OXFORD

MICROECONOMICS

Competition, Conflict, and Coordination

SAMUEL BOWLES & SIMON D. HALLIDAY

SAMUEL BOWLES & SIMON D. HALLIDAY

MICROECONOMICS: COMPETITION, CONFLICT, & COORDINATION

Contents

Preface

To its 18th and early 19th century founders, the subject of economics was the wealth of nations and people. This was no less true of Karl Marx, the most famous critic of capitalism, than it was of Adam Smith's whose *The Wealth of Nations* is considered the most powerful defence of the then emerging capitalist economic system.

Economics was at the time called political economy, and it sought to understand how and why society was being transformed as a result of capitalism, a novel way of organizing how people produce, exchange and distribute the things we live on. Capitalism continues to change the world, and the task of economics is to understand this process, and how our economies might be made to work better for people today and in the future.

Welcome to *Microeconomics: Competition, Conflict, and Coordination* and best wishes for your journey through its content. Let's begin by saying how we came to think that economics is important and then explaining our strategy for how you can best learn to do economics.

Economics engaged in the world

Contrary to its reputation among students for being remote from reality, economics has always been about changing the way the world works. The earliest economists – the Physiocrats in late 18th century France and the Mercantilists before them – were advisers to kings and queens of Europe. Today's macroeconomic managers, economic development advisors and advocates of competing policies concerning intellectual property rights or the global movements of goods and people continue this tradition of real world engagement. Economists have never been strangers to policy-making, constitution building and attempts at economic reform for the betterment of people's living conditions.

Alfred Marshall's (1842-1924) *Principles of Economics*, published in 1890 was the first great text in what came to be called neoclassical economics. It opens with these lines:

> “Now at last we are setting ourselves seriously to inquire whether .. there need be large numbers of people doomed from their birth to hard work in order to provide for others the requisites of a refined and cultured life, while they themselves are prevented by their poverty and toil from having any share or part in that life. ...[T]he answer depends in a great measure upon facts and inferences, which are within the province of economics; and this is it which gives to economic studies their chief and their highest interest.”

The hope that economics might assist in alleviating poverty and securing the conditions under which free people might flourish is at once economics’ most inspiring calling and its greatest challenge. Like many, both of us were drawn to economics by this hope.

One of us (Simon) grew up in Cape Town under the system of racial segregation called *apartheid*. He vividly remembers the demonstrations that finally brought that system down and the long lines of people waiting to vote in South Africa’s first democratic elections in 1994. He volunteered in the poor townships surrounding Cape Town teaching critical thinking and debating, skills required to make the new democracy work. . Having initially followed his passion for theater and poetry, he switched into economics to gain the analytical tools to understand and address his country’s challenges.

The other of your authors (Sam), having been a schoolboy in India and a secondary school teacher in Nigeria before turning to economics, naturally came to the field expecting that it would address the enduring problem of global poverty and inequality.

At age eleven Sam had noticed how very average he was among his classmates at the Delhi Public School – in sports, in school work, in just about everything. A question that he then asked his mother has haunted him since: "how does it come about that Indians are so much poorer than Americans, given that as people we are so similar in our abilities?" And so he entered graduate school hoping that economics might, for example, explain why workers in the United States produced almost as much in a month as those in India produce in a year, and why the Indian population was correspondingly poor.

We now know that the many conventional economic explanations for the gap in standards of living between the two countries are part of the answer but far from all of it: by any reasonable accounting, the difference in the amount of machinery, land and other capital goods per worker and in the level of schooling of the U.S. and Indian work forces explain much less than half of the difference in output per hour of work.

It seems likely that much of the unexplained difference results from causes that until recently have been less studied by economists but which are a central theme of this book. Chief among these are differences in institutions, that is differences in how the activities of the millions of actors in the two economies are coordinated by some combination of markets, private property, social norms, and governments.

What should economics be about?

We do not think that we are atypical – either among our economics colleagues, or our students, or for that matter among people generally – in our hope that economics can contribute to improving the way these institutions work. The CORE Team – a group of economic researchers and teachers who have created an open access introductory economics course (www.core-econ.org) – posed the following question to students around the world on the first day of their introductory classes: "what is the most pressing problem economists today should be addressing?" The results from a total of 4,442 students from 25 universities in twelve countries over the years 2016-18 are summarized in the word cloud in Figure 3.

The themes are remarkably consistent across universities and countries. Unemployment, inflation, and growth, all important topics in most macroeconomics courses, are on the minds of students. But inequality (along with "poverty") is the overwhelmingly dominant issue. Environmental sustainability (and "climate change"), the future of work (robots, digitalization), globalization and migration, innovation, financial instability; and how governments work ("corruption," "war") are also present. A few students also identified particular political events, like the "Brexit" referendum in 2016, that favored the U.K. leaving the European Union, as problems that economics should address. In word clouds based on more recent surveys "climate change" is as large a concern as "inequality" among students.

Figure 3: **Student replies to the question "What is the most pressing problem economists should be addressing?"** The size of the font is proportional to the frequency with which subjects mentioned the word or term. Surprisingly professional economists at the New Zealand Treasury and central bank and new hires at the Bank of England responded very similarly to students.The less frequently mentioned – smaller font– topics are more readable in the individual word clouds from each of the 25 samples of students that you can access at https://tinyco.re/6235473

The microeconomic theory that you will learn has a lot to say about these issues. Included are tried and true workhorse concepts that you have probably already encountered, like opportunity costs, mutual gains from exchange, constrained optimization and trade offs. Also essential in understanding issues like those in the word cloud are concepts that have more recently risen to prominence among economists. Examples are the importance of cooperation and social (rather than entirely selfish) motivations and modeling strategic interactions among people, including conflicts over the distribution of the mutual gains from exchange.

"If you are not doing something, you are not learning anything!"

The phrase just above is our motto when it comes to learning. Economics is not just something you learn. It is something you do. Think of studying economics as learning a new language. Mastering a large vocabulary and the grammatical rules is essential, but it is not the same as speaking the language.

The test of what you have learned after studying this book is not just what you know, but what you can do with it. Doing economics is what you can say or write – the case you can make for or against a proposed economic policy, the analysis of the reasons for some new development in the global economy – in other words what you can do as a result of what you know.

Like mastering a new language, doing economics is essential to learning the subject. And also like a language, you will learn to do economics more readily if you have a clear need to know.

We begin each chapter with a real world problem or example that can be better understood using the concepts and models to be introduced in the chapter. These opening paragraphs suggest the need to know what is to follow. The empirical examples also serve as a reminder that the point to the model is to understand the world; and as we proceed through chapters we will ask: how good a job does this particular model do in that respect?

In the margin at the beginning of each chapter is a set of learning objectives phrased as new capacities to do things that most likely you were unable to do before. We place great emphasis on your ability to solve problems in which there are right and wrong answers. But it is also important to learn how to formulate arguments and hypotheses about questions that are thus far unanswered, some of which may remain so, and to express economically informed opinions on issues that will continue to be debated due to the fact that people's values differ.

Interspersed with the contents of the chapters, but offset by boxes, are two important resources:

Mathematics Notes M-notes contain the details of mathematical derivations and other analysis as well as worked examples that illustrate the mathematical models in the text.

Checkpoints are self-tests to confirm that you understand the content of the section. The first step in "doing economics" is by checking your understanding of the passage you have just read.

At the end of each chapter you will find the following:

Important Ideas The main ideas in each chapter are provided in a table. At the end of the book, you will also find that all the definitions of the book are included as a glossary for you to consult and improve your understanding. Mastering the use of these terms is essential to doing economics. Try using each of them in a complete sentence of our own.

Making connections provides some guidance in seeing how the ideas in each chapter are connected to each other and to other themes in the book, so that you will be able to draw together the ‘big picture’ about the main messages and themes of the book. Try restating these connections making use of the terms in Important Ideas. Or better yet: make a mind map using the Important Ideas and Making Connections features.

Mathematical Notation The book contains a variety of important mathematical tools to help model the various economic ideas in the book. To assist you with your reading of each chapter and to understand better each model you encounter, we provide a table of the mathematical notation you will encounter in that chapter. There is also a complete list of the notation used at the end of the book.

We use the margins of the book for a variety of purposes:

Definitions We define important terms in the margins where they first are introduced. All of the definitions are collected in a glossary.

Reminders We put reminders in the text often to help you to see the connections of ideas throughout the book.

Example An example will often illustrate an idea with a relevant example of a person, firm, or country making decisions that are similar to those described in the text.

Fact Check When we need to verify or illustrate an idea with data or an empirical example we will do so with a Fact Check.

History These introduce you to some of those people who have contributed to economics or to relevant historical facts.

M-check If an idea requires a brief mathematical clarification that does not require it’s own M-Note, then we may convey that in a margin note.

Economics is an integrated body of knowledge, and it is best learned in a cumulative way, mastering a set of concepts and going on to use those concepts in mastering additional concepts. What this means, practically is that it is best to study earlier chapters before moving on to later ones. Sections labeled "application" however provide illustrations of how the ideas and models being taught in a particular chapter can be used, and these do not introduce new material that is essential to the chapters that follow.

Micoeconomics is waiting for you. Just do it!

Samuel Bowles and Simon Halliday

Santa Fe Institute, Santa Fe, New Mexico, U.S.A, and Smith College, Northampton, Massachusetts, U.S.A.

Part I

People, Economy and Society

> The man . . . enamored of his own ideal plan of government, . . . seems to imagine that he can arrange the different members of a great society with as much ease as the hand arranges the different pieces upon a chess-board . . . but . . . in the great chess-board of human society, every single piece has a principle of motion of its own, altogether different from that which the legislature might choose to impress upon it.
>
> If those two principles coincide and act in the same direction, the game of human society will go on easily and harmoniously, and is very likely to be happy and successful. If they are opposite or different, the game will go on miserably, and the society must be at all times in the highest degree of disorder.

Adam Smith,*Theory of Moral Sentiments*, 1759,Part VI, Section 1

As individuals, our physical capacities are hardly remarkable compared to other animals. But by coordinating with others – finding ways that our individual efforts can add up to a whole that is more than the sum of its parts – humans are unique as a species, engaging in common pursuits on a global scale and for better or worse, transforming nature and inventing previously unimagined devices and ways of life. Economics provides a lens for studying this social aspect of human uniqueness by analysing how people interact with each other and with our natural surroundings in producing and acquiring our livelihoods.

HISTORY What makes humans unique among all the animals is our capacity to cooperate in very large numbers and to adjust the ways that we cooperate to changing circumstances. Here (https://tinyurl.com/y3bpy4px) the Israeli historian Yuval Noah Harari explains why this is so.

ECONOMICS is the study of how people interact with each other and with our natural surroundings in producing and acquiring our livelihoods.

We begin (in Chapter 1) by developing a common framework for studying the various types of social interactions using game theory to pose a question older than economics. This is: how can a society's institutions – its laws, unwritten rules and social norms – harness individuals' pursuit of their own objectives to generate common benefits and to avoid outcomes that none would have chosen. The challenge is how to combine freedom – individuals' pursuit of their own objectives – with the common good, improving the livelihoods of all members of the society.

This challenge is called the problem of societal coordination: how we can coordinate – that is organize – our actions to yield desirable results for society? The example of societal coordination we use in Chapter 1 to illustrate this challenge is about the other aspect of economics: how we relate to our natural surroundings, illustrated by a problem of over-exploiting an environmental resource.

Adam Smith, considered by many to be the founder of economics, understood the challenge well. And he understood that economics – or "political economy" as it was then called –is fundamentally a *social* science: it is about how people interact. You can see this in his warning above about the disastrous consequences of treating people as if they were simply chess pieces who could be moved around on the chess board of life at the will of a government, more or less like an engineer might design a machine.

An adequate response to the challenge of combining freedom and the common good must therefore be based on knowledg of how individuals act depending on the situation they are in, and how changing the situation will change how they act. We therefore (Chapter 2) turn to individuals and their motives– whether self interested or generous, opportunistic or ethical – explaining how people do the best they can in given situations. In this chapter we consider individuals in situations where they act in isolation rather than interacting with other individuals.

But people rarely act in isolation: Economics allows us to understand the sometimes surprising or unintended society-wide effects of when we interact with others, whether it be directly with our own employer or indirectly with literally millions of people involved in producing and distributing the goods making up our livelihoods.

A basic insight for this understanding is that we are better off by interacting with others. But our interactions also give rise to conflicts. When people engage with others in buying and selling, working and investing there are mutual benefits potentially available to all parties involved. This must be the case if participation in these and other economic activities is voluntary. But unavoidably there are also conflicts over how these mutual gains are divided (Chapter 4).

We evaluate the outcomes of economic interactions by two standards:

- *Efficiency*: the extent to which all of the potential gains are realized (which is how economists use the term efficiency) and

- *Fairness*: whether the distribution of the gains and the process that determines who gets what is just.

And we study the various ways that exchanges and other economic activities may be carried out and how they may affect both the efficiency and fairness of the outcome.

In our interactions with each other and with nature we frequently fail to exploit all of the potential mutual gains. An example is when a person with the capacity and desire to produce goods and services needed by others cannot find a job. Another is over-exploitation of a fishery or some other environmental resource. These are called coordination failures because they result from inadequacies in the ways that our institutions coordinate the ways that we interact.

Coordination failures are often due to our conflicts over the distribution of potential mutual gains or to the fact that when we act we do not take account of the effects of our actions on others (Chapter 5). Markets, government policies, well-designed property rights, a concern for one another's well being, and communities can help address these coordination failures so that no potential mutual gains remain unexploited, and the distribution of gains is regarded as fair. The final chapter of this book will bring together the concepts, models and other ways of thinking that you have learned and apply them to the challenge of improving the way the economy works by both of these standards: efficiency and fairness.

1

Society: Coordination Problems & Economic Institutions

> Two neighbors may agree to drain a meadow, which they possess in common; because 'tis easy for them to know each others mind; and each must perceive, that the immediate consequence of his failing in his part, is the abandoning of the whole project.
>
> But 'tis very difficult and indeed impossible, that a thousand persons shou'd agree in any such action; it being difficult for them to concert so complicated a design, and still more difficult for them to execute it; while each seeks a pretext to free himself of the trouble and expense, and wou'd lay the whole burden on others.
>
> David Hume, *A Treatise of Human Nature*, Volume II (1967 [1742]: 304)

DOING ECONOMICS

This chapter will enable you to::

- Use game theory to analyze how people interact in the economy, each affecting the conditions under which the others decide how to act.
- Understand why the outcomes of interactions are often worse for people than they could be and how interactions might be better organized to improve the quality of people's lives.
- Recognize that unsatisfactory outcomes occur when people fail to coordinate with each other and to take account of the effect of their own actions on others.
- Explain how problems like environmental damage and global poverty can be the result of failed coordination.
- Represent institutions as "the rules of the game."
- Identify that economic institutions determine incentives for people's behavior and can affect how successfully we address coordination problems.
- Explain why when people have limited information and conflicts of interest they often fail to implement 'win-win' outcomes.

At the turn of the present century, the process of economic development had bypassed almost all of the two hundred or so families that made up the village of Palanpur in the Indian state of Uttar Pradesh. But for the occasional watch, bicycle or irrigation pump, Palanpur appeared to be a timeless backwater, untouched by India's cutting edge software industry and booming agricultural regions. Less than a third of the adults were literate, and most had endured the loss of a child to malnutrition or to illnesses that had long been forgotten in other parts of the world.

A visitor to the village approached a farmer and his three daughters weeding a small plot of land. The conversation turned to the fact that Palanpur farmers plant their winter crops several weeks after the date that would maximize the amount of grain they could get at harvest time. The farmers knew that planting earlier would produce larger harvests, but no one, the farmer explained, wants to be the first farmer to plant their seeds, as the seeds on any lone plot would be quickly eaten by birds.

Figure 1.1: **Palanpur farmers threshing and winnowing grain (separating grain from chaff)** . Photo courtesy of Nicholas Stern.

Curious, the visitor asked if a large group of farmers, perhaps members of the same extended family, had ever agreed to plant their seeds earlier, perhaps on the same day to minimize the individual losses. "If we knew how to do that," the farmer said, looking up from his hoe and making eye contact with the visitor for the first time, "we would not be poor."[1]

1.1 Societal coordination: The classical institutional challenge

For the Palanpur farmers, the decision when to plant is a **coordination problem**. A coordination problem is a situation in which people could all be better off, or at least some be better of and none be worse off, if they all jointly decided how to act – that is, if they coordinated their actions – than if they act individually.

The planting choice is a *coordination problem* because:

- the farmer does better or worse depending on what other farmers do,
- all the farmers would do better if they could *coordinate* their actions by jointly agreeing to all do what would be mutually beneficial namely, planting early, but
- it is a *problem* because the farmers may not be able to coordinate, and
- if they do *not* coordinate, then all of the farmers will do worse than they all could otherwise have done (that is, had they all planted late).

To stress the fact that coordination problems often affect an entire population (even though we explain them using two person examples) we sometimes use the expression *societal* coordination problems. Notice that one farmer cannot dictate the actions of the other farmers, nor can they come to a common agreement about what to do ("if we knew how to do that, we would not be poor") – the inability to come together and coordinate is at the heart of coordination problems.

Our example is about farming, but it could just as well have been about the owners of many firms each producing some different product deciding independently whether to invest in new buildings and equipment. Firms will only invest only if they anticipate that there will be sufficient demand for the resulting increase in their outputs. Each firm's investment – purchasing new machinery and construction materials, and hiring more employees – means greater demand for the *other* firms' products, including the products purchased by the newly employed workers constructing the new offices, factories and the like with their earnings.

If they all invest, then the other firms' investments will create sufficient demand to purchase the output from each of the firms' expanded capacity. But if one firm expands and the rest do not, then that firm is likely to find that it cannot sell all of what it is now capable of producing.

COORDINATION PROBLEM A **coordination problem** is a situation in which people could all be better off (or at least some be better of and none be worse off) if they jointly decide how to act – that is, if they coordinate their actions – than if they act independently. For example, deciding on which side of the road to drive is a coordination problem.

HISTORY In his address accepting the Nobel Prize for economics in 1979, University of Chicago economist T.W. Schultz said: "Most of the people in the world are poor, so if we knew the economics of being poor, we would know much of the economics that really matters." He was right then and he is right now.[2]

The best choice for each firm depends on the choices made by other firms. If they could coordinate – decide jointly that they would all invest – they would all profit, but coming to such an agreement may not be possible. The owners of firms therefore face a coordination problem similar to the problem faced by the Palanpur farmers. Their individual success depends on their ability to coordinate their actions with others. This is not a new problem: it was on a central concern of the founders of economics.

David Hume (the 18th century British philosopher and economist quoted at the start of this chapter) used an example – two landowners considering draining a meadow – to pose what he considered the most important problem facing society, namely, devising institutions that would reconcile the pursuit of individual objectives (avoiding the "trouble and expense" in his example of the meadow) with getting desired societal outcomes (improving the value of the meadow by draining it). His simple two-person example was meant to illustrate the need (in a society of "a thousand persons") for a government to address the broader societal coordination problems of his day.

Though the term was invented only two centuries after Hume, he was using what we now call game theory to make his case. Let's apply his reasoning to the farmers of Palanpur. Like Hume we will consider just two farmers as a way of representing the institutional challenge faced by the entire village.

Figure 1.2 shows the outcomes for two players, Aram and Bina, choosing when to plant their grain. The figure illustrates the values of the farmers' crops, which, in a poor village like Palanpur, is the main incentive for farmers, whether they consume the crop themselves or sell it for money income to spend on other things. Each farmer can either plant early or plant late and while (also as in Hume's example) two people could probably come to some agreement about what to do, remember that we are using this two-person example to illustrate the entire village of about 200 families of farmers, so we assume that they cannot coordinate on some agreed upon actions for the two jointly. There are four possible outcomes:

- If both players plant early, they each achieve their *best* possible harvest, because they grow the most grain through sharing the risk of having their seeds eaten by birds (outcome c in Figure 1.2).
- If Aram plants early while Bina plants late, Aram has his seeds eaten by birds and gets no harvest (the *worst* outcome for him), whereas the late planter gets a *good* (but not the best) harvest. While none of her seeds are eaten by the birds, planting late is not the best for growing the most grain (outcomes b and d in Figure 1.2). The same is true if Bina planted late when Aram planted early.
- If both plant late, they harvest a smaller crop while also sharing the risk of their seeds being eaten, a *bad* outcome (outcome a in Figure 1.2).

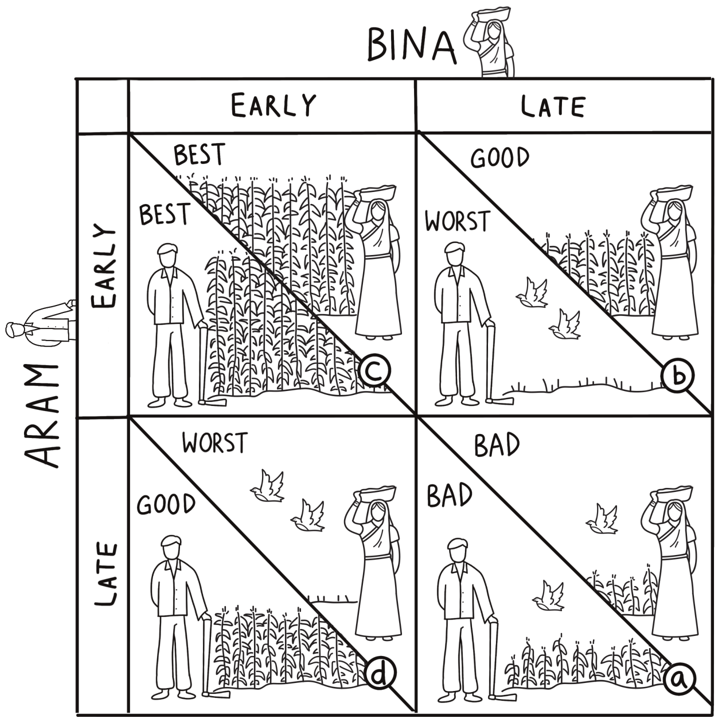


Figure 1.2: **Planting in Palanpur.** This figure shows "what-if" outcomes for planting in Palanpur. Each column represents a possible combination of Aram planting early or late and Bina planting early or late with the corresponding outcomes being worst, bad, good, or best in terms of how much grain they grow.

COORDINATION FAILURE A **coordination failure** occurs when people facing a coordination problem fail to coordinate their actions in a way to implement outcome that allows them all to be better off (or at least some to be better off and none to be worse off).

The people of Palanpur are stuck in the bad outcome even though they would all be better off if they all planted early (they would both move from a "bad" outcome to the "best" outcome in the figure). They are experiencing a **coordination failure**, namely a coordination problem that is not addressed by appropriate institutions. A modern day David Hume would point out that a government could simply impose a sufficient tax on those planting late to ensure that most farmers would plant early.

Adam Smith, a generation after Hume, would stress the value of the exchange of privately owned goods on competitive markets as a way of coordinating the actions of large numbers of people, who would be guided (even without knowing it) by what he termed "an invisible hand." Hume, Smith and the other founders of European political philosophy and political economy posed what we call the classical institutional challenge.

HISTORY Adam Smith wrote the following: "[E]very individual [...], indeed, neither intends to promote the public interest, nor knows how much he is promoting it [...] he intends only his own security; ... he intends only his own gain, and he is in this ... led by an invisible hand to promote an end which was no part of his intention ... By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it."[3]

These philosophers and economists wanted to know how *to design* institutions – *rules and practices governing peoples' behavior – so that people could be left free to make their own decisions, and at the same time avoid outcomes that were inferior for everyone.* More precisely, how do we design institutions which encourage coordination by free choice while avoiding poor outcomes such as planting late in Palanpur? The 18th and 19th century political economists and philosophers who founded the field of economics were attempting to provide solutions to coordination problems.

Checkpoint 1.1: Planting in Palanpur: A Coordination Problem

Imagine that you are Bina in the figure above, and that you did not know whether Aram would plant early or late. What would you do? Suppose you and Aram were neighbors and you could talk with him; what would you say?

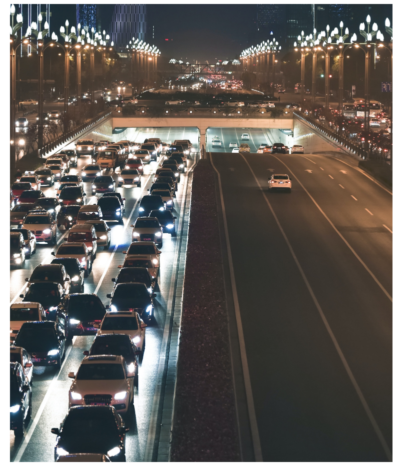


Figure 1.3: **Traffic headed out of a major city.** Image Credit: Photo by Preillumination SeTh (@7seth).

1.2 *The institutional challenge today*

The classical institutional challenge remains with us, although some of the forms that it takes today – including global climate change and the appropriate intellectual property rights for sharing digitized knowledge – were unknown to the great 18th and 19th century thinkers.

Consider the following coordination problems:

- How do we sustain the global environment? To avoid damaging climate change we need to coordinate our reduction of emissions. Many people and firms would prefer that someone *else* reduce *their* carbon footprint. How can we address climate change in a way that is both fair and imposes the least possible costs?
- How do we make the best use of our ability to create and use knowledge? If we all agree to share the knowledge we have with others we may all be better off: when I transfer my knowledge to you I do not lose the ability to continue using it. But each of us may profit by restricting others' use of our knowledge by means of patents, copyrights and other intellectual property rights.
- How do we move around a city without overcrowding streets and causing delays? My decision whether to drive, walk, or take public transport affects not only my own travel time, but also the degree of traffic congestion and delays experienced by everyone else. Everyone might be better off if the use of private vehicles was substantially reduced, but few will reduce their driving unless other people reduce theirs as well.

These are all coordination problems because an outcome that is better for all is possible if people find a way to jointly agree to a course of action. But for reasons we will explain in detail, people also routinely fail to coordinate and suffer bad consequences as a result, including the following:

- *over*use of some resources illustrated by pollution, over-grazing, traffic congestion, and climate change; and
- *under*use of other resources such as the productive capacities and creativity of people and the knowledge that we have created, illustrated by

unemployment and the enduring poverty of the people of Palanpur and villages like it around the world.

Checkpoint 1.2: Coordination Problems You Have Known

Think of a social interaction in which you have been involved that was a *coordination problem* and using the description of why planting in Palanpur is a coordination problem (the bulleted points above) explain why it was a problem and how coordination might have (or did) address the problem.

1.3 Anatomy of a coordination problem: The tragedy of the commons

The over use of environmental resources provides a good illustration of why coordination problems arise.

In 1968, Garrett Hardin, an ecologist, famously described what he called the **tragedy of the commons**, an example of a coordination failure.[4] He told a story about a group of herders who share a pasture. The pasture was common land – hence a "commons" – shared by many herders. But why was his story a tragedy?

TRAGEDY OF THE COMMONS The tragedy of the commons is a term used to describe a coordination failure arising when a shared resource available for all to use ('the commons') is over-used so that all users are worse off than they would have been if they had coordinated their actions so that use was restricted.

Each herder could put as many animals in the pasture as they wished, and overgrazing will lead to erosion and the ruin of the pasture. Hardin reasoned that if the land is common to all and no one herder owns it, each herder has no interest in limiting how many animals they put in the common pasture. A ruined pasture is of no value to any of the herders. But each herder's self-interest leads them to neglect the effect their actions have on others. The outcome is a tragedy.

With the term *tragedy of the commons*, Hardin gave social science one of the most evocative metaphors since Adam Smith's "invisible hand." Indeed Hardin called his tragedy a "rebuttal to the invisible hand." The two metaphors are powerful because they capture two essential yet contrasting social insights. When guided by an invisible hand, social interactions reconcile individual choice and socially desirable outcomes. By contrast, the actors in the tragedy of the commons pursue their private objectives to tragic consequences for themselves and others.

The natural setting for Hardin's tragedy was chosen for its imagery. The underlying problem applies to many situations where people typically cannot or do not take account of the effects of their actions on the well-being of others. You can think of a city's streets as a commons, and people deciding to drive rather than walk, bike, or use public transport as similar to the herders putting cattle on the common. The modern day "tragedy of the roadways" is a traffic jam.

What are the common elements in Hume's drain the meadow problem, the farmers in Palanpur planting late, Hardin's herders overgrazing their pasture and our modern city dwellers clogging the streets with their vehicles?

EXTERNAL EFFECT An external effect occurs when a participant's action confers a benefit or imposes a cost on other participants and this cost or benefit is not taken into account by the individual taking the action. External effects are also called simply *externalities*. External effects that result in costs to others are called *negative external effects* or *external diseconomies*. External effects that confer benefits on others are called *positive external effects*, *external benefits*, or *external economies*.

In each of these three cases, the reason why uncoordinated activities of people pursuing their own ends produce outcomes that are worse for all is that each participant's actions affect the well-being of others but these effects are not taken into account by the individual actors when they decide how to act. These impacts of our actions on others that we do not take account of in deciding what to do are termed **external effects**.

Here are the external effects that actors in our four examples do not take into account when deciding what to do:

- The person who lives in a city who drives to work, adds congestion to the streets, and therefore increases the travel time of others.
- Hume's farmer who does not drain the swamp and imposes the cost of doing so on the other farmer.
- The Palanpur farmer who plants late, imposes a cost on the other farmer who will have his seeds devoured by birds if he plants early. Likewise the farmer who plants early confers a benefit on the other farmer who can benefit by planting at the right time (early) without severe losses of seed to the birds.
- The herder who places additional cattle on the common pasture reduces the grass available to the other herders stock.

Addressing coordination problems by internalizing external effects

Simply abolishing these and other external effects that are the root of coordination problems is not an option. There is no way to organize society so that nothing that we do would affect others, each person on his or her self sufficient island.

Apart from not being much fun, life would be impossible in a society of total social isolates (just think about how the next generation would be born and raised!). So, to address the classical institutional challenge as to prevent or at least minimize coordination failures we need to find ways of inducing each participant *to take adequate account of the effects of their actions on others.*

INTERNALIZATION OF EXTERNAL EFFECTS in economics refers to any way that people can be brought to take appropriate account of the effects of their actions on others. In psychology the term internalization means to to adopt societies values or standards as one's own values.

This is called *internalizing* an *external effect*. We use the term *external effect* because the effect is *outside* of the individual's process of decision-making when taking the action. To internalize the external effect, you ensure that the person who acts bears the costs of their negative effects on others and reaps the rewards of their positive effects on others. In this way the otherwise "external" costs and benefits become part of the individual's decision-making process, leading them to "take adequate account of the effects of her actions on others."

If the "others" are our family, our neighbors, or our friends, our concern for their well-being or our desire to be well regarded by others might get us to take account of the effects of our actions on them. Reflecting this fact, an important response to the classical institutional challenge – one that long predates the classical economists – is that caring for the well-being of others need not be confined to friends and relatives but may extend to all of those with whom we interact. Ethical guides such as the "golden rule" are ways that people often internalize the effects of our actions on others, even when the others are total strangers to us.

HISTORY The "golden rule" is "to do unto others as you would have them do unto you" (Matthew, 7: 12). Or, treat others as you would like to be treated yourself. The same ethical principle is found in Islamic scriptures and in the teaching of other religions.

But, over the past five centuries, people have come to interact not with a few dozen people as humans have for most of our history and pre-history but directly with hundreds and indirectly with millions of strangers. The classical economists in the 18th century were responding to the fact that the generosity or ethical motivations that one might feel towards ones family or neighbors would not be sufficient to induce people to take account of the effect of their actions on others once these external effects spread across the entire network of global interactions.

An objective that economics has set for itself from that day until today, therefore, has been to design and implement institutions that would induce people to act *as if* they cared about those who were affected by their actions even when that was not literally true.

Checkpoint 1.3: External effects

a. Provide an example of a *negative external effect* that occurs in a social interaction. Explain *why* it is negative and why it is external.
b. Provide an example of a *positive external effect* that occurs in a social interaction. Explain *why* it is a positive external effect.

1.4 Institutions: Games and the rules of the game

Institutions

Institutions are the laws, norms, and beliefs that influence how people interact, and what the outcomes of these interactions will be. People adopt the behaviors prescribed by institutions (e.g. drive on the right if you are in the U.S.) because of some combination of

INSTITUTIONS Institutions are the laws, informal rules, and conventions which regulate *social interactions* among people and between people and the biosphere.

- *laws* enforced by a government (you will be arrested and fined for driving on the left in Brazil, the U.S, France, and other countries where driving on the right is the law.)

- *social pressures* – sometimes termed informal rules because they are not enforced by governments (your friends and neighbors will disapprove and think less of you if you drive on the left), and
- *information* that you have about what others will do (you expect others to drive on the right, so you will avoid accidents by doing the same.)

The late Nobel Laureate Douglass North called institutions the “rules of the game.”[5] People can *change* these rules, so institutions can themselves be outcomes of games that govern how the rules of the game can be changed.

Because institutions are the rules of the game for how people (and businesses, and trade unions, and governments) interact, we now introduce **Game theory**. Game theory uses mathematical models and verbal arguments to analyze how the outcomes of the interaction for the participants will depend on the rules of the game and the objectives of the players. It has been used extensively in economics and the other social sciences, biology, and computer science.

GAME THEORY Game theory is the study of strategic interactions using mathematical models and verbal arguments to analyze how the outcomes of the interaction for the participants will depend on the rules of the game and the objectives of the players.

EXAMPLE In 2020 under the pressure of popular protests, the government of Chile established a set of rules governing how the constitution of Chile would be amended. Another example of institutions is shown by football (soccer). FIFA governs how football can be played by what are called The Laws of the Game. These institutions also change: the corner kick was introduced in 1872 when the U.K. Football Association changed the rules.

Game theory focuses on *strategic* interactions where participants are interdependent and are aware of this interdependence: one player’s outcome depends on their own and other players’ actions and all players know this. We can contrast strategic with *non-strategic* situations in which the effect of your actions on the outcomes you will experience is independent of what others do. Your enjoyment of the program you are streaming at home alone is substantially independent of what others may be doing.

STRATEGIC INTERACTION An interaction is *strategic* when participants’ outcomes are *interdependent* – their well-being depends on the actions that both they and others choose, and this interdependence is known to the actors. An interaction is *non-strategic* when this interdependence of people’s outcomes is either absent or not recognized by the participants. A short-hand expression for the term strategic is: *mutual dependence, recognized*.

But many of our economic and social interactions are strategic:

- those considering driving to work know that their travel time will depend on how others decided to get to work that morning;
- the Palanpur farmer knows that how his crop will fare if he plants early will depend on how many others planted early

Checkpoint 1.4: Institutions

a. What are institutions?
b. What are "the rules of the game"?

Games

When we model strategic interactions using game theory we call the actors *players*. Players can be people, firms, social movements, governments and a variety of other entities. In biology, where game theory has been extensively used, even sub-individual entities are "players" such as viruses "trying to" spread or genes "trying to" get as many copies of themselves made as possible.

HISTORY John von Neumann (1903-1957) was a Hungarian-American mathematician, computer scientist, and physicist who is regarded as the father of game theory,[6] which he hoped would allow us to better understand the anti-Semitism and fascist political upheavals that he had witnessed in the early 20th century and provide the basis for understanding how groups interact.

Players may choose from a list of possible strategies. For example, if private property is an institution that is present and enforced, then a strategy set might include "Purchase a Trek bicycle for $850." But it would not include "Pick up any available Trek bicycle," without specifying the possible penalties for stealing. The Palanpur farmers' strategies are 'Plant Early' or 'Plant Late.' The strategies could also include a strategy based on what others did in the past (called a contingent strategy) such as: "Plant early as long as at least 5 others planted early last season."

The description of a *game* requires us to identify the following:

- *Players*: a list of every player in the game whether they be individuals (like the farmers in Palanpur), an organization such as Amazon or Alibaba, or some other entity that can be represented as choosing between alternative courses of action.
- *Strategy sets*: a list for each player of every course of action available to them at each point where they must make a choice (including actions that depend on the actions taken by other players, or on chance events). The strategies selected by each of the players is called the **strategy profile**.
- *Order of play:* a game can be *simultaneous* such that players make their choices without knowing the choices of others, as in the game of rock-paper-scissors. Or a game can be *sequential* such that players move in sequence, one after the other, as in chess, so that each player knows and responds to the choices of the previous players.
- *Information*: A game also specifies
  - who "knows" what,
  - when do they "know" it,
  - is what they "know" known to others as well,
  - can what they "know" be used in a court of law to enforce a contract, and
  - is what they "know" true (this is why we use the quotation marks)?
- *Payoffs:* Are numbers assigned to each possible outcome of the game (each strategy profile) for each player; a player chooses a strategy with the intention of bringing about the strategy profile with the highest number.

It is often useful to consider payoffs as something that the players actually get. For example, considering the farmers in Palanpur again, an outcome of the game is a strategy profile indicating who plants early and late, and the payoffs could be the amount of grain each farmer harvests. We say that the payoff associated with a particular outcome of a game is how much the player *values* that outcome. But that means nothing more than that a player

SET A set in mathematics is a collection of objects precisely defined either by enumerating the objects, or by a rule for deciding whether any particular object is in the set or not. For example, the set of positive, even integers less than or equal to 10 is, $\{2,4,6,8,10\}$.

EXAMPLE When we model the coordination problem of the Palanpur farmers as a game we assume they plant simultaneously. But when we model the interaction between a bank and a borrower we assume that the banks first makes an offer (the loan size, interest rate and schedule of repayment) and the prospective borrower responds.

COOPERATIVE GAME A game in which players can jointly agree upon how each will play the game (and the agreement will be respected or enforced) is a cooperative game. If no binding agreement on how to play the game is possible, then the game is non cooperative.

will choose a strategy resulting in an outcome with a higher payoff number if possible.

An important distinction concerning strategy sets is whether or not one of the strategies open to the players is to jointly agree on a strategy profile – that is to deliberately coordinate their actions. This is possible in what is called a **cooperative game.**

We use the set of *players*, their *strategy sets*, their *payoffs*, the *order of play*, and the *information* the players have to describe the institutions governing some economic interaction, whether it is between an employer and an employee, or a central bank like the U.S. Federal reserve and a commercial bank. But even this detailed description of the interaction does not give us enough information to predict how the game will be played.

The outcome of a game – how it will be played resulting in a particular strategy profile – also called a solution. To determine the solution as a way of predicting the outcome of a game we need what is called a solution concept. A solution concept for a cooperative game would include some rule for deciding on what the coordination would be, for example allowing one player selected at random to dictate the outcome, or a particular system of voting.

SOLUTION CONCEPT A solution concept is a rule for predicting the outcome of a game, that is, how a game will be played.

But by positing some way that people could jointly implement some outcome, cooperative game theory *assumes away* the problem of coordination. And the problem of how coordination is to be achieved is at the heart of the classical institutional challenge whether it takes the form of climate change or traffic jams.

So we need to see how players might coordinate in what is initially a non-cooperative setting – one in which coordination is not assumed at the outset – lets take a concrete example. We will use this example to illustrate a basic solution concept for non-cooperative games: the Nash equilibrium.

Checkpoint 1.5: Games

a. What is a game?
b. How do you describe the *outcome* of a game?

1.5 Over-exploiting nature: Illustrating the basics of game theory

Figure 1.4: Elinor Ostrom (1933-2012) was an American Political Scientist who won the Nobel Prize in economics for her work on social dilemmas, such as those encountered by Alfredo and Bob in the Fishermen's Dilemma, and on the institutions that promote cooperation in groups. Photo Credit: Holger Motzkau. Wikimedia Commons.

People who fish for a living interact with each other regularly. Each of them are aware that how much they benefit from fishing depends not only on their own actions, but on the actions of others. This is because the more others fish, the more difficult it will be for each to catch fish. The two fishermen therefore impose negative external effects on each other, and this is why they face a coordination problem. Given that they cannot jointly decide on how

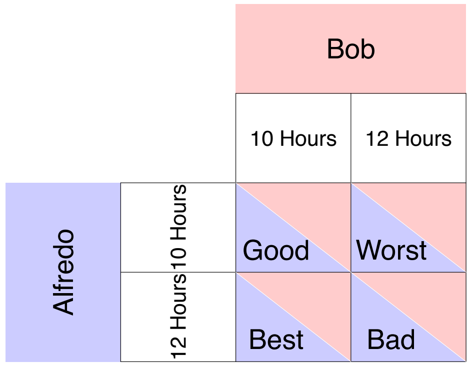


Figure 1.5: **Alfredo's payoffs to fishing more or less depend on how much Bob fishes.** Alfredo's payoffs are described using the words like we used for the coordination problem: Planting in Palanpur. Alfredo ranks his outcomes from best to worst: Best > Good > Bad > Worst. Alfredo's strategies and outcomes are highlighted in Blue. Bob's strategies and outcomes are highlighted in Red (but we have not put the words to describe Bob's outcomes in the figure).

much to fish, each faces a basic question: how much fishing to do given the time are fishing the same waters?

The game set up

Specifically, we consider two fictional fishermen, Alfredo and Bob, who share access to a lake, and catch fish, which they eat. There are no other people affected by their actions.

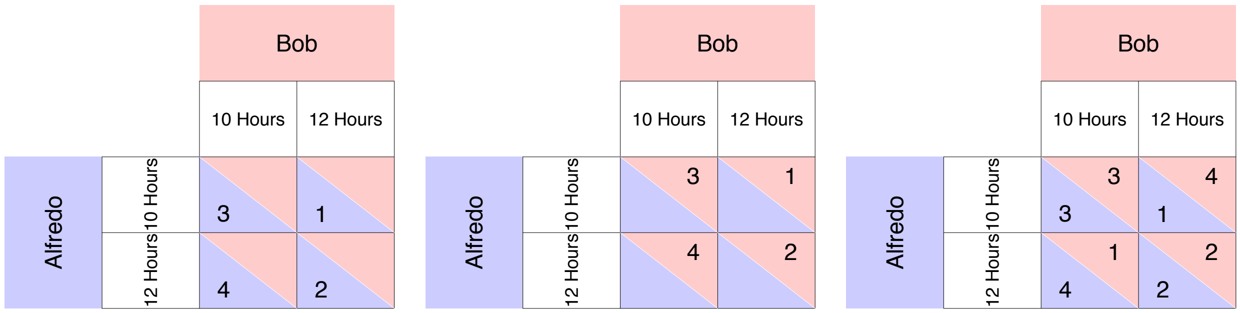
Here we illustrate the basic concepts of game theory in a game we call the Fishermen's Dilemma. We chose the name because it is an example of what is probably the most famous game, the Prisoners' Dilemma. As before we use a two-person example to illustrate a societal coordination problem among a much larger number of actors.

The Fishermen's Dilemma game is non-cooperative, which for two people fishing in the same lake may seem unrealistic because as neighbors they might be able to come to some kind of agreement about what each will do. We do not consider this option in the two-person case because the model illustrates a large number of people interacting. When many people interact arriving at and enforcing such a cooperative agreement would present serious challenges.

Here is the game.

