by far. Even with the benefit of some spectacular teachers, though, I felt that existing literature failed to convey an adequate understanding of the nature and significance of the project, and how the many different concepts and theories described as "behavioral" were tied together. When as an assistant professor I was offered the opportunity to teach my own course, I discovered that there were few texts available in the sweet spot between popular-science treatments, which do not contain enough substance for a university-level course, and scientific papers and more advanced textbooks, which are not easily readable and typically fail to provide sufficient background to be comprehensible to a novice reader.

This introduction to behavioral economics was written to be the book I wish I had had as a student, and the book that I want to use as a teacher. It aspires to situate behavioral economics in historical context, seeing it as the result of a coherent intellectual tradition; it offers more substance than popular books but more context than original articles; and while behavioral economics is a research program as opposed to a unified theory, the book not only describes individual concepts and theories but also tries to show how they hang together. The book was designed as a user-friendly, self-contained, freestanding textbook suitable for a one-semester course at the undergraduate level, but can easily be used in conjunction with books or articles in a variety of higher-level courses and programs.

In recognition of the fact that many students of behavioral economics come from outside traditional economics, the exposition was developed to appeal to advanced undergraduates across the social and behavioral sciences, humanities, business, public health, etc. The book contains no advanced mathematics and presupposes no knowledge of standard economic theory. If you are sufficiently interested to pick up a copy of this book and read this far, you have what it takes to grasp the material. Thorough battle-testing at two medium-sized state universities in the US over the course of several years has confirmed that the treatment is accessible to diverse audiences – including to economics majors and non-majors alike.

Serious economics does not need to be intimidating, and this book aims to prove it. Abstract, formal material is introduced in a progressively more difficult manner, which serves to build confidence in students with limited previous exposure. A wealth of examples and exercises help make the underlying intuitions as clear as possible. (Answers to the exercises are provided in an appendix.) In order to sustain the interest of readers with different backgrounds, and to illustrate the vast applicability of economic analysis, examples are drawn from economics, business, marketing, medicine, philosophy, public health, political science, public policy, and elsewhere. More open-ended problems encourage students to apply the ideas and theories presented here to decision problems they might come across outside the classroom.

The book is arranged in six main parts. The first five cover (1) choice under certainty, (2) judgment under risk and uncertainty, (3) choice under risk and uncertainty, (4) intertemporal choice, and (5) strategic interaction. Each of these parts contains two chapters: an even-numbered one outlining standard neoclassical theory and an odd-numbered one discussing behavioral alternatives. The unique structure makes it easy for instructors to teach the book at a more advanced level, as they can easily assign even-numbered chapters as background reading and supplement the odd-numbered chapters with more advanced material of their choosing. A final part (6) explores policy applications – including libertarian paternalism and the nudge agenda – and concludes. Additional material for general readers, students, and instructors is available via the companion website macmillanihe.com/angner-behavioral-economics-3e.

The non-trivial amount of neoclassical theory in this book may warrant explanation. First, because behavioral economics was developed in response to neoclassical economics, large portions of behavioral economics can only be understood against this background. Second, while behavioral economists reject the standard theory as a descriptive theory, they often accept it as a normative theory. Third, much of behavioral economics is a modification or extension of neoclassical theory, which remains useful under a wide range of conditions. Finally, to assess the relative merits of neoclassical and behavioral economics, it is necessary to understand both. Just as the study of a foreign language teaches you a great deal about your native tongue, so the study of behavioral economics can teach you a lot about standard economics.

As a textbook rather than an encyclopedia, this book does not aspire to be a complete record of contemporary theorizing in behavioral economics. Instead, it explores a selection of the most important ideas in behavioral economics and their interrelations. Many fascinating ideas, developments, and avenues of research have deliberately been omitted. No doubt every behavioral economist will disagree with some of my decisions about the things that were left out. But I think most will agree about the things that were included. The material presented in this book is, on the whole, uncontroversially part of the canon, and as such should be familiar to anyone who wishes to have a basic grasp of behavioral economics.

Like other introductory textbooks across the sciences, this book does not purport to describe the evidence supporting the theory in any detail. To keep the focus on theory and applications, the exposition is intentionally uncluttered by extensive discussion of data, standards of evidence, empirical (including experimental) methodology, and statistical techniques. Instead, theories are illustrated by reference to "stylized facts" and stories intended to elicit the intuition underlying the theory and to demonstrate that it is not entirely implausible. In this respect, the present book is no different from any of the standard introductions to microeconomics, to take one example.

For each new edition of the book, the publisher sought extensive feedback from current and prospective readers, including students and instructors, from across the world. As a result, the book has been updated with a wealth of new material, including entire new sections on Rabin's calibration theorem and the economics of happiness. Since the exercises (and answer key) turned out to be one of the most appreciated features of the original edition, I have added even more – and of a wider range of difficulty levels. Meanwhile, I have done my best to keep the book readable and to-the-point (and affordable too).

For readers who wish to continue their study of behavioral economics, or who want to know more about its evidential support, methodology, history, and philosophy, every chapter ends with a further reading section, which offers a selection of citation classics, review articles, and advanced textbooks. While this may be the first book in behavioral economics for many readers, my hope is that it will not be the last.

Erik Angner
Stockholm, Sweden

ACKNOWLEDGEMENTS

This book draws on many different sources of inspiration. I am particularly grateful to the teachers who set me off on this path (among them Cristina Bicchieri, Robyn Dawes, Baruch Fischhoff, George Loewenstein, Philip Reny, and Alvin Roth) and to the students (too numerous to name) who have kept me on the straight and narrow by catching errors and challenging me to continually improve the presentation.

This book project was originally conceived in conversation with Jaime Marshall and Aléta Bezuidenhout at what was then Palgrave Macmillan. I continue to be impressed with their wisdom and foresight in helping me articulate the details of the project. Kirsty Reade, Elizabeth Stone, and Jared Sutton played important roles in realizing the vision, just like Melanie Birdsall, Georgia Walters, and Lillie Flowers did with the second edition. Mallick Hossain kindly assisted me with some of the ancillary materials. I am grateful for all the careful work they put into this project, without which the text would have been far less effective. Luke Block, Amy Brownbridge, Jon Finch, Lloyd Langman, and Verity Rimmer at Red Globe Press worked with me on this third edition, offering invaluable suggestions and gentle encouragement. I am delighted about having the opportunity to refresh the text once more, and have benefited greatly from their thoughtful advice. A most sincere thank you to all.

Many others – including friends, colleagues, and anonymous reviewers – have offered invaluable feedback on earlier versions of this text. I am particularly grateful to Jörg Franke, Nick Huntington–Klein, Ramzi Mabsout, George MacKerron, Ivan Moscati, Norris Peterson, Mark Raiffa, Daniel Wood, and others who got in touch unprompted to share their thoughts on the text and its ancillary materials. Being able to leverage the intelligence and professional judgment of so many fantastically experienced colleagues is gratifying indeed, and I am most thankful.

The image of the mosaic of Ulysses and the Sirens is used with photographer Dennis Jarvis's kind permission. The Swedish Transport Agency (*Transportstyrelsen*) graciously granted the permission to use their images of traffic signs. Original illustrations by Cody Taylor are used with the artist's kind permission. The epigraph appears in Opus Sadicum (de Sade, 1889 [1791], p. 7). Every effort has been made to trace copyright holders, but if any have been inadvertently overlooked we will be pleased to make the necessary arrangements at the first opportunity. I gratefully acknowledge support from a Quality Enhancement Plan Development Grant from the University of Alabama at Birmingham. Most importantly, Elizabeth Blum's love and support every step along the way were essential to the completion of the project.

I continue to welcome suggestions for improvement via the companion website macmillanihe.com/angner-behavioral-economics-3e. As always, errors remain my own.

ABOUT THE AUTHOR

Erik Angner is Professor of Practical Philosophy at Stockholm University and Researcher at the Institute for Futures Studies in Stockholm, Sweden. He is affiliated with the Interdisciplinary Center for Economic Science at George Mason University, USA, where he previously taught Philosophy, Economics, and Public Policy. As a result of serious mission creep, he holds two PhDs – one in Economics and one in History and Philosophy of Science – both from the University of Pittsburgh. He is the author of a scholarly book and multiple journal articles and book chapters on behavioral and experimental economics, the economics of happiness, and the history, philosophy, and methodology of contemporary economics. He lives in Stockholm with his wife and their three children.

1 INTRODUCTION

Learning objectives

After studying this chapter you will:

- Know the difference between descriptive and normative theories of decision
- Understand how behavioral economics differs from standard (neoclassical) economics – and why
- Appreciate the variety of methods used by behavioral economists

1.1 Economics: Neoclassical and behavioral

This is a book about **theories of decision**. To use the language of the epigraph, such theories are about the negotiation of "the thorny career of life": they tell us how we make, or how we should make, decisions. Not that the Marquis de Sade would have spoken in these terms, living as he did in the eighteenth century, but the theory of decision seems to be exactly what he had in mind when he imagined "the master-piece of philosophy."

Developing an acceptable theory of decision would be an achievement. Most human activity – finance, science, medicine, arts, and life in general – can be understood as a matter of people making certain kinds of decisions. Consequently, an accurate theory of decision would cover a lot of ground. Maybe none of the theories we will discuss is the masterpiece of which de Sade thought so highly. Each theory can be, has been, and perhaps should be challenged on various grounds. However, decision theory has been an active area of research in recent decades, and it may have generated real progress.

Modern theories of decision (or **theories of choice** – I will use the terms interchangeably) say little about what goals people will or should pursue. Goals may be good or evil, mean-spirited or magnanimous, altruistic or egoistic, short-sighted or far-sighted; they may be Mother Teresa's or the Marquis de Sade's. Theories of decision simply take a set of goals as given. Provided a set of goals, however, the theories have much to say about how people will or should pursue those goals.

Theories of decision are variously presented as descriptive or normative. A **descriptive** theory describes how people *in fact* make decisions. A **normative** theory captures how people *should* make decisions. It is at least theoretically possible that people make the decisions that they should make. If so, one and the same theory can simultaneously be descriptively adequate and normatively correct. However, it is possible that people fail to act in the manner in which they should. If so, no one theory can be both descriptively adequate and normatively correct.

Exercise 1.1 Descriptive vs. normative Which of the following claims are descriptive and which are normative? (Answers to this and other exercises can be found in the Appendix.)

(**a**) On average, people save less than 10 percent of their income for retirement.
(**b**) People do not save as much for retirement as they should.
(**c**) Very often, people regret not saving more for retirement.

It can be unclear whether a claim is descriptive or normative. "People save too little" is an example. Does this mean that people do not save as much as they should? If so, the claim is normative. Does this mean that people do not save as much as they wish they did? If so, the claim is descriptive.