- *Players:* Alfredo and Bob, two fishermen.
- *Strategy sets*: Each may fish for either 10 or 12 hours.
- *Order of play*: They simultaneously select a strategy, resulting in the game's strategy profile
- *Payoffs*: The players catch and eat an amount of fish given by the strategy profile they have implemented.

This ends the game.



(a) A's payoffs only (b) B's payoffs only (c) Payoffs for both players

Figure 1.6: **Payoffs of players in the Fishermen's Dilemma.** Alfredo's payoffs are in the bottom-left corner of each cell and are shaded blue. We include Alfredo's payoffs in the right-hand and left-hand panels. Bob's payoffs are in the top-right corner of each cell and are shaded red. We include Bob's payoffs in the center panel and the left-hand panel.

Payoffs

The payoff of each player is composed of two parts:

- The amount of fish they are able to catch and consume, which they value and would like to increase; and
- The amount of time they spend fishing, which they find tiring and would like to decrease.

We can describe the fishermen's interaction in the form of a *payoff matrix*. (A *matrix* is a rectangular array of quantities or other quantitative information). We first present a version of the payoff matrix with words to represent Alfredo's payoffs (but not yet Bob's) in 1.5. Read the table this way: If Bob fishes 12 hours (the right hand column) and Alfredo fishes 10 hour (top row) this is the *worst* outcome for Alfredo. A payoff matrix presents hypothetical 'if-then' information; it presents all of the possible sets of payoffs, whether or not each is likely ever to occur.

The complete payoff matrix for the Fishermen's Dilemma is represented in Figure 1.6 with numbers indicating the two fishermen's evaluation of how good the outcome indicated is. So for example the payoff to each if they both fish ten hours (3) is fifty percent greater than if they both fish twelve hours (2).

NORMAL FORM GAME We will often describe games using payoff matrices in what are called **normal** or **strategic form**, like Figure 1.10. In normal or strategic form games, we do not explicitly represent the time sequence of the actions taken by each player. We assume that each player moves *without knowing* the move of the other players. Normal form games therefore often represent *simultaneous* move games, games where players move at the same time. Simple games in normal form are often presented in a *payoff matrix*, a table that includes all the relevant information about the players, strategies and payoffs in the game.

The convention we will use throughout this book is to list the row player's payoffs first and in the bottom left corner and the column player's payoffs second in the top right corner. So, in the Fishermen's Dilemma game, we list Alfredo's payoffs first and Bob's payoffs second. We shade each players payoffs to make them easier to differentiate: blue for the row player (Alfredo) and red for the column player (Bob).

Many of the games in this book involve two players and each player has two possible strategies. We often call a game like this a "2 x 2" game (a "two-by-two" game).

We now have all the elements we need for the complete description of the

Fishermen's Dilemma and its strategy profiles and associated payoffs.

- *Alfredo fishes 12 hours, Bob fishes 12 hours*: When both fishermen fish 12 hours, they each catch fewer fish per hour of work, while they also have a higher cost of effort because they've spent a lot of time fishing. Each fisherman ends up with 2.
- *Alfredo fishes 10 hours, Bob fishes 10 hours*: When both fishermen spend less time fishing they catch a decent amount of fish and they haven't fished so long that the other fisherman catches fewer fish. They also benefit from a lower cost of time spent fishing. Each gets a net benefit of 3.
- *Alfredo fishes 10 hours, Bob fishes 12 hours*: Because Bob fishes 12 hours, Al catches many fewer fish and because Bob still fishes for another two hours, he catches a lot of fish while Al doesn't fish. Consequently, with the cost of time and catching fewer fish, Al ends up with net benefits of 1 and Bob ends up with net benefits of 4.
- *Alfredo fishes 12 hours, Bob fishes 10 hours*: This is symmetrical to the previous description, so now Al gets net benefits of 4 and Bob gets net benefits of 1.

1.6 Predicting economic outcomes: The Nash equilibrium

As you already know, to predict a game outcome – the strategy profile that will result– we need more than the description of the game alone. We need to add what is termed **a solution concept** – a statement about *how* players will behave in the game – that can be the basis of a *prediction* of the game's outcome. Predicting the outcome of a game – based on the rules of the game and the solution concept – is especially important to understanding how changing the rules of a game can change the outcome of a game.

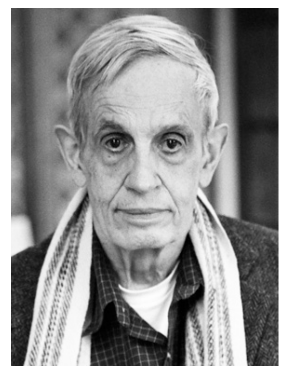


Figure 1.7: John F. Nash (1928-2015) was an American mathematician who contributed the idea of Nash equilibrium to game theory and won the Nobel Prize in economics in 1994. His life was documented in the book and movie *A Beautiful Mind*.[7] Source: Peter Badge, Wikimedia Commons.

The key idea on which a solution concept is based is **equilibrium**. An equilibrium is a state in which there is nothing in the situation that will cause the state to change. A *predicted outcome* will be an equilibrium, that is, an outcome that is stationary (not changing). To understand why, imagine this were not the case. You make a prediction, but then the outcome changes. Your prediction would *no longer be true* because the outcome had changed.

EQUILIBRIUM An equilibrium is situation that is stationary (unchanging) because, as long as the situation we are describing remains, there is nothing causing it to change.

Applying this reasoning to games, if we were to predict the outcome of a game to be a strategy profile under which one or more players would have reason to change their strategy, then the prediction would be falsified as soon as they carried out the change. So the status of stationarity – change-less-ness – is a property of a prediction; and this is why equilibrium is the fundamental idea of making predictions about game outcomes.

Think of a concrete example. Suppose you want to predict where a marble will be if all that you know is that it is going to be somewhere in a round bottomed salad bowl sitting on a table. If I predicted that the marble would be somewhere halfway up the side of the bowl you would doubt my prediction. The reason is that any marble in that position would move downward in the bowl, that is, its position would not be stationary so,if it ever were (for some reason) where I predicted it would be, it would not be there any longer. It is not that the prediction would necessarily be wrong. It could be true for a millisecond after I placed the marble in the bowl just above my predicted spot, for example.

The only predicted position in the salad bowl that would not immediately falsify itself in this sense is the bottom. So a reasonable prediction of the location of the marble would be "the bottom of the bowl."

There are some situations in which a prediction based on an equilibrium would be likely to be incorrect. Change the marble-in-bowl example by filling the bowl with very thick honey. Then if you were asked to predict where the marble would be found, you would want to know how long it had been in the bowl, did have time to reach the bottom? If if the marble had been placed in the bowl just a second ago, they you might be better off predicting that it would be where it had been placed, rather than the bottom of the bowl.

The marble-in-bowl-of-honey is often a better illustration of how economic processes work than the initial example. Markets are often out of equilibrium. Predicting things in motion is a much more challenging task than predicting them when they are stationary. We provide an example in our model of residential the segregation (below) where we are able to follow the process of change step by step. But for the most part we study equilibria and how to change them so as to improve outcomes.

In the marble-in-bowl illustration (without the honey) what is the solution concept that let us arrive the "bottom of the bowl" prediction? It is gravity, which is our understanding about a reasonable way for the marble to "behave." In modeling an economic interaction, the game structure is analogous to the salad bowl. What is the analogy to gravity? The answer is the players' *best response.*

Best-response strategies

BEST RESPONSE A player's *best response* is a strategy that results in the highest payoff *given* the strategies of the other players.

By far the most widely-used solution concept, the Nash equilibrium, is based on the idea that players choose **best-response** strategies; they do the best they can given the strategies adopted by everyone else

To understand better what a best response is, think about Alfredo's choices in the Fisherman's Dilemma. We will also introduce the "circle and dot" method for finding the Nash equilibrium on the basis of your analysis of his best responses. First, what strategy should Alfredo adopt in order to gain the highest

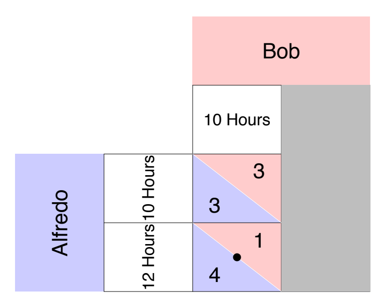


Figure 1.8: Hold constant Bob playing Fish 10 hours to assess Alfredo's best response to Bob's playing Fish 12 hours.

payoff if Bob were hypothetically to play Fish 10 hours (we say "holding constant" this strategy) as shown in Figure 1.8.

- Against Bob playing Fish 10 hours, Alfredo can get a payoff of 3 for fishing 10 hours or a payoff of 4 for Fishing 12 hours.
- $4 > 3$ therefore fish 12 hours is Alfredo's *best response* to Bob playing fish 10 hours.
- place a solid point in the cell (Alfredo plays Fish 12 Hours, if Bob plays Fish 10 hours) to indicate that it is Alfredo's best response.

Let's repeat analysis and hold constant Bob playing Fish 12 hours, as shown in Figure 1.9.

- Against Bob playing Fish 12 hours, Alfredo can get a payoff of 1 for playing Fish 10 hours or a payoff of 2 for playing Fish 12 hours.
- $2 > 1$ therefore Fish 12 hours is Alfredo's *best response* to Bob playing fish 10 hours.
- place a solid point in the cell (Alfredo plays Fish 12 Hours, Bob plays Fish 12 hours) to indicate that it is Alfredo's best response.

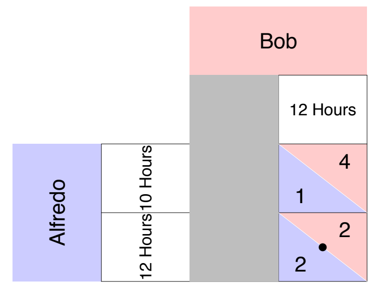


Figure 1.9: Hold constant Bob playing Fish 12 hours to assess Alfredo's best response.

Checkpoint 1.6: A best response for Bob

Repeat the process we went through for Alfredo, but do it for Bob instead. Notice that when you do so, you will blank out a *row* for Alfredo to hold his strategy constant, whereas you blanked out a *column* for Bob to hold his strategy constant. What are Bob's best responses? Show his best responses using a hollow circle.

M CHECK: STRONG AND WEAK BEST RESPONSE. A best response may be either *strong* or *weak*. A strong (also called strict) best response yields higher payoffs than any other: it is strictly "better" than any other strategy. There can be no strategy that is better than a weak best response but a weak best response need not be better than any other; it may be "as good as" (the payoffs to the strategy and some alternative strategy being equal.)

Nash Equilibrium and the outcome of a game

Using the best responses of the players we can now predict the outcome of a game using as our solution concept the **Nash equilibrium.** A Nash equilibrium is a profile of strategies – one for each player – each of which is a best response to the strategies of the other players. A Nash equilibrium is also called a *mutual* best response. Because at a Nash equilibrium all players are playing their best response to all of the others, it follows that no player has a reason to change his or her strategy as long as the other players do not change theirs. Some games do not have a Nash equilibrium and you will see shortly that some have more than one.

EXAMPLE The Rock,Paper, Scissors game (also called ro-sham-bo and by many other names in other languages) originated in China about two thousand years ago. It does not have a Nash equilibrium.

NASH EQUILIBRUM A Nash equilibrium is a profile of strategies – one strategy for each player – each of which is a best response to the strategies of the other players.

In Figure 1.10, Alfredo's best responses are shown by the solid black dot in the cell. Bob's best responses are shown by the hollow circle. Their best responses coincide at the Nash equilibrium (Fish 12 Hours, Fish 12 Hours) with payoffs (2, 2) shown in the cell where the solid point is inside the hollow circle. You can use the "dot and circle" method to find one or more Nash

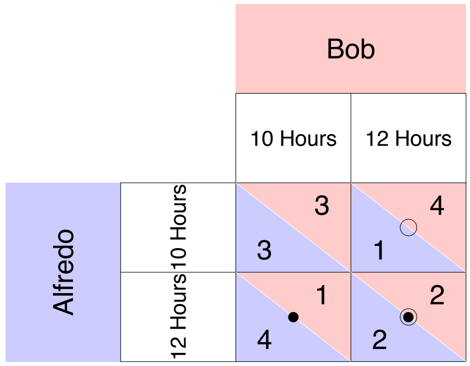


Figure 1.10: **Payoff Matrix for The Fishermen's Dilemma.** The solid dots indicate Alfredo's best responses. The hollow circles indicate Bob's best responses. A Nash equilibrium is a cell that contains both dots. In this case there is just one Nash equilibrium: both fishing 12 hours.

equilibria (if they exist) for games that can be represented by a payoff matrix like Figure 1.10.

In the Fishermen's Dilemma game described by Figure 1.10. Each player's best response to both of the others' strategies was Fish 12 hours. Therefore only one outcome is a Nash equilibrium: both fishermen fishing 12 hours. We say that (Fish 12 Hours, Fish 12 Hours) is the Nash equilibrium with payoffs (2, 2).

The outcome demonstrates how Nash equilibrium can initially seem counter-intuitive. Both would have had higher payoffs if they could have agreed to restrict their fishing to 10 hours (they could have had 3 each if they both fished 10 hours and $3 > 2$). But suppose both were for restricting their fishing to 10 hours; then both would an incentive to fish for 12 hours (because $4 > 3$) and unless they had a binding agreement to continue fishing less, both would choose to fish more.

The Fishermen's Dilemma is therefore a coordination problem and it returns us to the classical institutional challenge. Without institutions to align the individual interest of the participants with their shared interest, they get an outcome that is worse for both of them than other possible outcomes. We will later show how a change in the institutions reglating how Alfredo and Bob interact – that is, changing the rules of the game – might address this coordination.

Checkpoint 1.7: Nash Equilibrium

a. Explain why none of the other three outcomes (those that are not, (Fish 12 Hours, Fish 12 Hour) of the Fishermen's Dilemma satisfy the definition of Nash equilibrium.

b. At each of the other three outcomes, which player has an incentive to change strategy and in what way? Explain.

c. Explain why a game like Rock Paper Scissors would not be much fun if there was a Nash equilibrium.

Dominant strategies

In the Fisherman's Dilemma (and all Prisoners' Dilemmas) there is a single strategy that yields the highest payoffs to a player for both (or all of if there are more than two) of the strategies that the other player might adopt. A strategy is a player's **dominant strategy** if it is the player's best response to all possible strategy profiles of the other players. That is, a strategy is a dominant if by playing it the player's payoff is *greater than or equal to* the payoff playing any other strategy for every one of the other player's profiles of strategies.

DOMINANT STRATEGY A strategy is dominant if by playing it the player's payoff is *greater than or equal to* the payoff playing any other strategy for every one of the other players profiles of strategies. A strategy is *dominant* if it is the player's best response to all possible strategy profiles of the other players. A dominant strategy *dominates* all of the other strategies available to the player.

Likewise we say that strategy A is *dominated* by another strategy B if the payoff to playing B is at least as great or greater than playing A for every strategy profile of the other players. If there is a strategy that dominates all of the other strategies that an player may choose, then it is a dominant strategy. If each player in a game has a dominant strategy, then the strategy profile in which all players adopt their dominant strategy is called a *dominant strategy equilibrium*.

We can apply the concept of *dominant strategy* equilibrium to the Fishermen's Dilemma. To do so, we need to understand whether each player has a dominant strategy.

- When Alfredo fishes 10 hours, his payoff is 3 if Bob fishes 10 hours and 1 if Bob fishes 12 hours.
- When Alfredo fishes 12 hours, his payoff payoff is 4 when Bob fishes 10 hours and 2 when Bob fishes 12 hours.
- So, when Bob fishes 10 hours, fishing 12 hours gets Alfredo a higher payoff $(4 > 3)$ and when Bob fishes 12 hours, fishing 12 hours gets Alfredo a higher payoff $(2 > 1)$
- Therefore, Alfredo gets a higher payoff from fishing 12 hours against each of his opponent's strategies
- Fish 12 hours is therefore Alfredo's *dominant strategy*.

Fishing 12 hours is also Bob's dominant strategy. Because each player has a dominant strategy to fish 12 hours, the *dominant strategy equilibrium* is (Fish 12 hours, Fish 12 hours) with payoffs (2, 2).

DOMINANT STRATEGY EQUILIBRIUM. A *dominant strategy equilibrium* is a strategy profile in which all players play a dominant strategy.

The fact that the Fishermen's Dilemma has a dominant strategy equilibrium makes it a particularly simple problem (both for us, studying it, and for the players because what is best for each does not depend on what the other does). The dominant strategy equilibrium of a game is always a Nash equilibrium, so in the Fishermen's Dilemma, the outcome where both players fish 12 hours is the only Nash equilibrium.

Checkpoint 1.8: Dominance and Nash Equilibrium

a. Repeat the analysis we did for Alfredo for Bob and confirm that 12 hours is a dominant strategy for him too.

b. We said that a dominant strategy equilibrium is always a Nash equilibrium. But do you think that a Nash equilibrium is always a Dominant Strategy equilibrium? Why or what not?

1.7 Evaluating outcomes: Pareto-comparisons and Pareto-efficiency

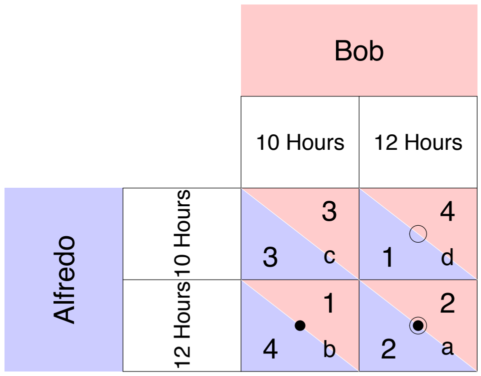
The Nash equilibrium can help us predict the result of a particular interaction. But it does not tell us anything about whether some outcome is good by any standard, or even better or worse than some other outcome. Economists, policy-makers and others would like to evaluate whether some outcomes are better or worse so that we can try to work out which rules of the game would make the better outcomes Nash equilibria, and therefore more likely to be what we observe. In Chapter 16 we show how economics deals with this for questions of public policy.

The challenge in making these comparisons is that whether some outcome is better than another depends on what you value, and there is no agreed upon standard of what makes one outcome better than another. Returning to our fishermen, here are some of the values that one could use in evaluating an outcome

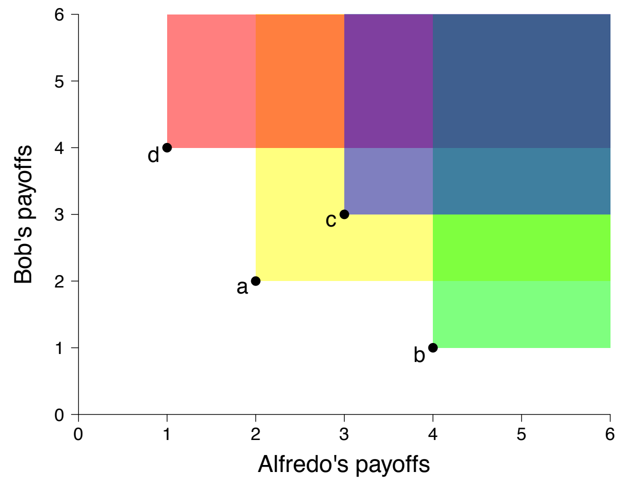
- *Fairness in the distribution of payoffs* among the players; is it fair that Alfredo receives 4 times what Bob gets when Alfredo does not limit his fishing hours and Bob does?
- *Are the rules of the game itself fair?* In the Fishermen's Dilemma the same rules applied to both players; but were the game a bit different, many would think it unfair if Alfredo could simply order Bob to fish 10 hours, or to hand over half of all the fish Bob caught.
- Setting aside fairness, is the outcome a *reasonable use of available resources* including the working time of the two fishermen and the sustainability of the lake itself and the living things that it supports.

There are many other standards that could be proposed. Questions of better and worse are called "normative question." With normative questions matters of ethics or morals are necessarily involved. We will introduce some experimental evidence on people's views concerning fairness in Chapter 2 and some analytical tools for studying normative issues in Chapter 13.

Here we introduce a concept that economists use to evaluate economic outcomes. The idea is simple: an objective of public policy and institutional design – the rules of the game – should be to avoid those outcomes – like traffic



(a) Fishermen's Dilemma with Labeled Points



(b) Analyzing points for Pareto efficiency

Figure 1.11: **Three Pareto-efficient outcomes of the Fishermen's Dilemma**. In 1.11 a, the outcomes are labeled in the bottom right corner as **a** for (2, 2), **b** for (4, 1), **c** for (3, 3) and **d** for (1, 4). These allow us to make Pareto comparisons for the outcomes of the Fishermens Dilemma. Alfredo's payoffs are plotted on the horizontal axis, increasing as you move rightward. Bob's payoffs are plotted on the vertical axis, increasing as you move upward.

In 1.11 b we show the Fishermen's Dilemma indicated by the four possible outcomes given by the same letters that appear in each of the cells of the payoff matrix. We use shaded colors indicating 90 degree angles to the northeast of the feasible outcomes (each of the lettered points).

jams, planting late in Palanpur, and over-fishing the lake – that are worse for everyone, compared to an alternative outcome that also would have been feasible.

Pareto comparisons

To compare outcomes when more than one person is involved economists use the concept of **Pareto-efficiency** based on **Pareto-comparisons** of outcomes.

An outcome is Pareto superior to another if it allows at least one of those involved to be better off without anyone being worse off. A Pareto-superior outcome is also called a *Pareto improvement* over the outcome it was compared to. This is a Pareto comparison. An outcome is Pareto efficient if no other feasible outcome is Pareto superior to it.

Figure 1.11 depicts the outcomes of the interaction between Alfredo and Bob. Alfredo's payoffs are on the horizontal axis (*x-axis*), so the outcomes get better for Alfredo as you move from the left to the right. Bob's payoffs are on the vertical axis (*y-axis*), so the outcomes get better for Bob as you move from bottom to top. The left panel of Figure 1.11 is the Fishermen's Dilemma payoff matrix with each outcome given a label **a**, **b**, **c**, or **d**. These same points appear in the right panel of Figure 1.11 where you can read on the vertical and horizontal axes the payoffs to the two players that you see in the payoff matrix.

The Pareto-comparison is geometrically easy to see in this type of plot. An outcome is Pareto-superior to another if the first outcome lies to the "north-east" of the second in the plot. "North-east" in this figure is "better for both." An outcome is Pareto-efficient if there is no other feasible outcome to the north-east. Or if you think of the colored areas whose lower left corners are points **a, b, c,** and **d** then a Pareto efficient point is one that has no other point in its "colored shadow" extending upwards and to the right of the point.

When two different outcomes are both Pareto-efficient, they cannot be Pareto-compared or Pareto-ranked. We could rank **c** above **a** because both players were better off, but with **b**, **c** and **d** we cannot move from one outcome to another without worsening outcomes for at least one of the players.

PARETO COMPARISONS AND PARETO EFFICIENCY An outcome is Pareto superior to another if it allows at least one of those involved to be better off without anyone being worse off. A Pareto-superior outcome is also called a *Pareto improvement* over the outcome it was compared to. This is a Pareto comparison. An outcome is Pareto efficient if no other feasible outcome is Pareto superior to it. If we can rank two outcomes such that one is Pareto superior to the other, then we say that these two outcomes can be *Pareto compared*, or *Pareto ranked*.

Here is a checklist to use when evaluating any given set of payoffs for both players. Consider a point called **x** with payoffs π_x^A for Player A and π_x^B for player B and compare it to another theoretical point, **y** with its corresponding payoffs:

- For any point, **x**, check whether there exists an alternative point **y** where at least one player gets a payoff that is greater than π_x^A or π_x^B without the other player being worse off.
- If at **y**, $\pi_y^A > \pi_x^A$ while $\pi_y^B \geq \pi_x^B$ or $\pi_y^B > \pi_x^B$ while $\pi_y^A \geq \pi_x^B$ then **y** is Pareto-superior to **x** (at least one player is better off while the other player is *not* worse off).
- If no other point exists where at least one player is better off with no other player being worse off (no Pareto-superior point exists), then **x** is Pareto-efficient.

Checkpoint 1.9: Pareto efficiency and Pareto improvements in the Fishermen's Dilemma

Referring to Figure 1.11, consider the following:

a. Is any point dominated by some other point? Say which, if any? item At which point is the total payoff of the two fishermen the greatest?
b. Would a change from any other point to that "total payoff maximum" point be a Pareto improvement?

1.8 *Strengths and shortcomings of Pareto efficiency as an evaluation of outcomes*

EXAMPLE We will see in Chapter 2 that people often reject a highly unequal division of some "pie" and are willing to sacrifice a substantial amount of money rather than to accept what they consider to be an unfair division. And we will see in Chapter 13 how people's desires for a fair society might allow for a choice among outcomes that differ in who gets what.

Pareto efficiency gives us a way to identify "lose-lose" outcomes we should seek to avoid, namely those " that are worse for all than they could be" . But, as we will now see, except in special cases, Pareto efficiency does not provide a rule to select what we might call in everyday speech "the best" outcome. To see why this is true, suppose we have a cake of given size and we are dividing it among people, all of whom equally enjoy eating cake.

An outcome in which one person gets the entire cake is surely Pareto-efficient because in any other allocation that lucky person would get less. Likewise an allocation in which everyone got an equal slice of the cake is Pareto-efficient, for in any other allocation at least one person would have to get less. The Pareto criterion can say nothing about such distributional fairness. All it says is "make sure there's no cake left on the table!"

When there are many Pareto-efficient outcomes there is always a *conflict of interest* among players over which outcome they would prefer we cannot say that one is "more Pareto-efficient" than the other.

It is also perfectly sensible to prefer an outcome that is not Pareto efficient but is more fair over an alternative Pareto efficient outcome that is unfair. To continue the cake example, if there are two people between whom the cake will be divided many people would reject the (Pareto efficient) outcome in which one person gets the entire cake in favor of a Pareto inefficient alternative in which each gets a third of the cake (the remaining third perhaps being thrown away or destroyed in the conflict over its distribution). But the Pareto comparison would remind us that each person getting half of the cake is preferable to each getting a third with the rest being wasted.

Pareto efficiency is a particular device for screening out those outcomes (like throwing away some of the cake in the above example, or planting late in Palanpur) that should not be among the list of candidate feasible outcomes among which the choice of better or best should be made or grounds of fairness or other bases.

Other ways of "screening" the list of candidate outcomes would give priority not to individuals' payoffs but to whether the rules of the games that produced the outcomes are themselves fair and consistent with other values such as individual dignity, respect and freedom.

Checkpoint 1.10: Pareto efficiency

Consider these questions about Pareto efficiency.

a. True or False (and explain): "The fact that an outcome is Pareto-efficient does *not* imply that it is preferred by all the actors to all the other outcomes."
b. Can two Pareto efficient outcomes be Pareto compared? Why or why not? Explain.
c. Imagine you are an impartial observer evaluating the possible outcomes that might occur for Bob and Alfredo. Are there any reasons why you might judge the Pareto-inefficient outcome **a** in the figure to be better than the Pareto efficient outcomes **b** and **d**, despite the fact that **a** is Pareto-inefficient?

1.9 Conflict and common interest in a Prisoners' Dilemma

The game the fishermen are playing is a particular case of the **Prisoners' Dilemma**. A Prisoners' Dilemma is a two-person interaction in which there is

PRISONERS' DILEMMA A Prisoners' Dilemma is a two-person social interaction in which there is a unique Nash equilibrium (that is also a dominant strategy equilibrium), but there is another outcome that gives a higher payoff to both players, so that the Nash equilibrium is not Pareto-efficient.

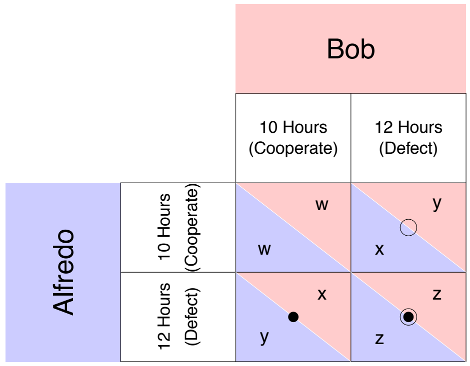


Figure 1.12: **A General Prisoners' Dilemma**. For the game to be a Prisoners' Dilemma, we require $y > w > z > x$ and $2w > y + x$ (this is like $4 > 3 > 2 > 1$ and $2 \times 3 > 4 + 1$ from the numerical example).

a unique Nash equilibrium (that is also a dominant strategy equilibrium), but there is another outcome that gives a higher payoff to both players, so that the Nash equilibrium is not Pareto-efficient. So, in the Prisoners' Dilemma both players get their second worst payoffs in the game by playing their strictly dominant best-response strategies.

MATH NOTE A third condition is sometimes added, namely $x + y < 2w$ which means that the sum of payoffs when both players cooperate is greater than the sum of payoffs when one cooperates and the other defects. This condition makes (Cooperate, Cooperate) preferable to any outcome in which one defects and the other cooperates.

We now point out some of the general characteristics of this particular kind of coordination problem. To do this, in Figure 1.12 we show the familiar payoff matrix for the Fishermen's Dilemma, but instead of the numbers indicating the payoffs of the players now we label the payoffs w, x, y, and z. We label the action of fishing 10 hours "Cooperate" because it is the action the two fishermen would take if they could coordinate their actions. The action of fishing 12 hours is labeled "Defect" because a fisherman who chooses to fish 12 hours is deviating from the mutual cooperate outcome on which the two fishermen might be able to coordinate.

The interaction is a Prisoners' Dilemma if two conditions hold:

- $y > w$ and $z > x$ means that fishing Defect is a strict dominant strategy
- $w > z$ means that mutual cooperation is Pareto superior to mutual defection.

For Alfredo, 12 Hours is a best response to Bob playing 10 Hours because $y > w$; 12 Hours is also a best response to 12 Hours because $z > x$ (both best responses are shown by the solid point). Similarly, for Bob, 12 Hours is a best response to Alfredo playing 10 Hours because $y > w$; 12 Hours is also a best response to 12 Hours because $z > x$ (both best responses are shown by the hollow circle). Therefore the Nash equilibrium is (12 Hours, 12 Hours) with payoffs (z, z).

If the players play the game non-cooperatively (they do not coordinate their actions) each will play their dominant strategy – defect – and get z when by cooperating they could have each received w.

Economic rent: The incentive to coordinate

Both players have a good reason to try to change the rules of the game so that they can agree on both cooperating. How much more they would get if they were to mutually cooperate than if they mutually defected – in this case $w - z$ – is called an **economic rent**, meaning the difference between the payoff that they would get if they cooperated and their next best alternative. Their next best alternative to cooperating, we assume, is mutual defection, also known as their **fallback position**.

FALLBACK POSITION A player's fallback position is the *payoff* they receive in their *next best alternative*.

Economic rents and the fallback position play a central role in microeconomic theory, so it is a good idea to master them. The meaning of fallback position is intuitive: it is what you *fall back* to if your current outcome is not possible, in this case if the mutual cooperation should not work out. A player's fallback position is the *payoff* they receive in their *next best alternative*.

The term "economic rent" may at first seem surprising, because the word "rent" also means a payment for the temporary use of something like a rent paid to a landlord or a rent or a car rental agency. The term economic rent means something entirely different. A participant's *economic rent* is the payoff they receive in excess of what they would get in their fallback position.

ECONOMIC RENT A participant's *economic rent* is the payoff they receive in excess of what they would get in their fallback position. When we use the term " rent" we mean economic rent. The sum of the economic rents received by the participants in an interaction is sometimes termed the economic surplus.

We shall use the idea of a fallback often, from social interactions like the Prisoners' Dilemma, to worker-employer relationships where a worker wants a job more than being unemployed, to a person applying to a bank for a loan rather than trying to get money from friends, family, or the government. As these examples indicate, in real life the fallback position will differ depending on the details of the situation that we set aside when we model interactions like fishing in a lake, employment and borrowing.

Impediments to coordination: Limited information and conflicts of interest

EXAMPLE The term economic rent is what is known in the study of language as a "false friend," a term that you think you know the meaning of but mean something entirely different in the new language you are now learning. "Sensible" in English means "reasonable" but in Italian it means "sensitive."

If $w - z$ is substantial – meaning substantial rents associated with cooperation for each player – then it might seem a simple matter for the players to agree to cooperate. But people often fail to reach or enforce such an agreement, for two main reasons:

- *Limited information.* The participants may lack the information needed to monitor and enforce an agreement. How can a participant know or verify what other participants do?
- *Conflict over distribution of the economic rents from cooperation* Disagreement about who gets what – for example who gets to fish more – may make it impossible for the two to agree.

Concerning the information problem, the fishermen, for example, may have no way of enforcing an agreement, or even knowing if the agreement has been violated. While each may know how many hours the other has fished on day with clear and sunny weather, on a foggy day it may be impossible to know. Even if one fisherman knows how much the other fished, that knowledge may be insufficient to enforce an agreement through a third party such as a court of law.

This is the problem of **asymmetric information** or **non-verifiable information**. Information is asymmetric if people know different things, or if what one person knows (for example how many hours he fished), the other person does not know. Information is not verifiable if people cannot use it to enforce an agreement or a contract. For example most courts will not accept "hearsay" (meaning "second hand") information, so if one of the fishermen had heard from someone else that the other had fished 12 hours, this would be non-verifiable information.

ASYMMETRIC INFORMATION Information is *asymmetric* if something that is known by one participant is not known by another.

NON-VERIFIABLE INFORMATION cannot be used to enforce a contract or other agreement

Asymmetric and non-verifiable information will play a central role in our analysis of how the labor market, the credit market and other markets work.

Concerning the conflicts over the distribution of the economic rents from cooperation, the Fishermen's Dilemma, the agreement to restrict fishing to 10 hours a day, is an agreement both to restrict fishing and to divide the benefits of restricting fishing in a particular way, namely equally. But the fishermen need not agree on 10 hours each. Alfredo might insist that he will fish 12 hours and Bob only 10 hours. Or Bob might insist on the opposite.

Or Bob might insist that both fish 10 hours, but that Alfredo give him most of Alfredo's catch, leaving Alfredo with just enough of his catch to be no worse off than had they both fished 12 hours, namely with a payoff of z. Unless they can find a mutually acceptable solution to the distribution problem they may end up having no agreement at all, and then simply fish at 12 hours each, at their fallback position.

The fishermen's distribution conflict highlights a challenge that arises in any voluntary economic interaction. Consider their possible agreement to limit their fishing time:

- The agreement is voluntarily entered into. This means that neither player can force the other to accept terms worse than their fallback position.
- The agreement therefore must allow each participant to achieve a payoff greater than (or at least not worse than) had the individual not agreed to cooperate. In other words, there must be some economic rents made possible by a voluntary cooperative outcome.
- This being the case, the participants have to find a way that the total rents will be divided.. If they are to agree to cooperate by restricting fishing, they must also agree on how these economic rents will be distributed.
- Conflict over the distribution of the economic rents (who gets what amount

of economic rent) may prevent the fishermen from coming to an agreement.

FACT CHECK In the next chapter we will see that across many cultures of the world, people would rather get nothing than get what they consider to be an unfair share of the economic rents, and as a result cooperation breaks down and nobody gets any rent at all.

We sometimes think of cooperation and conflict as opposites, as for example when members of a team cooperate in their efforts to win some conflict with another team. But the Prisoners' Dilemma is a scenario of *conflict and cooperation among the very same participants*. They have common interests in getting some share of the economic rents by cooperating; but they have conflicting interests in how the total will be divided into the rents received by each.

A catalogue of games: And their challenges to coordination

Some interactions present greater impediments to coordination than other; the Prisoners' Dilemma is in some respects the most challenging of all.

We can classify coordination problems and the challenges they present by the relation between Nash equilibria and Pareto-efficient outcomes of the games that represent them.

- In the Prisoners' Dilemma, you know, there is a unique Nash equilibrium that is Pareto-inefficient. Because this outcome is also a dominant strategy equilibrium, coordination on mutual cooperation will not occur even if one of the players insists (perhaps for moral reasons) on cooperating.

- In interactions like Planting in Palanpur, which are often called *Assurance Games*, there are two Nash equilibria, (both Plant Early and both Plant Late) one of which (Plant Early) is Pareto-superior to the other (Plant Late). In these games if one of the players plays the strategy making up the Pareto superior equilibrum (Plant Early) then the best response of the other will be to do the same. Finding institutions that will implement the preferred plant early outcome in a game like this will be a lot less challenging than in a Prisoners Dilemma.

- Another important class of coordination problems arise in what we call *Disagreement Games* where there are two Nash equilibria each of which is Pareto-efficient, so that they cannot be Pareto-ranked, and players disagree about which Nash equilibrium they would like occur. These are like the Planting in Palanpur game but with the additional challenge stemming from a conflict over which Nash equilibrium will be implemented.

We start with an even less challenging game in which players' self interests lead them to a Pareto-efficient Nash equilibrium.

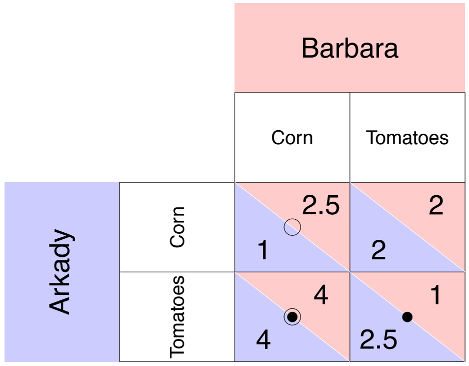


Figure 1.13: An Invisible Hand Game with best responses indicated by circles and dots. Arkady's payoffs are listed first in the bottom-left corner. Barbara's are listed second in the top-right corner. The game captures Adam Smith's ideas of *specialization* and *gains from trade* (that is, the opportunity to obtain economic rents from trade).

1.10 Coordination successes: An invisible hand game

INVISIBLE HAND GAME In an invisible hand game there is a Nash equilibrium that is Pareto-efficient.

The characteristic of an invisible hand game is that it has a Nash equilibrium that is Pareto-efficient. The **Invisible Hand game** illustrates Adam Smith's core insight that through the competitive buying and selling of privately owned goods on competitive markets, self-interested people can achieve outcomes to the benefit of all at least some conditions (that we spell out in Chapter 14). In modern economic language, we would say they avoid Pareto-inefficient outcomes.

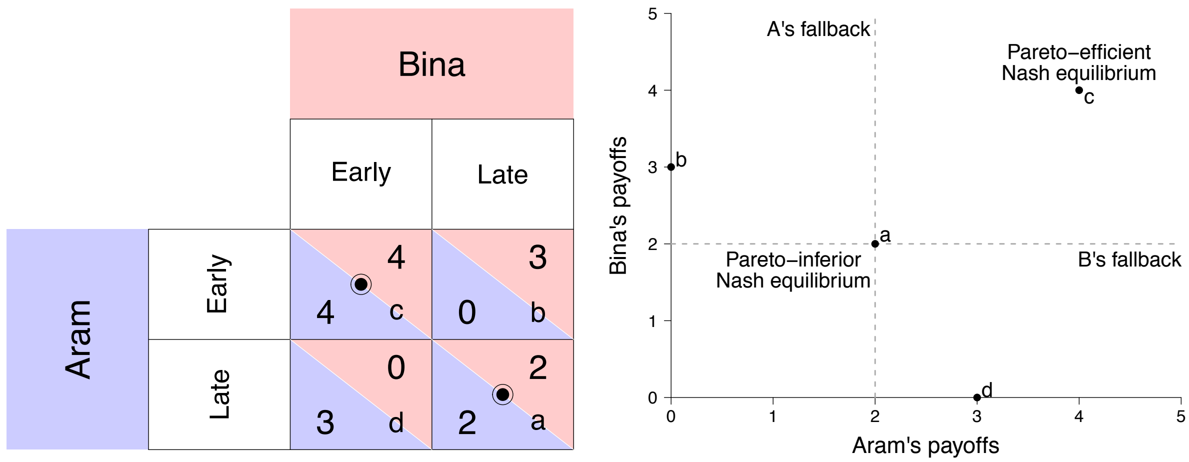
Though our game is much simpler than Smith's reasoning and Smith did not use ideas like Pareto efficiency, our game illustrates an aspect of Adam Smith's idea of how the invisible hand works. The participants, pursuing their self-interest, reach an outcome that beneficial for all of them. (We return to how the invisible hand is understood in contemporary economics in Chapter 14).

Consider a 2-by-2 game between two players, Arkady and Barbara, who are both farmers. Each player can choose between two strategies: planting corn or planting tomatoes. The payoffs that they achieve are provided in the payoff matrix in Figure 1.13, which we call the Corn-Tomatoes game.

The payoff matrix reflects two facts about the problem that the two farmers face.

- Either because of their skills or the nature of the land they own, Arkady is better at growing tomatoes; Barbara is better at growing corn
- They both do poorly when they produce the same crop because the increased supply of whichever good it is that they both produce drives down the price.

The Nash equilibrium of the Corn-Tomatoes game is (Tomatoes, Corn), that is, Arkady plants tomatoes, and Barbara plants corn, at which the players receive payoffs (4,4). (Tomatoes, Corn) is Pareto efficient as there is no alternative



(a) Assurance Game with labeled payoffs

(b) Payoffs plotted against each other

Figure 1.14: **Planting in Palanpur: An Assurance Game.** Aram's payoffs are listed first in the bottom-left corner. Bina's payoffs are listed second in the top-right corner. Aram's best responses are shown by the solid point and Bina's are shown by the hollow circle. The Nash equilibria of the game are (Plant Early, Plant Early) and (Plant Late, Plant Late), with payoffs (4, 4) and (2, 2). The Plant Early Nash equilibrium is Pareto-efficient. The Plant Late equilibrium is not. In the right-hand panel, the payoffs are plotted against each other. Aram's payoffs are plotted on the horizontal axis, increasing as you move rightward. Bina's payoffs are plotted on the vertical axis, increasing as you move upward.

outcome which is Pareto superior to it.

Just as in Adam Smith's reasoning about his invisible hand, Arkady and Barbara, in their interaction, through competition with each other and following their self-interest coordinate their economy to their mutual benefit. In the Invisible Hand Game, each player pursues self-interested objectives and benefits from the fact that the other does too. In an Invisible Hand game individual incentives lead people to act in ways that promotes mutual benefit.

Checkpoint 1.11: Invisible Hand Game

Which entries in the payoff matrix would you have to compare in order to show the follwing:

a. They each do better when Arkady specializes in tomatoes and Barbara specializes in corn then vice versa.
b. They each do worse when both produce the same crop.
c. Growing corn is Barbara's dominant strategy
d. Arkady growing tomatoes and Barbara growing corn is the dominant strategy equilibrium.
e. Explain why the Nash equilibrium of the game is Pareto efficient.

1.11 *Assurance Games: Win-win and lose-lose equilibria*

Return to the farmers in Palanpur. There are two Nash equilibria in this game, one in which both participants Plant Early and one in which both Plant Late. The best response to the other farmer planting early is also to plant early, while the best response to the other farmer planting late is also to plant late. The outcome where both farmers plant early is Pareto-superior to the outcome when both farmers plant late.

The players do not have any conflict of interest: both would share equally in the gains to cooperation, should they find a way to coordinate on planting early. The problem for the real life farmers of that village is that they are stuck in the Pareto-inefficient Nash equilibrium of an Assurance Game. Their challenge is how move to the Pareto-superior Nash equilibrium.