Example 1.2 Poker Suppose that you are playing poker, and that you are playing to win. Would you benefit from having an adequate descriptive theory, a correct normative theory, or both?

A descriptive theory would give you information about the actions of the other players. A normative theory would tell you how you should behave in light of what you know about the nature of the game, the expected actions of the other players, and your ambition to win. All this information is obviously useful when playing poker. You would benefit from having both kinds of theory.

Some theories of decision are described as **theories of rational choice**. In everyday speech, the word "rationality" is used loosely; frequently it is used simply as a mark of approval. For our purposes, a theory of rational decision is best seen as a **definition** of rationality, that is, as specifying what it means to be rational. Every theory of rational decision serves to divide decisions into two classes: rational and irrational. Rational decisions are those that are in accordance with the theory; irrational decisions are those that are not. A theory of rational choice can be thought of as descriptive or normative (or both). To say that a theory of rational decision is descriptive is to say that people in fact act rationally. To say that a theory of rational decision is normative is to say that people should act rationally. To say that a theory of rational decision is simultaneously descriptive and normative is to say that people act and should act rationally. Typically, the term **rational-choice theory** is reserved for theories that are (or that are thought to be) normatively correct, whether or not they are simultaneously descriptively adequate.

For generations now, economics has been dominated by an intellectual tradition broadly referred to as **neoclassical economics**. If you have studied economics but do not know whether or not you were taught in the neoclassical tradition, it is almost certain that you were. Neoclassical economics is characterized by its commitment to a theory of rational choice that is simultaneously presented as descriptively adequate and normatively correct. This approach presupposes that people by and large act in the manner that they should. Neoclassical economists do not need to assume that all people act rationally all the time, but they insist that deviations from perfect rationality are so small or so unsystematic as to be negligible. Because of its historical dominance, I will refer to neoclassical economics as standard economics, and to neoclassical economic theory as standard theory.

This is an introduction to **behavioral economics**: the attempt to increase the explanatory and predictive power of economic theory by providing it with more psychologically plausible foundations, where "psychologically plausible" means consistent with the best available psychology. Behavioral economists share neoclassical economists' conception of **economics** as the study of people's decisions under conditions of scarcity and of the results of those decisions for society. But behavioral economists reject the idea that people by and large behave in the manner that they should. While behavioral economists certainly do not deny that some people act rationally some of the time, they believe that the deviations from rationality are large enough, systematic enough, and consequently predictable enough to warrant the development of new descriptive theories of decision. If this is right, a descriptively adequate theory cannot at the same time be normatively correct, and a normatively correct theory cannot at the same time be descriptively adequate.

1.2 The origins of behavioral economics

Behavioral economics can be said to have a short history but a long past. Only in the last few decades has it emerged as an independent subdiscipline of economics. By now, top departments of economics have behavioral economists on their staff. Behavioral economics gets published in mainstream journals. Traditional economists incorporate insights from behavioral economics into their work. In 2002, Daniel Kahneman (one of the most famous behavioral economists) won the Nobel Memorial Prize "for having integrated insights from psychological research into economic science, especially concerning human judgment and decision-making under uncertainty." And then, in 2017, Richard Thaler (another leading figure) won the Prize for his contributions to behavioral economics. In spite of its short history, however, efforts to provide economics with plausible psychological foundations go back a long way.

The establishment of modern economics is marked by the publication in 1776 of Adam Smith's *The Wealth of Nations*. Classical economists such as Smith are often accused of having a particularly simple-minded (and false) picture of human nature, according to which people everywhere and always, in hyper-rational fashion, pursue their narrowly construed self-interest. This accusation, however, is unfounded. Smith did not think people were rational:

> How many people ruin themselves by laying out money on trinkets of frivolous utility? What pleases these lovers of toys is not so much the utility, as the aptness of the machines which are fitted to promote it. All their pockets are stuffed with little conveniences … of which the whole utility is certainly not worth the fatigue of bearing the burden.

Smith wrote these words 200 years before the era of pocket calculators, camera phones, iPads, and smartwatches. Nor did Smith think people were selfish: "[There] are evidently some principles in [man's] nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it." Smith and the other classical economists had a conception of human nature that was remarkably multi-faceted; indeed, they did not draw a sharp line between psychology and economics the way we do.

Early neoclassical economics was built on the foundation of **hedonic psychology**: an account of individual behavior according to which individuals seek to maximize pleasure and minimize pain. In W. Stanley Jevons's words: "Pleasure and pain are undoubtedly the ultimate objects of the Calculus of Economics. To satisfy our wants to the utmost with the least effort … in other words, to *maximise pleasure*, is the problem of Economics." The early neoclassical economists were inspired by the philosopher Jeremy Bentham, who wrote: "Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure* … They govern us in all we do, in all we say, in all we think." Because it was assumed that individuals have direct access to their conscious experience, some economists defended the principles of hedonic psychology on the basis of their introspective self-evidence alone.

After World War II, however, many economists were disappointed with the meager results of early neoclassicism in terms of generating theories with predictive power and so came to doubt that introspection worked. Similar developments took place in other fields: behaviorism in psychology, verificationism in philosophy, and operationalism in physics can all be seen as expressions of the same intellectual trend. Postwar neoclassical economists aimed to improve the predictive power of their theories by focusing on what can be publicly observed rather than on what must be experienced. Instead of taking a theory about pleasure and pain as their foundation, they took a theory of preference. The main difference is that people's feelings of pleasure and pain are unobservable, whereas their choices can be directly observed. On the assumption that choices reflect personal preferences, we can have direct observable evidence about what people prefer. Thus, postwar neoclassical economists hoped to completely rid economics of its ties to psychology – hedonic and otherwise.

In spite of the relative hegemony of neoclassical economics during the second half of the twentieth century, many economists felt that their discipline would benefit from closer ties to psychology and other neighboring fields. What really made a difference, however, was the cognitive revolution. In the 1950s and 1960s, researchers in psychology, computer science, linguistics, anthropology, and elsewhere rejected the demands that science focus on the observable and that all methods be public. Instead, these figures advocated a "science of cognition" or **cognitive science**. The cognitive scientists were skeptical of naive reliance on introspection, but nevertheless felt that a scientific psychology must refer to things "in the head," including beliefs and desires, symbols, rules, and images. Behavioral economics is a product of the cognitive revolution. Like cognitive scientists, behavioral economists – though skeptical of the theories and methods of the early neoclassical period – are comfortable talking about beliefs, desires, rules of thumb, and other things "in the head." Below, we will see how these commitments get played out in practice.

To some, the fact that behavioral economists go about their work in such a different way means that they have become economists in name only. But notice that behavioral economics is still about the manner in which people make choices under conditions of scarcity and the results of those choices for society at large – which is the very definition of economics. **Behavioral science** refers to the scientific study of behavior, which makes behavioral economics a kind of behavioral science.

Psychology and economics is also a broader category, referring to anything that integrates the two disciplines, and which therefore does not need to be about choice at all.

1.3 Methods

Before we explore in earnest the concepts and theories developed by behavioral economists in the last few decades, I want to discuss the data that behavioral economists use to test their theories and the methods they use to generate such data. I also want to assuage some skepticism that people may have about those methods.

Some of the earliest and most influential papers in behavioral economics relied on participants' responses to hypothetical choices. In such studies, participants were asked to imagine that they found themselves in a given choice situation and to indicate what decision they would make under those conditions. Here is one such question: "Which of the following would you prefer? A: 50% chance to win 1,000, 50% chance to win nothing; B: 450 for sure." Other early papers relied on readers' intuitions about how people might behave under given conditions. Thus, they offered scenarios such as: "Mr S. admires a $125 cashmere sweater at the department store. He declines to buy it, feeling that it is too extravagant. Later that month he receives the same sweater from his wife for a birthday present. He is very happy. Mr and Mrs S. have only joint bank accounts." These thought experiments were apparently inspired in part by the author's observations of the behavior of fellow economists, who argued that people were always rational but at times behaved irrationally in their own lives.

Soon enough, hypothetical choice studies were almost completely displaced by **laboratory experiments** in which laboratory participants make real choices involving real money. Such experiments have been run for decades. In the early 1970s, for example, psychologists Sarah Lichtenstein and Paul Slovic ran experiments at a Las Vegas casino, where a croupier served as experimenter, professional gamblers served as participants, and winnings and losses were paid in real money. More frequently, behavioral economists use college undergraduates or other easily accessible participants. When behavioral economists engage in experimental studies, they can be hard to distinguish from neoclassical experimental economists, that is, neoclassical economists who use experiments to explore how people make decisions. Experimentalists agree that decisions performed by laboratory participants must be real, and that actual winnings must be paid out.

Behavioral economists, during the last two decades, have increasingly relied on data gathered "in the field." In one famous **field study**, Colin F. Camerer and colleagues studied the behavior of New York City cab drivers by using data from "trip sheets" – forms that drivers use to record the time passengers are picked up and dropped off as well as the amount of the fares – and from the cabs' meters, which automatically record the fares. Researchers in this study simply observed how participants behaved under different conditions. In **field experiments**, researchers randomly assign participants to test and control groups, and then note how (if at all) the behavior of individuals in the two groups differs. In one prominent field experiment, Jen Shang and Rachel Croson tracked how voluntary donations to a public radio station varied when prospective donors were given different social information, that is, information about how much other people had given.

To some extent, behavioral economists use what psychologists call **process measures**, that is, methods that provide hints about cognitive and emotional processes underlying decision-making. Some rely on **process-tracing** software to assess what information people use when making decisions in games. Others employ brain scans, typically functional magnetic resonance imaging (fMRI), which allows researchers to examine, albeit crudely, which parts of an individual's brain are activated in response to a task or decision. Imaging methods have already been applied to a diversity of economic tasks, including decision-making under risk and uncertainty, intertemporal choice, buying and selling behavior, and strategic behavior in games. Even more exotic neuroscience methods are sometimes employed. For example, a tool called transcranial magnetic stimulation can be used to temporarily disable a part of participants' brains as they make decisions. The increasing use of methods borrowed from neuroscience is, not coincidentally, connected to the rise of **neuroeconomics**, which integrates economics with neuroscience.

The use of multiple methods to generate evidence raises interesting methodological problems. This is particularly true when evidence from different sources points in slightly different directions. Sometimes, however, evidence from multiple sources points in the same direction. When this is true, behavioral economists have more confidence in their conclusions. It can be argued that part of the reason why behavioral economics has turned into such a vibrant field is that it successfully integrates evidence of multiple kinds, generated by a variety of methods.

Recently, social and behavioral science has been thrown into something called the "replication crisis," as several well-known empirical results have proven difficult to replicate. It may turn out that these findings were mere experimental artifacts all along. The lack of reproducibility is obviously unwelcome news for the researchers invested in the results, and has fueled skepticism about the methods of social and behavioral science – and perhaps the entire enterprise of trying to understand human behavior with scientific methods. But it is important to note that (at least within bounds) the fact that some alleged findings are revised in light of new evidence is not as such devastating for social and behavioral science. In fact, what makes science different from other kinds of human activity is that *it is supposed to be* open to revision in light of new data. On statistical grounds alone, we should expect that some of the results generated by behavioral economists – and consequently some of the results discussed in the below – will not hold up. That said, systematic studies of reproducibility in psychology and economics suggest that economics is doing reasonably well by comparison. A 2016 report in the prestigious journal *Science* concludes that results from laboratory experiments in economics are at least as robust (and maybe more robust) than any other empirical result in economics, and moreover that laboratory experiments published in top economic journals have relatively high rates of replicability. The authors conclude on a positive note: "There is every reason to be optimistic that science in general, and social science in particular, will emerge much improved after the current period of critical self-reflection."

1.4 Looking ahead

As stated in the Preface, this book is arranged in six main parts: (1) choice under certainty, (2) judgment under risk and uncertainty, (3) choice under risk and uncertainty, (4) intertemporal choice, (5) strategic interaction, and (6) policy applications and conclusions. As suggested in Section 1.1, the ultimate goal of behavioral economics is to generate novel insights into people's decisions under conditions of scarcity and the results of those decisions for society. Behavioral and neoclassical economists alike try to attain this goal by building abstract, formal theories. In this book we will explore increasingly general theories, both neoclassical and behavioral.

Studying behavioral economics is a non-trivial enterprise. For one thing, the level of abstraction can pose an initial challenge. But as we will see below, it is the very fact that economics is so abstract that makes it so very useful: the more abstract the theory, the wider its potential application. Some readers may be prone to putting down a book like this as soon as they notice that it contains mathematics. Please do not. There is no advanced math in the book, and **numeracy** – the ability with or knowledge of numbers – is incredibly important, even to people who think of themselves as practically oriented.

Exercise 1.3 Numeracy In a 2010 study on financial decision-making, people's answers to three quick mathematics questions were strong predictors of their wealth: households where both spouses answered all three questions correctly were *more than eight times* as wealthy as households where neither spouse answered any question correctly. So if you have ever struggled with math, be glad that you did. You can try answering the three questions for yourself:

(**a**) If the chance of getting a disease is 10 percent, how many people out of 1000 would be expected to get the disease?

(**b**) If five people all have the winning numbers in the lottery, and the prize is 2 million dollars, how much will each of them get?

(**c**) Let us say you have $200 in a savings account. The account earns 10 percent interest per year. How much would you have in the account at the end of two years?

You will find the correct answers in the answer key at the end of the book.

There's also evidence that people who fall prey to the specific fallacies and mistakes that behavioral economists study are more likely to experience poor outcomes in their own lives. In a widely cited 2007 study, researchers assessed people's decision-making competence by checking to what extent they make mistakes such as honoring sunk costs (see Section 3.3) in pen-and-paper questionnaires. The study found that people with low decision-making competence were more likely to report poor real-world decision outcomes, such as having gotten a divorce, declared bankruptcy, lost one's driver's license, gotten oneself kicked out of a bar, and so on. The authors suggest that decision-making competence should be considered a separate cognitive skill that helps us avoid negative real-world outcomes.

To underscore the usefulness of behavioral economics, the book discusses a variety of applications. Among other things, you will learn how to choose a wingman or wingwoman, how to design a marketing scheme that works, how not to fall for such marketing schemes, how to compute the probability that your love interest is seeing somebody else, how to sell tires, and how to beat anyone at rock-paper-scissors. Ultimately, behavioral economics sheds light on human beings living in society – the way they really are, as opposed to the way great thinkers of the past have thought they should be – and on the nature of the human condition. Behavioral economics helps us live better lives – and to improve the world to boot.

Further reading

Kahneman's *Thinking, Fast and Slow* (2011) and Thaler's *Misbehaving: The Making of Behavioral Economics* are must-reads for anyone interested in behavioral economics, both for their unparalleled understanding of the theory and for their illuminating personal reminiscences. Angner and Loewenstein (2012) and Heukelom (2014) discuss the nature, historical origins, and methods of behavioral economics; Angner (2015a, 2019) explores further the relationship between behavioral and neoclassical economics. *The Wealth of Nations* is Smith (1976 [1776]); the quotations in the history section are from Smith (2002 [1759], p. 211) and Smith (2002 [1759], p. 11), Jevons (1965 [1871], p. 37), and Bentham (1996 [1789], p. 11). The sample questions in the methods section come from Kahneman and Tversky (1979, p. 264) and Thaler (1985, p. 199). The psychologists who went to Vegas are Lichtenstein and Slovic (1973). The study of NYC cabdrivers is Camerer et al. (1997); the one about social information is Shang and Croson (2009). Camerer et al. (2005) provide a widely cited overview of neuroeconomics, and Camerer, Dreber, et al. (2016, pp. 1435–6) examine the reproducibility of economics. The study on financial decision-making is Smith et al. (2010); the three numeracy questions were adapted from the University of Michigan Health and Retirement Study.

PART 1

CHOICE UNDER CERTAINTY

2 RATIONAL CHOICE UNDER CERTAINTY

Learning objectives

After studying this chapter you will:

- Know the theory of choice under certainty
- Understand the concept of rationality built into this theory
- Be able to prove theorems on the basis of axioms and definitions

2.1 Introduction

As promised, we begin by discussing the theory of rational choice. This theory forms the foundation of virtually all modern economics and is one of the first things you would learn in a graduate-level microeconomics class. As a theory of rational choice (see Section 1.1), the theory specifies what it means to make rational decisions – in short, what it means to be rational.

In this chapter, we consider **choice under certainty**. The phrase "under certainty" simply means that there is no doubt as to which outcome will result from a given act. For example, if the staff at your local gelato place is minimally competent, so that you actually get vanilla every time you order vanilla and stracciatella every time you order stracciatella, you are making a choice under certainty. (We will discuss other kinds of choice in future chapters.) Before discussing what it means to make rational choices under conditions of certainty, however, we need to talk about what preferences are and what it means to have rational preferences.

The theory of rational choice under certainty is an **axiomatic** theory. This means that the theory consists of a set of **axioms**: basic propositions that cannot be proven using the resources offered by the theory, and which will simply have to be taken for granted. When studying the theory, the first thing we want to do is examine the axioms. As we go along, we will also introduce new terms by means of definitions. Axioms and definitions have to be memorized. Having introduced the axioms and definitions, we can prove many interesting claims. Thus, much of what we will do below involves proving new propositions on the basis of axioms and definitions.

2.2 Preferences

The concept of **preference** is fundamental in modern economics, neoclassical and behavioral. Formally speaking, a preference is a **relation**. The following are examples of relations: "Alf is older than Betsy," "France is bigger than Norway," and "Bill is worried he may not do as well on the exam as Jennifer." Notice that each of these sentences expresses a relationship between two entities (things, individuals). Thus, "Alf is older than Betsy" expresses a relationship between Alf and Betsy, namely, that the former is older than the latter. Because these examples express a relation between two entities, they are called **binary** relations. The following relation is not binary: "Mom stands between Bill and Bob." This relation is **ternary**, because it involves three different entities; in this case, people.

For convenience, we often use small letters to denote entities or individuals. We may use a to denote Alf and b to denote Betsy. Similarly, we often use capital letters to denote relations. We may use R to denote the relation "is older than." If so, we can write aRb for "Alf is older than Betsy." Sometimes we write Rab. Notice that the order of the terms matters: aRb is not the same thing as bRa. The first says that Alf is older than Betsy, and the second that Betsy is older than Alf. Similarly, Rab is not the same thing as Rba.

Exercise 2.1 Relations Assume that f denotes France and n denotes Norway, and that B means "is bigger than."

(**a**) How would you write that France is bigger than Norway?
(**b**) How would you write that Norway is bigger than France?
(**c**) How would you write that Norway is bigger than Norway?

In order to speak clearly about relations, we need to specify what sort of entities may be related to one another. When talking about who is older than whom, we may be talking about people. When talking about what is bigger than what, we may be talking about countries, houses, people, dogs, or many other things. Sometimes it matters what sort of entities we have in mind. When we want to be careful, which is most of the time, we define a **universe** U. The universe is the set of all things that can be related to one another. Suppose we are talking about Donald Duck's nephews Huey, Dewey, and Louie. If so, that is our universe. The convention is to list all members of the universe separated by commas and enclosed in curly brackets, like so: {Huey, Dewey, Louie}. Here, the order does not matter. So, the same universe can be written like this: {Louie, Dewey, Huey}. Thus: U = {Huey, Dewey, Louie} = {Louie, Dewey, Huey}.

Exercise 2.2 The universe Suppose we are talking about all countries that are members of the United Nations. How would that be written?

A universe may have infinitely many members, in which case simple enumeration is inconvenient. This is true, for instance, when you consider the time at which you entered the space where you are reading this. There are infinitely many points in time between 11:59 am and 12:01 pm, for example, as there are between 11:59:59 am and 12:00:01 pm. In such cases, we need to find another way to describe the universe.

One relation we can talk about is this one: "is at least as good as." For example, we might want to say that "coffee is at least as good as tea." The "at least as good as" relation is often expressed using this symbol: $\succcurlyeq$. If c denotes coffee and t denotes tea, we can write this sentence as $c \succcurlyeq t$. This is the **(weak) preference relation**. People may have, and often will have, their own preference relations. If we wish to specify whose preferences we are talking about, we use subscripts to denote individuals. If we want to say that for Alf coffee is at least as good as tea, and that for Betsy tea is at least as good as coffee, we say that $c \succcurlyeq_{Alf} t$ and $t \succcurlyeq_{Betsy} c$, or that $c \succcurlyeq_A t$ and $t \succcurlyeq_B c$.

Exercise 2.3 Preferences Suppose d denotes "enjoying a cool drink on a hot day" and r denotes "getting roasted over an open fire."
(**a**) How would you state your preference over these two options?
(**b**) How would you express a masochist's preference over these two options?