This could happen if all the participants had confidence (were *assured*) that the other participants would follow their lead in moving to the superior outcome. This is why this type of game is often labeled as an "Assurance Game."

ASSURANCE GAME In an *Assurance Game*, there are two Nash equilibria, one of which is Pareto-superior to the other. The Planting in Palanpur Game is an example.

Figure 1.14 is the payoff matrix for two players, Aram and Bina, choosing when to plant their millet in the village of Palanpur, India. (It is the same as the earlier figure about the two farmers, except that we now have numbers representing the farmers' payoffs). Coordination failures arise in the Assurance Game because of *positive feedbacks*: the more people who plant late the more is the incentive for others to plant late, and vice versa. Each strategy exhibits **strategic complementarity**.

Checkpoint 1.12: Graphing Palanpur

Using the graphical method for identifying Pareto-efficient outcomes as shown in Figure 1.14, show which outcomes in the Palanpur game are Pareto-efficient. Can you explain why **a** and **c** are Nash equilibria?

Assurance game and strategic complementarity

STRATEGIC COMPLEMENTARITY *Strategic complementarity* exists when i) *A strategy is a strategic complement to itself*: The payoff to playing a particular strategy increases as more people adopt that strategy as a result of some form of *positive feedbacks*, or ii) *One strategy and another are strategic complements to each other.* The payoff to playing one strategy (say, A) is greater the more people adopt the other (B) in which case we say that strategies A and B are *strategic complements*.

Social media, dating platforms, and other matching services are examples of strategic complementarities. They are more valuable to for everyone if many people participate.

Strategic complementarity exists when either of two conditions hold.

1. *A strategy is a strategic complement to itself*: The payoff to playing a particular strategy increases as more people adopt that strategy as a result of some form of *positive feedbacks*. Dating platforms are an example. The strategy could be "Open a Tinder Account." The positive feedback arises because the more other people that are using Tinder the more people you will "meet." Plant Early in Palanpur is another example as we will see.

2. *One strategy and another are strategic complements to each other.* The payoff to playing one strategy (say, A) is greater the more people adopt the other (B) in which case we say that strategies A and B are *strategic complements*. An example is the Invisible Hand Game shown in Figure 1.13. The payoff to Arkady from planting tomatoes is greater if Barbara plants corn (instead of tomatoes), and the payoff to Barbara from planting

corn is greater if Arkady plants potatoes (instead of corn). Growing corn and growing tomatoes are strategic complements.

The farming in Palanpur problem is a case of the first, not the second. But nobody has any reason to participate if no others do. This is an example of what are called **network externalities** or *network external effects* which occur when the benefits to members of a social or physical network increase when more people join the network.

One example is a particular social network: if you're the only person on it, there is really no point. But, as more and more people join the network, then social networking site becomes more useful and your payoff increases with the number of users in the network. These so-called network externalities are a particular case of strategic complementarity. By joining a network each person confers an external benefit on the existing members of the network.

We predict and evaluate the possible outcomes of the Planting in Palanpur Game using the concept of best response (using the dot and circle method introduced earlier). And we see that the game has two Nash equilibria (Early, Early) with payoffs (4, 4) and (Late, Late) with payoffs (2, 2). (Early, Early) is Pareto-superior to (Late, Late) and it is Pareto-efficient because no alternative outcome is Pareto-superior to (Early, Early).

Even though there is a Pareto efficient Nash equilibrium, a population – like the people of Palanpur – may get stuck in the Pareto inferior Nash equilibrium. that does not guarantee players will actually play it. So we have two conclusions:

- The fact that a Pareto efficient outcome is a Nash equilibrium does not mean that it will be the one we observe and

- In cases where there is more than one Nash equilbrium, we need more information than is provided by the Nash equilibrium and Pareto efficiency concepts to make a prediction about the strategy profiles we will see in practice.

Checkpoint 1.13: Assurance Game

Which payoff table entries would you have to compare in order to show that:

a. Planting early is Pareto efficient.
b. Planting late is a Nash equilibrium.
c. The best response to the other planting early is to plant early.

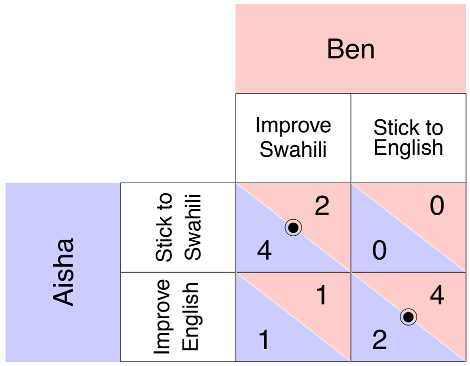


Figure 1.15: **A Disagreement Game**: The players need to coordinate on an equilibrium, but each prefers one equilibrium to the other, so there is a *conflict of interest*. If they fail to coordinate on one of the Nash equilibria because of the conflict of interest, the outcome will be a *coordination failure*.

1.12 Disagreement Games: Conflict about how to coordinate

We use the Language Game described in Figure 1.15 as an example of a **Disagreement Game**. A Disagreement Game illustrates coordination games in which there are two Pareto-efficient Nash equilibria (which are therefore Pareto-incomparable), and the players are in conflict over which Nash equilibrium each prefers. So, the players' problem is to manage to coordinate on one of the Nash equilibria, or alternate systematically between them, to ensure that no coordination failure results and they do not end up at an outcome neither would prefer.

Consider two players, a home-language Swahili-speaker (Aisha) and a home-language English-speaker (Ben) who have recently met. Each person can speak the other language, but prefers to speak their home language. They share many common interests but do not communicate as well as they would like. Each has two strategies: Stick to your home language or Improve the other language.

Among the possible outcomes are that he could learn better Swahili and they could routinely converse in that language; and she could learn better English and they could converse in English. They do not need to both be fluent in both languages. So for Aisha, if Ben becomes fluent in Swahili, then her best response is not to take the time and trouble to improve her English. For Ben, similarly, if Aisha were to become fluent in English, then there would be little point in taking the Swahili courses.

The result is two Nash equilibria (Stick to Swahili, Improve Swahili) with payoffs (4, 2) and (Improve English, Stick to English) with payoffs (2, 4). The two Nash equilibria are both Pareto-efficient because there are no alternative outcomes which are Pareto-superior to the Nash equilibria.

But, as shown in the payoff matrix 1.15, Aisha would prefer the (Stick to Swahili, Improve Swahili) Nash equilibrium and Ben would prefer the (Improve English, Stick to English) Nash equilibrium.

The Disagreement Game is similar to the Assurance Game in that:

- There are two Nash equilibria
- Both players do better if they coordinate (that is, speak the same language at one or the other of these equilbria)

The Disagreement game differs from the Assurance Game because:

- Each player in the Disagreement Game prefers one of the Nash equilibria while the second player prefers the other, while both prefer the Pareto-superior Nash equilibrium in the Assurance Game, so as a result
- the players in the Disagreement Game have a *conflict of interest* concerning which equilibrium gets selected.

HISTORY One of the first game theoretic studies of coordination problems – by David Lewis – was concerned with how we coordinate on a common language.[8]

Of course the English speaker would prefer to communicate in her home language and the Swahili-speaker feels the same way, but they would do much worse if they did not have any common language at all and if they failed to coordinate.

Disagreement Games highlight how there can be social interactions with multiple Nash equilibria, each of which is Pareto-efficient, but there may be no 'middle ground' to coordinate on and as a result conflict over who gets to benefit the most is unavoidable. Both players in the Disagreement Game would both be worse off out of equilibrium than at one of the Nash equilibria in the game. They have a common interest in coordinating somehow as opposed to *not* coordinating; but their interests conflict in *how* they coordinate.

Checkpoint 1.14: Language Game

Label the outcomes of the Language Game as in Figure **??**, plot them using axes with the players' payoffs, and determine which outcomes are Nash equilibria and which are Pareto-efficient.

1.13 Why history (sometimes) matters

As we have seen from Disagreement Games and Assurance Games, strategic complementarities in games may give rise to more than one Nash equilibrium. When this is the case we cannot say which Nash equilibrium is our prediction of how the game will be played. The best the Nash equilibrium concept could do is to say that the outcome of the game is likely to be one of the (perhaps many) Nash equilibria.

We need more information to make a prediction. Think about the Palanpur game, and imagine that all you know is the payoff matrix (not how the farmers played the game in recent years). Though you would be on solid ground predicting that it is likely that you'd see both farmers planting either early or late, you would not have much confidence in which it would be.

But now suppose you were told that last harvest they planted late. Then, unless they had discovered some way to coordinate a switch to planting early, you would be correct when you predicted that they would both be planting late this year too.

PATH DEPENDENCE A process is path dependent if the most likely state of something this period – the fraction of farmers planting early or late in the example – depends on its state in recent periods.

When history matters in this sense, we say that outcomes may be **path-dependent.** When the outcome of a game is **path-dependent**, without knowing the recent history of a social interaction we cannot predict which equilibrium will occur. So, quite different outcomes – poverty or affluence, for example – are possible for two interacting groups of participants with identical preferences, technologies, and resources but with different histories. This is how "history matters."

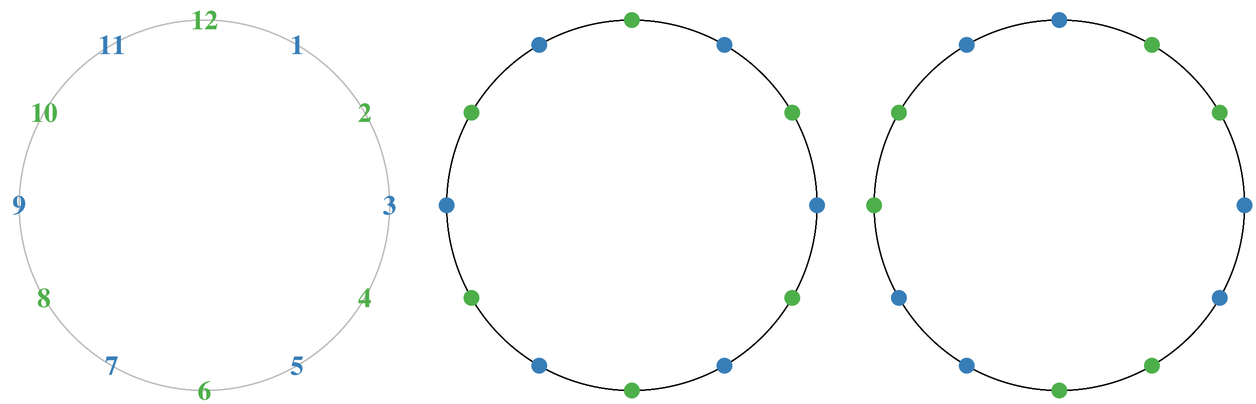
The Palanpur payoff matrix describes a **poverty trap**. A poverty trap occurs when identical people in identical settings may experience either an adequate living standard or poverty, depending only on chance events of their histories, for example were their parents rich or poor, or were they citizens of Norway or Nigeria. The possibility of poverty traps alerts us to the fact that people may be rich or poor not because of anything distinctive about their skills, hard work or other personal attributes, but because of the situation they find themselves in. Poverty may be inherited as it is in Palanpur not by anything that parents pass on to children but instead by the inheritance of a common history.

The same is true about other aspects of how we interact in society, for example in the ways our lives may be highly segregated in interacting with people who differ in the groups with which they are identified, whether that be ethnicity, or religion , or even loyalty in sports teams.

POVERTY TRAP A **poverty trap** occurs when identical people in identical settings may experience either an adequate living standard or poverty, depending only on chance events of their histories. Poverty in this case is a result of a person's circumstances.

Checkpoint 1.15: Drain the meadow: Name that game

Write down a payoff matrix for Hume's "drain the meadow" game, with the two actions open to farmers Adams and Brown being "drain" and "do not drain", and assuming that the value of the drained meadow (to each farmer) is 5, the value of the undrained meadow is 3, and if the two farmers jointly work on the draining it costs them 1 each, while if a single farmer does the draining alone it costs him 3. What kind of game is this? Explain how it might be solved if there were just two farmers, and why with many farmers (as Hume wrote) it would be " difficult and indeed impossible" for them to agree on a common course of action and avoid in a coordination failure.



(a) The circle as a clock (b) The baseline (c) The citizens' ideal integrated neighborhood

Figure 1.16: **The neighborhood and the citizen's ideal integrated outcome** Panel a is the "geography" of the neighborhood, showing that, for example, the citizen at position 2 on the circle has two immediate neighbors, the people at positions 1 and 3. Panel b shows that the person at position 2 is a Blue and her two immediate neighbors are both Greens. is just a starting point at which the neighborhood is as integrated as possible in the sense that the two immediate neighbors of each citizen are of the other type. Panel c shows the distribution of types across locations that the citizens prefer: each citizen has one immediate neighbor of each type.

1.14 Application: Segregation as a Nash Equilibrium among people who prefer integration

Segregated communities – whether on grounds of ethnicity, race, religion, or class – are often the basis of inter-group prejudice and hostility and the systematic denial of equal dignity to all citizens. Segregation often results from deliberate policies of discrimination by governments, lenders, and citizens, as illustrated by the apartheid system of enforced racial separation in South Africa that persisted until 1994 and legally mandated housing segregation in the U.S. – the so-called racial covenants the were finally outlawed in 1968.[9] But segregation can also result from the uncoordinated decisions of people who would actually prefer to live in integrated communities.

FACT CHECK In Seattle, Washington, what are termed "racially restrictive covenants" covering more than 20,000 homes prohibit sale or rental to particular groups. One stipulated that, "No person or persons of Asiatic, African or Negro blood, lineage, or extraction shall be permitted to occupy a portion of said property."[10] Racially restrictive covenants have been illegal since 1968 in the U.S. and are unenforceable.

This counter-intuitive result illustrates the use of the Nash equilibrium concept, underlining the lesson already learned from the interaction among the Palanpur farmers: There may be more than one Nash equilibrium – one Pareto superior to the other – and a society can find it it difficult to escape the inferior equilibrium. The example of segregation is also a reminder – like the case of the over-fishing Nash equilibria– that the fact that an outcome is a Nash equilibrium does not mean that it is something that the players would choose, if they could coordinate and decide jointly on the outcome.

The set-up of the model

Here is a model. There are two types of people, Greens and Blues, and they live in homes arrayed around a circle representing a **neighborhood**. The homes are identical except that they may differ in the types of the **immediate neighbors**. The neighborhood is the circle as a whole. A household's immediate neighborhood is made up of the two households on either side of it.

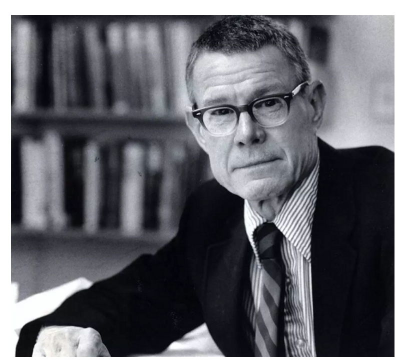


Figure 1.17: Thomas Schelling (1921-2016) was an American economist who won the Nobel Prize in economics in 2005 for his contribution to our understanding of conflict and cooperation. The model we propose here is based on his work.[11]

(a) Household B has one neighbor of each type, their *ideal* situation, is *not dis-satisfied*, and will play Do Nothing

(b) Household B has two neighbors of their own type, is *not dis-satisfied*, and will play Do Nothing

(c) Household B has two neighbors of the other type, is *dissatisfied*, and will play Signal Dissatisfied

Figure 1.18: **The preferences of a household depending on the kinds of neighbors that surrounds it**. Household B will either be satisfied or dissatisfied depending on the the types of neighbors they have. B's choice to play Signal Dissatisfaction or Do Nothing therefore depends on the composition of their immediate neighborhood.

If a citizen would like to live at some other location around the circle, they can switch with some other person currently occupying that position, as long as the other person is willing. The homes just change occupants with no money changing hands. We would like to know what the neighborhood will look like when all the switching that people can do has been carried out, so that the neighborhood's composition *stops changing*. The distribution of types among the houses around the circle when *no citizen can benefit* by moving is a Nash equilibrium.

Greens and Blues have identical preferences about the type of their two immediate neighbors only. All people in the neighborhood would prefer to have one neighbor of each type, as is shown in Figure 1.18. But they are "satisfied" as long as they either have an immediate neighbor of each type or if both are of their own type. People are "dissatisfied" if both immediate neighbors are of the *other* type. An ideal neighborhood, then is shown above in Figure 1.16 c: Each person has one neighbor of each type.

People have two strategies: "Do Nothing" or "Signal Dissatisfaction." Signalling dissatisfaction means being willing to switch positions with another person – anywhere in the neighborhood – who has also "signalled dissatisfaction." People are willing to switch only if they prefer the new location to their old location. For this reason people of either type will *never* switch with a person of the *same* type. This is because, for example, if a Green is dissatisfied with her current location and would like to move, all other Greens would be equally dissatisfied were they to take her position, so no other Green would agree to a switch. So all of the switches will be with different types: a green will switch with a blue, but a blue will never switch with a blue and a green will never switch with a green. This means that switches will change two things:

- *the switcher's new immediate neighborhood*: those on either side now experience having a neighbor of a new type given the switcher's arrival and the previously dissatisfied person's departure

- *the switcher's old immediate neighborhood*: those who were previously on either side of the switcher have a neighbor of a new type given the arrival of the person with whom the previously dissatisfied person switched.

A segregated Nash equilibrium

We begin with 6 Greens and 6 Blues occupying alternating positions in the 12 “houses” at the locations on the circle that are numbered as if from time on a clock (so, 12 is the top). The twelve homes on the circle are "the neighborhood." We call the assignment of different types to the the homes around the circle in Figures 1.16 b and c: an *allocation*. An allocation in this game is an assignment of homes to the types at a given stage of the game. The allocation before the game starts is the *initial* allocation. The allocation after the game ends is the *final* allocation.

The game proceeds as follows. At each step, each of the 12 people plays either Do Nothing, or Signal Dissatisfaction. Their choice of a strategy is known to all other players. Then, one of the twelve citizens is randomly selected and given the opportunity to make a switch if she can find another person willing who has also signalled dissatisfaction and is willing to make a switch.

At step 1, for example, we might ask the Green at position 10 o’clock if she would like to switch. She would, because both of her neighbors are Blues. Whether she is able to make a switch depends on whether there are others who have chosen the strategy Signal Dissatisfaction. Because everyone else is also dissatisfied, she has many choices.

Suppose she switches with her friend and immediate neighbor, the Blue at position 11, shown by the colors of position 10 and 11 changing from Panel a Start to Panel b Step 1. The two people are still friends and neighbors, but each now also has the same type of neighbors on the other side.

Suppose, next, that it is the Blue at position 7 who is picked to stay or switch. If he plays "Signal Dissatisfaction," he could switch with his friend and immediate neighbor at position 6. We continue this process until either no one is dissatisfied, or if someone is dissatisfied, there are no others playing the strategy Signal Dissatisfaction with whom a switch is possible.

This process could continue as shown in the figure, resulting at the end of 5 steps in the completely segregated neighborhood shown in Figure 1.19. Notice that over the course of the game, a particular home may change hands more than once. The home at position 7 for example started off occupied by a Blue who switched with a Green in Step 2, who then switched with a Blue in Step 4.

At step 6 (not shown), each of the 12 would choose the strategy Do Nothing, because 8 of them have same type as neighbors only and the other four have one neighbor of each type. So no one is dissatisfied. As a result there we observe no further moves: the allocation is *stationary* (meaning unchanging). It is a Nash equilibrium.

We could expand the strategies available to the players to allow those with

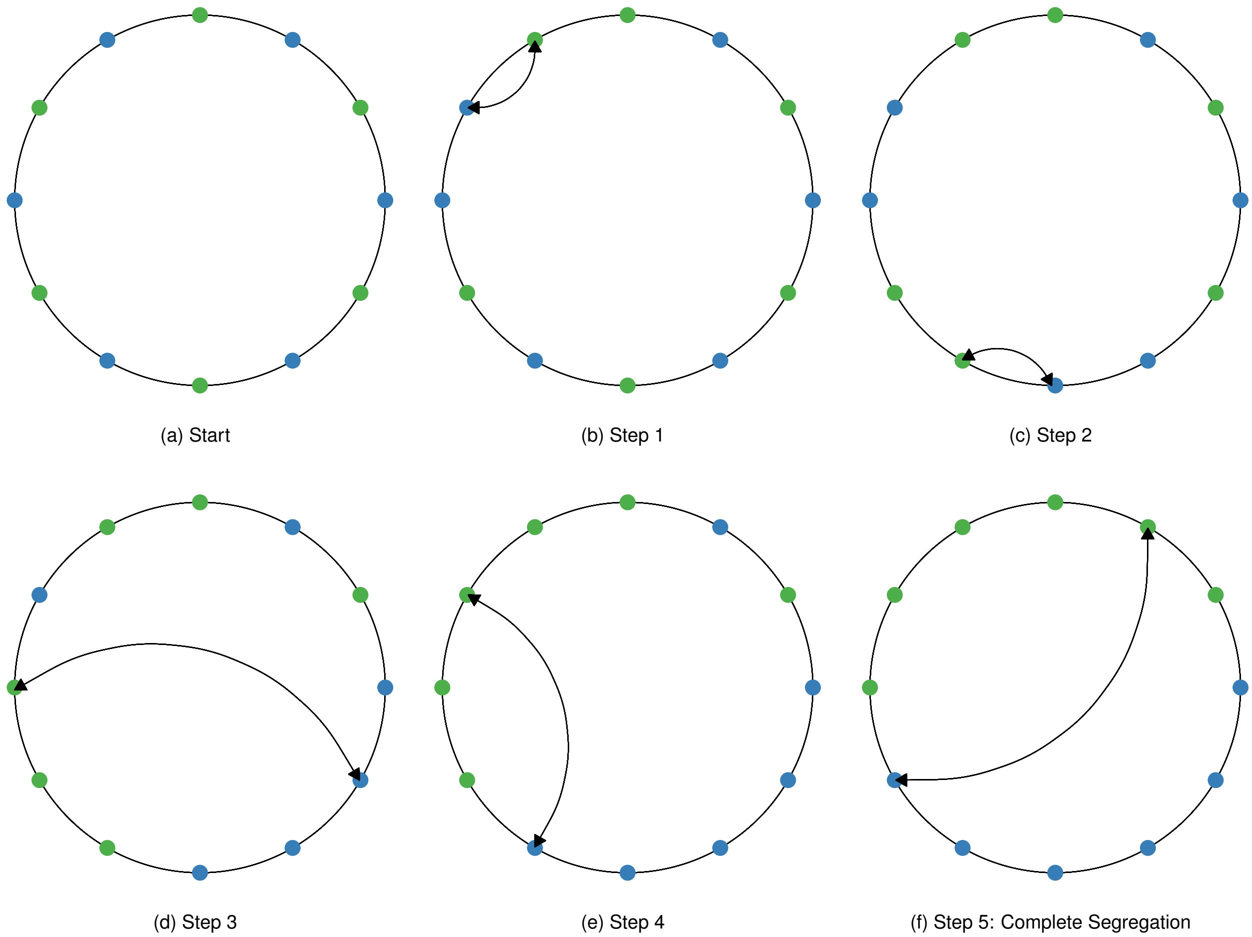


Figure 1.19: **From integration to a segregated Nash equilibrium.** The figure shows one of many possible progressions from an integrated non equilibrium situation to an entirely segregated Nash equilibrium. Panel a shows the starting point from the previous figure. In step 1 the Green at position 10 and the Blue at position 11 switch positions, shown by the double heded arrow, and resulting in the neighborhood shown in Panel b. The remaining panels show the successive steps to the final fully segregated Nash equilibrium.

both neighbors of their own type to signal dissatisfaction, and if a willing other citizen could be found, they could switch with this person so as to have one neighbor of each type. This will not disrupt the completely segregated neighborhood as long as people prefer to have both neighbors of their own type to having both neighbors of the *other* type.

Avoiding outcomes that nobody prefers

The conclusion is *not* that complete segregation will necessarily be the result. This is true for two reasons.

- *There is also a Nash equilibrium that is integrated rather than segregated.* In Figure 1.16 c, the allocation has each person's immediate neighborhood composed of both types. You can confirm that, like the completely segregated allocation, this integrated allocation is also a Nash equilibrium: every citizen has their ideal immediate neighborhood, so no citizen is dissatisfied and each are best responding with Do Nothing. This allocation could have come about by the same rules of the game that resulted in complete segregation. This is an example of implementing a desirable allocation within given set of rules of the game

- *The citizens could play the game cooperatively rather than non-cooperatively*. If the citizens had realized that playing the game non-cooperatively could lead them to a complete segregation outcome that nobody wanted, they could have acted cooperatively – that is jointly agreed – to implement their ideal allocation. This is an example of implementing a desirable allocation by changing the rules of the game: agreeing to act jointly was not an available strategy in the non-cooperative variant of the game above.

The outcome in the segregation model shares three features with a game from what would appear to be a very different situation: Planting in Palanpur.

- *A Pareto inferior Nash equilibrium:* There is a Nash equilibrium – planting late and a segregated community – in which everyone is worse off than they could be at some other allocation.
- *A path-dependent outcome*: History matters because an outcome that is preferred by all participants is also a Nash equilibrium, so if the preferred outcome were to occur, it could persist.
- *A change in the rules of the game can avoid the inferior outcome:* By coordinating their actions – changing the interaction to a cooperative game – they could escape the Pareto inferior outcome

In these three respects the two interactions – when to plant and where to live – are *not unique* or even unusual in these three respects.

Checkpoint 1.16: Segregation as a Nash equilibrium

a. Show that the segregated neighborhood in Figure 3 is a Nash equilibrium.
b. Show that the ‘ideal neighborhood’ in Figure 1 is also a Nash equilibrium.
c. Show that in Figure 1.19 if Step 3 had been different the equilibrium allocation could have been the citizens’ ideal integrated allocation. Which Step 3 switch would have accomplished this result?
d. Suppose that the game was changed slightly so that a dissatisfied person knows only the “dissatisfaction status” of her immediate neighbors. Show that, starting with the alternating types status quo (Panel a Start, in Figure 1.19) that the neighborhood would then evolve to the ideal distribution.
e. Suppose that in the fully segregated neighborhood case citizens decided to have a binding referendum to implement the ideal neighborhood (requiring whatever moves are necessary to bring that about), but because the question of where you live is a sensitive one, they adopt the rule that unanimous approval of the referendum is required for it to be implemented. Would it be implemented?

1.15 How institutions can address coordination problems

Game theory has given us a catalogue of coordination problems: Prisoners' Dilemmas, Invisible Hand Games, Assurance Games, and Disagreement Games (there are many more!). Knowing how the structure of these games differ will help to diagnose the nature of a coordination problem and to devise policies and constitutions – changes in the rules of the game – to avoid a coordination failure.

This is an example of using the concept of equilibrium to understand how to change an undesirable outcome. The idea is simple: a change in the rules of the game can eliminate an undesirable Nash equilibrium, so that it is no longer our prediction of how the game will be played. Instead it may be possible to make some preferable strategy profile a Nash equilibrium which then could be the outcome of the game.

A common approach to averting coordination failures is in a Prisoners' Dilemma to devise policies or institutions that transform the payoff matrix so that the game is no longer Prisoners' Dilemma. There are two possibilities to consider:

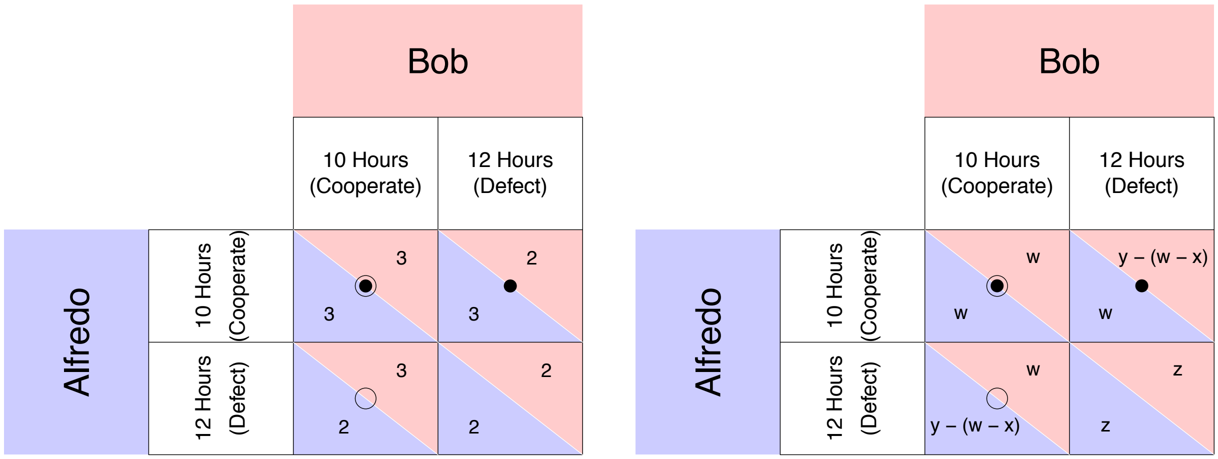
- Change the Prisoners' Dilemma to an Assurance Game
- Change the Prisoners' Dilemma to an Invisible Hand Game

Changing the Prisoners' Dilemma game into an Assurance Game means making the mutual cooperate outcome a Nash equilibrium (it was not in the Prisoners' Dilemma) even if mutual defect remains a Nash equilibrium.

The second options is a more ambitious policy objective: converting a social interaction from a Prisoners' Dilemma to an Invisible Hand Game. To see how this might work, remember that the coordination failure that results in the Prisoners' Dilemma is a consequence of the fact that in that game players can take actions that inflict costs on others – negative external effects – that are not part of their thinking when they decide what to do.

To see that internalizing these external effects can address the coordination failure, we examine the implementation of a liability rule in the Fishermen's Dilemma. Tort is a branch of law dealing with damages inflicted by one person on another (or another's property). Among other things, tort law establishes the responsibility – called *liability* – of the person inflicting the damages to compensate the harmed individual. The requirement to compensate the harmed individual internalizes the external effect.

How would a liability system work in the Fishermen's Dilemma? Look again at Figure 1.12, the general form of the Prisoners' Dilemma. Suppose Alfredo and Bob decided to jointly adopt "Cooperate" (fish 10 hours) as an agreement. In their agreement, they also choose to adopt a *liability rule* requiring compen-



(a) Liability rule with numbers

(b) General liability

Figure 1.20: **Fishermen's Dilemma with a liability rule**. Players can implement a desired outcome by Transforming Property Rights using a liability rule (the harm a player does to another player is deducted from their payoff). This payoff matrix is based on Figure 1.12 modified by the liability rule. Alfredo's payoffs are listed first in the bottom-left corners and shaded blue. Alfredo's best responses are shown by the solid point. Bob's are listed second in the top-right corners and shaded red; his best responses are shown by the hollow circle.

sation be paid to the other party if one's actions result in lower payoffs than would have occurred had the agreement to cooperate (fish only 10 hours) been observed.

With the liability rule the following will happen:

- If Alfredo defects on Bob (plays Fish 12 hours), Alfredo initially gets y as before
- But then Alfredo must compensate Bob for the costs his defection has inflicted, that is, Alfredo must pay compensation sufficient to give Bob a payoff of w (which is the payoff Bob would have obtained had Alfredo not violated the agreement)
- If both players defect, they both gain z
- But, then they must compensate the other player by a transfer of $w - z$.

We can use these changes to the payoffs to construct a transformed payoff matrix. The transformed payoff matrix for Alfredo's and Bob's payoffs is given by the entries in Figure 1.20.

Did the improved property rights succeed? Because $y - w + x < w$ by the definition of a Prisoners' Dilemma, Cooperate is now a best response to Cooperate and (Cooperate, Cooperate) is a Nash equilibrium. Cooperate is also a best response to Defect (because $w > z$), so Cooperate is the dominant strategy with the liability rule in force, and (Cooperate, Cooperate) is the dominant strategy equilibrium. The Prisoners' Dilemma game has become an Invisible Hand game through the adoption of a new set of property rights.

REMINDER For a Prisoners' dilemma we need $y > w > z > x$ and $x + y < 2w$. It's the second one that's important here.

Redefining property rights – to take account of liability for damages – can implement a Pareto-efficient outcome by inducing each player to account for how his actions effect the other player. By redefining property rights to include the liability of the damages (external effects) that one inflicts on others, we have transformed the game to an Invisible Hand game.

M-Note 1.1: The mathematics of the liability rule

Refer to Figure 1.20. For the original game to be a Prisoners' Dilemma requires:

- $y > w > z > x$,
- $x + y < 2w$,
- and has the Nash equilibrium (Defect, Defect) with payoffs (z, z).

The following inequalities must help us to think through the logic of the liability rule:

- With the transformed payoffs the Nash equilibrium must be (Cooperate, Cooperate) with payoffs (w, w).
- The condition for $x + y < 2w$ arises because we require that for $w > y - (w - x)$. We can re-arrange the condition because $w > y - w + x$, therefore, by adding w to both sides we get $2w > y + x$ as we said we need to assume.

Be sure that you can identify the intermediate step here that we have skipped in the matrix: For the combination (12 Hours, 10 Hours) and (10 Hours, 12 Hours) the payoff to the player playing 10 hours is w because it *was* $x + (w - x)$ (that is, the players started with x, but were then rewarded with the transfer for $(w - x)$ by the other player breaking the agreement. $x + (w - x) = w$! Notice too, that both of the players receiving z in the (12 Hours, 12 Hours) occurs because they *both* compensate the other with $(w - z)$ and $z - (w - z) + (w - z) = z$.

Checkpoint 1.17: Limited Liability by the Numbers

Use the model of the liability rule in Figure 1.20 to complete the following tasks.

a. Re-draw the payoff table, but substitute in the values for x, y, w and z from Figure 1.10. **Hint**: The payoffs should only be 2s and 3s.

b. Solve your new game using best response analysis (the circles and dots method) to find the Nash equilibrium of the game. What is it? Explain.

c. Does either player have a strictly dominant strategy? Is there a dominant strategy equilibrium? Explain.

1.16 Game theory and Nash equilibrium: Importance and caveats

We have started off this introduction to modern microeconomics with game theory. The reasons are that

- Many important economic relationships – in labor markets, families, credit and financial markets, between citizens and governments, among neighbors, between nations seeking to address climate change, and many more – are strategic, and require the tools of game theory.

- Focusing on people as actors often in conflict with each other, but also sharing common interests, is essential to economics as a social science, and game theory allows us to do this.

- How these interactions work out depends on the institutions that regulate them, and game theory allows us (even requires us) to be very specific about the varieties of possible rules of the game under which we now interact, and how we might change these rules for the better.

For game theory the Nash equilibrium is a key economic idea and it provides a way to answer the question: what will be the outcome if each of the actors adopts a strategy that will not lead any other actors to change what they do?

In many situations the Nash equilibrium among players who seek to maximize their own material payoffs provides a good prediction of what we observe in the real world. But not always.

- *Extreme individualism: Overlooked opportunities to coordinate:* If the two fishermen were fishing 12 hours each and the details of the situation were such that they could just agree to fish 10 hours instead, then we would be mistaken to use the Nash equilibrium of this game to predict what the players would do. In this case the rules of the game have been inadequately described: there was a third strategy that was overlooked, namely agree with the other to fish 10 hours as long as the other agreed to the same. With the game modified in this manner, the Nash equilibrium would provide a good prediction: both Bob and Alfredo would choose "Agree" as long as they believed that the other would do the same and that the agreement would be enforced.

- *Self-interest: Overlooked payoffs not in the form of one's own material gains:* Even if for some reason they could not agree on a common course of action, if Alfredo and Bob were brothers and each cared about how much fish the other had to feed his family, then the two would not define their payoffs in the game simply as amount of fish they each caught individually, but would include the consequences of their actions on the outcome for the other fisherman. So we would have to rethink the payoff matrix. In this case it is the payoff matrix that was mistaken. It no longer describes what the players are trying to maximize, that is the incentives that are shaping their behavior.

- *Selection among many equilibria:* As we have seen in the Planting in Palanpur Game, there may be more than one Nash equilibrium, so the prediction that the result of the interaction will be a Nash equilibrium is insufficient. We need to know more about the situation – such as the recent history of the people involved – to make a prediction.

EXAMPLE When people can bind themselves to common agreements or when they have values other than maximizing their own material gains, we need to modify the game as we will see later in this chapter, and in Chapter 2.

- *Dynamics*: Sometimes we are more interested in the process of change itself than in the stationary end point of this process. Whether studying how the people of Palanpur might make a transition from poverty to affluence, or how the polar sea ice might collapse due to climate change we are often interested in how the economy works when it is not in equilibrium.

1.17 Application: Cooperation and conflict in practice

If all that is needed to address a coordination failure is to require that people pay the costs that their actions impose on others then why are coordination failures so common?

Over-exploitation of fisheries is an international problem that humans as a world community have failed to solve. Many over-exploited fisheries will not recover for a long time. But local communities and groups of fishermen have found ways to combat over-fishing and we can learn many lessons from what they do. Many groups – from farmers to fishermen – face equivalent problems worldwide. These outcomes provide a concrete motivation to study the Prisoners' Dilemma Game and other coordination problems.

What we learn from these games is that an effective liability rule requires two things:

- The injured party or the courts have to have verifiable information (that is information sufficient to enforce the liability aspect of the property right) and
- There has to be a court or some other body willing to and capable of enforcing the contract.

When we turn from game theory to the study of real fishing communities we find that both conditions are unlikely to be met, which is why the over-exploitation of fisheries continues in many cases.

- *Limited information.* The lack of verifiable information is common in social interactions and this limits the policies that governments or private actors can design in response to the persistence of Pareto-inefficient Nash equilibria.
- *Conflicts of interest.* Governments may not have the capacity or the will to enforce the necessary policies especially in cases where doing this would impose costs on a powerful group. An example is the failure of many countries to address the problem of climate change, which is in part the result of the fossil fuel companies' opposition to putting a sufficiently high price on carbon emissions.

Fishing communities, of course, are not acting out a tragic script, as were the herders in Hardin's tale about the tragedy of the commons. They are not prisoners of the dilemma they face. Real fishermen are resourceful and seek solutions to the problem of over fishing.

- Lobstermen in the U.S. state of Maine limit how many lobster they catch using highly local restrictions on who can set traps where (the state government provides the legal framework for this).[12]
- Turkish fishermen allocate fishing spots by lottery and then rotate the spots so the distribution is fair.[13]
- The fishing community of Kayar in Senegal adopted the rule that only one trip to the fishing grounds per day is permitted (a bit like Alfredo and Bob limiting their hours of fishing) and appointed a committee to check that the rule was being observed. They also limited the number of boxes of fish that could be offloaded by a single canoe.[14]
- Shrimp fishermen in Toyama Bay, Japan have a rule that they offload their daily catch at the same time and place, so that the size of each boat's catch would no longer be asymmetric information.[15]

These rules and practices based on small local fishing communities are often disrupted by the entry of other groups who are not bound by the local rules. Conflicts of interest within the local community also sometimes limit the effectiveness of attempts to limit the catch. One reason is that restrictions on fishing are often supported as a way to raise the wholesale price of fish and hence the incomes of fishing families. But fish sellers – who buy the fish wholesale at the port and then sell to local consumers – would profit if they could pay less.

The rules regulating access to fishing that we see in existence are a small selection from a much larger set of rules that people have tried out at some point. What we see are the institutions that have succeeded well enough to allow the communities using them to persist and not to abandon their rules. The persistence of such rules does not require the rule to implement a Pareto-efficient outcome, it only requires that the rule be reproduced over time by people adhering to it. By this reasoning, even if the rules of the game do not implement Pareto efficient outcomes, we might expect a fishing community that has hit on a way of sustaining cooperation in the long run to do better in competition with groups that over-fish, and that successful groups may be copied by other groups.

Checkpoint 1.18: Institutions and Palanpur

Supposing that the only voters involved in approving the Palanpur village council's decision to require planting early were themselves farmers, explain why they would unanimously support the measure. What would happen if after implementing the law requiring early planting one year, the next year the law was taken away?

1.18 Conclusion

The classical institutional challenge which we stated originally was “How can social interactions be structured so that people are free to choose their own actions while avoiding outcomes that are worse for everyone than they could be?” With the terms you have learned this can now be re-phrased “How can social interactions be structured so to avoid Pareto-inefficient Nash equilibria resulting from people’s free choice of their own actions?” The Fishermen’s Dilemma is an example of a challenging coordination problem because an inefficient outcome is the unique Nash equilibrium.

To study a game and its likely outcomes and also how to improve these outcomes we have proceeded in three steps:

- First, use the Nash equilibrium concept to identify one or more likely outcomes of the game
- Second, use Pareto comparisons to identify outcomes that are "worse for everyone," and
- Third, devise changes to the relevant institutions – the rules of the game – or that would shift the population to a superior Nash equilibrium either pre-existing (as in the case of Planting in Palanpur, or the segregation case) or novel (as with the transformed Prisoners Dilemma Game).

We have illustrated the third step by a legal remedy, the introduction of tort liabilities for damages in the Prisoners’ Dilemma Game so as to internalize the external effects accounting for the coordination failure. But an adequate analysis of coordination problems and their possible mitigation must do two things that we have not yet taken up:

- Evaluate outcomes according to their fairness and other ethical criteria and
- illuminate the far broader range of how markets, governments, private associations like firms or neighborhoods and other bodies might jointly accomplish this task. We take up this challenge in Chapters 2, 4 and 5.

Making connections

Institutions and the rules of the game: To predict or explain the outcome of a social interaction, it is essential to know the “rules of the game” that determine who *knows* what and when, who *gets to do* what and when and as a result who *gets* what.

Equilibrium: Equilibrium describes an outcome that will persist until some aspect of the situation changes as a result of externally caused changes. A Nash equilibrium is a special kind of equilibrium widely used in economics.

External Effects: People often take actions without considering the effects of these actions on others. The resulting external effects – positive and negative – pervade social interactions.

Pareto inefficiency of Nash equilibria: A common result is that the outcomes of social interactions (the Nash equilibria of the games) are Pareto-inefficient meaning that opportunities for mutual gains remain unrealized.

Rents: When players interact they face opportunities for mutual benefit, or *common interest*. But this creates opportunity for rents and for a *conflict* over how the benefits resulting from the interaction will be distributed.

Policy and changes in the rules of the game: Improving property rights (such as making people legally responsible for the harms to inflict on others) can lead people to internalize external effects. Other institutions that would facilitate people making decisions to act jointly can also provide solutions to coordination problems. Policy may result in a shift to a Pareto-efficient equilibrium.

Positive feedbacks, increasing returns, and strategic complementarity: Often players strategies are *strategic complements* due to *positive feedback* and *increasing returns*. As a result, in some social interactions there may may be many equilibria as in the Assurance Game and the Disagreement Game.