In economics, we are typically interested in people's preferences over **consumption bundles**, which are collections of goods. You face a choice of commodity bundles when choosing between the #1 Big Burger meal and the #2 Veggie Burger meal at your local hamburger restaurant. In order to represent commodity bundles, we think of them as collections of individual goods along the following lines: three apples and two bananas, or two units of guns and five units of butter. When talking about preference relations, the universe can also be referred to as the **set of alternatives**. If bundles contain no more than two goods, it can be convenient to represent the set of alternatives on a plane, as in Figure 2.1. When bundles contain more than two goods, it is typically more useful to write $\langle 3, 2\rangle$ for three apples and two bananas; $\langle 6, 3, 9\rangle$ for six apples, three bananas, and nine coconuts; and so on.

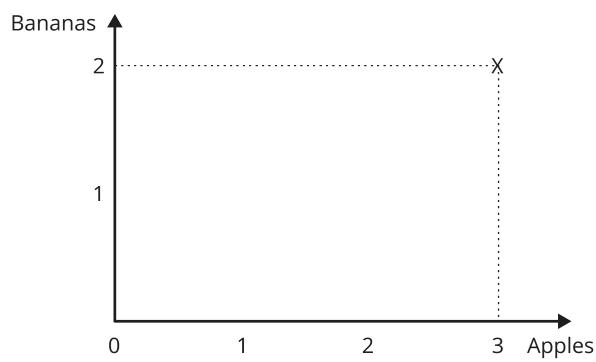

Figure 2.1 Set of alternatives

2.3 Rational preferences

We begin building our theory of rational choice by specifying what it means for a preference relation to be rational. A **rational** preference relation is a preference relation that is transitive and complete.

A relation R is **transitive** just in case the following condition holds: for all x, y, and z in the universe, if x bears relation R to y, and if y bears relation R to z, then x must bear relation R to z. Suppose the universe is the set of all the Marx brothers. If so, "is taller than" is a transitive relation: if Zeppo is taller than Groucho, and Groucho is taller than Harpo, then Zeppo must be taller than Harpo (Figure 2.2).

Figure 2.2 The Marx brothers. Illustration by Cody Taylor

Example 2.4 ***30 Rock*** Consider the following exchange from the TV show *30 Rock*. Tracy, Grizz, and Dot Com are playing computer games. Tracy always beats Grizz and Dot Com. When Kenneth beats Tracy but gets beaten by Grizz, Tracy grows suspicious.

Tracy: *"How were you beating Kenneth, Grizz?"*
Grizz: *"I don't know."*
Tracy: *"If Kenneth could beat me and you can beat Kenneth, then by the transitive property, you should beat me too! Have you been letting me win?"*
Dot Com: *"Just at some things."*
Tracy: *"Things? Plural?"*

Now you are the first kid on the block who understands *30 Rock*. You also know that the show had a former economics or philosophy student on its staff.

If the universe consists of all people, examples of **intransitive** relations include "is in love with." Just because Sam is in love with Pat, and Pat is in love with Robin, it is not necessarily the case that Sam is in love with Robin. Sam *may* be in love with Robin. But Sam may have no particular feelings about Robin, or Sam may resent Robin for attracting Pat's attention. It may also be the case that Robin is in love with Sam. This kind of intransitivity is central to the play *No Exit*, by the French existentialist philosopher Jean-Paul Sartre. In the play, which takes place in a prison cell, a young woman craves the affection of a man who desires the respect of an older woman, who in turn is in love with the young woman. Hence the most famous line of the play: "Hell is other people." To show that a relation is intransitive, it is sufficient to identify three members of the universe such that the first is related to the second, and the second is related to the third, but the first is not related to the third.

Formally speaking, a preference relation $\succcurlyeq$ is transitive just in case the following is true:

Axiom 2.5 Transitivity of $\succsim$ *If $x \succsim y$ and $y \succsim z$, then $x \succsim z$ (for all x, y, z).*

There are other ways of expressing the same thing. We might write: If $x \succsim y \succsim z$, then $x \succsim z$ (for all x, y, z). Using standard logic symbols, we might write: $x \succsim y \;\&\; y \succsim z \rightarrow x \succsim z$ (for all x, y, z). See the text box below for a useful list of logical symbols. Either way, transitivity says that if you prefer coffee to tea, and tea to root beer, you must prefer coffee to root beer; that is, you cannot prefer coffee to tea and tea to root beer while failing to prefer coffee to root beer.

A relation R is **complete** just in case the following condition holds: for any x and y in the universe, either x bears relation R to y, or y bears relation R to x (or both). If the universe consists of all people – past, present, and future – then "is at least as tall as" is a complete relation. You may not know how tall Caesar and Brutus were, but you do know this: either Caesar was at least as tall as Brutus, or Brutus was at least as tall as Caesar (or both, in case they were equally tall).

Given the universe of all people, examples of **incomplete** relations include "is in love with." For any two randomly selected people – your landlord and the current President of the US, for example – it is not necessarily the case that either one is in love with the other. Your landlord may have a crush on the President, or the other way around. But this need not be the case, and it frequently will not be. To show that a relation is incomplete, then, it is sufficient to identify two objects in the universe such that the relation does not hold either way.

Formally speaking, a preference relation $\succsim$ is complete just in case the following is true:

Axiom 2.6 Completeness of $\succsim$ *Either $x \succsim y$ or $y \succsim x$ (or both) (for all x, y).*

Completeness means that you must prefer tea to coffee or coffee to tea (or both); though your preference can go both ways, you cannot fail to have a preference between the two. The use of the phrase "(or both)" in the formula above is, strictly speaking, redundant: we use the "inclusive or," which is equivalent to "and/or" in everyday language. Using standard logical symbols, we might write: $x \succsim y \vee y \succsim x$ (for all x, y). If both $x \succsim y$ and $y \succsim x$, we say that there is a tie (see Section 2.4).

Logical symbols

Here is a list of the most common logical symbols:

$x \;\&\; y$	x and y
$x \vee y$	x or y
$x \rightarrow y$	if x then y; x only if y
$x \leftrightarrow y$	x if and only if y; x just in case y
$\neg p$	not p

The following exercise serves to illustrate the concepts of transitivity and completeness.

Exercise 2.7 Assuming the universe is the set of all people – past, present, and future – are the following relations transitive? Are they complete?
(**a**) "is the mother of"
(**b**) "is an ancestor of"
(**c**) "is the sister of"
(**d**) "detests"
(**e**) "weighs more than"
(**f**) "has the same first name as"
(**g**) "is taller than"

When answering questions such as these, ambiguity can be a problem. A word such as "sister" is ambiguous, which means that answers might depend on how it is used. As soon as the word is defined, however, the questions have determinate answers.

Exercise 2.8 The enemy of your enemy Suppose it is true, as people say, that the enemy of your enemy is your friend. What does this mean for the transitivity of "is the enemy of"? (Assume there are no true frenemies: people who are simultaneously friends and enemies.)

Exercise 2.9 Assuming the universe is the set of all natural numbers, meaning that U = {1, 2, 3, 4, …}, are the following relations transitive? Are they complete?
(**a**) "is at least as great as" ($\geq$)
(**b**) "is equal to" ($=$)
(**c**) "is strictly greater than" ($>$)
(**d**) "is divisible by" ($|$)

Exercise 2.10 Preferences and the universe Use your understanding of transitivity and completeness to answer the following questions:
(**a**) If the universe is {apple, banana, starvation}, what does the transitivity of the preference relation entail?
(**b**) If the universe is {apple, banana}, what does the completeness of the preference relation entail?

As the last exercise suggests, the completeness of the preference relation implies that it is **reflexive**, meaning that $x \succcurlyeq x$ (for all x). This result might strike you as surprising. But recall that completeness says that, any time you pick two elements from the universe, the relation must hold one way or the other. The axiom does not say that the two elements must be different. If you pick the same element twice, which you may, completeness requires that the thing stands in the relation to itself.

The choice of a universe might determine whether a relation is transitive or intransitive, complete or incomplete. If the universe were U = {Romeo, Juliet}, the relation "is in love with" would be complete, since for any two members of the universe, either the one is in love with the other, or the other is in love with the one. (This assumes that Romeo and Juliet are both in love with themselves, which might perhaps not be true.) Perhaps more surprisingly, the relation would also be transitive: whenever $x \succcurlyeq y$ and $y \succcurlyeq z$, it is in fact the case that $x \succcurlyeq z$.

The assumption that the weak preference relation is rational (transitive and complete) might seem fairly modest. Yet, in combination with a couple of definitions, this assumption is in effect everything necessary to build a theory of choice under certainty. This is a wonderful illustration of how science works: based on a small number of assumptions, we will build an extensive theory, whose predictions will then be confronted with actual evidence. The rest of this chapter spells out the implications of the assumption that the weak preference relation is rational.

2.4 Indifference and strict preference

As the previous section shows, the (weak) preference relation admits ties. When two options are tied, we say that the first option is **as good as** the second or that the agent is **indifferent** between the two options. That is, a person is indifferent between two options just in case, to her, the first option is at least as good as the second and the second is at least as good as the first. We use the symbol ~ to denote indifference. Formally speaking:

Definition 2.11 Definition of indifference *$x \sim y$ if and only if $x \succsim y$ and $y \succsim x$.*

Using logical symbols, we might write: $x \sim y \Leftrightarrow x \succsim y \;\&\; y \succsim x$.

Assuming that the "at least as good as" relation is rational, the indifference relation is both reflexive and transitive. It is also **symmetric**: if x is as good as y, then y is as good as x. These results are not just intuitively plausible; they can be established by means of **proofs**. (See the text box on page 20 for more about proofs.) Properties of the indifference relation are established by the following proposition.

Proposition 2.12 Properties of indifference *The following conditions hold:*

(i) $x \sim x$ *(for all x)*

(ii) $x \sim y \rightarrow y \sim x$ *(for all x, y)*

(iii) $x \sim y \;\&\; y \sim z \rightarrow x \sim z$ *(for all x, y, z)*

Proof.
Each part of the proposition requires a separate proof:

(i)	1.	$x \succsim x$	by Axiom 2.6
	2.	$x \succsim x \;\&\; x \succsim x$	from (1), by logic
	$\therefore$	$x \sim x$	from (2), by Definition 2.11 □

(ii)	1.	$x \sim y$	by assumption	
	2.	$x \succsim y \;\&\; y \succsim x$	from (1), by Definition 2.11	
	3.	$y \succsim x \;\&\; x \succsim y$	from (2), by logic	
	4.	$y \sim x$	from (3), by Definition 2.11	
	∴	$x \sim y \rightarrow y \sim x$	from (1)–(4), by logic	□
(iii)	1.	$x \sim y \;\&\; y \sim z$	by assumption	
	2.	$x \succsim y \;\&\; y \succsim x$	from (1), by Definition 2.11	
	3.	$y \succsim z \;\&\; z \succsim y$	from (1), by Definition 2.11	
	4.	$x \succsim z$	from (2) and (3), by Axiom 2.5	
	5.	$z \succsim x$	from (2) and (3), by Axiom 2.5	
	6.	$x \sim z$	from (4) and (5), by Definition 2.11	
	∴	$x \sim y \;\&\; y \sim z \rightarrow x \sim z$	from (1)–(6), by logic	□

These are the complete proofs. In what follows, I will often outline the general shape of the proof rather than presenting the whole thing.