Important Ideas

institutional challenge	fallback	Pareto-superior/inferior
coordination problem	next best alternative	Pareto efficient
player	best response (weak/strong)	(economic) rent
strategy	dominance (weak/strict)	payoff
Dominant strategy equilibrium	institution	Nash equilibrium
interdependence	positive external effect	negative external effect
Prisoners' Dilemma	(non-)cooperative games	optimization
Tragedy of the commons	Invisible Hand Game	Assurance Game
Disagreement Game	Increasing returns	Strategic complement/substitute
Path dependence	Poverty trap	Liability rule
positive feedback		

Mathematical Notation

Notation	Definition
π	a player's payoff

Note on superscripts: A, B: individuals.

Discussion questions

See supplementary materials.

Problems

See supplementary materials.

2

People: Self-interest and Social Preferences

> How selfish soever man may be supposed, there are evidently some principles in his 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.
>
> Adam Smith, *The Theory of Moral Sentiments* (1759) Chapter 1.

DOING ECONOMICS

This chapter will enable you to:

- Understand that people make decisions based the actions open to them (*constraints*), which of these possible actions they believe they must take (*beliefs*) to bring about the outcomes they most prefer (*preferences*).
- Use this approach to analyse economic situations involving risky outcomes, bargaining, and contributing to the public good.
- Analyse sequential games and games with multiple Nash equilibria, showing how being the first mover in these games may confer advantages on a player.
- Explain how the experimental method along with this "preferences, beliefs, and constraints approach" is used to study economic behavior.
- See how changes in the rules of the game can result in better outcomes for all.
- Describe the experimental and other evidence that people's preferences go beyond self-interest and include generosity, reciprocity, fairness and hostility toward others .
- Understand that these other-regarding preferences are as much part of what we consider to be rational action as is self interest.
- Give examples of the importance of social norms and culture for decision-making and economic policy-making.

Chicago, a city famous for its pizza, sports, jazz, and its skyline, built its fortune on the farming of the state of Illinois. Today Illinois farmers use high tech machinery and advanced business plans, some cultivating a thousand acres of land or more. But many of the farmers don't *own* the land they cultivate; they rent land and work it as a *sharecropper*. Sharecroppers are farmers – "tenants" – who pay the owners of land a share of the total harvest that they cultivate.

In the mid-1990's, over half of the contracts between farmers and owners were sharecropping agreements, and in Northern Illinois 95 percent of these contracts stipulated a fifty-fifty division of the crop between the owner and the sharecropper. An equal split of the crop means that a tenant on fertile land will have higher income for the same amount of effort and other inputs.[1] Because a tenant on fertile land will reap a larger harvest than a tenant on poor quality land, the fifty-fifty sharecropping contract presents us with a puzzle.

Figure 2.1: **Farming in Illinois is big business.**

Here's the puzzle: if half of the crop on poor quality land is sufficient to attract tenants, why should the owners of good quality land give up half of the crop to their tenants? Those tenants must be earning more than what was necessary to get them work the owner's land. So, why don't the owners of the better land propose a tenant's share less than half, giving the tenants just enough so that they are willing to farm the land?

We would expect owners to insist on lower tenant's shares to sharecroppers on higher quality land and offer higher shares to sharecroppers on low quality land. Because land varies in quality by small gradations, this would result in a pattern of sharecropping contracts with a range of shares depending the land quality. But they is not what we see. Almost all of the contracts are fifty-fifty.

Illinois sharecropping contracts allow the sharecroppers on good land to receive income attributable to the superior land quality, income the owners would otherwise have received if the owners had insisted on a lower tenant' s share on the high quality land. The fifty-fifty split effectively transfers millions of dollars annually from owners to sharecroppers simply because of the fifty-fifty division. This is not a peculiarity of Illinois. Fifty-fifty is the norm in sharecropping around the world.

Rice cultivation in West Bengal, India during the 1970s provides another example. There, poor illiterate farmers in villages isolated by impassable roads for much of the year and lacking electronic communication eked out a bare living on plots that average just two acres rather than the thousand-acre plots farmed in Illinois. The Indian farmers shared one similarity with farmers in Illinois: the division between sharecroppers and owners was fifty-fifty in over two-thirds of the contracts.[2]

Why was the contract the same in these distant parts of the world? The short answer is that where most contracts are fifty-fifty, that particular division is a social norm, something people feel they are morally obliged to respect. The fact that around the world land owners respect a social norm that overrides their material self-interest tells us that many people are committed to acting fairly, being treated fairly and conforming to ethical standards of appropriate behavior.

HISTORY In 1848 the British philosopher-economist John Stuart Mill noted the striking global pattern of equal division in crop sharing, calling it "the custom of the country" and "the universal rule."[3]

But the sharecroppers in Illinois and West Bengal, like farmers everywhere, are also trying to make a decent living, or even to become affluent. They are not simply following social norms. They carefully weigh alternative methods of cultivating their crops at the least possible cost and marketing their harvest at the highest possible price.

2.1 Preferences, beliefs and constraints

Understanding economic behavior requires a model that takes account of what people care about (for example, the farmers' incomes, and also their desire to uphold social norms) and how from the actions they are able to undertake, they adopt those that are they think will bring about desired results. We will develop a model of economic behavior based on:

- **constraints:** the feasible set of actions, meaning actions that are open to us,

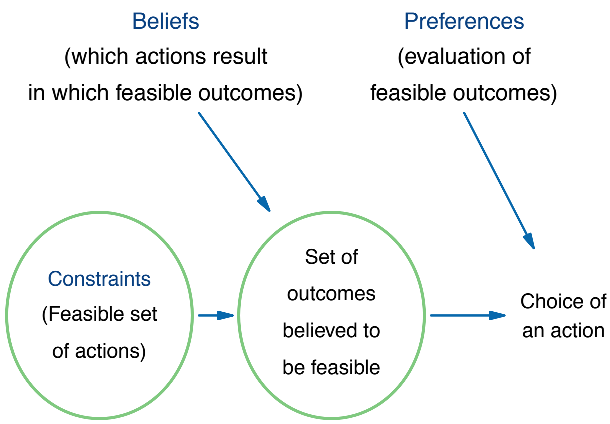


Figure 2.2: **Preferences, beliefs and constraints.** The actor may choose from a set of feasible actions (the constraint set on the left). Combining that set with her beliefs about the outcome produced by each of the actions in the the constraint set, she then has a set of outcomes that she believes are feasible, depending on her choice of an action. From all of these outcomes in the set believed to be feasible, she identifies the one that is ranked highest according to her preferences and then takes the action that she believes will bring about this outcome.

- **beliefs:** our understanding of the outcomes that will result from the actions that are open to us,and
- **preferences:** our evaluation of the outcomes that we believe will result from the actions we take.

This is called the preferences, beliefs and constraints approach.

The relationship between these three elements of the preferences, beliefs, and constraints approach is described below and is shown in Figure 2.2. Game theory, which you have already studied, is an important example of the preferences, beliefs and constraints approach.

EXAMPLE The preferences, beliefs and constraints approach is sometimes called rational choice theory or the rational actor model, but we prefer the more specific label that we use here as it identifies the the three important elements making it up.

PREFERENCES, BELIEFS, AND CONSTRAINTS APPROACH According to this approach, from the feasible set (which includes all of actions open to the person given by the economic, physical or other *constraints* she faces) a person chooses the action that she most values as given by *preferences*, in light of given her *beliefs* about the actions that will bring about the outcome.

Constraints: limits on action

From a long list of things she might consider doing, constraints define a *more limited possible set of actions,* namely the shorter list of all of those so called feasible actions she can carry out. In the game theory introduced in the previous chapter the constraint was the set of possible actions, that is, a list such as "Fish 10 hours", "Fish 12 hours" or "Plant early", "Plant late".

Constraints may be imposed by personal limitations, by laws of nature, or by the force of law. A constraint can also reflect a decision by the actor to eliminate some action from the feasible set of actions on moral grounds, irrespective of the payoffs. Examples are keeping promises, or not committing murder.

In Table 2.1 we give examples of how the preferences beliefs and constraints approach can be applied.

The list of feasible actions set by constraints need not be just a list of particular actions, like drive or take the bus. When marketing their output (first row of

Actor	*Constraints (feasible set of actions)*	*Beliefs (information about which actions will result in the preferred state)*	*Feasible Outcomes (states that could result from the actions)*	*Preferences (ranking of all outcomes)*
Firm owner	High or low prices	The demand curve (how quantity depends on price)	Various levels of profits	Maximize profits
Urban resident	Drive or take the bus	How many others will drive	Travel time	Minimize travel time
Ordering a meal	The menu; your budget	Simple: just order the best you can pay for	Meal quality, money left over	Maximize utility

Table 2.1: **Applications of the preferences, beliefs, and constraints framework**. Real choice situations are typically not as simple as Figure 2.2. The urban resident, for example, may care both about travel time to work and his carbon footprint.

the table), the owners of a firm, for example, can set any price they like (anywhere from 0.00 by penny increments up to some very high number).

Wealth, the availability of credit, and the prices of goods impose constraints on an actor's consumption. The institution of private property also imposes limits: it means that theft is not an option for increasing your consumption. Given private property and in the absence of gifts or other transfers from a government, the total amount of goods and services you can consume is limited by wealth and how much you can borrow. So when we study someone's consumption, their *budget constraint* is a critical factor as people have a certain budget determined by wealth, access to credit, and prices all limiting how much they can buy.

Beliefs: Translating actions into outcomes

Beliefs are a person's understanding of the outcomes that her actions will bring about.

BELIEFS A persons understanding of the relationship between her actions and the outcomes that will occur as a result of her actions are her beliefs. Beliefs are thus a causal mapping from the actions one can take to the outcomes that will occur. Where the outcomes of actions are not known with certainty, beliefs include probabilities of results occurring.

EXAMPLE The word "belief" is often used to refer to spiritual matters ("religious beliefs"); but in game theory a belief is a statement about how the world works, namely what action is required to bring about some particular outcome.

In many cases what I must do to get the outcome that I prefer depends on what other people do. I would like to spend the evening with friends, but where I should go to make it happen depends on where I think my friends will go. Given that I cannot communicate with my friends (the batteries to their phones have run out), my action (where I will go) will therefore depend on my belief about where I will find my friends.

In Table 2.1 the owners of firms are not constrained to set any particular price, but if they want to translate their choice of a price into what they care about – profits – they must form an opinion about the number of units they will be able to sell at each price. This is the demand curve, and it expresses the owners' beliefs about the relationship between their action (the price) and an outcome (how many goods they will sell).

In the game theoretic approach of Chapter 1 beliefs were expressed in the solution concept, that is a description of how the game would be played. The Nash equilibrium as a solution concept, for example, is based on the idea that players best respond to the play of the other players. This is the basis of players' beliefs about how their choice of an action will translate into an outcome of the game.

Preferences: Reasons for preferring one outcome over another

EXAMPLE While most widely used in economics, the preferences, beliefs, and constraints approach is also used in political science, for example, to understand the strategies followed by elected officials seeking to maximize their chances of re-election, in law to design criminal or civil penalties to effectively deter illegal activity, and even in biology to study the evolution of genes, modeled as if they are "trying to" increase their numbers.

PREFERENCES Preferences are evaluations of outcomes that provide motives for taking one course of action over another.

Preferences are evaluations of outcomes that provide motives for actions. A person's preferences are the reason why she takes the action that that she believes will bring about the outcome that is better than or at least as good as the others. In Chapter 1, preferences were represented by the payoffs in games that people played. For each player, a strategy profile was associated with a number – her payoff – and players chose actions that they believed would result in the strategy profile with their most preferred (highest payoff) outcome.

In many games preferences are represented by money payoffs. But more broadly, preferences represent the favorable (positive) or unfavorable (negative) feelings a person has about an outcome that lead them to try to make an outcome happen (high payoff) or that lead them to try to avoid an outcome (low payoff). Preferences include:

- tastes (food likes and dislikes, for example),
- habits (or even addictions),
- emotions (such as anger and disgust) often associated with visceral reactions (such as nausea or an elevated heart rate),
- social norms (for example, those that induce people to prefer to be honest or fair), and
- psychological tendencies (for aggression, extroversion, and the like).

The difference between preferences and beliefs is simple. A preference says: I like the outcome X more than the outcome Y. A belief says: I believe I can get X to happen if I do some action Q.

Self-regarding and other-regarding preferences

A feature of the preferences, beliefs, and constraints approach is that it allows us to model choices based on the entire range of preferences whether they be entirely self-regarding, caring for others (wishing them well or wishing to harm them), or reflecting religious commitments.

OTHER-REGARDING PREFERENCES A person with other-regarding preferences when evaluating the outcomes of her actions takes into account the effects of her actions on the outcomes experienced by others as well as the outcomes she will experience.

A key distinction about our preferences is whether in evaluating the results that we believe our actions will bring about (the right hand part of Figure 2.2) we think about the results that we ourselves experience only, or do we also consider the results that are experienced by others. This gives us two categories of preferences:

- If we think only about the results experienced by ourselves, we have **self-regarding preferences**
- If we *also* think about the results experienced by others, then we have **other-regarding preferences.**

Is the the same thing as "selfish" and "unselfish" preferences? No.

Abraham Lincoln is said to have remarked: “When I do good, I feel good. When I do bad, I feel bad. That is my religion.” Does this mean that Lincoln’s "good" acts were in fact self-regarding because they made him feel "good?" That does not follow. He had other-regarding preferences leading him to act differently than if he cared only about the outcomes that he personally experienced. In the preferences, beliefs and constraints model *all* actions are motivated by preferences, so doing a preferred thing cannot be termed "selfish" without making all behavior selfish by definition. That is why we use the term self-regarding rather than "self-interested" or "selfish." For example, if you (like Lincoln) enjoy helping others, and you act on these preferences, does this mean you are selfish (because, for example that’s what gives you a sense of leading a good life). No, it does not. You are acting on your preferences, but they are other-regarding because you enjoy trying to make the results that others experience be what they would want. Of course other-regarding preferences include feelings of altruism towards others, but they also include negative feelings about others, such as envy, spite, racism and homophobia.

SELF-REGARDING PREFERENCES When choosing an action, a self-regarding actor considers only the effect of her actions on the outcomes experienced by the actor, not outcomes experienced by others. A self regarding actor ignores the external effects of her actions on others.

HISTORY In 1977 Amartya Sen wrote "Rational fools" in which he pointed out that the preferences beliefs and constraint approach ignores the importance of promises, what he called commitments. The reason is that the approach seeks to explain behavior entirely on the basis of the actor’s anticipation of what her actions will bring about in the future. Honoring a past commitment – not because she would otherwise feel guilty in the future, but because it is the right thing to do – cannot be modeled in the preferences beliefs and constraints approach.

In sections 2.8, 2.9 and 2.12 we provide some evidence from experiments on the kinds of other-regarding preferences and how common they are across the world.

"Rationality"

The term rationality in economics means acting on the basis of:

- **Complete preferences** This means, that for any pair of possible outcomes that a person’s actions may bring about, A and B, it is the case that the person prefers A to B or B to A or is **indifferent** between the two. Preferences are not complete if there is some other pair, say A and D for which none of the above three comparisons can be made: to the three statements “I prefer A to D," “I prefer D to A" and “I am indifferent between A and D" the person responds "none of the above."
- **Consistent preferences** If an individual with consistent (also called *transitive*) preferences prefers a bundle of goods A to another bundle B, and bundle B to a third bundle, C, they also prefer A to C.

RATIONAL A rational person has complete and consistent (transitive) preferences and can therefore rank all of the outcomes that their actions may bring about (better, worse, equal) in a consistent fashion.

COMPLETE PREFERENCES Complete preferences specify for any pair of possible outcomes that a person’s actions may bring about, A and B, whether A is preferred to B, B is preferred to A or they are equivalent. Using the symbolic notation for preference: $A \succ B$, or $B \succ A$, or $A \sim B$.

INDIFFERENCE When a person is *indifferent* between two outcomes, it is because those outcomes provide them the same payoffs, or the same *expected* payoffs. As a result, a person will not care which of the two (or more) outcomes they obtain between (or among) which they are indifferent.

A person with complete preferences, which requires only that she can rank all pairs of outcomes, might nonetheless violate the consistency assumption. So she could prefer A to B, B to C, and C to A. All that matters for completeness is that she can rank each pair.

In the heading at the start of this section, we put quotation marks around rationality to underline the difference between how economists use the term and how it is generally used, that is to mean "based on reason." In everyday usage, the term "rational" often means something like the intelligent and perhaps even amoral pursuit of one's own interests. But in economics, as you can see from the above definition, it means something entirely different.

- *Rationality does not say anything about what it is that the person values*: A completely generous and ethical person is rational as long as her preferences are consistent and complete.

- *Rationality does not mean being intelligent or well informed*: The beliefs that (along with preferences) determine the choices a person makes need not be true.

CONSISTENT (OR TRANSITIVE) PREFERENCES Preferences are *consistent* (transitive) if whenever an individual prefers a bundle of goods A to another bundle B, and bundle B to a third bundle, C, they also prefer A to C. Using the symbol $A \succ B$ to mean "A is preferred to B" and the symbol $\Rightarrow$ to mean "implies", the condition for consistency can be written as: $A \succ B$ and $B \succ C \Rightarrow A \succ C$.

Moreover, people with incomplete preferences would hardly be called "irrational" in the ordinary meaning of that term, meaning "not logical" or "unreasonable." Ask yourself if your preferences are complete for the following outcomes: express preference or indifference over which of your two dearest friends will be tortured to death. If you were to say "I cannot rank those two outcomes, nor am I indifferent between them" you would not be "rational" by the economic definition, but nobody would think your behavior was bizarre either. We might be more inclined to worry about the person who would be able to make such a ranking.

Checkpoint 2.1: Why beliefs matter

Considering the coordination problems studied in Chapter 1

a. Explain why in the Assurance Game representing planting in Palanpur why the action a farmer takes to bring about the preferred outcome depends on the farmer's belief about what other farmers will do.
b. In the same game explain why the farmer who believes most other farmers will plant late, will also plant late.
c. Explain why Ben's belief about what Aisha will do matters for how he will play in the Disagreement Game.
d. Are there any games you have learned so far in which beliefs about what the other does did not affect the outcome of the game?

2.2 Taking risks: Payoffs and probabilities

Beliefs become especially important in cases where we have to take some action without knowing for sure what the outcome will be. You make many of this kind of choices every day, from the important choices of what to study at university, to more trivial choices like whether to take an umbrella to class. The theory of decision-making in these cases rests on the idea that the evaluation of how good a course of action is depends on

- how much the decision-maker values each of the possible but uncertain outcomes of the action and
- the decision-maker's beliefs about how likely each is.

Here we introduce a basic concept for decision-making with risk – expected payoffs – that will be used throughout the book. In Chapter 13 we return to the topic of risk including preferences about risk taking per se and the value of insurance.

The value of uncertain outcomes: Expected payoffs

There are two possible but uncertain outcomes of the action "take an umbrella to class," namely, "keep dry walking home in the rain" and "carry the umbrella to and from class without even opening it, because it does not rain." The feasible actions of the decision maker are just: take the umbrella or not.

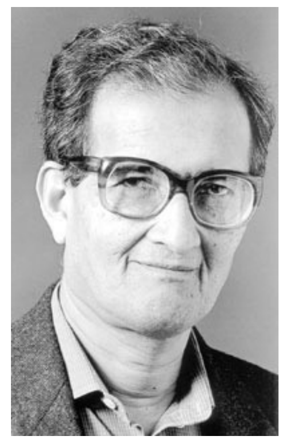


Figure 2.3: Amartya Sen (1933-). Image Credit: National Institutes of Health, Public Domain.

According to the preferences beliefs and constraints approach, the decision maker assigns numbers indicating how much she values each of the possible four outcomes shown in Table 2.2. These numbers give the ranking of the four possible outcomes: [Don't take the umbrella, No rain] is better than [Take the umbrella, Rain] and so on. But if they are to provide a framework for making a decision when you do not know for sure if it is going to rain or not, the numbers have to be more than a ranking. They have to indicate *how much* the actor values each of the possible four outcomes. So for example taking the umbrella when it rains is 5 times better than *not* taking the umbrella when it rains.

We call these numbers the *payoffs* to each of the four possible outcomes.

The likelihood of uncertain outcomes: Beliefs

CONTINGENCY The payoff to the outcome of a decision is said to be *contingent* if something affecting the payoff may or may not happen. The payoff in this case is said to depend on a contingency.

Only one of these two uncertain events will occur. Whether at the end of the day, it turned out to have been a good idea to have brought the umbrella said to be *contingent* on (meaning: depends on) whether it rains or not. The payoff to the two actions in this case is said to depend on a **contingency**. The contingency in this case is whether or not it rains and the payoff to taking the umbrella is contingent on (depends on) its occurrence.

PROBABILITY DISTRIBUTION A probability distribution for n contingent outcomes of a decision is a list of non-negative numbers $\{P_1, P_2, \ldots, P_n\}$ that add up to 1. These probabilities express the decision-maker's degree of belief about the likelihood that each of the of n *contingent outcomes* will occur.

		Uncertain event (contingency) Rain	No Rain
Action	**Take the umbrella**	15	8
	Don't take the umbrella	3	20

Table 2.2: **Two contingencies (rain or don't rain) and two actions (Take the umbrella, or Don't).** The payoffs correspond to the coincidence of an action and a contingency, so Anoushka receives 15 if she plays Take the umbrella when the contingency is *Rain*, and she receives 8 if she plays Take the umbrella and the contingency is *No rain*.

When you decide what to study at university before knowing what kind of work you'll do after, you're making choices about contingencies too: do you go risky and study in drama, or do you go safe and take in accounting? In this case, the contingencies include the the uncertainty about how good you will be at the field you choose and your chance of getting a job in your field.

RISK AND UNCERTAINTYThe term *risk* is conventionally used in economics to describe situations where the probabilities of the possible outcomes are known. The term *uncertainty* describes situations where the decision-maker does not know and cannot learn these probabilities.

The theory of decision-making about risky outcomes assumes that a decision-maker, call her Anoushka, has beliefs about the probabilities (P_i) that each of the contingencies $i = 1, \ldots, n$ will occur. Her beliefs can be based on observation, on empirical studies, guessing, experience, or superstition. They need not be correct.

For simplicity we assume contingencies with just two outcomes (like "it rains" or "it does not rain" above). The basic principles of decision-making are the same no matter how many contingent outcomes there are. In this case, we use the symbol P for the probability the contingency occurs, understanding that $1 - P$ is the probability the contingency does not occur.

The decision rule: Maximize expected payoffs

Often we must take an action prior to the realization of the outcome – you do not know with certainty what will happen, that depend on an uncertain events – called a contingency – that may or may not happen. But you have to make a choice anyway. To take account of the "action now, contingency later" aspect of the decision problem we distinguish between:

- *Expected payoff*: how much the actor values taking the action given her beliefs about the probability that the contingency will occur and
- *Realized payoff*: how much she values the possible outcomes that may happen, that is, after the contingency has been realized ("realized" here means really happened, or actually occurring).

The expected payoff of an action is the basis for her choosing one course of action over another: Anoushka chooses the action with the highest expected payoff. Here is how she can calculate expected payoffs.

HISTORY In 1947 John von Neumann and Oskar Morgenstern showed that how much we value some action that we can take, when the outcome of the action is subject to some risky contingency can be expressed as a weighted sum of how much we value the alternative outcomes of our actions that depend on the realization of the contingency, the weights being the probability of each outcome occurring if we take the action.

For each contingency, i, and each action she can take, x, Anoushka knows the payoff of taking action x conditional on i happening, which we write as $\pi(x|i)$. For example, if i is the contingency of rain in the afternoon, and x is the action of taking her umbrella with her in the morning, then her realized payoff is $\pi(x|i)$ associated with her having the umbrella when it rains. The vertical line

$|$ is read "conditional on", or "given", so $\pi(\text{umbrella}|\text{rain})$ is Anoushka's payoff to having the umbrella conditional on, or given, rain in the afternoon.

For a contingency with two outcomes – numbered 1 and 2 – we have to consider only two payoffs and the corresponding probabilities of each, $((\pi(x|1), P), (\pi(x|2), 1-P))$. For example, if Anoushka's payoffs for the four possible outcomes of her actions are as in Table 2.2, and the probability of rain in the afternoon as 0.6, her system of contingent payoffs for taking the umbrella is $(15, 0.6), (8, 0.4))$. These numbers can be interpreted as follows: since there is a 60% chance of rain, Anoushka has a 60% chance of receiving a payoff of 15 if she takes the umbrella. Further, this means that there is a 40% chance of no rain, therefore, if Anoushka takes the umbrella she has a 40% chance of having a 8 payoff.

The *expected payoff* to an action x given a system of contingent payoffs is the weighted average of the payoffs for each contingency where the weights are the the probability of each contingency being realized if the action x is taken. We abbreviate the expected payoff to choosing x given the probabilities (P_x) of contingencies 1 and 2 being realized as $E(\pi_x, P_x) = E((\pi(x|1), P), (\pi(x|2), 1-P))$

$$\text{Expected Payoff} \qquad E(\pi_x, P_x) = P\pi(x|1) + (1-P)\pi(x|2) \tag{2.1}$$

Equation 2.1 expresses the fact that the greater the probability of an outcome, the greater its weight in the weighted average calculated by the expected payoff. For example, Anoushka's expected payoff to taking the umbrella, assuming the payoffs and probabilities above, would be $0.6 \cdot (15) + 0.4 \cdot (8) = 9 + 3.2 = 12.2$.

Calculating expected payoffs with probabilities is essential to understanding strategic interactions, such as the games we introduced in Chapter 1. But in games – that is strategic interactions with other people – the contingencies are the strategies chosen by the other player, not something like whether it rains.

Checkpoint 2.2: Basis of probability assessments

a. Imagine that you are rolling two six-sided dice with sides corresponding to one of each of the numbers 1, 2, 3, 4, 5, and 6. You calculate the sum each time you roll the two dice simultaneously, for example, 1 + 2 = 3. Explain why the probability of getting a total of 7 from rolling the two dice is 1/6.
b. What is the expected payoff if you get paid $5 for rolling a sum of 6 or 8 on a roll of the two dice and $0 otherwise?
c. Go back to Table 2.2, what would Anoushka's expected payoff to *not* taking the umbrella be given the probability of rain being $P = 0.6$?

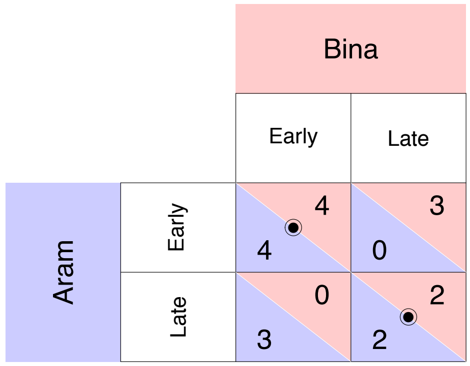


Figure 2.4: **Planting in Palanpur: An Assurance Game.** Aram's payoffs are listed in the blue bottom-left corner. Bina's payoffs are listed in the pink top-right corner. Aram's best response to Bina's choice of strategy is indicated by a black dot in the relevant cell, while Bina's best responses are indicated by hollow circles. Using the dot and circle method you can confirm that the the Nash equilibria of the game (any cells with both a dot and a circle) are (Plant Early, Plant Early) and (Plant Late, Plant Late), with payoffs (4, 4) and (2, 2). The Plant Early Nash equilibrium is Pareto-efficient. The Plant Late equilibrium is not.

2.3 *Expected payoffs and the persistence of poverty*

In games like the Prisoners' Dilemma which have a dominant strategy equilibrium, the action that will maximize your payoffs does not depend on what the other player does, so it does not matter that you do not know what the other will do.

REMINDER A *dominant strategy equilibrium* is a strategy profile in which all players play a dominant strategy.

But if – like most games – there is not a dominant strategy equilibrium, then your best response depends on what the others do, and we need to take account of this in our decision making rule. We can use expected payoffs to understand the choice of which strategy to play in an Assurance Game, like a farmer's choice between Planting Early or Planting Late in the Planting in Palanpur game. The game is shown in Figure 2.4 to remind you of the game's structure. The payoffs in each cell indicate how much the farmer values outcome resulting from the strategy profile given by the particular row and column.

As you know, the game has two Nash equilibria: (Early, Early) and (Late, Late). Recall that (Early, Early) is Pareto-superior to (Late, Late).

The Plant Early equilibrium is also the **payoff-dominant equilibrium**. An equilibrium is payoff dominant when no other equilibrium exists that is Pareto-superior to it. The Pareto-efficient Nash equilibrium in an assurance game is payoff-dominant. In our example, Plant Early is payoff-dominant because the payoffs in this equilibrium exceed the payoffs for both players in the Plant Late equilibrium.

PAYOFF-DOMINANT EQUILIBRIUM An equilibrium is payoff dominant when no other equilibrium exists that is Pareto-superior to it. The Pareto-efficient Nash equilibrium in an assurance game is payoff-dominant.

As we observed in Chapter 1, Palanpur farmers plant late even though a Pareto-superior alternative exists. To see why this occurs, think of what the other player will do as a contingency. We can then say that the degree of belief that other farmers will plant early can be expressed as a probability, P.

A farmer believing with probability P that the other farmer will plant early and probability $(1 - P)$ that the farmer will plant late is an example of decision-

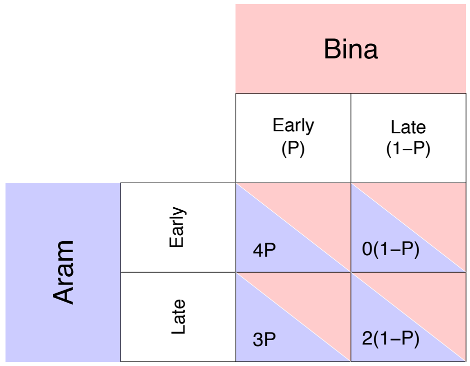


Figure 2.5: **Aram's view of Planting in Palanpur**. The figure shows Aram's payoffs only (and not the payoffs to any other farmers) and his belief about the probability they will occur. Aram's payoffs are multiplied by the probability of that cell occurring based on the probability that Bina plays that strategy. The other farmers play Plant Early with probability P and Plant Late with probability $1-P$. We can calculate Aram's expected payoffs to each of his strategies by adding the payoffs to a given strategy against each of the other farmers' strategies. Plant Early: $\hat{\pi}_{Early} = 4 \cdot P + 0 \cdot (1-P)$ Plant Late: $\hat{\pi}_{Late} = 3 \cdot P + 2 \cdot (1-P)$.

making under risk, since the farmer assigns probabilities to a contingency, in this case, the other farmer's behavior. We do not explore where these beliefs about probabilities come from, but we can imagine that he farmer will form beliefs based on what other farmers tell him or on the basis of their behavior in past planting seasons. We will include just Aram and Bina in the game, but remember we use only two players to simplify our analysis of what is really a much larger population of many people like Aram and Bina.

M-CHECK We label expected payoffs as $\hat{pi}$, which is sometimes written $E[\pi]$.

If Aram believes that the probability of Bina planting early is P we can construct his expected payoffs to each of his strategies, each part of which is shown by Figure 2.6.

INDIFFERENCE PROBABILITY In a two-by-two game, let P be the probability that player A attributes to B playing one strategy and $1-P$ the probability A attributes to B playing the other strategy. Then P_i is the value of P such that such that player A's expected payoffs to to playing each of her two two strategies are equal. In this case player A is therefore indifferent between playing the two strategies (which is why we use the letter i subscript.).

We use a "hat" on a variable to mean 'expected,' so $\hat{\pi}$ reads "π hat." Using these probabilities, Aram's expected payoff ($E(\pi) or \hat{\pi}$) to playing Plant Early is:

$$\begin{aligned} \hat{\pi} = \hat{\pi}(\text{plant early})) &= P\pi(\text{plant early}|\text{others plant early}) \\ &\quad + (1-P)\pi(\text{plant early}|\text{others plant late}) \end{aligned}$$

Aram's expected payoff to planting *late* is:

$$\begin{aligned} \hat{\pi}(\text{plant late}) &= P\pi(\text{plant late}|\text{others plant early}) \\ &\quad + (1-P)\pi(\text{plant late}|\text{others plant late}) \end{aligned}$$

An expected payoff-maximizing farmer will choose to plant early or late depending on which expected payoff is higher. As Figure 2.6 shows, for Aram, which action this will be depends on the probability that he thinks Bina will plant early. The vertical axis is the **expected payoff** to each strategy: Plant Early or Plant Late. The horizontal axis is the probability, P, that Bina plants early: from left to right P goes from $P = 0$ (the Bina plants late with certainty) to $P = 1$ (the Bina plants early with certainty).

The two upward-sloping lines plot the expected payoffs to the two strategies, Plant Early and Plant Late, for Aram the farmer making a decision and how

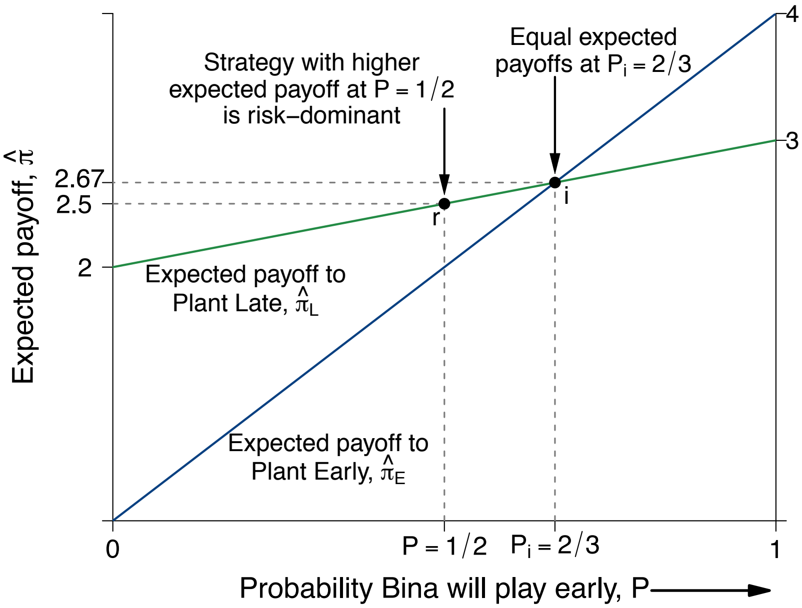


Figure 2.6: **Aram's expected payoffs to planting early or late depend on his belief about the probability that Bina will plant early.** Aram evaluated the expected payoffs to his strategies based on the probability that Bina will play Early. The indifference probability where the two strategies have the same expected payoff is $P_i = 2/3$, and the payoff to Planting Late is greater than the payoff to Planting Early for $P = 1/2$. The intercepts of the vertical axes are the payoffs in the payoff matrix for the planting game in Chapter 1 (Table 2.4).

they depend on the probability that he believes Bina will plant early (that is, for each value of P).

The blue line graphs the equation for the payoff to the strategy Plant Early which is $\hat{\pi}_{Early}(P) = P \cdot 4 + (1 - P) \cdot 0 = 4P$. When the probability the other farmer plants early is zero, i.e. $P = 0$, the payoff to Plant Early is zero. When the probability the other farmer will Plant Early is 1, i.e. $P = 1$, the payoff to Plant Early is 4.

We can draw the expected payoff line for Plant Late in the same way, where the expected payoff to Plant Late is $\hat{\pi}_{Late}(P) = P \cdot 3 + (1 - P) \cdot 2 = 2 + P$ depicted in green, and where $\pi_{Late}(P = 0) = 2$ and $\pi_{Late}(P = 1) = 3$. We can then interpret the expected payoffs as follows:

- Plant Late provides a higher expected payoff for all $P < 2/3$.
- Plant Early provides a higher expected payoff when $P > 2/3$.
- The expected payoffs to the strategies are equal at the indifference probability $P_i = \frac{2}{3}$ (where a farmer is indifferent between Plant Early and Plant Late).

The result is that Aram will choose Plant Late as long as he believes that the probability that Bina will Plant Early is less than two-thirds. Bina, facing the identical situation, has the same decision rule: Plant Late unless you think that Aram is going to Plant Early with a probability of at least two-thirds.

They will remain poor even though, had they somehow started of both planting early, they would have been much better off. The poverty trap in which they find themselves is not the result of rudimentary technology or infertile soil. What they lack is the "social technology" that would allow them to coordinate on the Pareto-superior strategy profile, planting early. Their poverty is due to the rules of the game, which make coordination difficult.

In this example we have assumed that both Aram and Bina had some idea (maybe a guess) of the likelihood that the other would plant early. They faced risk (they had some information on the probability of the contingent event), but not uncertainty (no information at all). Decision making under uncertainty is especially important in the field of climate change, where there are some contingencies for which there is no way to assign probabilities of their occurrence.

M-Note 2.1: Expected Payoffs for Planting in Palanpur

To understand the expected payoffs and the indifference probability in the game, we need to answer the following questions:

a. What are the **expected payoffs** to planting early and planting late (using payoff Table 2.4) for a farmer in Palanpur who believes the probability of Bina planting early is P, with $0 < P < 1$?
b. What value of P leads to an **equal** expected payoff to planting early and late?

We can work out these answers using the following steps:

- If Aram plants early and Bina plants early, Aram's payoff is 4 (with probability P)
- If Aram plants early and Bina plants late, Aram's payoff is 0 (with probability $1 - P$)
- If Aram plants late and Bina plants early, Aram's payoff is 3 (with probability P)
- If Aram plants late and Bina plants late, Aram's payoff is 2 (with probability $1 - P$)

The expected payoff to planting early is the weighted average of the two contingencies (Bina plays Plant Early or Plant Late) with the weights equal to the probability (P) of Bina playing Plant Early and ($1 - P$) of Bina playing Plant Late:

- Expected payoff for planting early $\hat{\pi}_{Early}(P) = 4 \cdot P + (0) \cdot (1 - P) = 4P$.
- Expected payoff for planting late: $\hat{\pi}_{Late}(P) = 3 \cdot P + 2 \cdot (1 - P) = 2 + P$.
- To find the indifference probability at which expected payoffs are equal: $\hat{\pi}_{Early}(P) = \hat{\pi}_{Late}(P)$
- Which is the condition: $P = P_i = \frac{2}{3}$, the *indifference probability*.

2.4 Decision-making under uncertainty: Risk-dominance

What is the farmer facing uncertainty to do? Economics does not have a very good answer.

HISTORY The "principle of insufficient reason" due to the Swiss mathematician Jakob Bernoulli (1655-1705) states that if we have no information on which to estimate the probability that one of two contingencies will occur, we should consider them to be equally likely. Not everyone finds this satisfactory. John Maynard Keynes found it "paradoxical and even contradictory."[4]

A two-person risk-dominant equilibrium

Economists often use what is called the "principle of insufficient reason" when a player has no information on which to place a probability on some contingency. This principle holds that the farmer who has no information on likely strategy choice of his neighbor will assign equal probability to the two events and hence use the probability $P = \frac{1}{2}$ that the other will plant early. What is termed the **risk dominant strategy** is that which yields the highest expected payoff when a player attributes equal probability to the two actions of the other player.

RISK DOMINANT STRATEGY The strategy that yields the highest expected payoff when the player attributes equal probability to the two actions of the other player.

Using this definition in the Planting in Palanpur game, a farmer who assigns the probability $P = \frac{1}{2}$ to the outcome that the other farmer will Plant Early will himself Plant Late because his expected payoffs are $2 = \frac{1}{2} \cdot 4$ to Plant Early, and $2.5 = \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 2$ to Plant Late.

Thus Plant Late is the **risk-dominant** strategy, that is, the strategy that maximizes the farmer's expected payoffs when $P = 1/2$. You can confirm this by going back to Figure 2.6: at $P = 1/2$ the green line (expected payoff to planting late) is above the blue line (expected payoff to planting early). Because this is true for the other farmer as well, both farmers Planting Late is the *risk-dominant equilibrium*.

Planting late in the Planting in Palanpur game is risk dominant because planting early when the other plants late is much worse (you get zero rather than the payoff of two you would have received had you also planted late) than planting late when the other plants early (you get three rather than the four you would have received had you also planted early).

Checkpoint 2.3: Risk dominance and the worst case outcome

a. Redraw the expected payoff line for planting early with the payoff to planting early when the other plants late to be even worse than shown in the figure, e.g. -2 instead of 0.
b. In this case what is the indifference probability?
c. What is the least payoff to planting early when the other plants late, that would make planting late no longer risk dominant?

A risk dominant equilibrium in a large population

Instead of thinking about only two farmers, we can interpret as portraying a population of farmers in a village like Palanpur itself, who all face the same set of incentives for planting early and late. Like Aram and Bina, all the farmers face a coordination problem: doing well if they all plant early and doing poorly if they all plant late. How well each farmer does depends on what the others do, so if a minority of farmers plants early while the majority plants late, then those who planted early while others planted late will have their seeds eaten while the others will get an adequate harvest. The farmers are therefore involved in a many-player coordination problem.

We can re-purpose Figure 2.6 such that the horizontal axis is the fraction

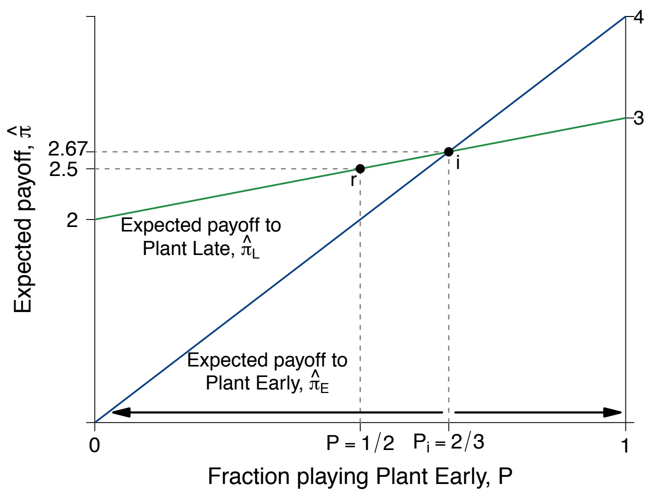


Figure 2.7: **Fraction of farmers planting early.** P is the fraction of farmers playing Plant Early. $1 - P$ is the fraction of farmers playing Plant Late. In the case of the population as a whole, the indifference probability (or in the case of a population of players, the indifference fraction) shown at point **i** with fraction P_i corresponds to the fraction of the population at which the players are indifferent between the strategies (Plant Early or Plant Late). In the case of the whole population, point **i** is also the *tipping point*: when a fraction of the population less than P_i plays Plant Early all farmers will want to play Plant Late; when a fraction of the population greater than P_i plays Plant Early all of the farmers will want to play Plant Early.

of the population going from 0 to 1 who choose Plant Early, P (reading left to right) as shown in Figure 2.7. Reading the horizontal axis from right to left, it measures the fraction of the population who Plant Late $(1 - P)$. The payoff lines in the figure have the same interpretation as before: They are the expected payoffs for any one of the large number of identical farmers in the village. The probabilities translate to population fractions too:

- $P < \frac{2}{3}$: When less than two-thirds of the population choose Plant Early (i.e. more than one-third play Plant Late), the Plant Late strategy will get him a higher expected payoff. At any fraction $P < \frac{2}{3}$ *all* farmers will Plant Late and all farmers will end up with a payoff of 2.

- $P > \frac{2}{3}$: When more than two-thirds of the population select Plant Early (i.e. less than one third select Plant Late), the Plant Early strategy has a higher expected payoff. At any fraction $P > \frac{2}{3}$ *all* farmers will Plant Early and all farmers will end up with a payoff of 4.

- $P = \frac{2}{3}$: At two-thirds Planting Early and one-third Planting Late, the expected payoffs are equal. The point at which the expected payoffs are equal is a **tipping point** as a small change will drive all players to adopt one or the other strategy: Plant Early or Plant Late.

TIPPING POINT An intersection of the expected payoffs to strategies shows a tipping point when a small change in population fractions playing a strategy results in a feedback loop driving the game to one of the extremes, either $P = 0$ or $P = 1$. We describe it as a *tipping point* since a small push either way will "tip" the outcome to one extreme equilibrium or the other.

Now imagine that as in the village of Palanpur virtually all of the farmers have been planting late year after year (maybe even generation after generation). There would not be much uncertainty about what fraction of the population would plant late the next planting season. Each of the farmers would hold the belief that P is close to zero and as a result they all would plant late, confirming their beliefs. The belief that almost nobody would plant early sustains both the low income of the farmers, and perpetuates the belief itself, which year after year turns out to be correct.

In the Fishermen's Dilemma the best outcome for one of the players is the worst for the other, so there is a conflict of interest between the two. And this contributes to the difficulty of finding some way of coordinating so as to avoid over-exploitation of the fishing stock. This is not the problem in the Assurance Game. There is no conflict of interest: All of the Palanpur farmers prefer the outcome when they all Plant Early to any other outcome. Their failure to implement the mutually desired outcome is the result of their inability to coordinate on planting early, for example when all are planting late to make a joint decision to all change to planting early.

What may seem to be a minor tweak to the rules of the game under which the farmers are interacting can help them escape their poverty trap.

2.5 Sequential games: When order of play matters

When we looked games involving risk and uncertainty, we saw that players could end up selecting risk-dominant strategies that implement Pareto-inefficient Nash equilibria. But the game we introduced to model the coordination problem facing Aram and Bina was unlike many real world social interactions, they were total strangers who had no way of coordinating their actions, and they acted simultaneously (or at least, without knowledge of what the other had done.)

But it might be that rather than playing simultaneously, they play sequentially. Playing sequentially is a change in the rules of the game; it represents a change in the institutions governing their interaction. We will see that this seemingly small change makes it into an entirely different kind of game possibly even allowing Pareto-efficient outcomes.

BACKWARD INDUCTION Backward induction is a procedure by which a player in a sequential game chooses a strategy at one step of the game by anticipating the strategies that will be chosen by other players in subsequent steps in response to her choice. (Induction here means *causation*.)

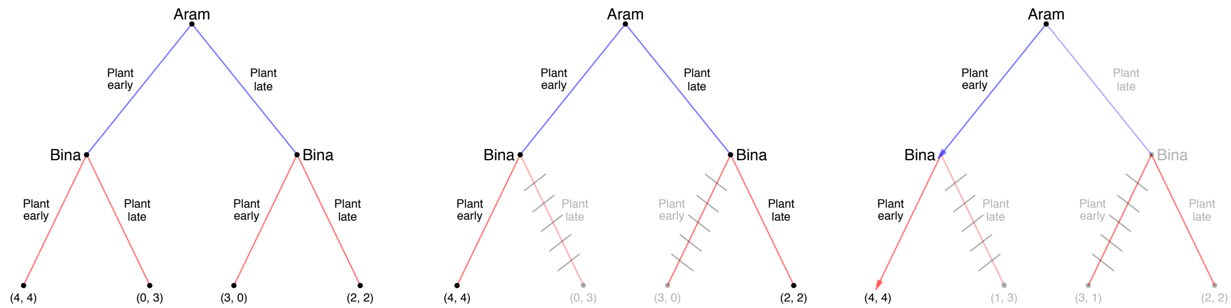
To see how this could work, suppose the Planting in Palanpur game the Aassurance Game) is now sequential. Aram moves first (he is called the **first mover**) and Bina moves second. How will Aram reason?

He has to think about what Bina will do in response to his planting early or late. He knows that:

- Bina's best response to his planting late is to plant late and the best response to his planing early is to plant early, and
- his payoff is greater if they both plant early.

So he will announce that he will plant early, and Bina will respond with planting early. Rather than being stuck planting late with a small harvest, they have now solved their coordination failure. How did they manage it?

The answer is that the sequential nature of the game gave them a way of acting together even if they had no way of actually coordinating. By looking ahead to what Bina would do later Aram he was able to act so that they would



(a) Full Game Tree

(b) Aram's Reduced Choice of Actions

(c) Fully Solved Game

Figure 2.8: **Game tree of the Planting in Palanpur (assurance) game**. 2.8 a presents the full game tree for both players. 2.8 b shows the reduced set of action that Aram considers while using backward induction to solve the game. 2.8 c shows the solved game tree with the arrows indicating the path to the Nash equilibrium (Plant Early, Plant Early). Aram's actions are shown by the blue branches and Bina's by the red branches. Aram's actions are reduced because he has projected forward in time and used backward induction to work out what Bina will do: planting early if he plants early and planting late if he plants late, therefore reducing Aram's choices to a payoff of 4 if he plants early and a payoff of 2 if he plants late. So backward induction leads to the Nash equilibrium of the game being (Plant Early, Plant Early) with payoffs (4,4)

together implement the single Pareto efficient outcome. What Aram did is called backward induction, which is is a procedure by which a player in a sequential game chooses a strategy at one step of the game by anticipating the strategies that will be chosen by other players in subsequent steps in response to her choice.

To see how this works when order of play matters (and where backward induction gets its name), instead of using a payoff matrix (or as we have done for normal form simultaneous move games) we will use a *game tree* and refer to the game as being an **extensive form game**. *Game Trees* have the same basic structure as a normal form games represented by a payoff matrix – they show the strategy set and the payoffs associated with each strategy profile – except that the tree-like structure tells us something about who moves when; and a strategy profile is now a path through the branches of the tree.

In the game tree structure, the players move in sequence, with the player on the top of the tree moving first and the player at the bottom moving last. A game tree for the sequential version of the Planting in Palanpur game is shown in Figure 2.8. Aram is the first player, so he is at the top of the game tree. Bina is the second player, so she is shown acting *after* Aram. Each player's action – planting early or planting late – is shown alongside a *branch* of the tree to indicate which action the player chooses as they move along that branch. Each player's payoffs are shown at the *end* of a branch of the game tree that indicates a specific path to that end point, Aram planting early, then Bina planting early; Aram planting early, then Bina planting late; and so on. The payoffs correspond to those we used in Chapter 1. Because of the branching tree-like structure of the figure there is only one path to each of the end points.

EXTENSIVE FORM GAME A game portrayed by a game tree in which the sequence of actions by the players is made explicit. The player at the top of the tree moves first, with subsequent players moving in sequence after the first player. Payoffs are shown at the end of the game tree in player order, e.g. (Player A's Payoff, Player B's Payoff). Refer to Chapter 1 for the definition of normal form games.

In Figure 2.8 on the left-hand side we have the full game tree, showing all the potential payoffs for the game. Bina is the second-mover and she needs to decide what to do at each point where she could move. If Aram plants early, Bina can get a payoff of 4 for planting early, or a payoff of 3 for planting late. So if Bina is self-interested, then she will plant early when Aram plants early

$(4 > 3)$.

Bina also has to make a choice between her actions if Aram plants late. Bina can get a payoff of 0 if she plants early given Aram planting late or a payoff of 2 if she plants late given that Aram plants late $(2 > 0)$. So if Bina is self-interested, she will plant late when Aram plants late.

We now know what Bina will do, but what will Aram choose to do knowing this? Using backward induction and having a belief that Bina is self-interested, Aram will have a choice between a reduced set of payoffs, shown in the central panel of Figure 2.8: either 4 if he plants early or 2 if he plants late. So if he is self-interested, then he will choose to plant early. As a result, the only Nash equilibrium of the game is (Plant Early, Plant Early) with payoffs (4, 4).

EXAMPLE The timing of a sequential game does not depend on the exact actions being taken in that sequence, but can depend on *commitments* to those actions being taken. For example, in the second quarter of the year a company could commit to the pricing strategy it will follow in the third quarter and if they believed the company's commitment, then other companies have to respond to that commitment, even if it's not the third quarter yet. Similarly, a professor commits to a policy in her syllabus even if her students haven't written a midterm exam or solved a problem set yet. The student must respond to the professor's committed actions and work out what he will do in response to her commitment. To design the syllabus the professor using backward induction thought through what a student would do in response to her commitments in the syllabus.

Checkpoint 2.4: Back in Palanpur

Making the game sequential solved the problem for Aram and Bina. But would that work for the couple of hundred families in Palanpur? Suppose some order of play was determined and that the first family had announced that they would plant early. Would the second family then follow? And the third? Explain why or why not?

2.6 First-mover advantage in a sequential game

Being first mover did not give Aram any particular advantage in the Planting in Palanpur Game, it just allowed him and Bina to coordinate on the Pareto-efficient Nash equilibrium. The result would have been the same had Bina been first mover.

But sometimes it is advantageous for a player to move first; this person then has what is called **first-mover advantage.**

FIRST-MOVER ADVANTAGEA player who can commit to a strategy in a game before other players have acted is a first mover. This limits the outcome of the game to a strategy profile made up of his chosen strategy and to the other players' best responses to it, which may result in higher payoffs for the first mover. This is called first-mover advantage.

Think about the Disagreement Game from Chapter 1. Recall that two players, Aisha and Ben, have a disagreement over which (or perhaps both) of them should study to improve the language spoken by the other. Both prefer when they are good at speaking *some* language. But, Aisha prefers that it be Swahili and Ben prefers that it be English. What happens in this game when Aisha is the first mover rather than when they both move simultaneously?

Considering the game tree in Figure 2.9, we can solve the game by backward induction and see that the Nash equilibrium of the game is (Stick to Swahili (for Aisha), Improve Swahili (for Ben)) with payoffs (4, 2). The outcome (Improve English, Stick to English) which was one of two Nash equilibria in the simultaneous version of the game is no longer a solution in the sequential version of the game. Aisha does better as a *first-mover* because she obtains her

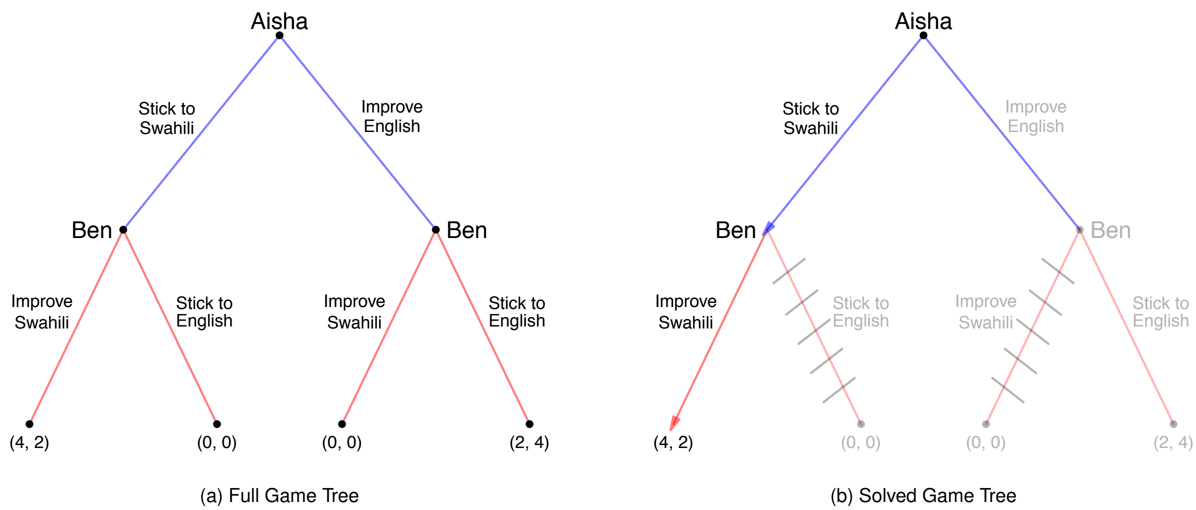


Figure 2.9: **Game tree of the Language (Disagreement) game**. The left-hand side presents the full game tree for both players. The right-hand side figure shows the solved game tree with the arrows indicating the path to the Nash equilibrium (Stick to Swahili, Improve Swahili). Aisha's actions are shown by the blue branches and Ben's by the red branches. Aisha's actions are reduced because she has projected forward in time and used backward induction to work out what Ben will do: Stick to Swahili if he plays Improve Swahili and Improve English if he plays Stick to English, therefore reducing Aisha's choices to a payoff of 4 if she plays Stick to Swahili and a payoff of 2 if she plays Improve English. So backward induction leads to the Nash equilibrium of the game being (Stick to Swahili, Improve Swahili) with payoffs (4,2). The outcome favors Aisha over Ben and therefore conveys a first-mover advantage.

preferred outcome. Ben would have benefited in the same way had he been first mover.

The reason why being first mover gave Aisha an advantage is that the normal form game has two Nash equilibria – one preferred by Aisha and the other by Ben. In the sequential game the first mover determines which of the two Nash equilibria will occur. Once Aisha has moved and has established that she will Stick to Swahili (and *not* try to Improve English), Ben needs to choose his action. Ben needs to take Aisha's move as *given*. He must therefore choose his best response to Aisha choosing Stick to Swahili. Given that he would like to communicate with Aisha, his best response is to Improve Swahili.

We have not asked why Aisha rather than Ben was first mover in the Disagreement Game and why it was Aram rather than Bina in Planting in Palanpur. We showed only that differences in a person's position in the game gave them advantages. In this case Aram and Aisha – the first movers – had strategy sets that gave them the capacity to commit to a strategy in advance. Aram's first mover status did not allow him to benefit at Bina's expense; but this was not the case with Aisha, her first mover status gave her an advantage over Ben.

First movers in a modern economy are more like Aisha:

- *Employers*: they commit to the wage, job requirements and working conditions; workers – actual and prospective – best respond to that.
- *Banks and other lenders*: they set to the interest rate, repayment schedule and other aspects of a loan contract. Borrowers and would be borrowers best respond to that.
- *Owners of major companies*: in the U.S. Walmart, Amazon, Apple – com-

mit to prices and delivery schedules. Consumers best respond.

The fact that people occupy different positions in our economy – Employers and workers, lender and borrowers – interacting under rules of the game that give some first mover status and other special advantages is an important part of the explanation of inequality of wealth and income as we will see in Chapters11,12, 13, and 15.

Checkpoint 2.5: Ben has the first-mover advantage

a. Consider the sequential Disagreement Game shown in Figure 2.9. Re-draw the game tree, but with Ben as the first mover rather than Aisha. Show that (Improve English, Stick to English is the Nash equilibrium of the game.
b. Assuming that the payoffs in the Disagreement Game are in hundreds of dollars and that you are Ben, how much would you pay for the privilege of being first mover a) if otherwise Aisha would be first mover, and b) if the game were to be played simultaneously (so that there is no first mover)?

2.7 Social preferences: Blame Economic man?

The characters in our economics episodes – Aisha and Ben, Aram and Bina – care exclusively about their own payoffs. For them a "best response" is simply a "best-for-me response." Is that why they have had difficulty overcoming the coordination failures they face? The answer, we will see, is that being concerned about how your actions affect others will help to address coordination failures, but will not be sufficient.

Homo economicus or "economic man" is the term economists have used to designate an entirely self-regarding and amoral actor, a person who is not motivated by either a concern for others, or a desire to conform to any ethical principles. The term is often put in italics to parallel the biological terminology for a species (like *Homo sapiens*). *Homo economicus*, however is a fictional character or ideal type representing one possible variety of human behavior.

HISTORY The idea of basing economics on the assumption that people are entirely self-regarding –"solely as a being who desires to possess wealth" goes back to the last of the great classical economists, John Stuart Mill author of Principles of Political Economy (1848), considered to be the first economics textbook in the English language. He considered this view of people to be "an arbitrary definition of man."[5]

Models based on *Homo economicus* have provided predictions about behavior that are borne out by empirical studies that range from how American windshield installers and Tunisian sharecroppers respond to different work incentives to the effect of taxes on cigarette consumption. But, as we shall see, Homo economicus is not an accurate depiction of how people behave:

- People volunteer for fire fighting, delivering food to the sick during a pandemic, and other dangerous but socially beneficial tasks, and contribute substantial sums to charity.
- People participate in joint activities such as strikes or protests even knowing that their individual participation is unlikely to affect the success of

the event and that, if successful, the benefits would be widely shared, not confined just to those people participating in the protest.

- People donate blood to strangers.
- In public opinion polls and in voting, people support taxes that transfer incomes to the poor even when they are sufficiently rich and unlikely ever to benefit directly from these policies.

Motivated by these and similar observations a augmented by controlled experiments about human behavior (that we will review below), economists have revised our assumptions to recognize that people are capable of ethical, generous, and other motivations as well as self-regarding motives. This is important because as you learned in the first chapter, coordination failures occur because we fail to take adequate account of the effect that our actions have on others. Our concern for others can help to internalize these external effects whether it be our willingness to curb our carbon footprint or willingness to protest for causes whose benefits would be widely shared.

REMINDER Remember that in Chapter 1 we saw how 'internalizing the external effects' means getting people to pay for the external costs they imposed on others and this resulted in the fishermen choosing to cooperate and fish less in the Fishermen's Dilemma.

But coordination failures cannot be blamed entirely on people seeking to maximize their own payoffs. Think again about the real farmers in Palanpur, all planting late when they could all do better if they all switched to planting early. Suppose one of those farmers was deeply concerned about the poverty of his entire village, and wished to improve living standards for everyone. He could not do this by individually planting early.

Now suppose that every villager shared his concerns for all members of their community. Each one would know that their own decision to plant early would change nothing (except that their seeds would be eaten by the birds). What has captured the people of Palanpur in a poverty trap is not that they care only about their own harvest (they surely care about others'), but their inability to come to a common agreement to plant early. Their poverty stems from a problem of institutions, not motivation.

To understand individual behavior and its social consequences we need an approach that allows for the full range of human motivation.

Checkpoint 2.6: Homo economicus goes to the polls

a. Given that it costs time to cast a vote (going to the voting station, standing in line, and the opportunity cost of your time), do you think a person with *Homo economicus* preferences would vote in most elections? Why or why not?
b. In what circumstances do you think someone would someone with the preferences of a *Homo economicus* vote?

While answering these questions, think about the *beliefs* the person with *Homo economicus* preferences would have about the probability his vote will be important to the outcome of the election.

And so, while *Homo economicus* is among the kinds of actors this approach considers, there are other characters, representing other sides of human behavior such as generosity, fairness, reciprocity and spite.

What these four aspects of behavior have in common is that they are other-regarding: the outcomes that the actor considers in choosing an action include things experienced by others, not just outcomes affecting the actor herself. Here are some common forms of other-regarding preferences:

- Those with *altruistic* preferences, such as basic generosity, are motivated to help others even at a cost to themselves, they place a positive value on the well-being or payoffs of others.
- *Inequality-averse* or fairness-based preferences motivate people to seek to reduce unjust or unfair economic differences even if the actor is herself a beneficiary of these differences.
- A person with *reciprocal* preferences is motivated to help others who have themselves behaved generously or upheld other social norms, and also to punish those who have treated others badly.
- *Spite* and *'us versus them'* distinctions that place a negative value on outcomes experienced by others, often motivate hostility towards members of religious, racial, ethnic and other groups. Therefore a *negative* outcome another person experiences, can result in a *positive* value for someone who feels *spiteful*.

The term **"social preferences"** is used to describe all types of other-regarding preferences.

INEQUALITY AVERSION is a preference for more equal outcomes and a dislike for both *disadvantageous inequality* that occurs when others have more than the actor and and (to a lesser extent typically) *advantageous inequality* that occurs when the actor has more than others

Checkpoint 2.7: Social Preferences & Social Norms

a. Give an example of a preference you have that is *not* self-regarding.
b. Can you think of any social norms that lead you to act in an other-regarding way?
c. Suppose that Aram and Bina (in the Planting in Palanpur Game) were of different religions between which there is hostility, so that each would gain some pleasure from the misfortunes of the other. Can you show how this could change the game so that instead of having the Pareto-efficient mutual early planting as one of its two Nash equilibria, it becomes a prisoners dilemma with planting late as the dominant strategy equilibrium.

2.8 Experiments on economic behavior

Suppose you wanted to know if someone has altruistic preferences. How would you find out? Would you ask her? Well, that could provide some information, but merely asking might not be entirely convincing , because many people would like others to think they are altruistic even when they are not, so they might lie.

What about observing her behavior and comparing her behavior to how others behave? Such observation might be informative, but if we see people behaving differently that could be because the people we observe have different beliefs or different constraints, not different preferences.

Economists use experiments to study preferences because at least ideally this allows us to control for (hold constant) the constraints and beliefs of the individuals to focus on the nature of preferences. Experiments allow economists to implement the ***ceteris paribus*** – all else equal – assumption that we think is so important when we are trying to identify causes and consequences of some change or difference.

CETERIS PARIBUS is a Latin term that means "other things equal." When we held another player's strategy constant in Chapter 1 to find a player's best response we were using the *ceteris paribus* assumption. Similarly, when we use calculus and mathematically hold other variables constant we are employing the *ceteris paribus* assumption.

To understand how common different types of preferences are, and how they affect our behavior, economists use laboratory experiments in which subjects play games designed to elicit the nature of their motivations.

Experiments play a central role in science: they allow predictions made from theories to be tested empirically. This has has been done, for example, with the prediction that players in a Prisoners' Dilemma experiment choose the dominant strategy equilibrium. (You will see what happens in this experiment below.)

FACT CHECK Behavioral experiments are a recent addition to economists' empirical tool kits; but they have been used in psychology for almost a century and a half. The main innovations that economists have made to experimental social science are the use of game theory to clarify the role of beliefs and preferences and nature of incentives and the common use of monetary payoffs.

In some sciences, experimenters can control almost all relevant conditions in their environment, the laboratory. Their subjects can be anything from yeast cells for a biochemist, to fruit flies for a zoologist.

In economics, however, our subjects are people asked to make *decisions* or *choices* in the experiment. It is much harder to control for the various factors that affect human behavior than to control the chemical environment of a colony of yeast cells. Experimental evidence carries little weight unless the experiment can be *replicated*, that is, repeated by different researchers reaching the same results.

REPLICATION When other researchers independently repeat an experiment and reach the same results they have *replicated* the experiment. When an experiment can be replicated we know that its results are *reproducible*. Science is founded on reproducible evidence.

We use a specific vocabulary when we talk about behavioral experiments in economics. The following terms will come up often:

- *Subject/participant*: A *subject* or *participant* is a person who participates in an experiment.
- *Endowment*: The *endowment* is an initial amount of money or tokens later converted to money that subjects receive at the beginning of the experiment, and later make decisions about in the experiment.
- *Incentives*: The fact that players stand to win material rewards in varying degrees depending on how they play the experimental game means that the experiment mimics many real economic interactions..

- *One-shot vs. repeated*: A *one-shot* experiment occurs once and subjects make one decision in the experiment as a whole and are paid for that one decision. A *repeated* experiment involves subjects making repeated decisions often with information about the play of others on previous rounds, sometimes with the same other subjects in a group or sometimes with different subjects.

Here is an example of the importance of using results from experiments to test a theory. Subjects with self-regarding preferences are predicted to defect in a one-shot Prisoners' Dilemma game because defection is the dominant strategy. But in Prisoners' Dilemma experiments, the proportion of players who cooperate rather than defect is commonly between 40 and 60 percent.[6] This means the *prediction* based on the assumption that people are entirely self regarding was borne out for some but far from all of the subjects. The finding therefore provoked some rethinking of the predictions based on the assumption that people are entirely self-regarding.

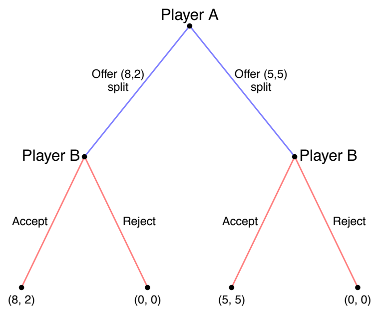
Many subjects prefer the mutual cooperation outcome and are willing to take a chance on the other player also not defecting, rather than the higher material payoff they can obtain by defecting when the other cooperates. When subjects defect, experimental evidence suggests it is because they dislike being taken advantage of, not because defection is the payoff maximizing strategy independently of the other participant's actions.

2.9 The Ultimatum Game: Reciprocity and retribution

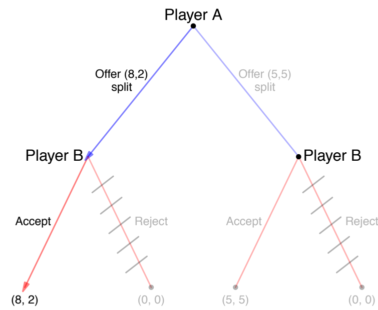
Observing substantial levels of cooperation in the Prisoners' Dilemma game was a shock to the standard *Homo economicus* assumptions. But the experiment that has sparked the perhaps the greatest reconsideration of the *Homo economicus* model is the *Ultimatum Game*.

Here is the game with its basic treatment:

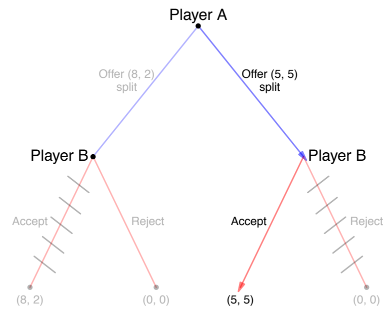
- Subjects are anonymously paired for a "one-shot" interaction with another person.
- The role of "Proposer", is randomly assigned to one of the subjects; the other is then the "Responder".
- The Proposer is given an endowment, the "pie" (e.g. $10), by the experimenters and the Responder knows the size of the pie.
- The Proposer then proposes how to divide the endowment between Proposer and Responder, transferring to the Responder any amount between nothing and the entire endowment, e.g. the Proposer chooses to keep $ 8 and give $2 to the Responder.



(a) Full Game Tree



(b) Self-regarding players



(c) Reciprocal Player B

Figure 2.10: **Game tree of the ultimatum (bargaining) game**. 2.10 a presents the full game tree for both players regardless of type. Player A is the Proposer and their actions are shown by the blue branches. Player B is the Responder and their actions are shown by the red branches.

- If the Responder accepts the proposed division, the Responder gets the proposed portion, and the Proposer keeps the rest.
- If the Responder rejects the offer both get nothing and the game ends.

Figure 2.10 presents a *game tree* for a variant of the Ultimatum Game where the Proposer chooses between two offers: divide the pie equally and each person gets $5 for an outcome (5, 5) or keep $8 and offer the Respondent $2 for an outcome of (8, 2). The Responder then chooses whether to accept or reject the offer. The payoffs to each player are listed in the order of play (Player A, Player B), so (8, 2) means Player A gets 8 and Player B gets 2.

If the Proposer cares only about her monetary payoffs in the game and believes that the Respondent is similarly self-regarding, then the Proposer (Player A) will reason backwards as follows:

- Player A predicts that the Responder (Player B) will accept the offer of $2 (because A believes that B is also self-regarding and because $2 is greater than his fallback of $0 which is what he gets if they reject the offer.
- A will propose the $(8, 2)$ split to maximize A's payoff.
- The Responder (B) will accept.

2.10 b 2.10 shows how the game unfolds if Player A and Player B are both entirely self-regarding and maximize their monetary payoffs. Player B will then always prefers a positive money amount over zero, and so they will never reject a positive offer. Player A knows how Player B will respond, and therefore, has a choice between a payoff of 8 and a payoff of 5; they prefer 8 and so offer a split of (8, 2). So if the two player's are self-regarding then backward induction leads to the Nash equilibrium of the game being (Offer (8, 2) Split, Accept) with payoffs (8, 2).

But A, the Proposer, who has followed through this reasoning now understands that she could have offered B the smallest possible positive amount, say $1, and that the offer would have been accepted. Figure 2.10 c 2.10

shows how the game might be played if Player B is a Reciprocator, that is, Player B cares both about monetary payoffs and also about reciprocating how Player A treats them. In this case, Player B views an offer of (8, 2) as unfair or demonstrating bad intent, and they would rather get a payoff of zero dollars than be treated poorly, so they would *reject*.

If, on the other hand, Player A offers (5, 5) then Player B views that as fair or demonstrating good intent and they would prefer a payoff of 5 in that context to a payoff of 0, so they would *accept*. Player A prefers a payoff of 5 to a payoff of 0 (this is true regardless of whether player A is self-interested or reciprocal) and so the Nash equilibrium of the game is (Offer a (5, 5) Split, Accept) with payoffs (5, 5).

FACT CHECK Did the subjects not understand the game? It is not that complicated a game, and later experiments in which subjects played the game many times with different partners showed this wasn't true. Their behavior remained consistent with the one-shot experiments and their results continued to be reproduced with many people making 50-50 splits (or nearly so) and rejecting low offers.

The Ultimatum Game has been played anonymously for real money in hundreds of experiments with university student subjects and other populations – businessmen, fishermen, farmers, civil servants – in all parts of the world.[7] The prediction based on the assumption that people are entirely self-regarding and believe that others are too invariably fails as a description of how people behave. For example:

- Modal offers – the most common offers in the experiments – are typically half of the pie, and average offers generally exceed 40 percent of the pie, and

- Offers of 20 percent of the pie or less are often rejected with frequencies; people in the position of Responder choose to reject and get zero rather than accept and get a payoff of, say, \$2 offered from the Proposers \$10 pie.

FACT CHECK Some have suggested that the results were due to the relatively low stakes in the game, such as the \$10 mentioned earlier. But subsequent experiments conducted among university students in Indonesia for a 'pie' equal to three months average expenditures replicated the results as did experiments with U.S. students with a 'pie' ranging in size up to \$100. Evidence from France showed similar behavior by proposers with stakes ranging from 40 French francs (\$7.20) to 2000 French francs (\$360) (this was prior to the adoption of the Euro). A further study in India observed stakes that varied by a magnitude of over 1000: From 20 rupees (\$0.41) to 20,000 rupees (\$410) as the stakes. [8]

These rejections of small but positive offers from the Proposer are interpreted as evidence for *reciprocity* motives on the part of the Responder. Why? Because the Responder is willing to pay a price (giving up a positive payoff) to punish the Proposer for making an unfair offer (an offer the Responder considers too low). Responders apparently consider a low offer to be a violation of a norm of fairness, and a person with reciprocal preferences responds by depriving the proposer of any payoffs at all.

Explaining the behavior of Proposers is more complicated. The outcomes of the experiments are not sufficient to say whether the large number of even splits (and other seemingly fair or near-fair offers) is explained by adherence to fairness norms or altruism by the Proposer or to self-regarding preferences informed by fear that the Responder will reject an unfair offer. The evidence for reciprocity motives therefore, comes from the Responders' behaviors, not the Proposers' behaviors.[9]

Figure 2.11: **Small-scale societies where the Ultimatum Game experiments were conducted.** A map of the world showing the locations of the small-scale societies where the Ultimatum Game experiments were conducted. Source: Henrich et al. (2005).

2.10 A global view: Common patterns and cultural differences

Anthropologists and others were surprised that the results of experiments with the Ultimatum Game have been so similar across the many countries in which they have been conducted. One observed that in virtually all of the early experiments the subjects were from WEIRD countries, meaning Western, Educated, Industrialized, Rich, and Democratic.[10] A team of anthropologists and economists (including one of your current authors) designed a series of experiments to explore whether the results reported so far are replicable in societies with quite different cultures and social institutions and whether results differed across the different societies.[11] These societies included hunter-gathers, herders, and farmers (some using modern methods, others not even having cattle, horses, or plows). In their Ultimatum Game experiments the pie was substantial, approximately a day's average wages or other income.

Figure 2.11 shows the location of the 15 small-scale societies around the globe. The team was wondering if they would find cultural differences, and they found them.

Among the Au and Gnau people in Papua New Guinea offers of more than half of the pie were common, and many of these high offers were rejected. In fact Responders among the Au and Gnau peoples were as likely to reject a offer of much more than half as an offer of much less than half.

Though this seemed odd to the economists on the team, it did not surprise the anthropologists who study New Guinea. They know that people in New Guinea compete with each other to see who can give more or better gifts. Gift-giving conveys status in their society and people use giving gifts as a way to obtain status over others. Refusing a gift suggests that you are *not* subordinate to the gift-giver, while accepting it means their status is higher than yours.

By contrast, among the highly individualistic Machiguenga slash and burn farmers in Amazonian Peru, almost three quarters of the offers were a quarter of the pie or less and there was just a single rejection, a pattern strikingly different from other experiments. The Machiguenga came as close to acting like Homo economicus as any population yet studied. Even among the Machiguenga, however, the mean offer was still 27 percent of the pie, more than the zero we'd expect if they all were consistently self-interested.

The researchers who analyzed the experiments in the 15 small scale societies made the following conclusions:

- Although behaviors vary greatly across societies, not a single society approximated the behaviors that would be observed if everyone cared only about their own payoffs and believed others were the same.
- Between-society differences in behavior seem to reflect differences in the kinds of social interaction people experience in everyday life.

Here is some evidence that the experimental game behavior reflected the lived experiences of the people.

- The Ache hunter gatherers in Paraguay share meat and honey equally among all group members. Ache Proposers contributed half of the 'pie' or more.
- Among the Lamalera whale hunters of Indonesia, who hunt in large crews and divide their prey according to strict sharing rules, the average proposal was 58 percent of the pie.

Given the evidence from small-scale societies like the Lamalera and the Ache, we might ask whether we find other-regarding behavior in real-world situations elsewhere in the industrialized world. A different team of researchers were interested in exactly this question and designed an experiment that mirrors a real life dilemma: what would you do if you found a wallet someone had lost: would you return it?

The team distributed a total of 17,303 "lost" wallets some with money in them some without, in 355 cities across 40 countries.[12] In each country, the researchers targeted big cities to ensure that there was a good sample of subjects (and to ensure anonymity). Using transparent wallets with a business

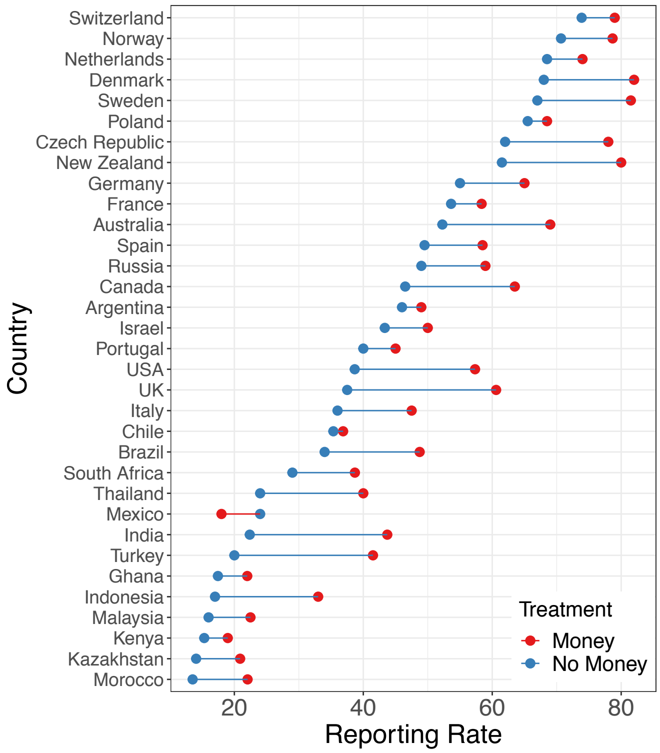


Figure 2.12: **Wallets with details of their owners were more likely to be given back to their owners when they contained money than when they did not.** The "Reporting Rate" is the fraction of wallets that were "returned"

card, grocery list, key and cash, the researchers could check how many people contacted the "owner" of the wallet given in the email address listed on the business card to return the wallet.

Before reading on, ask what you think would happen in your community: how many people would try to return the wallet? Would more people return the wallet if it had money in it, than if it did not?

The results of people's choices are shown in Figure 2.12. Though there were differences across countries, with just one exception among the 33 countries people were more likely to contact the "owner" if the wallet contained money ($13.45, the treatment) in it than if it did not ($0, the control). In a subset of cases – in the US, UK, and Poland – the researchers added a treatment with even more money in the wallet ($94.15). With a really substantial sum of money in the wallet, people were as likely, if not more so, to contact the listed email address on the business card in the wallet.

In interpreting the results keep in mind that the countries differ greatly in how much an additional $13.45 would make to a person standard of living. Per capita income in the richest countries in the sample (Norway for example) are ten and even in some cases 20 times the per capita income in others (Kenya for example), even when account is taken of the differing purchasing power of each national currency at domestic prices.

The evidence from both the Ultimatum Game and the wallet experiments suggests two important take-aways:

- *Culture matters*: people from different parts of the world live by different social norms and mutual expectations – what we can loosely call "culture." People from different cultures differ in what they consider fair offers and whether think it's acceptable to make a self-regarding offer. They also differ substantially in whether they will return a lost wallet.

- *People are similar in many important respects*: people across the world have other regarding motives including altruism, fairness, and reciprocity. In the "lost wallet" experiment in most countries a substantial fraction of people attempted to return the wallet.

The Ultimatum game and the lost wallet experiment provide valuable information, but they are lacking in one respect. Most of the coordination problems we face involve large numbers of people, like the people of the world making decisions about their carbon footprint, or owners of businesses across the entire economy deciding whether or not to invest, or the herders placing more cows on the commons, or the people deciding whether to drive to work, or the farmers of Palanpur deciding when to plant.

2.11 The Public Goods Game: Cooperation and punishment

A **public good** is one which more people can enjoy without reducing the amount available to others, and from which others cannot be excluded from access to the good. An example is global climate: it is experienced by everyone, and efforts to address the problem of climate change contribute to a public good: that is a more sustainable environment. Another example is the rules of calculus: if you learn how to differentiate that does not deprive others of the knowledge of the same rules of differentiation.

This sounds like a good thing. But there is a problem. Why do people produce or contribute to the provision of a public good? If nobody can be excluded from enjoying the good, its hard to see how it would be possible to make money by providing it. (Imagine trying to make a living by selling the rules of calculus!)

We return to the problem of public goods in Chapter 5

It shares with the Prisoner's Dilemma Game the feature that everyone could do better if they agreed on a common course of action (i.e they all contribute) but the dominant strategy for a self-regarding player is not to contribute. For this reason a Public Goods Game is sometimes called an n-person Prisoners' Dilemma because it has the same incentive structure.

Rules of the Public Goods Game

The *Public Goods Game* experiment is designed to understand how people will play in a game with this structure.

Here are the rules of the game:

- n players are each given an endowment of z.
- Each player simultaneously selects an amount $e^i, 0 \leq e^i \leq z$ to contribute to the public good (think of e^i as the player's "effort" in contributing to the public good).
- The amount of the public good produced depends on the level of contributions. For example it could be half of the sum of all of the contributions. In this case the average productivity of contributions would be one-half.
- Each player, regardless of whether they contribute or not, obtain the entire benefit of the total amount of the public good produced. ,

As a result of the rules, each player's payoff can be read as follows:

$$\text{Own payoff} = \text{Endowment} - \text{Contribution} + \text{Average productivity} \times \text{Total Contributions}$$

Figure 2.13 illustrates the benefits of the public good minus the costs of contributing to a public good in a 4-person public goods game. In the version of the game we depict, they can each contribute \$10 or \$0: which we call "Contribute" or "Don't." Now compare how a player does if they Contribute (red line) or Don't (blue line) if they are the only one who contributes, or there are 1, 2, or all 3 others contributing. You can see that in every case she will earn higher payoffs by not contributing. Therefore, if all players are self-regarding, the dominant strategy equilibrium is Pareto-inefficient and an alternative outcome, full contribution by all, which is not a Nash equilibrium, is Pareto-efficient.

A self-regarding player who cares about only their own costs of contributing and the benefits they get from the public good will choose to contribute nothing (analogous to Defection in the Prisoners' Dilemma). And this is the case no matter how much or little the other players contribute. So contributing nothing is the dominant strategy for each player. And, like in the Prisoners' Dilemma, everyone contributing nothing is the dominant strategy equilibrium.

As a result economists expected that when this game is played for real money that no player would contribute. They were in for a surprise.

M-Note 2.2: Why the dominant strategy in the Public Goods Game is to contribute nothing

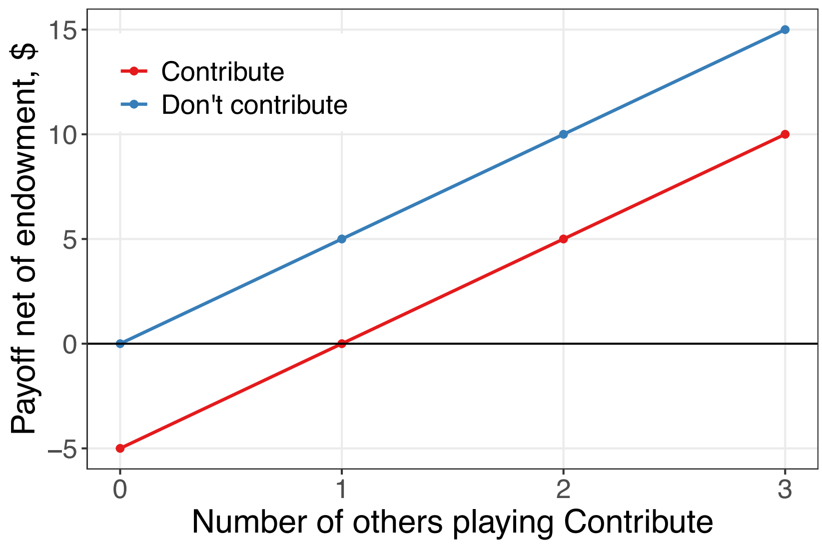


Figure 2.13: **A 4-player public goods game with choices to contribute or not.** Each player can play either Contribute or Don't contribute, and as there are 4 players, this means that the number of others contributing can be any of the numbers 0, 1, 2, or 3 players playing either of the strategies. Playing Don't contribute yields a higher payoff for the player regardless of how many players play Contribute or Don't. Therefore, Don't contribute is a strictly dominant strategy.

In Figure 2.13, there were four players and we limited their actions to either contributing \$10 or contributing \$0, but in a standard public goods game players can contribute any amount up to and including their entire endowment (such that $e^1 = z$). In the public goods game in which players can contribute any amount from their endowment, a player's payoff is given by Equation 2.2:

$$\pi^i = z - e^i + M \sum_j e^j \text{ for } j = 1, \ldots n \tag{2.2}$$

As earlier, we can break down this equation:

- z is the endowment of money the player receives from the experimenters.
- e^i is the contribution a player makes at a cost to themselves.
- M is the multiplier, or the average productivity of contributions < 1.
- $\sum_j e^j$ is the total amount contributed by *all* players.