The indifference relation is not complete. To show this, it is enough to give a single counterexample. Any rational preference relation according to which the agent is not indifferent between all options will do (see, for instance, Figure 2.3).

Exercise 2.13 Prove the following principle: $x \succsim y \;\&\; y \sim z \rightarrow x \succsim z$.

In your various proofs, it is always acceptable to rely on propositions you have already established. The following exercise shows how useful this can be.

Exercise 2.14 Iterated transitivity In this exercise you will prove the following principle in two different ways: $x \sim y \;\&\; y \sim z \;\&\; z \sim p \rightarrow x \sim p$.

(**a**) First prove it by applying the transitivity of indifference (Proposition 2.12(iii)).

(**b**) Then prove it without assuming the transitivity of indifference. (You may still use the transitivity of weak preference, since it is an axiom.)

If you have difficulty completing the proofs, refer to the text box on page 20 for hints.

Heavenly Bliss

$\curlyvee$

Coke $\sim$ Pepsi

$\curlyvee$

Eternal Suffering

Figure 2.3 Preference ordering with tie

When a first option is at least as good as a second, but the second is not at least as good as the first, we say that the first option is **better than** the second or that the agent **strictly** or **strongly** prefers the first over the second. We use the symbol $\succ$ to denote **strict** or **strong preference**. Formally speaking:

> **Definition 2.15 Definition of strict preference** *$x \succ y$ if and only if $x \succsim y$ and it is not the case that $y \succsim x$.*

Using logical notation, that is to say: $x \succ y \Leftrightarrow x \succsim y \;\&\; \neg y \succsim x$. For clarity, sometimes the "is at least as good as" relation will be called **weak preference**.

Assuming (still) that the weak preference relation is rational, it is possible to prove logically that the strict preference relation will have certain properties. The following proposition establishes some of them.

> **Proposition 2.16 Properties of strict preference** *The following conditions hold:*
>
> (i) $x \succ y \;\&\; y \succ z \rightarrow x \succ z$ *(for all x, y, z)*
> (ii) $x \succ y \rightarrow$ *not* $y \succ x$ *(for all x, y)*
> (iii) *not* $x \succ x$ *(for all x)*

Proof.

(i) Suppose that $x \succ y \;\&\; y \succ z$. In order to establish that $x \succ z$, Definition 2.15 tells us that we need to show that $x \succsim z$ and that it is not the case that $z \succsim x$. The first part is Exercise 2.17. The second part goes as follows: suppose for a **proof by contradiction** that $z \succsim x$. From the first assumption and the definition of strict preference, it follows that $x \succsim y$. From the second assumption and Axiom 2.5, it follows that $z \succsim y$. But from the first assumption and the definition of strict preference, it also follows that $\neg z \succsim y$. We have derived a contradiction, so the second assumption must be false, and therefore $\neg z \succsim x$.

(ii) Begin by assuming $x \succ y$. Then, for a proof by contradiction, assume that $y \succ x$. Given the first assumption, Definition 2.15 implies that $x \succsim y$. Given the second assumption, the same definition implies that $\neg x \succsim y$. But this is a contradiction, so the second assumption must be false, and therefore $\neg y \succ x$

(iii) See Exercise 2.19. □

Proposition 2.16(i) says that the strict preference relation is transitive, 2.16(ii) that it is **anti-symmetric**, and 2.16(iii) that it is **irreflexive**.

Exercise 2.17 Using the definitions and propositions discussed so far, complete the first part of the proof of Proposition 2.16(i).

Notice that the proofs of Proposition 2.16(i) and (ii) involve constructing proofs by contradiction. Such proofs are also called **indirect proofs**. This mode of reasoning might look weird, but it is actually quite common in mathematics, science, and everyday thinking. For example, when mathematicians prove that $\sqrt{2}$ is an irrational number, they can proceed by assuming (for a proof by contradiction) that $\sqrt{2}$ is a rational number (meaning that $\sqrt{2}$ can be expressed as a fraction p/q of natural numbers p and q) and then use this assumption to derive a contradiction.

Exercise 2.18 The enemy of your enemy, cont. Use a proof by contradiction to establish that "is the enemy of" is not transitive, as in Exercise 2.8 on page 15.

In future exercises, you will see just how useful proofs by contradiction can be.

Exercise 2.19 Prove Proposition 2.16(iii). Prove it by contradiction, by first assuming that there is an x such that $x \succ x$.

Exercise 2.20 Prove the following principle: $x \succ y \; \& \; y \succcurlyeq z \rightarrow x \succ z$ (for all x, y, z). Notice that this proof has two parts. First, prove that $x \succcurlyeq z$; second, prove that $\neg z \succcurlyeq x$.

Exercise 2.21 Establish the following important and intuitive principles. (For the record, some of them are logically equivalent.)

(**a**) If $x \succ y$ then $x \succcurlyeq y$
(**b**) If $x \succ y$ then $\neg y \succcurlyeq x$
(**c**) If $x \succcurlyeq y$ then $\neg y \succ x$
(**d**) If $x \succ y$ then $\neg x \sim y$
(**e**) If $x \sim y$ then $\neg x \succ y$
(**f**) If $\neg x \succcurlyeq y$ then $y \succcurlyeq x$
(**g**) If $\neg x \succcurlyeq y$ then $y \succ x$
(**h**) If $\neg x \succ y$ then $y \succcurlyeq x$

If you run into trouble with parts (f) and (g), note that you can always play the completeness card and throw in the expression $x \succcurlyeq y \vee y \succcurlyeq x$ any time. Also note that $p \vee q$ and $\neg p$ implies that q. If you find part (h) difficult, feel free to invoke the principle known as **de Morgan's law**, according to which $\neg (p \; \& \; q)$ is logically equivalent to $\neg p \vee \neg q$. Also note that $p \vee q$ and $p \rightarrow q$ implies that q.

For the next exercise, recall that it is acceptable to rely on propositions already established.

Exercise 2.22 Prove that if $x \sim y$ and $y \sim z$, then $\neg x \succ z$.

Exercise 2.23 Negative transitivity Prove the following two principles. You might already have been tempted to invoke these two in your proofs. But remember that you may not do so before you have established them.

(**a**) If $\neg x \succcurlyeq y$ and $\neg y \succcurlyeq z$, then $\neg x \succcurlyeq z$
(**b**) If $\neg x \succ y$ and $\neg y \succ z$, then $\neg x \succ z$

The last two exercises illustrate some potentially problematic implications of the theory that we have studied in this chapter. Both are classics.

Exercise 2.24 Vacations Suppose that you are offered two vacation packages, one to California and one to Florida, and that you are perfectly indifferent between the two. Let us call the Florida package f and the California package c. So $f \sim c$. Now, somebody improves the Florida package by adding an apple to it. You like apples, so the enhanced Florida package f^+ improves the original Florida package, meaning that $f^+ \succ f$. Assuming that you are rational, how do you feel about the enhanced Florida package f^+ compared with the California package c? Prove it.

How to do proofs

The aim of a **proof** of a proposition is to establish the truth of the proposition with logical or mathematical certainty (see the proofs of Proposition 2.12(i)–(iii) for examples). A proof is a sequence of propositions, presented on separate lines of the page. The last line of the proof is the proposition you intend to establish, that is, the conclusion; the lines that come before it establish its truth. The conclusion is typically preceded by the symbol ∴. All other lines are numbered using Arabic numerals. The basic rule is that each proposition in the proof must follow logically from (a) a proposition on a line above it, (b) an axiom of the theory, (c) a definition that has been properly introduced, and/or (d) a proposition that has already been established by means of another proof. Once a proof is concluded, logicians like to write "QED" – Latin for "quod erat demonstrandum," meaning "that which was to be shown" – or add a little box. □

There are some useful hints, or rules of thumb, that you may want to follow when constructing proofs. **Hint one**: if you want to establish a proposition of the form $x \rightarrow y$, you typically want to begin by assuming what is to the left of the arrow; that is, the first line will read "1. x by assumption." Then, your goal is to derive y, which would permit you to complete the proof. If you want to establish a proposition of the form $x \leftrightarrow y$, you need to do it both ways: first, prove that $x \rightarrow y$, and second, that $y \rightarrow x$. **Hint two**: if you want to establish a proposition of the form $\neg p$, you typically want to begin by assuming the opposite of what you want to prove for a proof by contradiction; that is, the first line would read "1. p by assumption for a proof by contradiction." Then, your goal is to derive a contradiction, that is, a claim of the form $q \, \& \, \neg q$, which would permit you to complete the proof.

Exercise 2.25 Teacups Imagine that there are 1000 cups of tea lined up in front of you. The cups are identical except for one difference: the cup to the far left (c_1) contains one grain of sugar, the second from the left (c_2) contains two grains of sugar, the third from the left (c_3) contains three grains of sugar, and so on. Since you cannot tell the difference between any two adjacent cups, you are indifferent between c_n and c_{n+1} for all n between 1 and 999 inclusive. Assuming that your preference relation is rational, what is your preference between the cup to the far left (c_1) and the one to the far right (c_{1000})?

Your findings from Exercise 2.21 are likely to come in handy when answering these questions.

2.5 Preference orderings

The preference relation is often referred to as a **preference ordering**. This is so because a rational preference relation allows us to order all alternatives in a list, with the best at the top and the worst at the bottom. Figure 2.3 shows an example of a preference ordering.

A rational preference ordering is simple. Completeness ensures that each person will have exactly one list, because completeness entails that each element can be compared with all other elements. Transitivity ensures that the list will be linear, because transitivity entails that the strict preference relation will never have cycles, as when $x \succ y$, $y \succ z$, and $z \succ x$. Here are two helpful exercises about cycling preferences.

Exercise 2.26 Cycling preferences Using the definitions and propositions discussed so far, show that it is impossible for a rational strict preference relation to cycle. To do so, suppose (for the sake of the argument) that $x \succ y$ & $y \succ z$ & $z \succ x$ and show that this leads to a contradiction.

Exercise 2.27 Cycling preferences, cont. By contrast, it is possible for the weak preference relation to cycle. This is to say that there may well be an x, y, and z such that $x \succcurlyeq y$ & $y \succcurlyeq z$ & $z \succcurlyeq x$. If this is so, what do we know about the agent's preferences over x, y, and z? Prove it.

In cases of indifference, the preference ordering will have ties. As you may have noticed, Figure 2.3 describes a preference ordering in which two items are equally good. Assuming that the universe is {Heavenly Bliss, Coke, Pepsi, Eternal Suffering}, this preference ordering is perfectly rational.

In economics, preference orderings are frequently represented using **indifference curves**, also called **indifference maps**. See Figure 2.4 for an example of a set of indifference curves. You can think of these as analogous to contour lines on a topographic map. By convention, each bundle on one of these curves is as good as every other bundle on the same curve. When two bundles are on different curves, one of the two bundles is strictly preferred to the other. Insofar as people prefer more of each good to less, bundles on curves to the top right will be strictly preferred to bundles on curves to the bottom left.

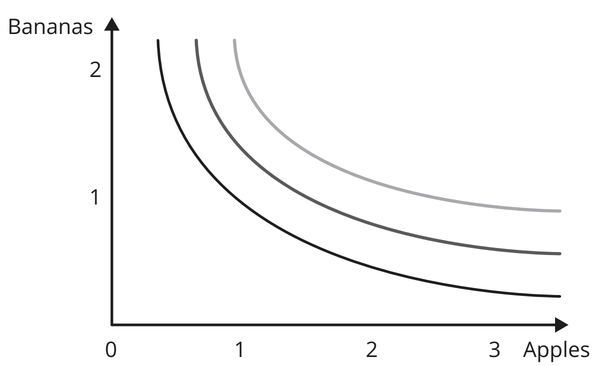

Figure 2.4 Indifference curves

Exercise 2.28 Indifference curves Represent the following sets of indifference curves graphically:
(**a**) Suppose that an apple for you is always as good as two bananas.
(**b**) Suppose that one apple is always as good, as far as you are concerned, as a banana.
(**c**) Suppose that you do not care for tea without milk or for milk without tea. However, every time you have two units of tea and one unit of milk, you can make yourself a cup of tea with milk. You love tea with milk, and the more the better, as far as you are concerned.

2.6 Choice under certainty

To make a **choice under certainty** is to face a menu. A **menu** is a set of options such that you have to choose exactly one option from the set. This is to say that the menu has two properties. First, the items in the menu are **mutually exclusive**; that is, you can choose at most one of them at any given time. Second, the items in the menu are **exhaustive**; that is, you have to choose at least one of them.

Example 2.29 The menu If a restaurant offers two appetizers (soup and salad) and two entrées (chicken and beef) and you must choose one appetizer and one entrée, what is your set of alternatives?

Since there are four possible combinations, your set of alternatives is {soup-and-chicken, soup-and-beef, salad-and-chicken, salad-and-beef}.

Exercise 2.30 The menu, cont. If you can also choose to eat an appetizer only, or an entrée only, or nothing at all, what would the new menu be?

There is no assumption that a menu is small, or even finite, though we frequently assume that it is.

In economics, the menu is often referred to as the **budget set**. This is simply that part of the set of alternatives that you can afford given your budget, that is, your resources at hand. Suppose that you can afford at most three apples (if you buy no bananas) or two bananas (if you buy no apples). This would be the case, for instance, if you had \$6 in your pocket and bananas cost \$3 and apples \$2. If so, your budget set – or your menu – is represented by the shaded area in Figure 2.5. Assuming that fruit is infinitely divisible, the menu is infinitely large. The line separating the items in your budget from the items outside of it is called the **budget line**.

Exercise 2.31 Budget sets Suppose that your budget is \$12. Use a graph to answer the following questions:
(**a**) What is the budget set when apples cost \$3 and bananas cost \$4?
(**b**) What is the budget set when apples cost \$6 and bananas cost \$2?
(**c**) What is the budget set when apples always cost \$2, the first banana costs \$4, and every subsequent banana costs \$2?

So what does it mean for a person **to be rational**? To be rational, or **to make rational choices**, means (i) that you have a rational preference ordering, and (ii) that whenever you are faced with a menu, you choose the most preferred item, or (in the case of ties) one of the most preferred items. The second condition can also be expressed as follows: (ii') that … you choose an item such that no other item in the menu is strictly preferred to it. Or like this: (ii") that … you do not choose an item that is strictly less preferred to another item in the menu. *This is all we mean when we say that somebody is rational in the context of choice under certainty*. If you have the preferences of Figure 2.3 and are facing a menu offering Coke, Pepsi, and Eternal Suffering, the rational choice is to pick either the Coke or the Pepsi option. When there is no unique best choice, as in this case, the theory says that you have to choose one of the best options; it does not specify which one.

The rational decision can be determined if we know the agent's indifference curves and budget set. If you superimpose the former (from Figure 2.4) onto the latter (from Figure 2.5), you get a picture like Figure 2.6. The consumer will choose the bundle marked X, because it is the most highly preferred bundle in the budget set. As you can tell, there is no more highly preferred bundle in the budget set.

It is important to note what the theory of rationality does *not* say. The theory does not say why people prefer certain things to others, or why they choose so as to satisfy their preferences. It does not say that people prefer apples to bananas because they

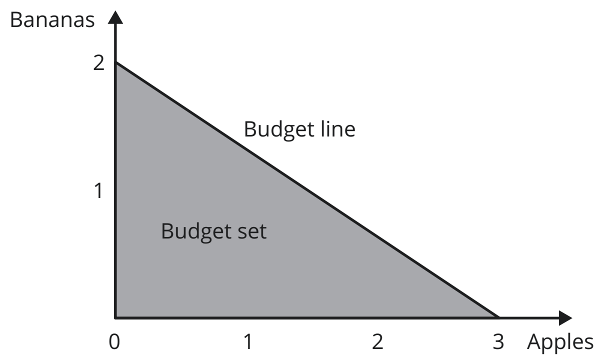

Figure 2.5 Budget set

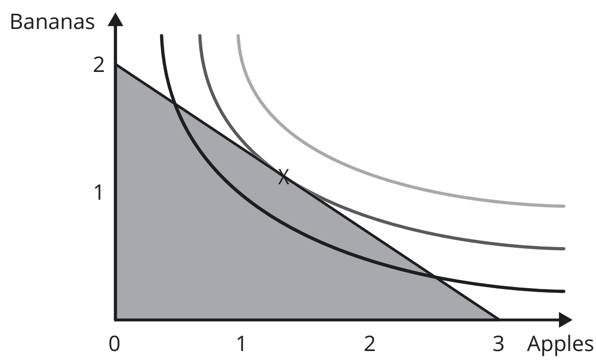

Figure 2.6 Consumer choice problem

think that they will be happier, feel better, or be more satisfied if they get apples than if they get bananas (although that may, in fact, be the case). This theory says nothing about feelings, emotions, moods, or any other subjectively experienced state. As far as this theory is concerned, the fact that you prefer a cool drink on a hot day to being roasted over an open fire is just a brute fact; it is not a fact that needs to be grounded in an account of what feels good or bad, pleasant or unpleasant, rewarding or aversive. Similarly, the theory does not say why people choose the most preferred item on the menu; as far as this theory is concerned, they just do.

Moreover, the theory does not say that people are selfish, in the sense that they care only about themselves; or that they are materialistic, in the sense that they care only about material goods; or that they are greedy, in the sense that they care only about money. The definition of rationality implies that a rational person is **self-interested**, in the sense that her choices reflect her own preference ordering rather than somebody else's. But this is not the same as being selfish: the rational individual may, for example, prefer dying for a just cause over getting rich by defrauding others. The theory in itself specifies only some formal properties of the preference relation; it does not say anything about the things people prefer. The theory is silent about whether or not they pursue respectable and moral ends. Rational people may be weird, evil, sadistic, selfish, and morally repugnant, or saintly, inspiring, thoughtful, selfless, and morally admirable; they can act out of compulsion, habit, feeling, or as a result of machine-like computation. This conception of rationality has a long and distinguished history. The Scottish eighteenth-century philosopher and economist David Hume wrote:

> 'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. 'Tis not contrary to reason for me to choose my total ruin, to prevent the least uneasiness of an Indian or person wholly unknown to me. 'Tis as little contrary to reason to prefer even my own acknowledge'd lesser good to my greater, and have a more ardent affection for the former than the latter.

Rational people cannot have preferences that are intransitive or incomplete, and they cannot make choices that fail to reflect those preferences.

2.7 Utility

The notion of **utility**, which is central to modern economics, has generated a great deal of confusion. It is worth going slowly here. Suppose that you want to use numbers to express how much a person prefers something, then how would you do it? One solution is obvious. Remember that a rational person's preferences allow us to arrange all alternatives in order of preference. Consider, for example, the preference ordering in Figure 2.3. The preference ordering has three "steps." In order to represent these preferences by numbers, we assign one number to each step, in such a way that higher steps are associated with higher numbers. See Figure 2.7 for an example.

A **utility function** associates a number with each member of the set of alternatives. In this case, we have associated the number 3 with Heavenly Bliss (HB). That number is called the utility of HB and is denoted $u(\text{HB})$. In this case, $u(\text{HB}) = 3$. The number

Heavenly Bliss	—	3
$\succ$		
Coke $\sim$ Pepsi	—	2
$\succ$		
Eternal Suffering	—	1

Figure 2.7 Preference ordering with utility function

associated with Eternal Suffering (ES) is called the utility of ES and is written $u(\text{ES})$. In this case, $u(\text{ES}) = 1$. If we use C to denote Coke and P to denote Pepsi, then $u(\text{C}) = u(\text{P}) = 2$. Because we designed the utility function so that higher utilities correspond to more preferred items, we say that the utility function $u(\cdot)$ **represents** the preference relation $\succcurlyeq$.

As the example suggests, two conditions must hold in order for something to be a utility function. First, it must be a function (or a mapping) from the set of alternatives into the set of real numbers. This means that every alternative gets assigned exactly one number. If Figure 2.7 had empty spaces in the right-hand column, or if the figure had several numbers in the same cell, we would not have a proper utility function. While the utility function needs to assign some number to every alternative, it is acceptable (as the example shows) to assign the same number to several alternatives. Second, for something to be a utility function, it must assign larger numbers to more preferred alternatives; that is, if x is at least as good as y, the number assigned to x must be greater than or equal to the number assigned to y. To put it more formally:

Definition 2.32 Definition of $u(\cdot)$ *A function $u(\cdot)$ from the set of alternatives into the set of real numbers is a utility function representing the preference relation $\succcurlyeq$ just in case $x \succcurlyeq y \Leftrightarrow u(x) \geq u(y)$ (for all x and y).*

A function $u(\cdot)$ that satisfies this condition can be said to be an **index** or a **measure** of the preference relation $\succcurlyeq$. Historically, the word "utility" has been used to refer to many different things, including the pleasure, happiness, and satisfaction of receiving, owning, or consuming something. Though most people (including economics professors) find it hard to stop speaking in this way, as though utility is somehow floating around "in your head," this usage is archaic. Utility is nothing but an index or measure of preference.