Whatever the other players do, you can see from Equation 2.2 if you differentiate π with respect to $hate_i$ that for person i contributing one unit (say, penny) more changes the her payoff by $-1 + M$ which is the cost of contributing minus the public good that the contributor herself enjoys as the result of her contribution. So as long as $M < 1$ contributing anything reduces the contributor's payoffs. This is why not contributing is the dominant strategy

Checkpoint 2.8: Two-action Public Goods Game

a. Draw a payoff table with two players, A and B, playing the Public Goods Game. Limit their actions to $e = 10$ and $e = 0$ with $M = 0.5$. Check which is the dominant strategy and explain why. What happens if $M = 0.75$?

b. Revise your payoff table and check what would happen if the strategies were $e = 1$ and $e = 0$ with $M = 0.5$? Would anything change? What happens if $M = 0.75$?

c. Think about the condition $M < 1 < Mn$. Why must this be true for the game

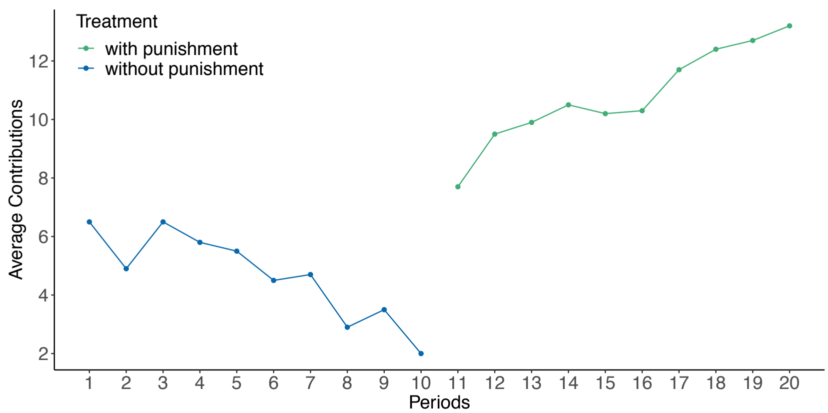


Figure 2.14: **Public Goods Game with Punishment. Average contributions over periods 1 to 10 decrease without punishment. Over periods 11 to 20, subjects can be punished by their peers and average contributions are higher on average than in the first 10 rounds. Source: Fehr and Gächter (2000).**

The vertical axis is the average contribution each round. The horizontal axis is the period. At period 11 the subjects are given the opportunity punish each other. There are three treatments in this Public Goods Game experiment. This figure portrays the behavior in the "Strangers" treatment where players are randomly re-matched each round, but could have some players re-enter the group. The two other treatments, which show similar results, are "Partners" where players are in the same group for all the rounds; and "Perfect Strangers" where players are re-matched, but no player will encounter any other more than once during the experiment.

to be an n-person Prisoners' Dilemma game? (**Hint:** Think about what would happen if it were *not* true. What would happen if $M > 1$? What would happen if $Mn < 1$?)

2.12 Application: Evidence from Public Goods Games

The prediction of the self-regarding model that all players choose to contribute nothing ($e = 0$) is consistently contradicted by the experimental evidence. The evidence we have comes from people playing one-shot (single-period) games and from people playing repeated games with as few as 5 rounds and as many as 50 rounds.[13]

In one-shot games, contributions average about half of the endowment, while in repeated games contributions begin around half and then decline so that a majority of players contribute nothing in the final round of a ten-round game.

Researchers have interpreted the decline in the first half as a reflection of people getting disappointed about the expectations they had that other people would contribute more, along with the desire people have to punish low contributors (or at least not to be taken advantage of) in a situation in which one person can punish a low contributor only by reducing their own contributions.

In this interpretation it is the reciprocity motives of the higher contributing subjects, disappointed or angry about their free-riding fellow subjects that explains why cooperation unravels. So the decline in contributions becomes a vicious circle: only by reducing how much they contribute can people punish others, but in so doing other people might want to punish them for their low contributions by contributing yet less again.

The Public Goods with Punishment game

The idea of that the decline in contributions is due to the fact that in the standard game contributing less is the only way to punish low contributors is supported by an ingenious experiment. This has the same public goods structure but with what turned out to be a major difference: after subjects contributed, their contributions were then made public to all the group members, and members then had the opportunity to punish others in the group. Subjects could punish them by imposing a cost, therefore, reducing the defectors total payoff. In order to impose this cost, however, the Punisher also had to suffer a cost themselves.

The change in the rules of the game – adding the punishment option – represents a change in the institutions governing contributions to the public good. In the language of experiments the new rules are termed a new different treatment. So the standard game is one treatment and the game with punishment is a second treatment.

In the experiment subjects engaged in extensive punishment of low contributors. At the start of the game people contributed over half of the endowment and then, apparently in response to punishment of low contributors, they contributed more over the course of the game. The change in institutions modeled by adding the punishment option altered the result dramatically.

To see if subjects' willingness to punish could be based on the expectation that they would benefit in subsequent rounds of the game, a slightly different experiment was tried. The researchers adopted what they called a “perfect strangers” treatment: after each round of the ten-round experiment the groups were re-shuffled, so that no player ever encountered any other player more than once. The "perfect strangers treatment" turned the experiment into a series of one-shot games.

Since every player would only encounter every other player once, if low-contributors responded to punishment by contributing more in subsequent rounds, they would raise the payoffs of others but not the punisher (who would never again be in the same group with the target of her punishment).

In this way, punishment itself became a public good. This is because a punisher incurs a cost, yet the benefits of punishing a low contributor and getting them to increase their contributions are non-rival and non-excludable. Even in the perfect stranger treatment subjects avidly punished low contributors.

Further evidence comes from the fact that people punish low contributors even in the last round of the game when punishment cannot be motivated by the expectation that the punisher will benefit from their targets improved behavior in the future. There is no future (the game ends after they punish).

So reciprocal preferences – the pleasure of punishing a someone who is violating a social norm – are most likely involved.[14]

Culture matters

When low contributors who are punished why do they subsequently contribute more? You may think that the answer is obvious: they contribute more to avoid the future reduced payoffs that being punished imposes on them.

But there must be something else going on. In two similar experiments – one in the laboratory in the U.S. and one in the field among farmers in Zimbabwe "punishment" was not in reduced payoffs, it was just a purely verbal expression of displeasure (e.g "selfish guy") by a fellow subject.[15]

But the targets of purely verbal punishment contributed more in subsequent rounds. This occurs most likely because in many societies there is a norm that people should contribute to the public good, and when a person is criticized for violating it, they feel shame and try to make amends.[16]

Culture affects experimental play in other ways. The anthropologist Jean Ensminger conducted public goods experiments with the Orma, a herding people in Kenya. Members of the Orma regularly voluntarily contribute their labor to producing some public good – for example, the repair of a road – a system they call in Swahili, "Harambee." Families that have more cattle – more wealth – are expected to contribute more to the project.

When Ensminger explained the Public Goods Game to the Orma participants, they promptly called it the "Harambee game." Those with more cattle contributed more in the experiment, just as would have been the case in a real "Harambee."

When the Orma subjects played the Ultimatum Game, however, they did not compare it to "Harambee" and wealth did not predict any aspect of their experimental play.[17] This difference in how the wealthy played the Public Goods game and the Ultimatum Game probably would not have surfaced among the farmers you have already met from, Palanpur, Illinois or West Bengal.

2.13 Social preferences are not "Irrational"

People sometimes think of other-regarding and ethical preferences as something special – different from the taste for ice cream, for example – and requiring a model different from the preferences, beliefs, and constraints approach. But the desire to contribute, to punish those who do not contribute, and otherwise to act on the basis of social preferences, like the desire to consume conventional goods and services, can be represented by preferences that conform to standard definitions of rationality.

What we know from experiments is that whether its ice cream or contributions to the public good, people respond to trade-offs, taking account of the costs of and how much they value the activity in question: the higher the cost of helping others, the less its frequency. In other words, other-regarding preferences are consistent with rationality, namely consistency (transitivity) and completeness.

Researchers tested the rationality of seemingly altruistic choices by asking 176 subjects to play a version of what is called the Dictator Game.[18] In what is called the “Dictator Game", one player (the Dictator), Alice, is given a sum of money by the experimenter, and asked to transfer whatever proportion of the money that she wishes to an other (anonymous) subject, Bob. Alice is told that that for every dollar that Bob receives from her, she will have to pay p dollars. So p is the price of altruism:how much she has to pay for every dollar that Bob gets. After Alice makes her decision, the money is transferred, and the game is over.

In this experiment, 75 percent of the Dictators gave away some money, demonstrating altruistic preferences. The average amount given away was a quarter of the endowment when the price $p = 1$ (a dollar for-dollar transfer).[19] Moreover, the higher the price of generosity, the less money was transferred. For instance, when each dollar transferred to Bob cost Alice two dollars $(p = 2)$, only 14.1 percent of the endowment was given away on average, and when each dollar transferred cost four dollars, only 3.4 percent of the dictator's endowment was transferred. The higher the price of altruism, the less did Alice "purchase."

FACT CHECK In a Public Goods Game with Punishment experiment researchers found that the level of punishment that subjects inflicted on others was less when each dollar subtracted from the payoffs of the target cost more in foregone payoffs to the punisher.[20]

It may be, as the old saying goes, that "virtue is its own reward." But that does not mean that people will act virtuously no matter what the price. This finding is perfectly consistent with the fact that people respond to the price of virtuous behavior just as the preferences, beliefs, and constraints model predicts.

Checkpoint 2.9: Dictator Game?

Is the "Dictator Game" a game? Think about how we've defined games (check back in Chapter 1 to ensure you remember).

2.14 Application: The lab and the street

EXTERNAL VALIDITYResults of experiments or other scientific research that can be generalized to circumstances outside (external to) the laboratory or other setting in which the research was produced, are said to be externally valid.

Do people behave in the real world the way they do in experiments? The experimental evidence for reciprocity or related forms of other-regarding behavior would not be interesting if it did matched by similar behavior outside the lab. We therefore need to check whether laboratory evidence is **externally valid**, that is, consistent with behavior observed outside of the laboratory in similar circumstances to those found in the lab. External validity is particularly important for policy questions because policy-makers and governments need to know whether a policy will work outside of the controlled conditions of the laboratory.

Generalizing directly from experiments to behavior in other contexts is often unwarranted. For example, in the Dictator Game typically more than 60 percent of the Dictators allocate a positive sum to the recipient, and the average given is about a fifth of the endowment.[21] But we would be sadly mistaken if we predicted on the basis of this experimental result that 60 percent of people would spontaneously give money to an anonymous person passing them on the street, or that the same subjects would offer a fifth of the money in their wallet to a homeless person asking for help.

FACT CHECK In an experimental game about trust and reciprocity played by groups of students and groups of chief executive officers of Costa Rican businesses, the businessmen were both more trusting of others and also reciprocated the generosity of their game partners to a far greater degree than did the students.[22] Based on existing experimental evidence, students are not particularly other-regarding.

Many researchers have tried to see whether behavior in lab experiments predicts behavior outside the lab.

Along the coast of northeastern Brazil, for example, shrimpers catch shrimp in large plastic bucket-like contraptions. The shrimpers cut holes in the bottoms of the traps to allow the baby shrimp to escape, thereby preserving the stock of shrimp for future catches.

The shrimpers face a real-world coordination problem: the expected income of each would be greatest if he were to cut smaller holes in his traps (increasing his own catch) while others cut larger holes in theirs (preserving future stocks). In Prisoners' Dilemma terms, small trap holes are a form of defection that maximizes the individual's material payoff irrespective of what others do (it is the dominant strategy). But a shrimper might resist the temptation to defect if he were both public spirited toward the other fishers and sufficiently patient to value the future opportunities that they all would lose were he to use traps with smaller holes.

Ernst Fehr and Andreas Leibbrandt implemented both a Public Goods game and an experimental measure of impatience with the shrimpers. They found that the shrimpers with both greater patience and greater cooperativeness in the experimental game punched significantly larger holes in their traps, thereby protecting future stocks for the entire community.[23]

Additional evidence of external validity comes from a set of experiments and field studies with 49 groups of herders of the Bale Oromo people in Ethiopia, who were engaged in forest-commons management. Devesh Rustagi and his coauthors implemented public-goods experiments with a total of 679 herders, and also studied the success of the herders' cooperative forest projects.[24]

The most common behavioral type in their experiments, constituting just over a third of the subjects, were reciprocators who responded to higher contributions by others by contributing more to the public good themselves. The authors found that groups with a larger number of reciprocators were more successful – they planted more new trees – than those with fewer reciprocators. This was in part because members of groups with more reciprocators spent significantly more time monitoring others' use of the forest. As with the Brazilian shrimpers, differences in the fraction of reciprocators in a group were associated with substantial increases in trees planted or time spent monitoring others.


Figure 2.15: A shrimping bucket with holes in it

2.15 Application: A fine is a price

How might a policy-maker or CEO of a business make use of the fact that people care about what happens to others and they value behaving ethically?

Think about a set of rules for compensating employees. The rules typically specify pay and provision for time off, sick days and the like. But problems arise with using purely material incentives to influence how people behave.

Having noticed a suspicious bunching of sick call-ins on Mondays and Fridays, the Boston Fire Commissioner on December 1, 2001 ended the Department's policy of unlimited paid sick days. Instead, the commissioner imposed a 15-day sick day limit. The pay of firefighters exceeding that limit would be cut. The firefighters responded to the new incentives: those calling in sick on Christmas and New Year's Day increased ten times over the previous year's sick days.

The Fire Commissioner retaliated by cancelling their holiday bonus checks. The firefighters were unimpressed: the next year they claimed 13,431 sick days; up from 6,432 the previous year.[25]

Many of the firefighters, apparently insulted by the new system, abused it, or abandoned their previous ethic of serving the public even when injured or not feeling well. In the language of the Ultimatum Game, they responded reciprocally to an offer they disliked by rejecting it. They were trying to *punish* the Commissioner at a cost to themselves.

The Commissioner's difficulties are far from exceptional.

Consider the following experiment in Haifa, Israel.[26] Parents everywhere are sometimes late in picking up their children at day care centers.

- *Treatment*: At six randomly chosen day care centers, a fine was imposed for parents picking up their children late.

- *Control*: In a control group of day care centers no fine was imposed.

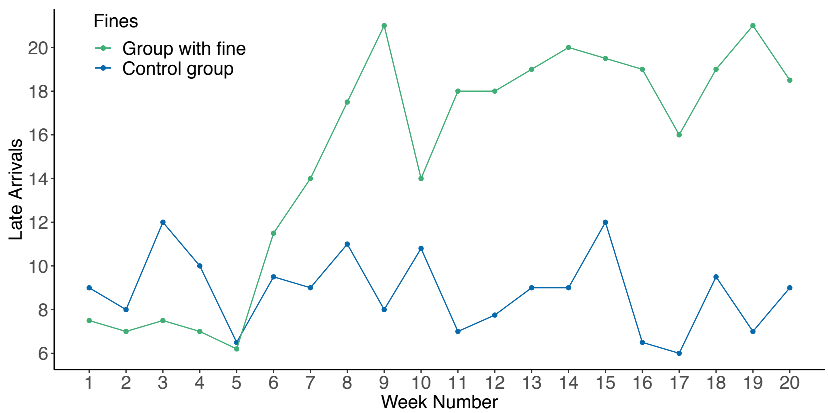


Figure 2.16: **The effect of a fine for lateness in Haifa's day care centers.** Source: Gneezy and Rustichini (2000). The fine was imposed in week 5 and retracted in week 17.

Researchers expected parents to arrive on time because of the fine. But parents responded to the fine by arriving late more often: the fraction of parents picking up their kids late more than doubled. when the fine was taken away after 16 weeks, the parents continued to arrive late, showing no tendency to return to the status quo prior to the experiment. Over the entire 20 weeks of the experiment, there were no changes in the degree of lateness at the day care centers in the control group.

The researchers reason that the fine was a contextual cue, unintentionally providing information about appropriate behavior. The effect was to convert lateness from the violation of a social norm or obligation that the parents were to respect, to a choice with a price that many were willing to pay. They titled their study "A Fine is a Price" and concluded that imposing a fine labeled the interaction as a market-like situation, one in which parents were more than willing to buy lateness for money. Revoking the fine did not restore the initial context.

When monetary incentives undermine social preferences as they did among the Boston firefighters and Haifa parents, this is called **crowding out**. These two cases of crowding out are cautions that the use of monetary incentives may be inappropriate where the targets of the incentives are motivated by other regarding preferences. But they are not reasons to think that incentives are ineffective, as we will see in many examples to follow. We have no doubt that had the fine for lateness in Haifa been 500 New Israeli Shekels rather than 10 the parents would have found a way to pick up their kids on time.

CROWDING OUT is said to occur when monetary or other material incentives undermine other regarding or ethical preferences.

2.16 Complexity: diverse, versatile, and changeable people

The experimental and observational evidence suggests an adequate understanding of preferences should recognize four aspects in human social behavior.

- *Diversity*: people differ in their preferences .
- *Versatility*: even a single person has a diversity of preferences, and which of these is salient for making a decision depends on the situation, for example, shopping as opposed to spending time with friends.
- *Changeability*: people learn new preferences under the influence of their experiences.

These three aspects of our preferences contribute to a fourth attribute of how human beings interact:

- *Complexity or "the whole is not the sum of its parts"*: the outcome of an interaction of many people cannot be deduced in any simple way from the characteristics of the individual people involved.

Diversity

What motivates people differs, both locally and across different cultures and across time. Using data from a wide range of experiments, researchers estimate that between 40 and 65 percent of people exhibit other-regarding preferences of some kind. The same studies suggest that between 20 and 35 percent of the subjects exhibit conventional self-regarding preferences.[27] The authors of another study (in the U.S.) termed 29 percent of their experimental subjects as "ruthless competitors" (presumably resembling Economic Man) and 22 per cent as "saints."[28]

DIVERSITY AND HETEROGENEITY We use the term *heterogeneous* to describe a group of actors with different preferences or some other attributes, for example, wealth gender, nationality, first-mover status in game, and so on.

Versatility

A common observation about human behavior made by psychologists is that the same person can act differently depending on the situation. As a result, we say that people are *versatile*: they change in response to what their situation seems to require of them, for example, being self regarding while shopping and other regarding with one's neighbors.

VERSATILITY How people act depends very much on the situation, resembling *Homo economicus* in some contexts (say, in business), but other-regarding social preferences in other contexts (say, around their family). Psychologists explain how a situation can frame a decision so as to suggest appropriate attitudes (or as economists would say, preferences) towards the possible actions an individual might make. We refer to this aspect of our behavior as versatility.

In the Ultimatum Game, Proposers often offer amounts which maximize their expected payoffs. But Responders rarely do. Researchers have also found this in experiments where subjects play both roles: Proposer at one stage in the experiment and Responder in another stage. The same person when in the role of Responder typically rejects positive offers if they appear to be unfair, even if they had made a similar low offer when in the role of Proposer. They therefore act as if they had reciprocal or inequality averse preferences as Responders, despite exhibiting self-regarding preferences when they are Proposers. The fact that in the role of proposers people are more like "ruthless competitors" while in the role of responder are more like "saints" is evidence of our versatility.

Changeability

Some preferences are part of our genetic makeup, having a taste for sweet and fatty foods, for example. But most preferences are *learned* rather than *given* by our genetic inheritance. Durable changes in an individual's evaluations of outcomes often take place as a result of experience. When this occurs we say that preferences are endogenous, meaning that they change as a result of influences such as where a person lives, how they make their living or the rules of the game that govern how they interact with others.

FACT CHECK In experimental games about dishonesty, people who grew up in Communist Party ruled East Germany are more likely to cheat than those who grew up in West Germany.[29]

Over a lifetime or even generations, migrants to a new country, or those moving from a rural to an urban area often adopt new preferences (for example concerning food tastes). The fact that preferences are learned may account for the fact that, as we saw from the experiments in small scale societies, people who hunt large animals tend be generous with the meat they acquire; and they seem to generalize these habits to other realms of life. Preferences are exogenous of they do not change or change only in response to influences that considered to be external.

ENDOGENOUS PREFERENCESPreferences are endogenous if they change as a result of influences such as where a person lives, how they make their living or the rules of the game that govern how they interact with others.

A consequence: Complexity

In everyday language the word "complexity" refers to the state of being intricate or complicated. The term is used in quite a different way in the study of interactions of a large number of independent entities – whether particles or people. A key idea is that the results of these interactions for the system as a whole cannot be predicted in any simple way from even the most detailed knowledge of the interacting entities.

The best example of complexity in the social sciences is Adam Smith's invisible hand. What Smith suggested two and a half centuries ago, and modern economics has shown (as seen in Chapter 15) is that under some conditions uncoordinated interactions among entirely self-regarding total strangers through competition in markets among private property owners can (unwittingly) create an outcome that is better for all than many of the alternatives.

The idea of complexity is often expressed the adage: the whole is different from the sum of the parts. The key here is not that the whole may be greater or less than the sum; it is that summing the parts is not the right way to calculate the whole. Averaging the components of some interacting system will not give what their interactions will actually add up to. The results of the interaction – called their *emergent property* – may be surprising given the nature of the interacting entities.

Here are some examples of surprises (with which you are already familiar) in the properties that emerge from people with heterogeneous and versatile preferences interacting.

- Small differences in the distribution of types – the presence in a population of a few people willing to punish those who do not contribute in a Public Goods Game, for example – can have large effects on how everyone behaves, getting the self-regarding people to act as if they were cooperators. You have seen this in Figure 2.14.

- Seemingly small differences in institutions can make large and surprising differences in outcomes. Why did adding the punishment option so radically change the outcomes in the Public Goods Game? We know that cooperation – contributing to the public good – unravels in in the absence of the punishment option. But the incentives to punish would seem identical to the incentives to contribute to the public good in the first place: everyone would like someone else to bear the cost of punishing the free riders. So not contributing and not punishing should be the dominant strategy in this game. But we now know that that is not what we observe.

- While imposing a fine or other cost on socially undesirable behaviors may create socially desirable outcomes in certain circumstances such as getting people to stop using plastic grocery bags , a fine on parents arriving late to pick up their kids backfired. We saw that the nominal fine decreased parents' willingness to pick up their children on time when the viewed the fine as a price to pay for additional day care: the fine changed what they viewed as socially acceptable behavior.

- Letting a self-regarding player be the first mover in a Prisoners Dilemma game when she knows that the other player has strong reciprocity motives can avert the coordination failure resulting in mutual cooperation. Letting the Reciprocator be the first mover would have the opposite result: both players would defect, resulting in the Pareto inefficient outcome. You can confirm this by doing Checkpoint 28

Checkpoint 2.10: Sequential Prisoners' Dilemma: Self-interest versus reciprocity

For a sequential Prisoners' Dilemma game where the first player is self-interested and the second player is reciprocal draw a game tree in which the Nash equilibrium may be (Cooperate, Cooperate) and explain why could occurs.

2.17 Conclusion

We have explored various social interactions represented as games and also as studied empirically using experiments, such as the Ultimatum game and the Public Goods Game. While self-regarding preferences are represented an essential and powerful motivator for human behavior, we have also found that people behave cooperatively – viewing it as *the right thing to do* – and they enjoy behaving cooperatively.

People dislike unfair treatment and enjoy punishing those who violate norms of fairness or cooperation. The evidence that social preferences are common does not, however, suggest that people are irrational. Indeed, as we have seen, the experimental evidence suggests strongly that when individuals give to others their behavior conforms to the requirements of rational choice. People respond to the price of giving, giving more when it costs them less to benefit the people who receive their money.

The importance of other-regarding preferences thus, does not challenge the assumption of rationality or the preferences, beliefs and constraints approach. However, it does suggest that for many applications we should take account of people's concerns for others and for doing the right thing.

Making connections

Preferences, beliefs, and constraints: This framework for analyzing decisions will be used throughout the rest of the book.

Risk and uncertainty: Many, maybe most, of the important decisions that people make are risky because the resulting outcome depends on contingencies the probability of which occurring the actor does not know.

The rules of the game and coordination problems: Sequential rather than simultaneous play may result in a better outcome in an Assurance Game (or even a Prisoners Dilemma). The reason is that the first mover can help to coordinate play in the game. Another example: allowing players to punish low contributors in a public goods game dramatically changes the outcome.

External effects and Pareto-inefficient Nash equilibria: The public goods game illustrates an extreme form of positive external effects (each person's contribution benefits everyone equally).

Evidence: Economists have recruited novel experimental evidence – from the laboratory and the field – to examine our theories about how people behave. Economists have used the evidence to modify and improve existing models and to develop entirely new models of how people behave.

Heterogeneity: People differ in their preferences (self-regarding, other regarding) and in the advantages associated with their positions (first mover, second mover)

Important ideas

preferences	beliefs	constraints
rationality	self-interest	social preferences
fairness	altruism	reciprocity
spite	endogenous (preference)	exogenous (preference)
institutions	*ad hoc*	external validity
laboratory experiment	field experiment	endowment
ultimatum game	public goods game	punishment
group membership	complexity	inequality aversion
diversity	changeability	learning

Mathematical Notation

Notation	**Definition**
x	a contingency
P	probability of a contingency to occur
$\pi()$	a player's payoff
$E(), \hat{\pi}$	a player's expected payoff
z	individual endowment in Public Goods Game
e	individual contribution in Public Goods Game
M	return factor to contributions in Public Goods Game
n	number of participants in Public Goods Game
p	price of altruism in Dictator Game

Note on super- and subscripts: i: an individual; $Early$: the strategy planting early; $Late$: the strategy planting late.

Discussion questions

See supplementary materials.

Problems

See supplementary materials.

3

Doing the best you can: Constrained optimization

“What a useful thing a pocket-map is!” I remarked.

“That’s another thing we’ve learned from your Nation,” replied Mein Herr, “map-making. But we’ve carried it much further than you."

“What do you consider the largest map that would be really useful?”

“About six inches to the mile.”

“Only six inches! “ he exclaimed.

“We very soon got to six yards to the mile. Then we tried a hundred yards to the mile. And then came the grandest idea of all! We actually made a map of the country, on the scale of a mile to the mile!”

“Have you used it much?” I enquired.

“It has never been spread out, yet,”

“The farmers objected: they said it would cover the whole country and shut out the sunlight!"

Lewis Carroll, *Sylvie and Bruno Concluded*, 1893

Lewis Carroll, the author of this dialogue (not to mention *Alice in Wonderland*) was also a mathematician and a philosopher. The point Carroll made about maps also goes for economic models. Maps are useful because they convey the necessary information, not because they are an exact representation of the territory, as the people from Mein Herr’s country discovered. Carroll’s point? The map is not the territory.

A good model is not reality, but it’s a helpful guide.

What qualifies a map or a model as useful depends on what we need it for: six inches to the mile might be adequate for a map of hiking trails, but such a hiking map would not be much use to an airplane pilot. The same is true of economic models.

DOING ECONOMICS

This chapter will enable you to:

- See how the preferences, beliefs and constraints framework from Chapter 2 forms the basis for mathematical models of economic behavior.
- Recognize how preferences – whether entirely self regarding or altruistic – can be represented both in mathematical form (a utility function) and graphical form (an indifference curve map).
- Understand that constrained optimization is a method that economists use to explain the actions that people take; it is not a description of the thoughts or feelings making up individuals’ decision making processes (e.g.studied by a psychologists).
- Explain how people are constrained – for example by limited time – and how these constraints give rise to opportunity costs and, along with our preferences, to trade offs.
- Use the preferences, beliefs and constraints framework to analyze difficult choices concerning in policy-making, including how much of society’s resources should be devoted to the abatement of environmental damages.
- Use the concepts of ordinal and cardinal utility explain how they differ and how cardinal utility provides a way to represent the societal cost of economic inequality.
- Understand the shortcomings and limits as well as the insights of these models.

Figure 3.1: **The London underground transit system:** the map represents, but is not the same as, the territory to which it corresponds. The map is a helpful model.

Think of a model as a lens. An economic model is a way of focusing on what is important given the question that one wants to address without complicating the picture with things that do not matter for the question at hand.

A key component of many economic models – those using the preferences, beliefs and constraints approach – is that we can understand the actions people take by assuming that they are doing the best they can under the circumstances that they are in. When implemented using mathematical reasoning, this process is called "constrained optimization, a process by which a person determines a course of action to accomplish a goal (reflecting the person's preferences), given the information that the person has (beliefs) and the actions they may feasibly takes (a constraint).

CONSTRAINED OPTIMIZATION is the mathematical representation of a process by which a person determines a course of action in order to accomplish a goal (reflecting the person's preferences), given the information that the person has (beliefs) and the actions they may feasibly takes (a constraint.)

We illustrate a model and the process of constrained optimization by something that matters to all of us: Time, and how we use it.

3.1 Time: A scarce resource

Benjamin Franklin (1706-1790) – the American politician and inventor – once said, "Time is money." Franklin was referring to the presence of trade-offs in how people choose to spend their limited time. His three-word sentence is therefore a constrained optimization model: people choose their daily actions to achieve their goals under the constraint of limited time.

Spending an hour or minute on an activity provides us value of some kind: we enjoy the activity itself (e.g. eating) or the results of the activity (e.g. being paid a wage). But, since time is limited, choosing one activity also means we give up that time to do something else. We incur a cost of doing an activity because we forfeit the value of the next best thing we could have spent our time on instead: this is the **opportunity cost** of our time.

OPPORTUNITY COST The opportunity cost of x in terms of y is the marginal rate of transformation: how much y a person must give up to get a unit more of x .

Unless we have time to spare, and are wondering how we will fill up our day there is an opportunity cost to our use of time. As a result, we can model how we use our time as the result of our evaluating the benefits and costs (including opportunity costs) of pursuing one set of activities rather another. To do this we use constrained optimization.

Before developing the concepts on which constrained optimization is based, let's look at the kinds of facts that a model of time use should be able to explain.

Figure 3.2 shows how men and women from the USA used their time each day during the year 2013. The largest time use is for the categories sleep, work (meaning for pay), leisure, and house work. Men and women differ typically in the hours they devote to paid work and house work and care work, often reflecting differing social norms about the kinds of activities that it is "appropriate" or "natural" for men and women to do.

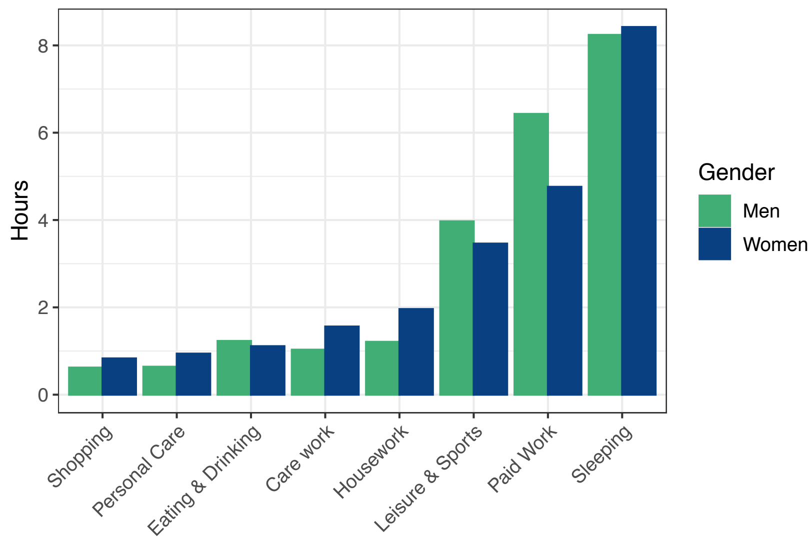


Figure 3.2: **Daily time use of American men and women.** These data – for hours in each activity measured on the horizontal axis for all adults – differ from data restricted to those with small children, or retired people, or students.

Source: Hofferth, Flood, and Sobek (2013).

But these social norms also change, sometimes in ways that show that the differences in the distribution of work time between men and women are far from determined by "nature" but instead reflect changed economic conditions. During the second half of the 20th century in the rich countries the fraction of women doing paid work outside the home dramatically increased. While we do not have detailed information like that shown in Figure 3.2 for the mid-20th century on how men and women spent their time there almost certainly has been a decline in the amount of time doing housework.

Part of the change in the distribution of women's time between house work and work for pay is due to the availability at affordable prices of new technologies – household appliances – that reduced the amount of time required to clean house, wash clothes, and the carry out the other housework tasks. These include washers, refrigerators, and vacuum cleaners which in the U.S. became common from the late 1940s onward, and dryers, dishwashers and microwaves somewhat later.

FACT CHECK For a long historical view of why the washing machine was a "miracle" have a look at this video by Hans Rosling: <LINK HERE>

Evidence that these new technologies contributed to the change in the distribution of women's work time comes from a comparison across countries of increases in the fraction of women working outside the home – called the labor force participation rate – and decreases in price of these labor saving household appliances (compared to other prices).[1] The results are in Figure 3.3, which shows that in countries such as the U.S. where the prices of these appliances fell the most, women's labor force participation rate rose the most. By contrast, in Germany where prices of household appliances fell the least, the increase in labor force participation was half as great as in the U.S.

Other factors contributed, of course, most importantly the reduction in the

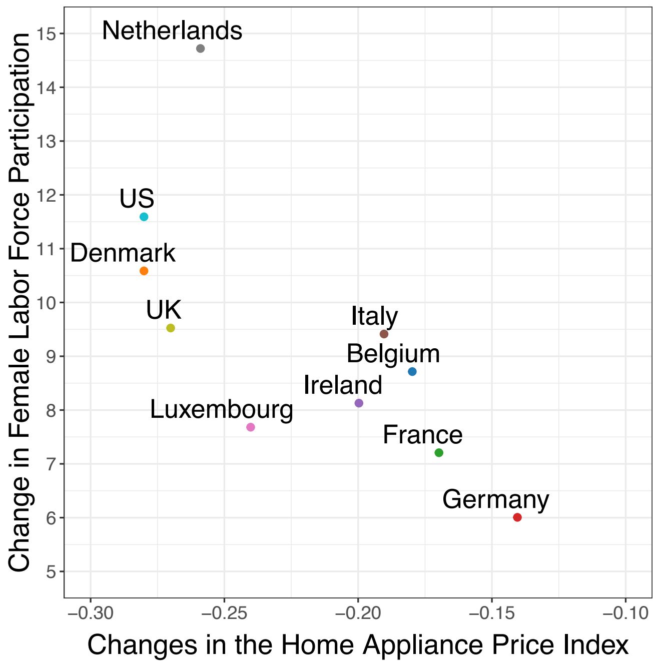


Figure 3.3: **The relative price of home appliances and the female and male labor force participation rates.** The vertical axis represents an index that records the change in the fraction of adult women working outside the home, termed the female labor force participation rate (FLFP), as well as the change in the home appliance price index (HAPI) on the horizontal index. The figures shows that as the price of labor saving household appliances decreases, the female labor force participation rate increases. Notice that a *bigger* price decrease would be shown by a larger *negative change* (further to the left on the x-axis) so the US, Denmark and the Netherlands had big *decreases* in the prices of home appliances and a big *increase* in the female labor force participation rate. Household appliances – like TVs – that did not reduce the amount of time necessary to perform housework tasks are *excluded*. Source: de V. Cavalcanti and Tavares (2008).

number of children born per woman. But the fall in the prices of appliance, the study concluded, was of approximately equal importance. It appears that economic changes – the new household appliances and their falling prices – changed how women spent their time – more working outside the home . This in turn may have been a both a result and a cause for the changing social norms about "women's work" and the decreased adherence to the ideal of a family with a husband income earner and a wife raising (many) children and taking care of the home. This is an example of how preferences – for example, these norms – change as economic conditions – the prices of home appliances – change.

We begin with these examples because methods of constrained optimization – the preferences, beliefs, and constraints framework – provide a way of posing and in some cases answering questions like: Why do men and women spend the time they do on the various activities shown? Or why did work hours fall so dramatically in some countries over the 20th century?

We begin with with preferences, before turning to constraints later in the chapter. Because we are not considering strategic interactions or other situations in which the relevant facts are not known, we do not treat beliefs until the next chapter.

Checkpoint 3.1: Labor-saving household appliances and women's labor force participation

Imagine a conversation around the year 1970 between a husband and wife who just learned that a very effective clothes washing machine is available at a low price. How might the conversation have led to the woman taking up paid work outside the home?

3.2 Utility functions and preferences

In Chapter 1, we represented preferences – our evaluations of the outcomes our actions may bring about – as *payoffs* that is, numbers indicating how much the decision maker values each of the possible outcomes. We discussed, as an illustration, the choice of whether to take an umbrella or not, with a decision [Don't take the umbrella, It rains] resulting in a payoff of 3. The payoff to [Take the umbrella, It rains] was 15, meaning that if it rains the person valued having the umbrella by 5 times as much as not having it.

In that example we simplified things by limiting the actions and the outcomes to just a few, for example, it either rained or it did not. The simplification allowed us to focus on 2 by 2 payoff matrices with just four possible outcomes.

But most of the economic interactions that we study are not that simple: we can contribute any amount to the public good (not just $10 or nothing), the farmers in Palanpur have the choice to plant a little bit earlier, or much earlier, and so on. Or, to return to the question of time: how we divide up our day among the activities in Figure 3.2 could be measured in variations of minutes devoted to each of the nine activities, giving us trillions of "outcomes" to choose amongst. We need a way of representing preferences when there are a great many outcomes, without expanding our payoff matrices to the unusable size of the 1:1 maps in the Lewis Carroll fable at the beginning of the chapter.

REMINDER: PREFERENCES represent the favorable (positive) or unfavorable (negative) feelings that could lead a person to choose one outcome over another. Included are tastes (food likes and dislikes, for example), habits (or even addictions), emotions (such as anger and disgust) often associated with visceral reactions (such as nausea or an elevated heart rate), social norms (for example, those that induce people to prefer to be honest or fair), and psychological tendencies (for aggression, extroversion, and the like). Do not think about a preference or the number a utility function assigns to some bundle as "how much a person likes" the bundle.

UTILITY FUNCTION A utility function is an assignment of a number $u(x,y)$, to every outcome bundle (x,y) representing a person's valuation of that bundle. This means that if given the choice between two bundles (x,y) and (x',y'), the individual will choose the first if $u(x,y) > u(x',y')$.

Why we use utility functions to represent preferences

To do this we use a **utility function**, a mathematical expression that translates the full range of possible outcomes into a person's valuation of the outcome – her payoffs.

The word "utility" (in ordinary language, "usefulness") is used to mean the same thing as "payoff." It is a number assigned to a particular outcome bundle that has the property that when choosing between alternative bundles, a person will select the one with the highest (utility) number.

Both "utility" and "payoff" sound like some monetary or other amount of something you take home as the outcome of a game. But in economics utilities, like payoffs are not something you get or even experience. You don't take them home, they are nothing more than numbers that indicate the course of action you will take.

For simplicity, we call this number "how much the person values the outcome" but the utility function tells us nothing about *why* the bundle has a higher number. It could be any of the reasons for the collection of pro or con evaluations that make up our preferences for some bundle, ranging from food tastes to addictions to ethical norms.

What the function allows us to do is to take account of more complex outcomes than "Don't take the umbrella" and "It rains." The decision maker, as before, will choose the actions the she believes will result in the highest utility outcome. Suppose that our decision-maker, Annette, an Uber driver, is deciding how much time to work, x, and what fraction of the resulting income to spend on food, y. The utility function then assigns a number – the level of utility – to each possible combination of x and y, say, work for 4 hours and 15 minutes and spend 35 percent of the resulting pay on food. Any other combination, say, work four hours and spend 40 percent of the resulting income on food, will be assigned another number, representing Annette's valuation of that particular outcome.

M-CHECK We read x' as 'x prime" and x'' as "x double prime". We usually denote a bundle other than (x, y) as (x', y') to indicate a different composition of the underlying x and y.

CONSISTENCY Consistency (or transitivity) requires that when considering three bundles (x, y), (x', y'), and (x'', y''), if (x, y) is preferred to (x', y') and (x', y') is preferred to (x'', y''), then (x'', y'') *cannot* be preferred to (x, y). Consistent preferences can never lead someone to make *contradictory* choices.

COMPLETENESSCompleteness requires that all possible outcomes can be ranked. For any two bundles (x, y) and (x', y') either the person prefers (x, y) to (x', y')) or the person prefers (x', y') to (x, y)) or the person is *indifferent* between (x', y') and (x, y)).

This assignment of numbers is a *utility function*, $u(x, y)$: for every outcome (x, y) the value of the utility function is the number representing a person's valuation of the outcome. Then if we know what combinations of x and y are available to Annette based on the relevant constraints, then we can predict the choice Annette will make, namely the combination with the highest utility.

What do the utility numbers measure?

We measure how much a person values various outcomes in two ways, either:

- by indicating how valuable each is on some absolute scale, or
- by simply ranking them in order.

If Annette compares two bundles (or outcomes), namely (x, y) and (x', y') with $u(x, y) = 3$ and $u(x', y') = 9$ there are two statements we could make about Annette, one much more informative than the other:

- Annette values (x', y') three times as much as (x, y) and
- Annette values (x', y') more than (x, y)

In the first case above, utility is a number indicating *by how much* Annette prefers (x', y') to (x, y). Utility is therefore termed a **cardinal** measure (cardinality in mathematics refers to the size of something). In Chapter 2 we represented people's preferences by the payoffs associated with particular outcome bundle of games like (x', y') or (x, y). When we defined the expected payoffs to some course of action we added up the payoffs of each possible outcome (weighting them by the probability of each outcome occurring). Doing this required that utility is a measure of size. The numbers representing payoffs in Chapter 2 are cardinal utilities In the second case the utility function gives us an ordering of better-worse for the pair of outcomes. When the utility function is measured in this way, we say that Annette has **ordinal preferences** or that utility is ordinally measured. Ordinal utility says nothing about how much better the preferred outcome. Instead of assigning numbers to the outcomes, in the case of ordinal utility, it would be clearer if we just assigned ranks, like instead of 1,2,3,4 and so on, we used 1st, 2nd, 3rd, 4th (and in cases of indifference: for example, tied for 7th). In the cartoon figure about the Planting in Palanpur game (Figure 1.2), we listed the four possible outcomes as "Best, Good, Bad" and "Worst": this is an example of ordinal utilities. There is no way that we can say that the top ranked bundle is twice as good as the second-ranked bundle or ten times as good as the tenth-ranked bundle. Nor could we add up the ranks, saying, for example, that getting your second ranked bundle and your third ranked bundle with equal probability is as good as getting your first and fourth ranked bundle with equal probability. None of these statements make any sense. This is why when dealing with decisions involving risk, we used a cardinal measure.

ORDINAL PREFERENCES Ordinal preference rank outcomes: e.g. $(x,y) \succ (x',y') \succ (x'',y'')$, without specifying *how much* (x,y) is preferred to (x',y') or (x',y') is preferred to (x'',y''). The assignment of numerical utilities representing ordinal preferences is meaningful only to express the ordering: $u(x,y) > u(x',y')$ implies only that the first bundle is preferred to the second but not by how much.

CARDINAL PREFERENCE A cardinal utility function assigns a number to each outcome, with the property that the ratio of the numbers assigned to alternative bundles expresses the relative degree of the preference for the alternative bundles. For example, with a cardinal utility function, $u(x,y) = 10u(x',y') = 5u(x'',y'')$ means that (x,y) is preferred ten times as much as (x',y') which is preferred five times as much as (x'',y''), and that (x,y) is preferred fifty times as much as (x'',y'').