Given a rational preference relation, you may ask whether it is always possible to find a utility function that represents it. When the set of alternatives is finite, the answer is yes. The question is answered by means of a so-called **representation theorem**.

Proposition 2.33 Representation theorem *If the set of alternatives is finite then $\succcurlyeq$ is a rational preference relation just in case there exists a utility function representing $\succcurlyeq$.*

Proof.
Omitted. □

When the set of alternatives is infinite, representing preference relations gets more complicated. It remains true that if a utility function represents a preference relation, then the preference relation is rational. However, even if the preference relation is rational, it is not always possible to find a utility function that represents it.

As you may suspect, a utility function will associate strictly higher numbers with strictly preferred alternatives, and equal numbers with equally preferred alternatives. That is, the following proposition is true:

Proposition 2.34 Properties of $u(\cdot)$ *Given a utility function $u(\cdot)$ representing the preference relation $\succcurlyeq$, the following conditions hold:*

(i) $x \succ y \Leftrightarrow u(x) > u(y)$
(ii) $x \sim y \Leftrightarrow u(x) = u(y)$

Proof.

(i) First, assume that $x \succ y$, so that $x \geq y$ and $\neg y \geq x$. Using Definition 2.32 twice, we can infer that $u(x) \geq u(y)$ and that not $u(y) \geq u(x)$. Simple math tells us that $u(x) \succ u(y)$. Second, assume that $u(x) \succ u(y)$, which implies that $u(x) \geq u(y)$ and that not $u(y) \geq u(x)$. Using Definition 2.32 twice, we can infer that $x \succcurlyeq y$ and $\neg y \succcurlyeq x$, which in turn implies that $x \succ y$.

(ii) See Exercise 2.35. □

Recall (from the text box on page 20) that if you want to prove something of the form A ⇔ B, your proof must have two parts.

Exercise 2.35 Prove Proposition 2.34(ii).

It is easy to confirm that the proposition is true of the utility function from Figure 2.7.

One important point to note is that utility functions are not unique. The sequence of numbers ⟨1, 2, 3⟩ in Figure 2.7 could have been chosen very differently. The sequence ⟨0, 1, 323⟩ would have done as well, as would ⟨–1000, –2, 0⟩ and ⟨$-\pi$, e, 1077⟩. All these are utility functions, in that they associate higher numbers with more preferred options. As these examples show, it is important not to ascribe any significance to absolute numbers. To know that the utility I derive from listening to Justin Bieber is 2 tells you *absolutely nothing* about my preferences. But if you know that the utility I derive from listening to Rihanna is 4, you know something, namely, that I strictly prefer Rihanna to Justin Bieber. It is equally important not to ascribe any significance to ratios of utilities. Even if the utility of Rihanna is twice the utility of Justin Bieber, this does not mean that I like Rihanna "twice as much." The same preferences could be represented by the numbers 0 and 42, in which case the ratio would not even be well defined. In brief, for every given preference relation, there are many utility

A final word about proofs

While the proofs discussed in this chapter may at first blush seem intimidating, notice that the basic principles are fairly simple. So far, we have introduced only two axioms, namely, the transitivity of the weak preference relation (Axiom 2.5 on page 14) and the completeness of the weak preference relation (Axiom 2.6 on page 14); three definitions, namely, the definition of indifference (Definition 2.11 on page 16), the definition of strict preference (Definition 2.15 on page 18), and the definition of utility (Definition 2.32 on page 25); and two hints (see text box on page 20). In order to complete a proof, there are only seven things that you need to know.

functions representing it. Utility as used in this chapter is often called **ordinal utility**, because all it does is allow you to order things.

How do utilities relate to indifference curves? A utility function in effect assigns one number to each indifference curve, as in Figure 2.8. This way, two bundles that fall on the same curve will be associated with the same utility, as they should be. Two bundles that fall on different curves will be associated with different utilities, again as they should be. Of course, higher numbers will correspond to curves that are more strongly preferred. For a person who likes apples and bananas, $u_1 < u_2 < u_3$.

How does utility relate to behavior? Remember that you choose rationally insofar as you choose the most preferred item (or one of the most preferred items) on the menu. The most preferred item on the menu will also be the item with the highest utility. So to choose the most preferred item is to choose the item with the highest utility. Now, **to maximize utility** is to choose the item with the highest utility. Thus, you choose rationally insofar as you maximize utility. Hence, *to maximize utility is to choose rationally*. Notice that you can maximize utility in this sense without necessarily going through any particular calculations; that is, you do not need to be able to solve mathematical maximization problems in order to maximize utility. Similarly, you can maximize utility without maximizing feelings of pleasure, satisfaction, contentment, happiness, or whatever; utility (like preference) still has nothing to do with subjectively experienced states of any kind. This is a source of endless confusion.

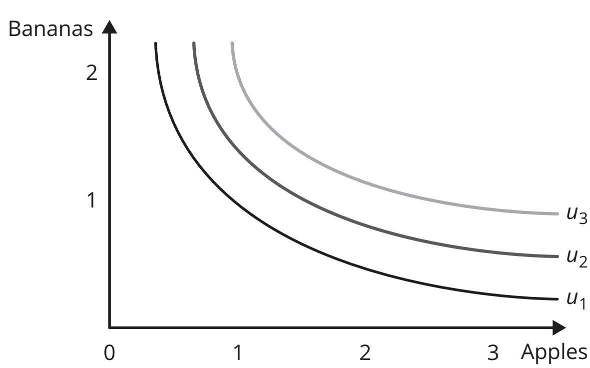

Figure 2.8 Indifference curves and utility

Consumer-choice theory

The theory of rational choice under certainty is for practical purposes often supplemented with various other, **auxiliary**, assumptions. The following three assumptions tend to appear under the heading of **consumer-choice theory** in microeconomics textbooks, and are worth knowing because they are so common. It is important to notice, however, that these auxiliary assumptions are not strictly speaking part of the theory of rational choice; they are useful add-ons.

Non-satiation Non-satiation says that no matter what you have, there is always a nearby bundle that you would rather have. Consider any bundle x in the set of alternatives (as in Figure 2.1) and draw a circle around it. Non-satiation says that no matter how small the circle around x, there is always another bundle within the circle that is strictly preferred to x. Diners who exhibit non-satiation, for example, will never be perfectly satisfied with their meal, no matter how good: they could always use a little more salad, or a little more champagne. Non-satiation is a convenient assumption, because it guarantees that the solution to the consumer choice problem (Figure 2.6) lies on the budget line. How do you know? Suppose, for a proof by contradiction, that the most highly preferred bundle in your budget set, call it x, is not in fact on the budget line. By non-satiation every circle drawn around x contains some bundle that you strictly prefer to x. If that circle is small enough, the preferred bundle will also be inside the budget set – but then x cannot in fact be the most highly preferred bundle in your budget set. QED.

Convexity Convexity in preferences captures a preference for variety or combination. Start with any two points on one indifference curve and draw a straight line between them. Convexity requires that points on the chord (excluding the end points) are preferred to the end points. Whenever this condition is satisfied, indifference curves will bulge toward the origin (the point where the axes intersect) in the manner of Figures 2.4 and 2.8. A person with convex preferences will always prefer one unit of gin & tonic to either one unit of gin or one unit of tonic. Convexity excludes snake-shaped indifference curves and guarantees that the consumer choice problem in Figure 2.6 will have a unique solution.

Continuity Continuity in preferences says that a person has similar preferences for similar bundles. Suppose that x is weakly preferred to y, and that there is another bundle x^n which is similar to x. Continuity requires that when x^n becomes ever more similar to x, in the limit, x^n must also be weakly preferred to y. This assumption guarantees that there are no "jumps," where a person has radically different preferences over very similar bundles. Continuity excludes so-called **lexicographic** preferences. You have lexicographic preferences whenever you always prefer a bundle with the largest amount of a and only in cases of ties prefer the one with more b. Lexicographic preferences involve jumps, since they would make you strictly prefer $\langle a + \varepsilon, b \rangle$ to $\langle a, b \rangle$, no matter how small $\varepsilon > 0$ is. The continuity assumption is especially useful in the context of utility theory, where it guarantees that representation theorems (analogous to Proposition 2.33) go through even when the set of alternatives is infinitely large.

2.8 Discussion

The first thing to notice is how much mileage you can get out of a small number of relatively weak assumptions. Recall that we have made only two fundamental assumptions: that preferences are rational and that people choose so as to satisfy their preferences. As long as these two assumptions are true, and the set of alternatives is not too large, we can define the concept of utility and make sense of the idea of utility maximization. That is the whole theory. The second thing to notice is what the theory does not say. The theory does not say that people are selfish, materialistic, or greedy; it says nothing about why people prefer one thing over another; it does not presuppose that people solve mathematical maximization problems in their heads; and it makes no reference to things like pleasure, satisfaction, and happiness. The fact that the theory is relatively noncommittal helps explain why so many economists are comfortable using it: after all, the theory is compatible with a great deal of behavior.

Though brief, this discussion sheds light on the nature of economics as some economists see it. Nobel laureate Gary Becker defines the economic approach to behavior in terms of three features: "The combined assumptions of maximizing behavior, market equilibrium, and stable preferences, used relentlessly and unflinchingly, form the heart of the economic approach as I see it." Because this part focuses on individual choice, I have little to say about market equilibrium here. However, what Becker has in mind when he talks about maximizing behavior and stable preferences should be eminently clear from what has already been said. In this analysis, preferences are **stable** in the sense that they are not permitted to change over time.

Exercise 2.36 Misguided criticism Many criticisms of standard economics are quite mistaken. Explain where the following critics go wrong.

(**a**) An otherwise illuminating article about behavioral economics in *Harvard Magazine* asserts that "the standard model of the human actor – Economic Man – that classical and neoclassical economics have used as a foundation for decades, if not centuries … is an intelligent, analytic, selfish creature."

(**b**) A common line of criticism of standard economics begins with some claim of the form "the most fundamental idea in economics is that money makes people happy" and proceeds to argue that the idea is false.

Is this a plausible theory of human behavior under conditions of certainty? To answer this question we need to separate the descriptive from the normative question. The first question is whether the theory is descriptively adequate, that is, whether people's choices *do as a matter of fact* reflect a rational preference ordering. This is the same as asking whether people maximize utility. Though both transitivity and completeness may seem obviously true of people's preferences, there are many cases in which they do not seem to hold: a person's preference relation over prospective spouses, for example, is unlikely to be complete. The second question is whether the theory is normatively correct, that is, whether people's choices *should* reflect a rational preference ordering. This is the same as asking whether people *should* maximize utility. Though transitivity and completeness may seem rationally required, it can be argued that they are neither necessary nor sufficient for being rational.

Next, we explore what happens when the theory is confronted with data.

Additional exercises

Exercise 2.37 For each of the relations and properties in Table 2.1, use a check mark to identify whether or not the relation has the property.

Table 2.1 Properties of weak preference, indifference, and strong preference

	Property	Definition	$\succcurlyeq$	$\sim$	$\succ$
(**a**)	Transitivity	xRy & $yRz \rightarrow xRz$ (for all x, y, z)			
(**b**)	Completeness	$xRy \vee yRx$ (for all x, y)			
(**c**)	Reflexivity	xRx (for all x)			
(**d**)	Irreflexivity	$\neg\, xRx$ (for all x)			
(**e**)	Symmetry	$xRy \rightarrow yRx$ (for all x, y)			
(**f**)	Antisymmetry	$xRy \rightarrow \neg\, yRx$ (for all x, y)			

Exercise 2.38 More properties of the preference relation Here are two relations: "is married to" and "is not married to." Supposing the universe is the set of all living human beings, which of these is…

(**a**) reflexive
(**b**) irreflexive
(**c**) symmetric
(**d**) asymmetric
(**e**) antisymmetric

Exercise 2.39 As part of your answer to the following questions, make sure to specify what the universe is.

(**a**) Give an example of a relation that is complete but not transitive.
(**b**) Give an example of a relation that is transitive but not complete.

Exercise 2.40 Irrationality Explain (in words) why each of the characters below is irrational according to the theory you have learned in this chapter.

(**a**) In the drama *Sophie's Choice*, the title character finds herself in a Nazi concentration camp and must choose which one of her children is to be put to death. She is not indifferent and cannot form a weak preference either way.
(**b**) An economics professor finds that he prefers a $10 bottle of wine to an $8 bottle, a $12 bottle to a $10 bottle, and so on; yet he does not prefer a $200 bottle to an $8 bottle.
(**c**) Buridan's ass is as hungry as it is thirsty and finds itself exactly midway between a stack of hay and a pail of water. Unable to decide which is better, the animal expires.

Exercise 2.41 Consumer-choice theory Do the following quotations agree or disagree with the assumptions of consumer-choice theory (see box on page 28)? (**a**) "Variety is the spice of life" – William Cowper. (**b**) "(I Can't Get No) Satisfaction" – The Rolling Stones. (**c**) "People will always choose more money over more sex" – Douglas Coupland. (**d**) "Greed, for lack of a better word, is good" – Gordon Gecko.

Exercise 2.42 Consumer-choice theory and indifference curves For each of the sets of indifference curves in Figure 2.9, what assumption rules it out?

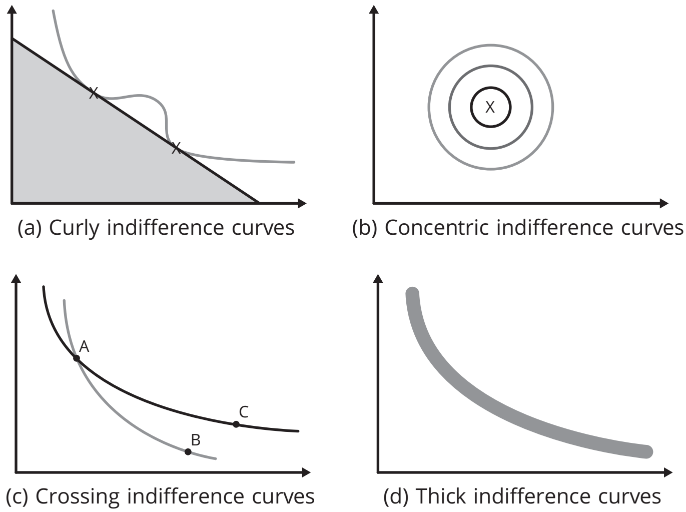

Figure 2.9 Impossible indifference curves

Further reading

A nontechnical introduction to decision theory is Allingham (2002). More technical accounts can be found in Mas-Colell et al. (1995, Chapters 1–2). The paragraph from David Hume comes from Hume (2000 [1739–40], p. 267). The Becker quotation is from Becker (1976, p. 5). The *Harvard Magazine* article is Lambert (2006), and the critics talking about happiness Dutt and Radcliff (2009, p. 8). The quotations in Exercise 2.42(a) and (c) are from Cowper (1785) and Coupland (2008).

3 DECISION-MAKING UNDER CERTAINTY

> **Learning objectives**
>
> *After studying this chapter you will:*
>
> - Be able to identify common behavior patterns that violate the theory of rational choice under certainty
> - Know some building blocks of the (descriptive) behavioral theories, including the biases-and-heuristics program and prospect theory
> - Apply the (normative) theory of rationality in the real world – but also appreciate how demanding the theory is

3.1 Introduction

The previous chapter showed how an extensive theory of choice under certainty can be built upon the foundation of a modest number of assumptions. Though the assumptions may seem weak, their implications can be challenged on both descriptive and normative grounds. In this chapter, we confront the theory with data. We explore some of the phenomena that behavioral economists argue are inconsistent with the theory of choice under certainty, as we know it. We focus on a couple of different phenomena, beginning with the failure to consider opportunity costs. Moreover, we will begin discussing what behavioral economists do when they discover phenomena that appear inconsistent with standard theory. In particular, we will discuss some of the building blocks of prominent behavioral alternatives, including prospect theory and the heuristics-and-biases program.

3.2 Opportunity costs

Imagine that you invest a small amount of money in real estate during a period when it strikes you as a safe and profitable investment. After you make the investment, the markets become unstable and you watch nervously as prices rise and fall. Finally, you sell your assets and realize that you have made a profit. "Wow," you say to yourself, "that turned out to be a great investment!" But when you boast to your friends, somebody points out that you could have earned even more money by investing in the stock market. At some level, you knew this. But you still feel that investing in real

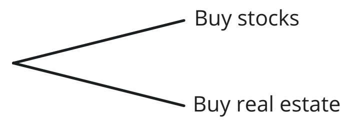

Figure 3.1 Simple decision tree

estate was a good choice: at least you did not lose any money. This is a case where you may have been acting irrationally because you failed to consider **opportunity costs**.

In order to analyze this kind of situation, let us stand back for a moment. An agent's decision problem can be represented using a **decision tree**: a graphical device showing what actions are available to the agent. Given that you only have two available actions – buying stocks and buying real estate – your decision problem can be represented as a decision tree (see Figure 3.1). Because this chapter is about choice under certainty, I will pretend that there is no uncertainty about the consequences that follow from each of these choices. (We will abandon this pretense in our discussion of choice under risk and uncertainty in Part 3.)

Suppose that you are tempted to buy real estate. What is the cost of doing so? There would be an out-of-pocket or **explicit cost**: the seller of the property would want some money to give it up. The real cost, however, is what you forgo when you buy the real estate. The opportunity cost – or **implicit cost** – of an alternative is the value of what you would have to forgo if you choose it. In dollar terms, suppose that stocks will gain $1000 over the next year and that real estate will gain $900. If so, the opportunity cost of buying real estate is $1000 and the opportunity cost of buying stocks is $900. If you buy real estate, then, your economic profit will be $900 − $1000 = −$100. If you buy stock, your economic profit would be $1000 − $900 = $100. If there are more than two options, the opportunity cost is the value of the *most valuable* alternative option. Suppose that you can choose between stocks, real estate, and bonds, and that bonds will gain $150 over the next year. The opportunity cost of buying stocks would remain $900, and the economic profit would still be $100.

Exercise 3.1 Investment problem
(**a**) Draw a decision tree illustrating this decision problem.
(**b**) What is the opportunity cost of buying real estate?
(**c**) What is the opportunity cost of buying bonds?

Decision trees make it clear that you cannot choose one alternative without forgoing another: whenever you choose to go down one branch of the tree, there is always another branch that you choose not to go down. When you vacation in Hawaii, you cannot at the same time vacation in Colorado; when you use your life savings to buy a Ferrari, you cannot at the same time use your life savings to buy a Porsche; when you spend an hour reading sociology, you cannot spend the same hour reading anthropology; when you are in a monogamous relationship with *this* person, you cannot at the same time be in a monogamous relationship with *that* person; and so on. Consequently, there is an opportunity cost associated with every available option in every decision problem.

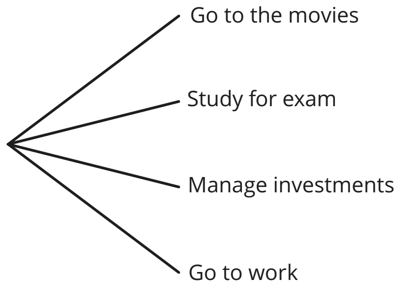

Figure 3.2 Everyday decision tree

For another example, imagine that you are considering going to the movies. On an ordinary evening, the decision that you are facing might look like Figure 3.2. Remember that the opportunity cost of going to the movies is the value of the most valuable option that you would forgo if you went to the movies; that is, the opportunity cost of going to the movies is the greatest utility you could get by going down one of the other branches of the decision tree, which is the utility of the most valuable alternative use for some $20 and two hours of your time.

As a matter of notation, we write $a_1, a_2, \ldots, a_n$ to denote the n different acts available to you; $u(a_1), u(a_2), \ldots, u(a_n)$ to denote the utilities of those acts; and $c(a_1), c(a_2), \ldots, c(a_n)$ to denote the opportunity costs of those acts. The opportunity cost $c(a_i)$ of act a_i can then be defined as follows:

Definition 3.2 Opportunity cost

$$c(a_i) = \max\left\{u(a_1),\ u(a_2),\ \ldots,\ u(a_{i-1}),\ u(a_{i+1}),\ \ldots,\ u(a_n)\right\}$$

This is just to say that the opportunity cost of act a_i equals the maximum utility of the other acts.

Figure 3.3 represents a decision problem in which utilities and opportunity costs of four acts have been identified. The number on the left is the utility; the number in parentheses is the opportunity cost. You can compute the profit (in utility terms) by subtracting the latter from the former.

Exercise 3.3 Opportunity costs This exercise refers to Figure 3.3. Suppose that a fifth act (call it a_5) becomes available. Assume that a_5 has a utility of 9.
(**a**) What would the tree look like now?
(**b**) What would happen to the opportunity costs of the different alternatives?

There is a tight connection between opportunity costs, utilities, and the rational thing to do. As it happens, you are rational – that is, you maximize utility – just in case you take opportunity costs properly into account.