So, when we introduced the Palanpur farmers' uncertainty about when the other farmer or farmers would plant their crops, we needed to think about expected payoffs, which requires adding up the values that each farmer attaches to an outcome. Because you cannot add up ordinal measures, we gave the payoffs numeric values (the numbers in the payoff matrix) representing cardinal utility.

M-CHECK For simplicity, we generally restrict our analysis to outcomes that can be described in terms of two variables x and y, though it is straightforward to generalize this model to outcomes described by more than two variables. The actor therefore makes choices among "bundles" that combine different amounts of x and y.

For some questions in economics the ordinal – better or worse – meaning of utility is all we need to understand and predict the actions that people will take. But in many situations, those involving risk and uncertainty, as we have just seen, or in evaluating the effects of differing rules of the game – policies to ensure competition in markets or concerning fairness, for example – addressed in Section 3.13, the cardinal measure is required.

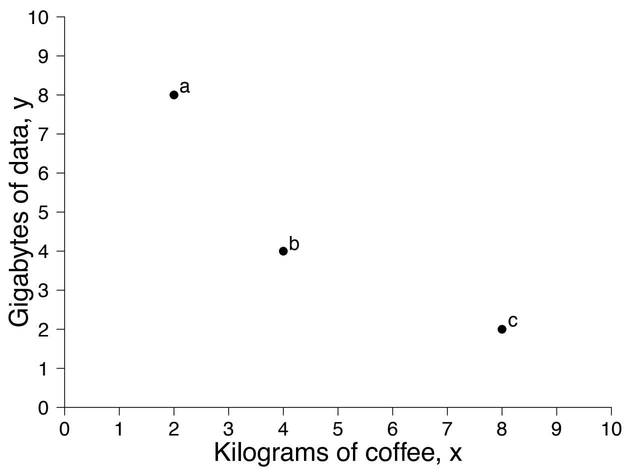
Checkpoint 3.2: Utility and payoffs

Give examples of preferences that might lead people to act in ways that they would regret.

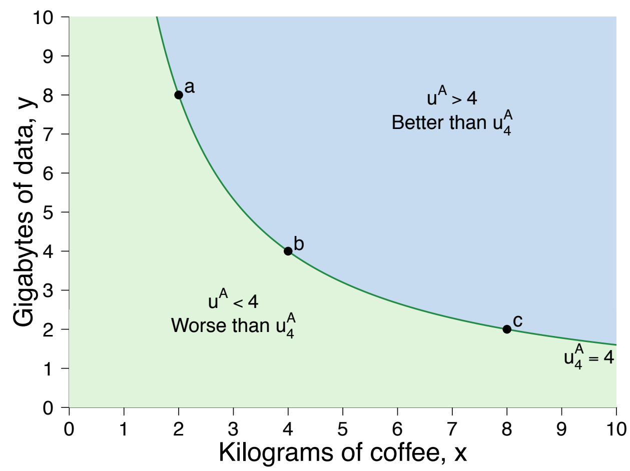
3.3 Indifference curves: Graphing preferences

Indifference curves are a useful way to visualize a person's preferences. Let's illustrate the concept of an indifference curve by Annette, who is choosing among differing amounts of kilograms of coffee (x) and gigabytes of data (y).

INDIFFERENCE CURVE The points making up an individual's indifference curve are bundles – indicated by (x,y), (x',y') and so on – among which the person is indifferent, so that $u(x,y) = u(x',y')$ and so on.This means that all of the bundles indicated by points making up an indifference curve are equally valued by the person.



(a) Consumption Bundles



(b) An indifference curve

Figure 3.4: **One of Annette's indifference curves: coffee and data**. The dark green indifference curve u_1^A represents all the combinations of x and y that provide Annette (A) with the same level of utility, 4. The blue area above and to the right of Annette's indifference curve shows combinations of the amounts of coffee and data that provide her with utility greater than 4. The light green area beneath her indifference curve shows the bundles of x and y that she values at less than 4. She would therefore rather choose a combination of x and y on the indifference curve shown than any point to the left or below it.

Every point given by the coordinates (x, y) in Figure 3.4a is a pair of the quantities of the two goods, called a bundle. Points **a, b**, and **c** therefore represent three bundles of differing amounts of coffee and data. Suppose that Annette ranks the points **a**, **b**, and **c** equally – she is indifferent among the three bundles – then these three points lie on the same indifference curve, as shown in Figure 3.4 b. Her indifference curve represents the combinations of bundles among which she is indifferent. This means that for either bundle **a** – 8 gb of data and 2 kg of coffee – or bundle **b** – 4 gb of data and 4 kg of coffee– or bundle **c** – 2gb of data and 8 kg of coffee, $u(2,8) = u(4,4) = u(8,2) = 4$.

Figure 3.4 b shows the indifference curve made up of all bundles for which Annette's utility is equal to 4. Her indifference curve is labeled by a u with a subscript which represents the level of *utility* that is the same for all points on that indifference curve. Annette prefers to consume *more* of both data and coffee, so she would like to be anywhere in the blue-shaded area where her utility would be *greater* than 4). She would rather not consume *less* of both data and coffee, so she would not like to be down to the area shaded in green where her utility would be *less* than 4.

BUNDLE A bundle is a particular allocation, in the case of two goods given by (x, y). A bundle that results from the choice of one or more decision makers is call an **outcome bundle.**

The single indifference curve shown in Figure 3.4 b divides the space of all possible bundles of x and y into three categories: bundles that are respectively better or worse than any of the bundles making up u_4 and bundles that are equally valued with a utility of 4.

To understand a decision-maker's choice we proceed in steps:

- *Step 1*: In this and the next section we use many such indifference curves to evaluate all of the possible outcome of the decision.
- *Step 2*: In Section 3.6 we then limit the decision maker's choices to those

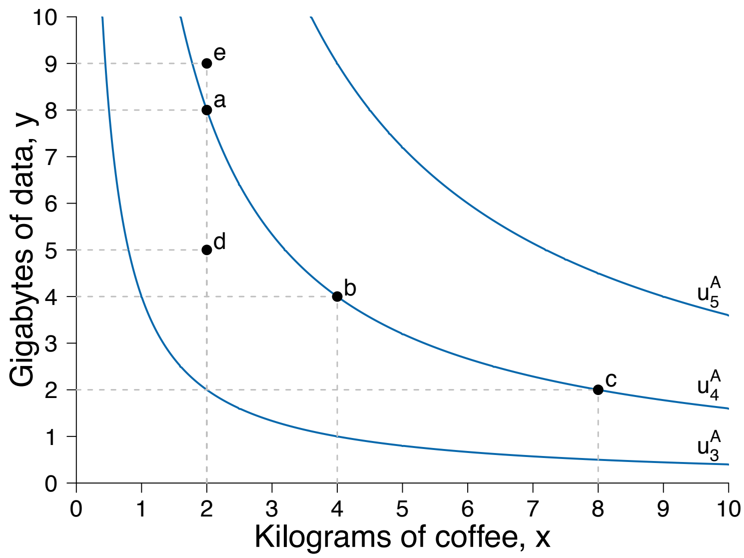


Figure 3.5: **An indifference map for kilograms of coffee, x, and gigabytes of data, y.** The quantity of good x is on the horizontal axis and the quantity of good y is on the vertical axis. Three indifference curves are shown: u_3^A, u_4^A, and u_5^A, where the rank of the utilities is $u_5^A > u_4^A > u_3^A$. The constant level of utility for $u_4^A = 4$. Points **a**, **b**, and **c** all lie on u_2 and give Annette the same utility of 4. Point d would give Annette lower utility and point **e** would give Annette higher utility (because every bundle is associated with some utility number, we could draw indifference curves through those points, and through any point in the figure).

that are feasible for (that is, choices that are actually open for the decision maker to take).

- *Step 3*: Finally, use the evaluations in Step 1 to rank all of the feasible outcomes, showing us the one the decision-maker ranks the highest.

To take Step 1, an individual's utility function allows us to rank all of the outcome bundles – all combinations of x and y – that the decision maker considers. We do this using what is called a set of indifference curves (also termed an indifference map) as shown in Figure 3.5.

INDIFFERENCE MAP An indifference map is a set of indifference curves selected so as to illustrate some concept or result. For example to compare two bundles or to identify an outcome bundle that is the outcome of the decision-making process.

Figure 3.5 shows three indifference curves, u_3, u_4, and u_5, part of Annette's indifference map. Annette prefers more of both goods – that's why they are called "goods." Therefore, indifference curves to the upper right, like u_5, are higher, (corresponding to the blue-shaded area in Figure 3.4). Indifference curves representing less preferred combinations, like u_1 are to the lower left (corresponding to the green-shaded area in Figure 3.4). Of the three indifference curves plotted on the indifference map of Figure 3.5, u_1^A provides Annette with her lowest utility, whereas u_3^A provides Annette with her highest utility. A different person, one who valued coffee more than Annette would have a different indifference map.

If you think of her indifference curves as a kind of contour map, Annette can be pictured standing somewhere on a mountain wanting to get to the top. She might, for example be in the lower left corner of the contour map of a hill shown in Figure 3.6 wanting to reach the 800 meter top of the hill.

Her utility is the altitude where she is standing, say, at a point on the 720 meters above sea level contour. Her indifference curves are the numbered contour lines on a map of the mountain she is climbing, each indicating locations on the mountain the same height above sea level.

A map, as the quotation at the beginning of this chapter reminds us, is a representation of territory. The territory represented by Annette's indifference map is her evaluation all possible outcomes she might experience. An indifference curve runs through every point in the (x, y) plane, but just like maps that could not possibly show *every* contour line, we can plot only a selected number of them in any case.

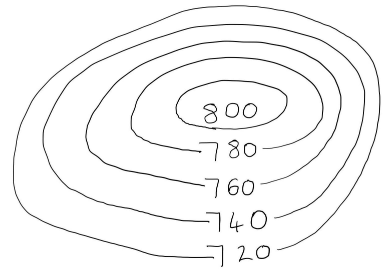


Figure 3.6: **A contour map of a hill showing altitudes.** Indifference curves are similar to contour lines, which are composed of all the points in the landscape which are at the same altitude. The lower left quarter of the contour map resembles the indifference map in Figure 3.5.

Annette wants to climb as high as she can up the utility-mountain as possible, given whatever limitations she faces, including her own physical capacities and possibly impassible cliffs blocking her way. As Annette advances up the mountain, she crosses contour lines, moving from lower to higher indifference curves. She is engaging in a constrained optimization problem.

Checkpoint 3.3: Maps, Points and Bundles

Sketch your own version of the indifference map in Figure 3.5. Add two new points to your graph:

a. A bundle, labeled **f**, where Annette holds the same amount of *y* as she does at point **b,** but Annette prefers bundle **b** to **f**.
b. A bundle, labeled **g**, where Annette holds the same amount of *y* as she does at bundle **b**, but which Annette prefers to bundle **b.**
c. Having manipulated the graphs and thought through the ideas of indifference curves, explain why the following is true: Consistency of preferences implies that indifference curves cannot cross.

3.4 Marginal utility and the marginal rate of substitution

Indifference maps are used to summarize the values that an individual places on differing bundles of goods. But goods need not be things like Annette's coffee or data. Goods can be anything a person values, such as free time. (Indifference curves, as we will show later in the chapter, can also summarize the preferences people have about "bads" such as environmental degradation, that, unlike goods, are things that people would prefer to avoid).

To see this, we will move from the choice about coffee and data, and think instead about a new person, Keiko (KAY-i-ko), who is a student making a choice about the use of her time. As Keiko progresses through her studies (no doubt fueled by coffee and using data), she has two important priorities, which she thinks of as "Living" and "Learning."

- *Learning* comprises all the aspects of her life as a student that contribute to her goals of becoming an educated person and becoming qualified for an interesting career.

- *Living* comprises everything else, including keeping up with friends, meeting new people, and taking care of herself.

As there are only so many hours in a day, and because Learning takes time, Keiko faces what is called a trade-off between Learning and Living, the more she has of one the less she will have of the other. So she is facing another constrained optimization problem.

TRADE-OFF A trade-off is a situation in which having more of something desired (a "good") requires having less of some other "good" or more of something that the actor would like to have less of (a "bad").

We explain in this chapter's last but one section that constrained maximization is not a description of the the mental and emotional processes by which we adopt one course of action over another. It is a research strategy that we use to understand what people do, not how they come to do it. But to illustrate the method we will suppose the Keiko consciously maximizes her utility function subject to her only-24-hours-in-the-day constraint, by comparing the utility associated with each of the combinations of Learning and Living that are open to her. (OK only a student in economics would actually do this!)

Keiko is a systematic and quantitatively oriented person, and decides to measure her Learning quantitatively with a number. In calculating her Learning, she takes account of feedback from her teachers, such as grades (marks), but also evaluates this feedback in terms of her own estimation of how much she has learned, such as how much her study is improving her writing skills and general understanding.

Keiko measures her the amount of Living by the hours she can spend *not studying*, x, and the amount of her Learning by her personal rating, y.

DIMINISHING MARGINAL UTILITY What is sometimes called the "Law of diminishing marginal utility" holds that the marginal utility of any thing that we value is less the more of it that we have.

Key to how the preferences, beliefs and constraints approach works is the fact that for most of the things that we may value, if we have little of it, we highly value having more of it, but the more of the thing we have, the less valuable will be the next additional unit that we *could* have. This is called diminishing marginal utility, where the new idea here is "marginal."

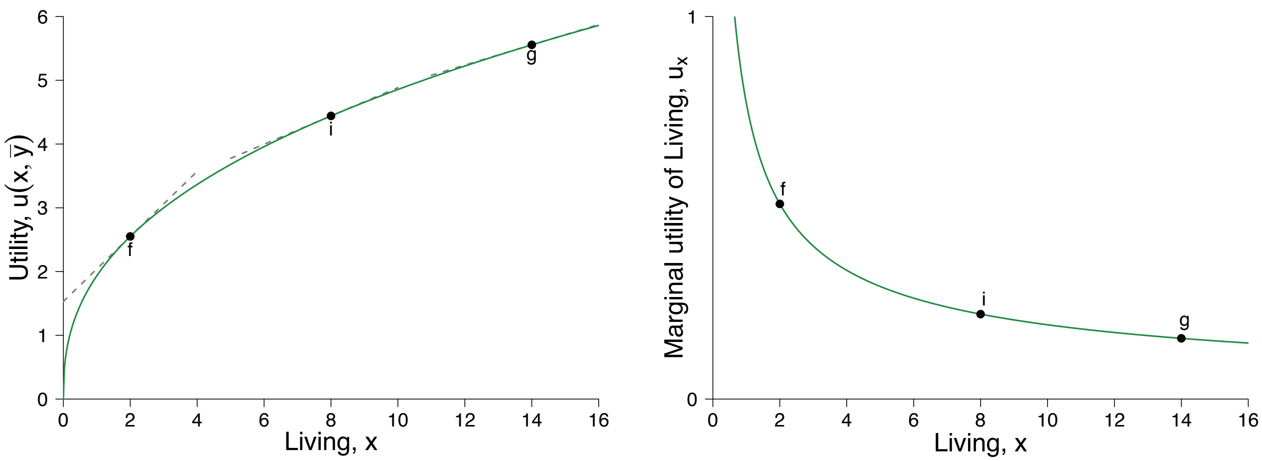
FACT CHECK Diminishing marginal utility in economics is often based on the psychological principle of satiation of wants, which states that satisfying our wants is pleasurable, that our wants (for example hunger) are limited, when the resources allowing satisfaction of wants are limited we satisfy our most urgent wants first, and that the more satisfied is the want (by eating) the less pleasure do we derive from further satisfying the want.

M-Note 3.1: The meaning of marginal

The change in the value of a function – like utility, $u(x,y)$ – when just one argument of the function x or y changes is a basic concept in calculus. The partial derivative of the function with respect to an argument – that is either $u_x(x,y)$ or $u_y(x,y)$ – is an approximation of the effect of a small change in the argument on the value of the function, holding constant the other arguments. If the decision-maker increases her consumption of x by a small amount Δx, then her utility is $u(x+\Delta x,y) \approx u(x,y)+u_x(x,y)\Delta x$.

The marginal effect on utility of a change in some element in the bundle a person is consuming (x) is calculated as the size of the change in u relative to the size of some small change in x with no other changes in the bundle. This is the size of the effect (Δu) divided by the size of the cause (Δx), so $u_x(x,y) = \frac{\Delta u}{\Delta x}$ where Δx is small. Conventionally this is expressed as the effect on u of a one unit change in x.

If the marginal utility of any thing that we value positively is less, the more of it that we have – diminishing marginal utility – then this means that:



(a) Utility of Living holding Learning constant

(b) Marginal utility of living holding Learning constant at 3

Figure 3.7: **Diminishing marginal utility**. In panel a, utility is an increasing and concave function of Living, meaning that the curve it is positively sloped, but with a decreasing slope for higher levels of Living. The slope of the curve is the marginal utility, and this is shown in panel b. The points the in panel a, correspond to the same points in panel b. For example, the height of point **f** in panel a shows the level of utility when Keiko experiences 2 just hours of Living and the slope of a tangent to the curve at that point is the marginal utility of Living that point. The height of point **f** in panel b shows the value of that slope, that is the marginal marginal utility of increased Living when Keiko experiences 2 hours of Living.

- the *first* partial derivative of the utility function with respect to good x is positive, $u_x > 0$, (because more x is better than less) and
- the *second* partial derivative of the utility function with respect to x, $u_{xx} < 0$ (because as x increases, utility is increasing ($u_x > 0$,) but at a diminishing rate, therefore giving us diminishing marginal utility).

Diminishing marginal utility

A change of one variable – like Keiko's Living – by one very small unit while holding constant everything else, including her Learning, is a *marginal change*, meaning the the change is very small and in only one variable. The change in utility corresponding to a marginal change in x or y is called the **marginal utility** of x or y. Keiko's marginal utility of Living which we denote as u_x, like her utility itself, depends on how much Living and Learning she is currently experiencing. So we write u_x as a function of x and y: $u_x(x, y)$.

M-CHECK We also use the symbol for partial differentiation $\frac{\partial u(x,y)}{\partial x}$ to mean the marginal utility of x. When it is not necessary to be reminded of the other variables (held constant) that the marginal utility depends on, we eliminate the (x, y) and just use $\frac{\partial u}{\partial x}$ or u_x .

Similarly, Keiko's marginal utility of Learning, $u_y(x, y)$ (or using the alternative notation $\frac{\Delta u(x,y)}{\Delta y}$, is how much her utility changes as she changes her Learning (y) by one unit, holding constant the amount of Living she does (x).

To understand marginal utility, let us compare points **f**, **i** and **g** in Figure 3.8. By comparing these three points we can see how the marginal utility of Living (x) changes while holding Learning (y) constant. Keiko's amount of Learning (y) is the same as she compares **f**, to **i** to **g** and the increase in her Living results in Keiko's utility increasing from u_1^A to u_3^A to u_3^A. As her Living increases, however, each additional increase in x is associated with a smaller increase in utility, as reflected by her indifference curves getting flatter as she increases the amount of Living (that is, moving to the right along a horizontal line in the figure). This shows that the marginal utility of Living is decreasing as she gets more Living (increased x).

Figure 3.7 shows just a slice of Keiko's preferences, namely how they vary

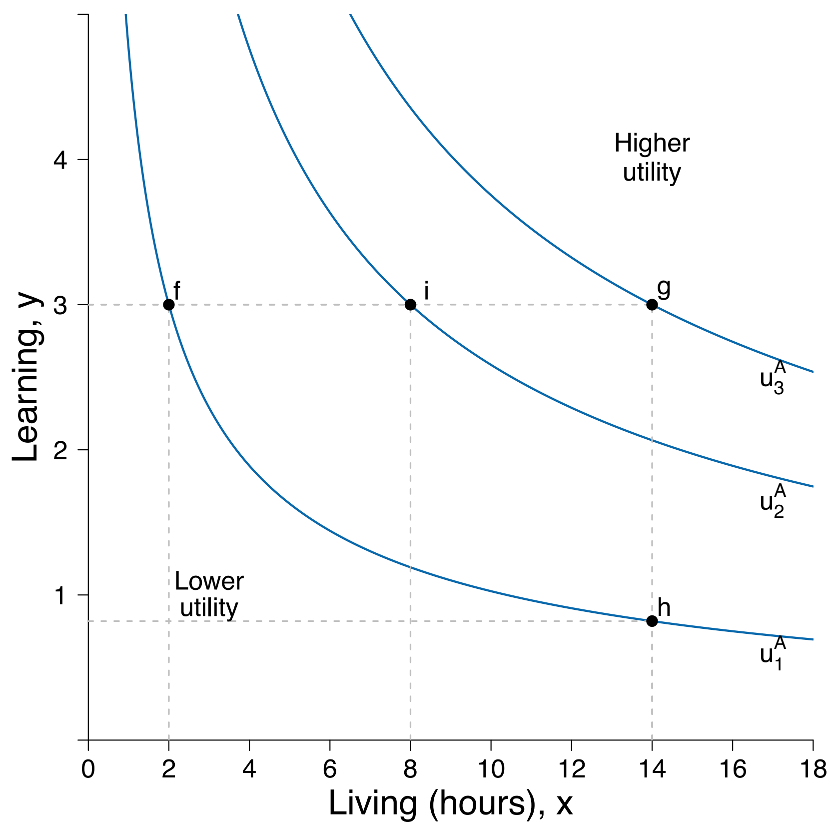


Figure 3.8: **An indifference map portraying the choices between Living (x) and Learning y.** The negative of the slope of the indifference curve is the *marginal rate of substitution* of Learning (y) for Living (x), $mrs(x, y)$, capturing the trade-offs of Keiko's preferences for the two goods. At **f**, Keiko has a high level of Learning (3) and little Living (2 hours) and she is willing to give up a lot of Learning to get more Living (her slope is steep at point **f**, therefore her marginal rate of substitution is large). At **h**, Keiko has a low level of Learning (0.82) and a lot of Living (14 hours) and she is willing to give up very little Learning to get more Living (her slope is relatively flat at point **h**, therefore her marginal rate of substitution is small). Keiko has a Cobb-Douglas utility function with $u(x, y) = x^{0.3}y^{0.7}$.

with the level of Living she experiences, when the level of Learning she experiences is fixed at $y = 3$. We can study the full range of her preferences when the values of both goods varies by looking at her entire indifference map.

The marginal rate of substitution

This is shown in figure 3.8 where points **f**, **i** and **g** correspond to the same points in the previous figure.

At point **f**, Keiko spends 14 hours studying and attending classes and has 2 hours left over for Living, with the result of a lot of Learning and not so much Living. At point **h**, Keiko spends 2 hours studying and attending classes and has 14 hours left over for Living, but her Learning is lower than at point **f**. Comparing between points **f** and **h**, we can see that Keiko sacrifices some of one good to get more of the other. Comparing **f** to **h**, Keiko gives up Learning to get more Living, but her utility remains the same.

If we apply the same reasoning to very small differences in the quantity of the two goods, we can see that at point **f** the largest amount of Learning that

Keiko would be willing to give up in order to get one more unit of Living is the negative of the slope of her indifference curve at that point (0.64 at point **f**), which is the *marginal rate of substitution* at that point or $mrs(x,y)$.

The fact that the indifference curve is steep at that point means she would be willing to give up a substantial amount of Learning to get a little more Living. This is because – as is clear from the previous figure – at point **f** she has an ample amount of Learning and not much Living, so her marginal utility of Living is high.

MARGINAL RATE OF SUBSTITUTION The *marginal rate of substitution* is the *negative of the slope of the indifference curve.* It is also the *willingness to pay* for a small increase in the amount x expressed as how much of y the person would be willing to give up for this. This is sometimes called the *offer price.*

Or, to put it a different way: the fact that at point **f** she has a lot of Learning means that the marginal utility of the Learning that she would give up to get some more Living is not very large. So the opportunity cost to Keiko of trading some Learning or some Living is low.

The marginal rate of substitution is the maximum amount of y that Keiko can give up to get a small unit more of x without lowering her utility. The marginal rate of substitution is also the amount of y that Keiko would view as *substitute* for losing a small unit of x. The marginal rate of substitution should be read as "units of good y per unit of good x."

We show in M-Note 3.2 that the marginal rate of substitution is equal to the ratio of the marginal utilities of the two goods:

$$mrs(x,y) = \frac{u_x(x,y)}{u_y(x,y)} \tag{3.1}$$

This is true because the amount of y that compensates Keiko for a small loss of x is the ratio of her marginal utility of x, which tells us how much she misses the x she has lost, to the marginal utility of y, which tells us how much she appreciates the compensating gain in y.

Equation 3.1 also tells us something more about Figure 3.8. We can use the idea of marginal utilities to understand Keiko's marginal rate of substitution at **f**, **i** and **g** in Figure 3.8.

At point **f** on indifference curve u_1^A:

- The marginal utility of x is high because Keiko has very little x.
- Conversely, the marginal utility of y is low because Keiko has a lot of y.
- Therefore, her $mrs(x,y) = \frac{u_x}{u_y}$ is *large* and as you have already seen she is willing to give up a lot of Learning to get a bit more Living.

But at point **g** on u_3^A the opposite is true:

- The marginal utility of x relative to the marginal utility of y is lower because Keiko has a lot of x than she had at point **f**.

- Therefore, her $mrs(x,y) = \frac{u_x}{u_y}$ is *smaller* and she is willing to give up very little Learning to get more Living. This is why the indifference curve is flatter at point **g** than at point **f**

This shows that because of diminishing marginal utility, for a given amount of Learning, the more hours of Living Keiko has, the less amount of Learning she is willing to give up to get another unit of Living.

The same reasoning shows (and Figure 3.8 confirms) that if Keiko's preferences exhibit *diminishing marginal utility* for both *x* and *y*, her marginal rate of substitution of y for x declines starting at point **f** as we consider points on an indifference curve having more x and less y.

M CHECK When considering two *goods* – things that people value positively, like data and coffee, or living and learning – the indifference curves are downward sloping. That is, they have a negative slope. The negative of the slope of an indifference curve is just its slope with the sign changed.

M-Note 3.2: The mrs is the ratio of marginal utilities

To derive the marginal rate of substitution using calculus, we use the method of *total differentiation* (covered in the Mathematical Appendix). First of all, along an indifference curve the amount of utility is a constant, $u(x,y) = \bar{u}$.

So, to find the slope of the indifference curve we ask what changes in the quantities of x and y (one increasing the other decreasing) are consistent with $u(x,y)$ not changing. This is what total differentiation tells us. The reason is that when we totally differentiate the utility function with respect to its arguments we express the change in Keiko's utility as the sum of the changes due to changes in her consumption of each good.

The total derivative of this equation is:

$$du = u_x(x,y)dx + u_y(x,y)dy = d\bar{u} = 0$$

Since her utility is constant on an indifference curve by definition, the change in her utility is zero.

Recall, too, that the derivative of a constant like $\bar{u}$ is 0.

We can now re-arrange equation 3.2 to find the $mrs(x,y)$:

$$\begin{aligned} &\text{Subtract } u_y(x,y)dy \text{ from both sides} & -u_y(x,y)dy &= u_x(x,y)dx \\ &\text{Divide by } u_y(x,y) \text{ and } dx & mrs(x,y) &= -\frac{dy}{dx} = \frac{u_x(x,y)}{u_y(x,y)} \end{aligned} \quad (3.2)$$

As a result, the negative of the slope of the indifference curve $-\frac{dy}{dx}$ is equal to the ratio of the marginal utilities of the goods, $\frac{u_x(x,y)}{u_y(x,y)}$. But the negative of the slope of the indifference curve is the marginal rate of substitution of y for x, so we have shown that the marginal rate of substitution is the ratio of the marginal utilities.

The mrs has the dimensions of an amount of good y per unit of good x because the marginal utility of y has the dimensions utility per unit y, and the marginal utility of x has the dimensions utility per unit x.

The mrs and the willingness to pay

The marginal rate of substitution provides us with an essential piece of information. Imagine that Keiko had some bundle (x,y) and she were offered the following exchange – trade away some of her y in order to get more x. The mrs tells us the greatest amount of y that she would be willing to give up to get one more unit of x in such a trade. This why we call the mrs the **willingness to pay** y to get more x.

Why does the mrs tell us her *maximum* willingness to pay? The answer is that if in return for another unit of x she gave up an amount of y equal to the mrs she would be moving from one point on her indifference curve to another point *on the same indifference curve*. This is how we constructed the mrs. So she would be no better off after the trade than before.

She would happily pay less than the mrs to get one more unit of x because this would increase her utility (put her on a higher indifference curve). But she would *not* pay more. This is why we call the mrs the maximum *willingness to pay*.

Before going on to the constraints facing Keiko we will now show how what you have learned so far can be used with an explicit mathematical function.

3.5 *Application: Homo economicus with Cobb-Douglas utility*

In Chapter 2, we saw that people may be some combination of preferences including self-regarding altruistic, fair-minded, reciprocal, spiteful, and so on. Representing these preferences mathematically requires knowledge of what Keiko values including:

- How important to her are Learning and Living?
- Is her own Living and Learning all she cares about, or does she value other people's Living and Learning.

In this section we study the preferences of a self-regarding Keiko: she does not care about the Living and Learning of others. We use what is called a Cobb-Douglas function utility function to illustrate how we can model the difference it makes what value she places on the two elements in her choice bundle.

HISTORY The Cobb Douglas function is named after the economist and later U.S. Senator Paul Douglas and his then Amherst College colleague, mathematician and economist Charles Cobb, who jointly came up with the function in 1928 for an econometric study of the contributions of labor and capital goods to output in the U.S. economy.

Here is a Cobb-Douglas utility function:

$$u(x,y) = x^{\alpha} y^{(1-\alpha)} \tag{3.3}$$

The size of α, which is a positive number less than 1, is a kind of baseline measure of how much the individual values x independently of how much x and y she has. In M-Note 3.4 we show that if Annette is consuming the same amount of x and y then the maximum number of units of y that she would be willing to pay for one unit of x is $\frac{\alpha}{1-\alpha}$. So if $\alpha = 0.4$ she her willingness to pay for a unit of x would be $\frac{0.4}{0.6}$ or two-thirds of a unit of y. We also show in Chapter 7 that the fraction of a utility-maximizing consumer's budget that will be spent on good x is α. The fraction spent on y will be $1 - \alpha$.

When a person's preferences are described by a Cobb-Douglas utility function then as long as the Keiko has some of each good, $x > 0$ and $y > 0$, the following will be true:

- her utility $u^{CD}(x,y) > 0$ is positive, and
- her utility increases as she consumes more of either good x or y, meaning that the marginal utility of both goods is positive.

Because the marginal utilities for both goods is positive, Keiko will select a bundle with more of each over one with less of either if both bundles are available to her.

Here is an example of a Cobb-Douglas utility function where a consumer, Annette from earlier, has a stronger preferences for y than for x because $\alpha = 0.4$ and $(1-\alpha) = 0.6$.

$$u^{CD}(x,y) = x^{0.4}y^{0.6} \tag{3.4}$$

Let's assume that x is kilograms of coffee and y is gigabytes of data as we did earlier. Because of the values of α and $(1-\alpha)$, Annette has a stronger preference for data than for coffee because $\alpha = 0.4 < 0.6 = (1-\alpha)$.

When $\alpha > 0$ and $(1-\alpha) > 0$, Cobb-Douglas utility functions have the property that a bundle must include some of both goods to be assigned a positive utility, so we consider cases in which the person has $x > 0$, $y > 0$.

M-Note 3.3: Cobb-Douglas Diminishing Marginal Utility

How do we check that marginal utility is diminishing? Let us examine the marginal utility of Living in the Cobb-Douglas utility function. For the moment, we keep the function general with α:

$$\text{Utility Function} \qquad u(x,y) \quad = \quad x^{\alpha}y^{1-\alpha} \tag{3.5}$$

To find the marginal utility of x we differentiate Equation 3.5 with respect to x:

$$\text{Marginal utility of x} \qquad u_x = \frac{\partial u}{\partial x} \quad = \quad \alpha x^{\alpha-1}y^{1-\alpha} \tag{3.6}$$

For $0 < \alpha < 1$, the marginal utility of x is positive, that is $u_x > 0$. Why? x and y are both positive, as is the parameter α, as is the exponent $1-\alpha$. The exponent $\alpha - 1 < 0$, but this simply means that x can be read as being in the denominator of the marginal utility. For example, for $\alpha = 0.6$, the marginal utility of x is:

$$u_x \quad = \quad 0.6\frac{y^{0.4}}{x^{0.4}} = \alpha\frac{y}{x} \tag{3.7}$$

You can see from Equation 3.7 that the larger is x the smaller will be the marginal utility of x. To confirm that the marginal utility of x is diminishing, we need to differentiate the marginal utility of x with respect to x. That is, we need to find the second derivative of the utility function with respect to x, $\frac{\partial u^2}{\partial^2 x}$, that is, to partially differentiate Equation 3.6 with respect to x.

$$\text{Change in } u_x \qquad \frac{\partial u^2}{\partial^2 x} = (\alpha)(\alpha-1)x^{(\alpha-2)}y^{(1-\alpha)} < 0$$

Because $0 < \alpha < 1$, $\alpha - 1 < 0$. Therefore, $\alpha(\alpha-1) < 0$. Therefore, the rate of change of the marginal utility with respect to x is negative (marginal utility is diminishing), or what is the same thing: utility increases at a decreasing rate as x increases.

Checkpoint 3.4: Positive utility for x and y

Consider Annette's consumption of coffee and data as described by Equation 7.14.

a. Sketch a map of three indifference curves for Annette based on her utility function.

b. Confirm that any bundle with positive consumption of coffee and data ($x > 0$ and y> 0), is assigned a positive utility.

c. How would you confirm whether consuming one more coffee or data increases Annette's utility?

d. If either (both) coffee and data increase her utility, at what rate does Annette's utility change for changes in her consumption of coffee or data? (**Hint:** Mathematically, think through how you would find a "change in a change").

M-Note 3.4: Cobb-Douglas Coffee & Data

We can derive the marginal rate of substitution for the general Cobb-Douglas utility function we defined earlier for coffee and data. Remember that along an indifference curve utility is a constant, such as $\bar{u} > 0$. When we find the "change" of a constant like $\bar{u}$, that change is *zero*, therefore $du = 0$ as in the set of equations below.

$$\begin{aligned} u^{CD}(x,y) &= x^{\alpha}y^{(1-\alpha)} = \bar{u} \\ du = u_x(x,y)dx + u_y(x,y)dy &= 0 \end{aligned}$$

To find the marginal rate of substitution, we need to find the marginal utilities of x and y. Consequently, we differentiate the utility function with respect to x to find u_x, the marginal utility of coffee, and with respect to y to find u_y, the marginal utility of data.

$$u_x = \alpha x^{\alpha-1}y^{(1-\alpha)} \tag{3.8}$$

$$u_y = (1-\alpha)x^{\alpha}y^{(1-\alpha)-1} \tag{3.9}$$

We substitute the marginal utilities from equation 3.8 and 3.9 into the definition of marginal rate of substitution, $mrs(x,y)$ to find the formula for the marginal rate of substitution.

$$\begin{aligned} mrs(x,y) = -\frac{dy}{dx} &= \frac{u_x(x,y)}{u_y(x,y)} \\ &= \frac{\alpha x^{\alpha-1}y^{(1-\alpha)}}{(1-\alpha)x^{\alpha}y^{(1-\alpha)-1}} \\ \text{Factorize out } x^{-1} \text{ and } y^{-1} \qquad &= \frac{\alpha x^{\alpha}x^{-1}y^{(1-\alpha)}}{(1-\alpha)x^{\alpha}y^{(1-\alpha)}y^{-1}} \end{aligned}$$

Remember that $x^{-1} = \frac{1}{x}$ and $\frac{1}{y^{-1}} = y$ and cancel the terms $\frac{x^{\alpha}}{x^{\alpha}}$ and $\frac{y^{(1-\alpha)}}{y^{(1-\alpha)}}$:

$$mrs(x,y) = \frac{\alpha}{(1-\alpha)}\frac{y}{x} \tag{3.10}$$

For example, equation 3.10 shows that if Annette is consuming the same number of gigabytes of data and kilograms of coffee (say, 5 each) she will evaluate them at ratio $\frac{\alpha}{(1-\alpha)}$. The preferences for each good (α and $(1 - \alpha)$) determines the ratio at which Annette is willing to trade data for coffee, together with the amount of coffee and data she is actually consuming. You can see that if Annette had a different level of current consumption of the two goods, say, more x and less y her mrs would be lower.

Checkpoint 3.5: Diminishing mrs as $\frac{y}{x}$ falls

Go back to Figure 3.5 and explain why the mrs is lower at point **c** than at point **b**, and lower at point **d** than at point **b**. Can you say if the mrs is lower at point **a** than at point **d**? Or at point **e** than at point **a**? Equation 3.10 may help you answer this question.

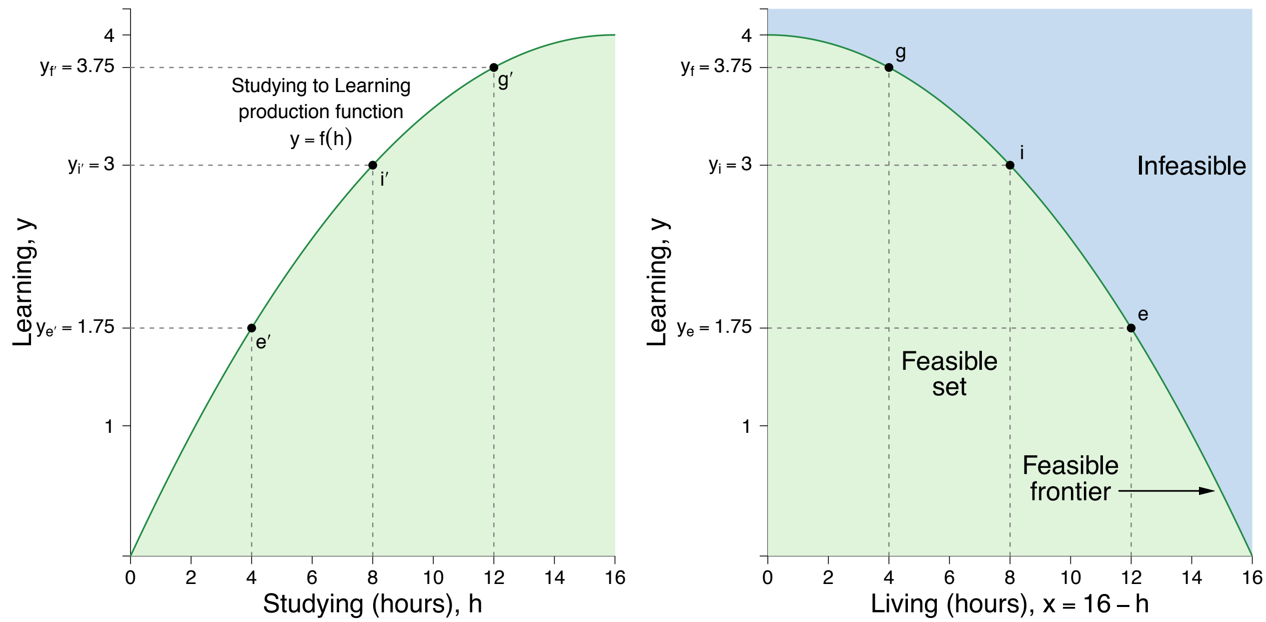
3.6 The feasible set of actions

Keiko's preferences and the resulting utility numbers she assigns to each bundle are a reflection of what she wants to achieve, what her goals are. But her preferences do not tell us what she can *feasibly* obtain. To understand the bundles that are feasible for her, we need to know how she obtains Learning from spending her time Studying. We suppose that Keiko sleeps 8 hours every night and she is not considering changing that. Her choice is what she will do with the 16 hours in the rest of the day .

The relationship between the time Keiko spends studying and the amount of learning she achieves is given by a function that shows for the time (in hours) spent Studying (h), how much Learning (y) results, $y = f(h)$. This is what economists call a **production function** – a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand and the maximum quantity of output that the given amount of input allows. Production functions are more often used to study things other than success in coursework, that is outputs such as meals served, lines of code written, or bushels of corn harvested.

PRODUCTION FUNCTION A production function is a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand and the maximum quantity of output that the given amount of input allows.

Keiko's production function is depicted in Figure 3.9 a. From it you can see that to obtain Learning a Keiko must spend hours (h) Studying. Up to a maximum of 16 hours she can increase her learning by studying more. But starting from studying just a few hours, doubling the amount of studying she does does not double her Learning. We can see this by comparing points **e'**, **i'** and **g'**. Four hours of study ($h = 4$), gets Keiko $y = 1.75$ points of Learning, as



(a) The production of Learning by Studying

(b) The feasible frontier of Living and Learning

Figure 3.9: **Production of Learning by Studying and the feasible frontier of Living and Learning.** Points **e'**, **i'** and **g'** on the production function show combinations of hours of Study and the maximum amount of Learning she could accomplish in that time. Point i' for example shows that if she studies 8 hours she could attain learning equal to 3 (she could also attain less if she did spent the "studying" time texting with friends.) The amount of Living that she can have is her 16 hours minus the time she spends learning, i.e. $x = 16 - h$, as shown in panel b. Panel b. shows the feasible frontier (dark green curve), which is the border of the feasible set (shaded in green). The feasible frontier is just a flipped version of her production function. Points beyond the feasible set (shaded in blue) are *infeasible* or *infeasible* given the number of ours in the day and her Learning production function. In this figure the equation for the feasible frontier is given by: $y = 4 - \frac{1}{64}x^2$.

shown by point **e'**. But 8 hours gets her just 3 units of Learning, far from a proportional increase. This is because if she has just 4 hours she focuses on the really important key points, while if she has 8 hours she gets into the details, which add to her Learning, but not as much as the key ideas.

Keiko's learning production function illustrates an important common economic phenomenon: diminishing marginal productivity. The marginal productivity of hours studying is the effect of a small increase in studying time on the resulting Learning. As you can see from the fact that the production function in figure 3.9 is flatter for more hours of study, this marginal productivity of Studying hours is diminishing.

This is similar to diminishing marginal utility. Just as the person satisfies her most pressing needs if she has very limited expenditures, but can turn to frills if she has more to spend, Keiko focuses on the essential points if her study time is limited but can turn to the examples and further illustration if she has more time to spend.

Because Keiko has two ways to use her time – Studying or Living – and her waking hours are just 16 we know that:

$$\text{Hours of Living} = 16 \text{ Hours } - \text{Hours of Studying}$$

This makes it clear that

- She has just one decision to make not two: if she chooses hours of Studying, that also determines her Hours of Living.

- Because more time living means less time studying this means that the opportunity cost of living more is some amount of learning less.

FEASIBLE FRONTIER The feasible frontier is the border of the feasible set, showing for any value of x the maximum value of y that is feasible, meaning, that the decision-maker can obtain.

To see what this opportunity cost is see Figure 3.9 Panel b, showing the feasible set of outcomes that Keiko might experience. The feasible frontier shown there is the mirror image of the production function in the panel a. The horizontal axis is no longer Studying hours but instead 16 minus Studying hours, which is the amount of Living she can have for each level of Studying she chooses.

At **e**, Keiko can Study for 4 hours, which means she is Living for 12 hours at **e** and her Learning is 1.75. Or (point **g**) she could study for 12 hourw and have learning of 3.75. All of the points like e, g, i and the rest of the feasible frontier are choices that she could make.

The **feasible set** is the area bounded by the feasible frontier and the x and y axes composed of all combinations of Living and Learning that she could experience.

M-Note 3.5: A Living-Learning feasible frontier

A mathematical expression for the feasible frontier is:

$$\text{Feasible frontier} \qquad y = \bar{y} - c(x) \tag{3.11}$$

The parameter $\bar{y}$ is the maximum amount of y when $x = 0$, the y-intercept of the feasible frontier: if $c(0) = 0$ then $y = \bar{y}$. The term $c(x)$ is the cost of x, that is, how many units of y (Learning) one must give up to get the value of x (Living) that she chooses.

Suppose Keiko's feasible frontier between Living, x, and Learning, y, is described by the relation:

$$y = 4 - \frac{1}{64}x^2 \tag{3.12}$$

The negative of the slope of the feasible frontier, $-\frac{dy}{dx} = \frac{1}{32}x$.

Checkpoint 3.6

Redraw Figure 8.3 to show a new situation in which either:

a. Keiko discovers that by changing her diet she can get by perfectly well on 7 hours of sleep.
b. She transfers to a new university where it's more difficult to get high grades.

3.7 The marginal rate of transformation and opportunity cost

Turning to Figure 3.9 we can also contrast two points on the feasible frontier in, such as points **a** and **b**. At point **a**, Keiko spends 14 hours studying and attending classes and has 2 hours left over for Living, with the result of a lot of Learning and not so much Living. At point **b**, Keiko spends 2 hours studying

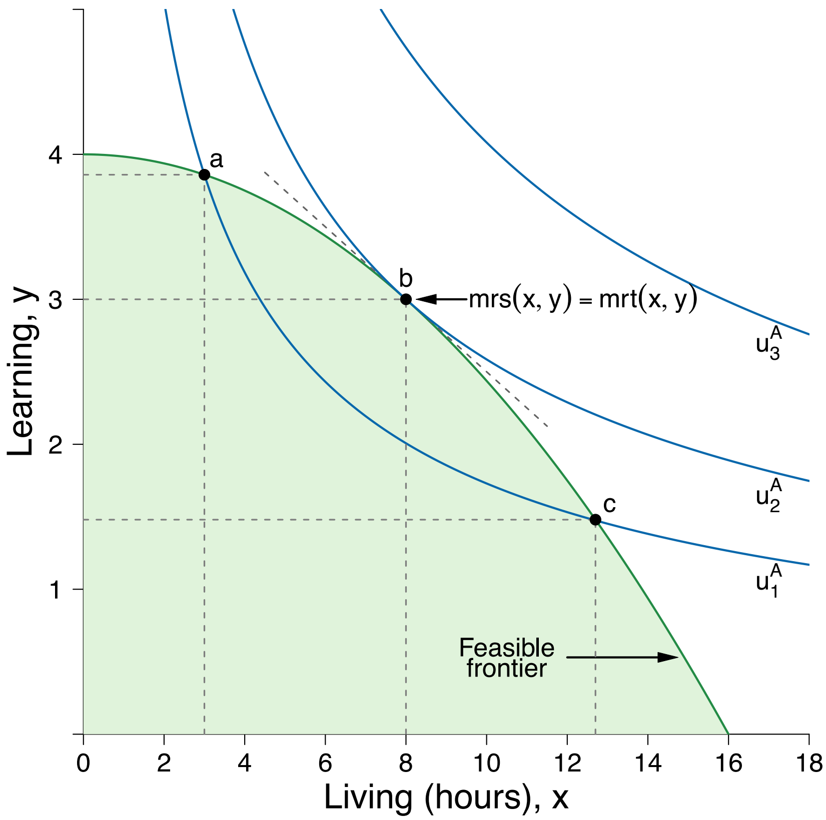


Figure 3.10: **Utility maximization: Living and Learning.** Keiko's feasible frontier for Living and Learning is shown in green. Three of her indifference curves are shown by u_1^A, u_2^A and u_3^A in blue ($u_3^A > u_2^A > u_1^A$). She maximizes her utility at the point on her feasible frontier on the highest indifference curve, that is, at point **b** (where the two are tangents). At point **b**, she maximizes her utility where the marginal rate of substitution equals the marginal rate of transformation by choosing to spend 8 hours Living which gives her a subjective Learning score of 3.

and attending classes and has 14 hours left over for Living, but her Learning is lower than at point **a**. The difference between the two points on the feasible frontier illustrates another trade-off that is central to Keiko's choice: more living means less learning. And vice versa.

If we apply the same reasoning to very small differences of the two goods, we can see that the opportunity cost in less learning that is required to get more living is the negative of the slope of her feasible frontier at that point, namely $-\frac{\Delta y}{\Delta x}$. This is termed the **marginal rate of transformation** or $mrt(x, y)$.

HISTORY: FREE LUNCH The 1975 collection of essays by Nobel Laureate Milton Friedman titled *There's No Such Thing as a Free Lunch: Essays on Public Policy* popularized the idea that there is an opportunity cost to having more of anything that we value.[2]

MARGINAL RATE OF TRANSFORMATION The marginal rate of transformation is the negative of the slope of the feasible frontier. It measures the sacrifice of the y-good necessary in order to get more of the x-good. It is therefore the opportunity cost of the x good in terms of the y good.

The marginal rate of transformation is the smallest amount of y that Keiko has to give up to get a small unit more of x. The mrt is therefore Keiko's *opportunity cost* of x in terms of y or the minimum amount of y she has to sacrifice in order to get a small unit of x. The interpretation of the mrt as the opportunity cost of the x-good plays a major role in the reasoning in this book.

$$\text{Opp. cost of x} \quad = \text{ - Slope of feasible frontier} = -\frac{\Delta y}{\Delta x} = mrt(x, y)$$

The marginal rate of transformation should be read as "units of good y per unit of good x." The term "transformation" is used because we think of a movement *downwards* and to the *right* along the feasible frontier as hypothetically transforming (giving up) the y-good into (having more of) the x-good. There is nothing actually being transformed.

You can also interpret the negative of the slope of the feasible frontier as how much Learning you can get by giving up one unit of Living.

The opportunity cost of the x-good in therms of the y-good differs depending how much of each good Keiko has. If her feasible frontier exhibits an *increasing marginal rate of transformation* as is shown in the figure, then her marginal rate of transformation of y for x increases as she moves along the feasible frontier toward having more x and less y. In this case, Keiko has to sacrifice more y for x the more x and the less y she has.

As Figure 3.9 showed, between the y-intercept and point **g**:

- She needs to give up relatively little Learning (0.25 points) to get four hours of Living.
- Keiko's feasible frontier is relatively flat.
- Her marginal rate of transformation is therefore small.
- Therefore the opportunity cost of Learning for Living is low.

Between point **e** and the x-intercept, however, the opposite is true:

- She must give up a large amount of Learning (1.75 points) to get an additional 4 hours of Living.
- Her slope is steeper.
- Her marginal rate of transformation is higher.
- The opportunity cost of Learning for Living is greater.

Keiko's feasible frontier demonstrates an increasing marginal rate of transformation, which is to say increasing opportunity costs of Learning, moving from the y-intercept down the curve (left to right) to the x-intercept. This occurs because of diminishing marginal productivity of studying time in producing learning.

M-Note 3.6: The Marginal Rate of Transformation

Using Figure 3.9, we saw that the marginal rate of transformation was the negative of the slope of the feasible frontier. Suppose the feasible frontier is described by the equation:

$$y = 4 - \frac{1}{64}x^2$$

$$mrt = -\frac{dy}{dx} = \frac{1}{32}x$$

In this case the mrt increases with x, and exhibits the property of increasing marginal rate of transformation, or increasing opportunity cost. If you substitute the values we used earlier, for four hours, eight hours and 12 hours of living, we can evaluate the $mrt(x,y)$ at each point:

- $mrt(x=4) = \frac{1}{32}(4) = \frac{1}{8}$
- $mrt(x=8) = \frac{1}{32}(8) = \frac{2}{8} = \frac{1}{4}$
- $mrt(x=12) = \frac{1}{32}(12) = \frac{3}{8}$

As x increases, the marginal rate of transformation increases, illustrating the idea of increasing opportunity cost.

3.8 Constrained utility maximization: The $mrs = mrt$ rule

CONSTRAINED UTILITY-MAXIMIZATION
An outcome bundle (x,y) is constrained utility-maximizing if there is no point in the feasible set that is on a higher indifference curve.

From the feasible frontier we know that when maximizing her utility, the limited time in Keiko's day creates a *trade-off*. By combining the insights of feasible frontiers and indifference curves – as in Figure 3.10 – we can understand how Keiko will manage this tradeoff

- *Constraints:* She can choose some point on or within her feasible frontier given by her production function and the limits of her time.
- *Preferences*: From among the points in her feasible set, she will prefer the outcome bundle with the highest utility, meaning on the highest indifference curve.

To understand Keiko's **constrained utility-maximizing** problem, we contrast points **a**, **b**, and **c** in Figure 3.10. An outcome bundle (x,y) is *constrained utility-maximizing* if there is no other point in the feasible set with a higher utility.

Point **a** is on Keiko's feasible frontier and lies on indifference curve u_1^A. But, **a** is not constrained utility-maximizing because Keiko could increase her utility by increasing her Living time and decreasing her Learning, by moving along the feasible frontier to the southeast. By similar reasoning point **c** cannot be the highest indifference curve she can reach.

Keiko's constrained utility-maximizing point is **b** in Figure 3.10, the point on the feasible frontier that is on the highest indifference curve. We label it **b** because it is the point where Keiko does the **b**est she can.

Figure 3.10 suggests a useful way to think about Keiko's constrained utility-maximization problem. In the figure, we see that the constrained utility-maximizing bundle is the point where Keiko's indifference curve is tangent to her feasible frontier. This means the indifference curve and the feasible frontier have the same slope at the constrained utility-maximizing point.

The slopes of the indifference curve and the feasible frontier express *trade-offs* between the two goods. This is the basis of what we call the $mrs = mrt$

rule.

M-Note 3.7: Equating mrs to mrt to find the constrained maximum

Suppose Keiko's utility for Living (x) and Learning (y) is described by a Cobb-Douglas utility function with parameter $\alpha = 0.4$ and $(1-\alpha) = 0.6$:

$$u(x,y) = x^{0.4}y^{0.6}$$

We find her marginal rate of substitution by finding the marginal utilities and substituting them into the equation $mrs(x,y) = \frac{u_x^A}{u_y^A}$

$$\begin{aligned} u_x^A &= 0.4(x^A)^{-0.6}(y^A)^{0.6} \\ u_y^A &= 0.6(x^A)^{0.4}(y^A)^{-0.4} \\ mrs(x,y) = \frac{u_x^A}{u_y^A} &= \frac{2}{3}\frac{y}{x} \end{aligned} \tag{3.13}$$

Suppose her feasible frontier is described by the equation:

$$y = 4 - \frac{1}{64}x^2$$

Keiko's constrained maximum must be on her feasible frontier. We find her marginal rate of transformation by differentiating y with respect to x:

$$-\frac{dy}{dx} = mrt(x,y) = \frac{1}{32}x$$

To find Keiko's constrained maximum, we use the two expressions above for mrs and mrt, equating them to find a point on the feasible frontier consistent with the $mrs = mrt$ rule:

$$\frac{2}{3}\frac{y}{x} = \frac{1}{32}x \tag{3.14}$$

Then multiplying through by $\frac{3}{2}x$:

$$\begin{gathered} y = \frac{3}{64}x^2 = 4 - \frac{1}{64}x^2 \\ x^2 = 64 \\ x = 8 \\ y = 3 \end{gathered}$$

Keiko spends 8 hours on Living, studies 8 hours, and achieves a Learning level of 3.

Doing the best you can: The $mrs = mrt$ rule

Summarizing the results so far, in Figure 3.10

1. The negative of the slope of the feasible frontier is the *opportunity cost of getting a unit more more of the x good*, in terms of the amount of the y good forgone.

2. The negative of the slope of an indifference curve is a measure of the person s *willingness to pay for a little more of the x good* in terms of how much of the y-good she would be willing to give up to get an additional unit of the x good.

Using these two statements we can see why point **a** in in Figure 3.10 could not be the utility maximizing outcome bundle. The indifference curve is steeper than the feasible frontier, so the value of getting more living exceeds the associated opportunity cost (2 above is grater than 1) So she could do better by giving up some Learning in favor of more Living.

REMINDER: mrs AND mrt

- mrs, the *marginal rate of substitution*, is the negative of the slope of an indifference curve.
- mrt, the *marginal rate of transformation*, is the negative of the slope of the feasible frontier.

The opposite is true at point **c**: the feasible frontier is steeper than the indifference curve, so the opportunity cost of having more Living falls short of the value of an additional unit of Living. So she definitely would not want to give up more Learning to get more Living.

In fact, it means the opposite. By giving up a unit of Living she would get a substantial increase in Learning (that is what the steep feasible frontier means). Giving up a unit of Living could be compensated by a *modest* increase in Learning (that is what the flatter indifference curve means). So the *benefits* of giving up some Living in return for more Learning *outweigh the cost*. So any point like a and c where the feasible frontier and the indifference curve intersect cannot be the constrained utility maximizing output bundle. This gives us the $mrs - mrt$ rule: The the utility maximizing output bundle is a point where

THE $mrs = mrt$ RULE In many of the models that we consider in the remainder of this book, the constrained utility-maximizing outcome is a point on the feasible frontier at which an indifference curve representing the trade offs between the decision maker's objectives is tangent to the feasible frontier representing the opportunity costs of having more of one good in terms of the amount of the other good foregone. This is the point where the marginal rate of substitution is equal to the marginal rate of transformation.

$$\begin{aligned}
\textit{Slope of feasible frontier} &= \textit{Slope of indifference curve} \\
\text{which requires that:} & \\
\textit{Marginal rate of transformation} = mrt &= mrs = \textit{Marginal rate of substitution} \\
\text{Or, what is the same thing} & \\
\textit{Opportunity cost of x} &= \textit{Willingness to pay for x}
\end{aligned}$$

The rule expresses a simple and true idea: if the opportunity cost of something is less than your willingness to pay you should choose more of it (if you can) and if the opportunity cost is greater than your willingness to pay, you should choose less of it (if you can). But there are cases in which the utility-maximizing outcome bundle is not a tangency of the feasible frontier and an indifference curve:

- It may be that an difference curve is steeper than the feasible frontier, but there is no way to get more of the x good. In this case the slope of feasible frontier does *not* measure the opportunity cost of getting more of the x-good; that is impossible (its cost is infinite). The utility maximizing outcome bundle at point **b** in Panel b of Figure 3.14 an example of a case – called a corner solution – there the $mrt = mrs$ rule does not work.
- We show in M-Note XX that there are conditions under which a bundle such that $mrt = mrs$ can also be a minimum not a maximum. We provide an example of this in Chapter 6.

Checkpoint 3.7: Changes in Keiko's preferences

Find Keiko's constrained utility-maximizing level of Living and Learning when $\alpha = 0.6$ and $(1-\alpha) = 0.4$, so that she values Living more than Learning.

M-Note 3.8: When the $mrs = mrt$ rule fails

The rule can fail to identify the constrained utility maximum under two conditions: when the maximum is a corner solution (so the rule is not satisfied) and when the rule is satisfied at a minimum rather than a maximum. Positing a case with diminishing opportunity cost of obtaining one good in terms of the other good foregone will illustrate both cases

Setup. Assume that a person's utility varies with the amount of goods x and y:

$$u(x,y) = x+y$$

and the feacible amount of good y is a function of good x:

$$y(x) = (1-x)^2 \tag{3.15}$$

The rule may select a minimum, not a maximum.The marginal rate of substitution and marginal rate of transformation are:

$$\begin{aligned}
mrs(x,y) = \frac{u_x}{u_y} &= 1 \\
mrt(x,y) = -\frac{dy}{dx} &= 2(1-x) \\
\text{Equyating the } mrs \text{ and } mrt \quad 2(1-x) = &= 1 \\
x* &= \frac{1}{2} \\
\text{Using Equation 3.15} \quad y* &= \frac{1}{4}
\end{aligned}$$

Note that using $x*$ and $y*$, the utility is $u = \frac{3}{4}$. Alternatively, we could set $(x,y) = (1,0)$, or $(x,y) = (0,1)$: both allocations are in the feasible set. In both cases, $u = 1$, which is higher than the one that we have reached using the condition $mrs = mrt$.

The condition $mrs = mrt$ will not give the utility maximum if the second order condition is violated: the second derivative of the utility function with respect to the variables must be negative. Let's calculate it, replacing Equation 3.15 into the utility function:

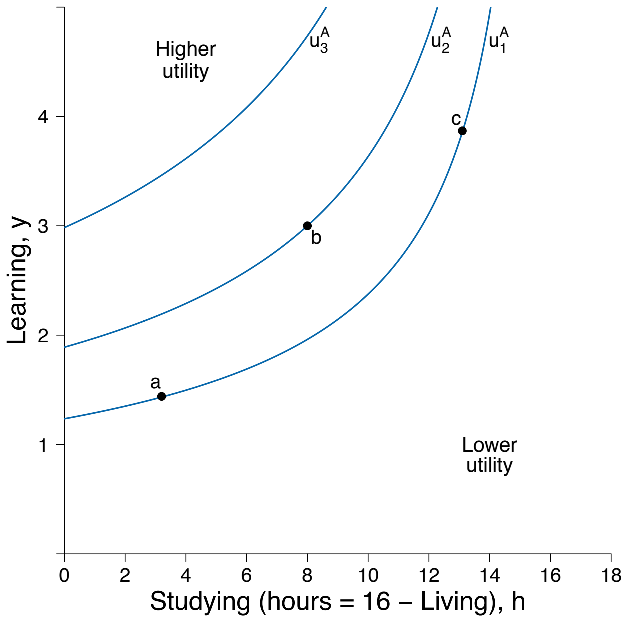
$$\begin{aligned}
u &= x+(1-x)^2 \\
\frac{du}{dx} = u_x &= 1-2(1-x) \\
\frac{d}{dx}\left(\frac{du}{dx}\right) = u_{xx} &= 2 > 0
\end{aligned}$$

The utility maximum may be a corner solution. In the example the utility maximums at both $x = 1$ and $y = 1$ are corner solutions (only one of the goods is consumed.)

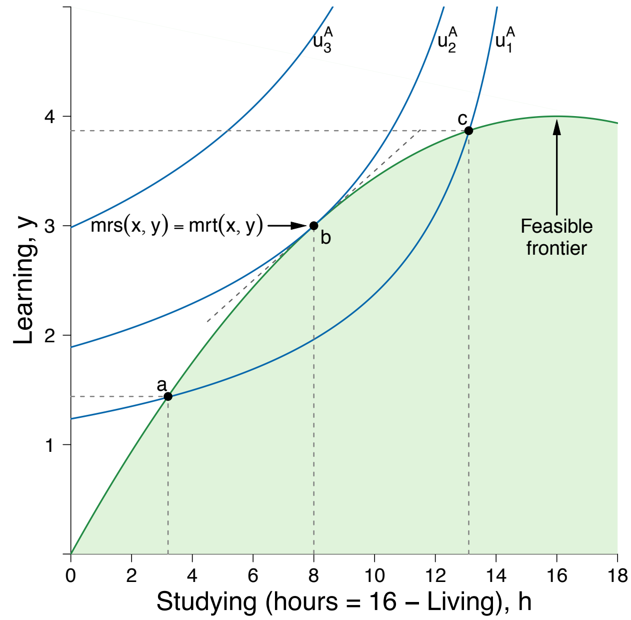
The rule may be inapplicable. Where either the indifference curves or the feasible frontier are not smooth but instead are kinked (are not differentiable), the derivatives on which the *mrs* and *mrt* are based will not exist at some points.

Trade-offs between goods and bads

In many situations it is easier to understand decisions in terms of a trade-off between a good and a bad rather than a trade-off between two goods. Recall that a bad is something that you would prefer to have less of, such as working



(a) Indifference curves of Learning and Studying



(b) The utility-maximizing choices with bads

Figure 3.11: **The $mrs = mrt$ rule:.** Keiko's problem of choosing (h, y) when $h = 16 - x$ = time Studying, is a bad. Studying time, h is plotted on the horizontal axis, and Keiko's Learning, y is plotted on the vertical axis. Keiko's feasible frontier is shown in green in the right-hand panel. Three of her indifference curves are shown by u_1, u_2 and u_3 in blue in both panels. The points **a**, **b**, and **c** are the same as in Figure 3.10. Keiko maximizes her utility at the point on her feasible frontier on the highest indifference curve, that is, at point **b**, choosing to spend 8 hours on Living and 8 hours on Learning.

harder than is comfortable or safe.

For example, Keiko might think of her decision in terms of a trade-off between her time studying time that she does not enjoy, $h = 16 - x$, and her Learning, y. The more time Living the better for Keiko, therefore x is a good. The more time Studying the worse for Keiko, therefore h is a bad. But as before since $x = 16 - h$, choosing (h, y) to maximize utility, $u(16 - h, y)$ is the same thing as choosing x to maximize utility $u(x, y)$. These are just different ways of posing the same problem.

Figure 3.11 shows Keiko's indifference curves and feasible frontier plotted in terms of Study time, h and Learning y. Her indifference curves slope upward because an increase in Studying, h, *lowers* Keiko's utility, and requires an *increase* in Learning, y to compensate in order to stay at the same level of utility. Utility increases as we move to the northwest and decreases as we move to the southeast in this plot.

Similarly, Keiko's feasible frontier slopes *upward*, because an increase in Study time, h, leads to *more* Learning, y. This is her "learning production function" introduced earlier. So the slope of the feasible frontier is the marginal productivity of studying time or $\frac{\Delta y}{\Delta h}$ and this is also the marginal rate of transformation of Study time into Learning. (In this case "transformation" actually describes the process underlying the feasible frontier).

As was the case for tradeoffs between two goods, a bundle in the feasible set is the utility maximizing output bundle if there is no other feasible bundle with greater utility. And this is the bundle for which the $mrs = mrt$ rule holds, namely the point on the feasible frontier where the marginal rate of substitution equals the marginal rate of transformation ($mrs(x,y) = mrt(x,y)$).

M-Note 3.9: The marginal utility of the bad

The utility function for Studying (h) and Learning (y) is given by:

$$u^A(h,y) = (16-h)^{0.4}y^{0.6} \tag{3.16}$$

To find the marginal utility of the "bad," Studying, we need to partially differentiate Equation 3.16 with respect to h. Remember that when we partially differentiate we treat the other variable as a constant, so the term $y^{0.6}$ will simply remain where it is. We only have to think about the h term.

$$\begin{aligned}\frac{\partial u^A}{\partial h} = u_h^A &= (0.4)(-1)(16-h)^{0.4-1}y^{0.6}\\ u_h^A &= \underbrace{-0.4}_{<0}\underbrace{(16-h)^{-0.6}}_{>0}\underbrace{y^{0.6}}_{>0} < 0\end{aligned}$$

The first term is negative whereas the second and third terms are positive. So the marginal utility of hours of study is negative. We call such a utility a *disutility* and will often talk about the *disutility of work* or the *disutility of effort*.

Checkpoint 3.8: Understanding goods and bads

Find Keiko's constrained utility-maximizing level of Study time and Learning, using the $mrs = mrt$ rule.

3.9 *The price-offer curve, willingness to pay, and demand*

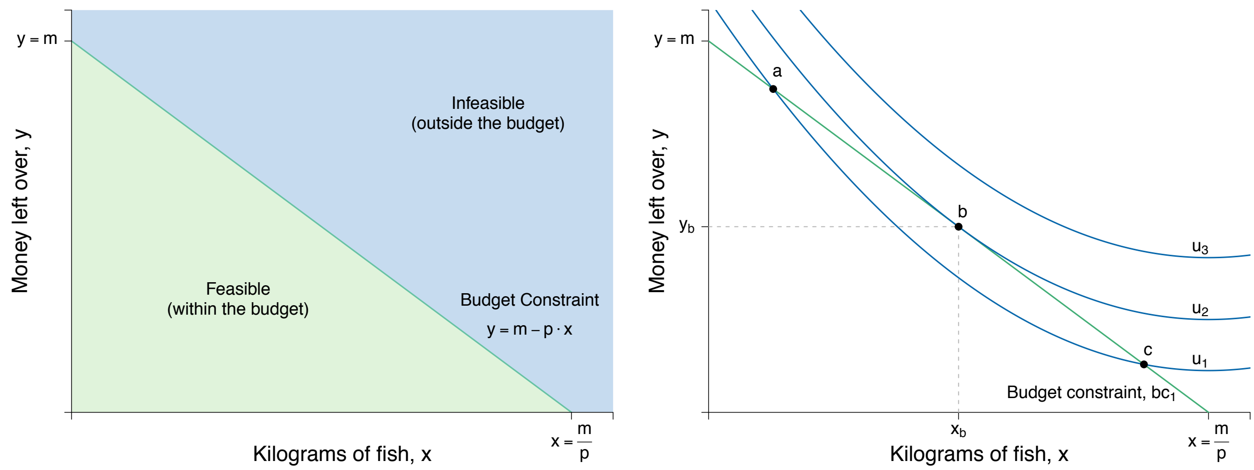
We often want to know how people respond to different options for exchange in the form of prices. We may be interested in knowing, for each price at which she can purchase any amount of the good she pleases, how much Keiko will purchase, namely the utility maximizing amount. This is Keiko's **individual demand curve**.

Remember that in explaining Keiko's indifference curves we asked what is the maximum amount of Learning she would be willing to give up in exchange for more Living. The answer is given by her maximum *willingness to pay*, or what is the same thing her marginal rate of substitution of Learning for Living.

INDIVIDUAL DEMAND CURVEAn individual demand curve (or demand function) indicates for each price that might hypothetically be offered at which a buyer can purchase any amount that they please, the quantity that an individual will purchase.

We now ask almost the same question except that rather than giving up Learning to get more Living, Keiko is now giving up *money* – that is paying for a good according to its *price*.

For each offered price she faces another constrained utility-optimization problem. The demand curve is constructed by a series of hypotheetical constrained optimization problems, one for each possible offered price. Each



(a) The budget constraint

(b) Utility-maximizing choice

Figure 3.12: **Budget constraint and utility-maximizing choice for fish and money for other goods.** The budget set is shaded in green and the budget constraint (feasible frontier) is the dark green line on the border of the budget set (feasible set). Consumption bundles (x, y) in the budget set and on the budget constraint can feasibly be obtained with the current budget (m) at the price, p, for kilograms of fish, x, Outside the budget constraint, in the shaded green area, the bundles of x and y cannot feasibly be obtained with the existing budget. Harriet maximizes her utility subject to her budget constraint bc_1. She maximizes her utility at **b** where her marginal rate of substitution, $mrs(x, y) = \frac{u_x}{u_y}$, equals her marginal rate of transformation or the price ratio of x to y, $mrt(x, y) = p$.

price defines a *feasible set*; its boundary, the feasible frontier, defines the bundles of goods Keiko has access to. For each each of these feasible sets there is a bundle that maximizes her utility. This is a single point on her demand curve.

Indifference curves tell us the utility number that Keiko assigns to each possible consumption bundle. Using this logic, her choice will be the point on the feasible frontier with the greatest utility, which will be the bundle in the feasible set that is on the highest indifference curve. This is a standard constrained utility maximization problem.

The budget constraint and feasible utility-maximizing choices

We shall use one particular kind of feasible frontier to think this through: the **budget constraint**. The budget constraint defines an amount of money m that a person has or has access to, through wealth and credit markets, which constitutes their budget to spend on goods and services. People can use their budget to spend on goods at prices that are given to them. Imagine that you want to buy the fish that Alfredo and Bob were trying to catch in Chapter 5 at a fish market. The price (p) is measured in dollars per kilogram.

BUDGET CONSTRAINT A person's budget constraint gives the bundles (x, y) that just exhausts some given budget at a set of market prices (p) of the goods. The feasible set includes all purchases bundles that do not exhaust the budget, so the budget constraint is the feasible frontier.

Figure 3.12 a. shows the budget constraint for Harriet, someone deciding on how much fish to buy from Alfredo or Bob at price p. The budget set is shaded in green and the budget constraint (feasible frontier) is the dark green line on the border of the budget set (feasible set). Consumption bundles (x, y) in the budget set and on the budget constraint can feasibly be obtained with the current budget (m) at the price, p, for kilograms of fish, x, Outside the budget constraint, in the shaded green area, the bundles of x and y cannot feasibly be obtained with the existing budget.

We know how to find the utility-maximizing bundle for a given feasible frontier – or the budget constraint – with given indifference curves: we apply the $mrs = mrt$ rule finding the bundle where the marginal rate of substitution equals the marginal rate of transformation. We can combine these insights and calculate what the consumer's utility-maximizing bundle will be for every potential price of the good given a fixed budget and when the other good, y, is money for other goods.

Figure 3.12 b shows Harriet maximizing her utility subject to her budget constraint bc_1. To find her utility-maximizing choice, we must apply the $mrs - mrt$ rule to find where her marginal rate of substitution (her willingness to pay in money for kilograms of fish) equals her marginal rate of transformation, here the price for a kilogram of fish. At point **a** she consumes too little of x and too much of y (her marginal utility of money for other goods (y) is much lower than her marginal utility of kilograms of fish (x), or her $mrs(x,y)$ is too high, and she would be better off if she consumed less y and more x. Conversely, at **c**, she consumes too little of y and too much of x (her marginal utility of x is much lower than her marginal utility of y, or her $mrs(x,y)$ is too low, and she would be better off if she consumed less x and more y. She maximizes her utility at **b** where her marginal rate of substitution, $mrs(x,y) = \frac{u_x}{u_y}$, equals her marginal rate of transformation or the price ratio of x to y, $mrt(x,y) = p$.

The demand curve: Utility-maximizing choices at difference prices

PRICE-OFFER CURVE The price-offer curve shows every utility-maximizing consumption bundle at each price of good x. It demonstrates the *principle of demand* by connecting every point where a consumer's indifference curve is tangent to every possible budget constraint for a change in the price of x at given income m. We will use the price-offer curve in 4.

With every change in price, the consumer's budget constraint will pivot.

The budget constraint will pivot upwards as a good's price *decreases*, because a consumer can buy *more* of the good with the same budget. The opposite is true for price increases. As the price of a good *increases*, the same budget buys less of the good, pivoting the budget constraint *inward*.

With every pivot of the budget constraint, at the utility-maximizing point, the new budget constraint will be tangent to a new indifference curve which will be either higher if the price of the good decreases or lower if the price of the good increases.

Because we can calculate the utility-maximizing consumption bundle for each possible price, we can find a curve that records every utility-maximizing consumption bundle for each price, connecting up points **a**, **b**, and **c** in the left panel of the figure. That curve is called the **price-offer curve**. Sometimes, for individual consumers, it is called the *price-consumption curve* because it indicates what the consumer will *consume* at different *prices*.

M-CHECK We can find the equation for the price-offer curve by using the equation for the budget line and combining it with the equation for the marginal rate of substitution. We do not derive it here as it is not required to understand the intuition of the demand curve.

Figure 3.13 maps three different utility-maximizing consumption bundles at three prices of x. With each price decrease, the budget constraint pivots outward from p_1 to p_2 to p_3. With each change in the price of x, the utility-maximizing bundle – the point at which the marginal rate of substitution is

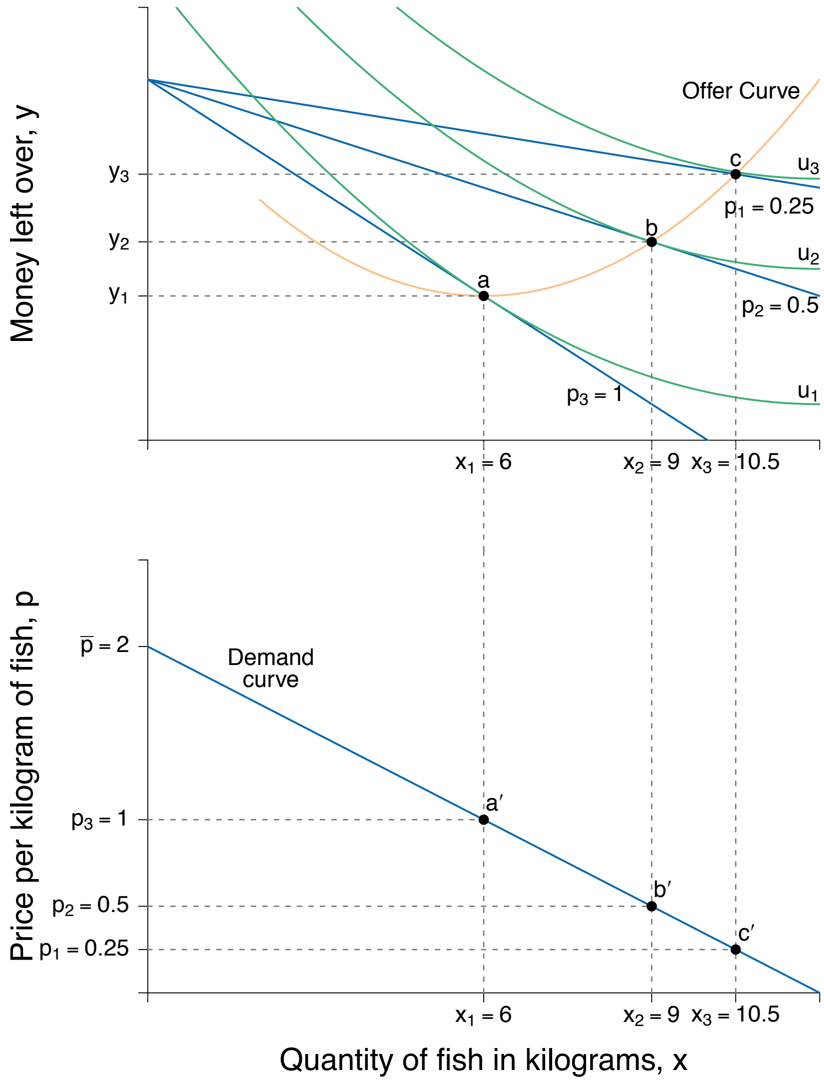


Figure 3.13: **Offer curve and demand curve for fish:** The price of x in the top panel is in terms of the money Jane sacrifices to get more fish. Similarly, in the lower panel the amount of money Keiko must sacrifice to get more fish – the price per unit of fish – determines Keiko's quantity of fish demanded along the demand curve. Points **a**, **b**, and **c** in the top panel correspond to points **a'**, **b'** and **c'** in the lower panel.

equal to the opportunity cost – changes. At $p_3 = 1$, the bundle includes $x = 6$, at $p_2 = 0.5$, the bundle includes $x = 9$, and at $p_1 = 0.25$, the bundle includes $x = 10.5$. With each price change, there is a new bundle for both x and y.

The different bundles suggest a price-quantity relationship between the quantity demanded of x and different prices of x. As the price of x decreases, the quantity demanded increases. In fact, we can take each price-quantity combination and map a demand curve to it. In the lower panel of Figure 3.13, we have taken each utility-maximizing consumption bundle from the different consumption bundles at each price and identified their coordinates on price-quantity axes. The price-quantity combinations provide a downward-sloping demand curve where quantity demanded, x, decreases as its price, p, increases.

Measured horizontally from the vertical axis, it tells us the amount that can be sold to the consumer at each particular price. Measured vertically from the horizontal axis, it also tells us what is the consumer's maximum willingness to pay for each amount on the horizontal axis.

EXAMPLE We demonstrate in Chapter 7 how to find the equation for the demand curve and how to see that a reduction in the price of fish has two effects. First, the lower price leads the person to buy less meat and more fish; this is the *substitution effect*. Second, the lower price also allows the person to buy more of *everything* if she chooses (fish, meat or whatever); this is called the *income effect*.

3.10 Social preferences and utility maximization

The preferences we have looked at so far have been entirely self regarding, depicting a person who is concerned with their choices among bundles that they alone will experience. But people often make choices where they are not the only person affected, where what they choose can benefit or harm someone else. Consider the Dictator Game that we mentioned in Chapter 2. In that game, a person, the Dictator has an endowment of money, y, that they can choose to split between themselves and another person in any way they choose.

Imagine that Annette is the Dictator and she is able to choose whether or not to give to some amount to Ben. As a result, Annette must choose some split of of her endowment $z = \pi^A + \pi^B$, where π^A is the amount in dollars that Annette keeps for herself and π^B is the amount that she gives to Ben. As a result, we can re-arrange the equation to find the equation to the feasible frontier for $y = 10$ dollars:

$$\text{Feasible Dictator Allocations} \qquad \pi^B = 10 - \pi^A \tag{3.17}$$

Looking at Equation 3.17, we can see that the feasible frontier is a straight line with a slope of -1. This tells us that the feasible frontier slopes downward. Remember that the negative of the slope of the feasible frontier is the marginal rate of transformation: so a player in the Dictator Game who wishes to give \$1 to someone else has an opportunity cost of \$1 for doing so. If Annette is like *Homo economicus*, she is purely self-regarding. She sets $\pi^B = 0$ and keeps everything or herself and therefore $z = \pi^A = 10$. What happens when the Proposer is an *altruist* who believes in making an offer of more than zero to a partner?

REMINDER A game is a mathematical representation of a *strategic* interaction, which means one in which players recognize that their payoffs depend on the actions taken by other players. So the so-called Dictator Game is not really a game at all, because the Dictator's payoffs do not depend at all on anything that the other player does.

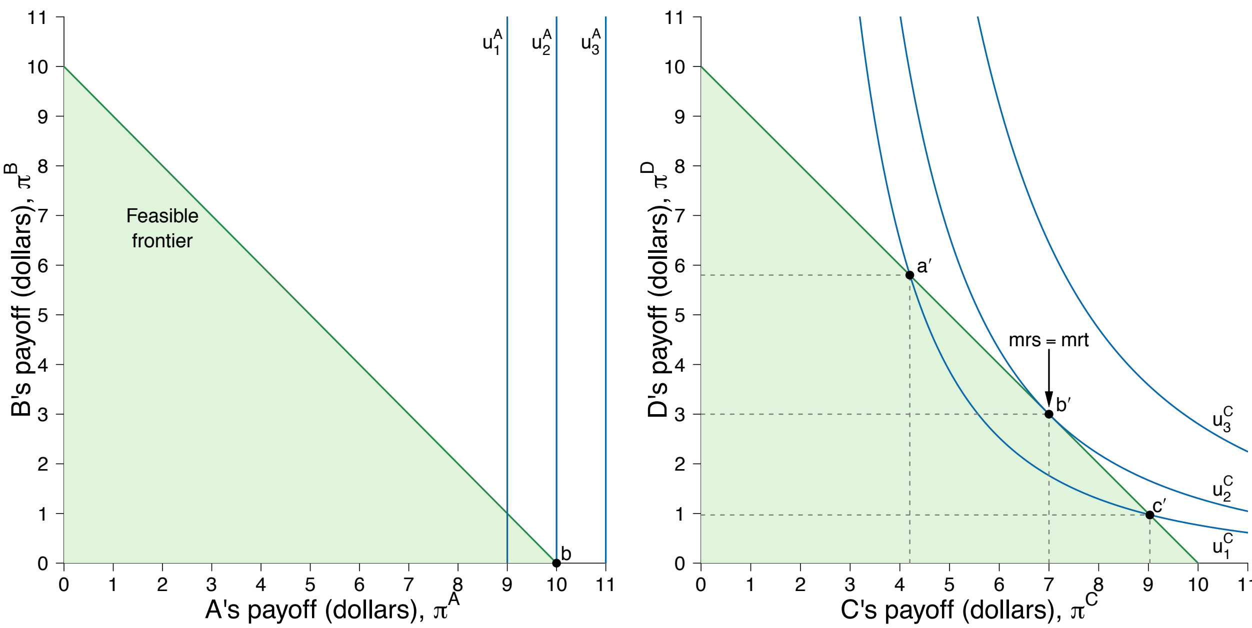
To see what happens in these cases, let us contrast two pairs of people:

- Annette (A) is paired with Ben (B). Annette makes choices about how much money she gets and how much money Ben gets.
- Chen (C) is paired with Diane (D). Chen makes choices about how much money he gets and how much money Diane gets.

To think about the choices that Annette and Chen make, let us consider two different kinds of Cobb-Douglas utility functions that Annette and Chen might

M-CHECK Two things to remember when thinking about Equations 3.18 and 3.19.

- The exponents in the Cobb Douglas utility function Equation 3.19 mean that if they both had the same payoff, then Chen would value increasing his own payoffs more than he would value increasing Diane's.
- Any number raised to a zero exponent is equal to 1, so because Annette does not value Ben's payoffs at all (the exponent is zero) her utility is unaffected by the amount that he gets (her utility is simply how much she keeps for herself).



(a) Self-interested Annette offers a (10, 0) split

(b) Chen is altruistic, offering a (7, 3) split

Figure 3.14: **Utility maximization: Self-interested offer vs. altruistic offer**. Annette offers a split to Ben of (10, 0), whereas Chen offers Diane a split of (7, 3). Annette's indifference curves are vertical because she gives no weight in her utility function to Ben getting any money ($(1-\alpha)=0$), therefore she gets \$10 and Ben gets \$0. Between Chen and Diane, Chen gives some weight to Diane getting money ($(1-\alpha)=0.3$), therefore his indifference curves are shaped like indifference curves we've looked at previously and at his constrained utility maximum Chen gets \$7 and Diane gets \$3. Notice that if Chen gives any less or any more to Diane, then he would be on a lower indifference curve, such as at points **b'** and **a'** on u_1^C.

have.

$$\text{Annette's Utility Function} \qquad u^A(\pi^A,\pi^B) = (\pi^A)^1(\pi^B)^0 \tag{3.18}$$

$$\text{Chen's Utility Function} \qquad u^C(\pi^C,\pi^D) = (\pi^C)^{0.7}(\pi^D)^{0.3} \tag{3.19}$$

Equations 3.18 and 3.19 represent the utility functions of two different people. Chen is other-regarding, he cares about Diane's payoff as is indicated by the positive exponent on her payoff in his utility function, though not as much as he cares about his own (compare the two exponents). Annette is entirely self-regarding, placing a zero weight on Ben's payoff and therefore her choice will depend entirely on the outcome that she experiences.

We display indifference curves for Annette and Chen in Figure 3.14. The indifference curves in panel a are unusual: they are vertical because the only thing that Annette values is what is on the horizontal axis, namely, her payoff. Using the $mrs = mrt$ rule, we find the constrained utility-maximizing point for each person where their highest indifference curve touches the feasible frontier.

In this case, though, the feasible frontier is given by a straight line because it represents a split of money. The maximum amount of money that Annette or Chen can keep is \$10 and they can offer splits in 1 cent increments between themselves and their partners. The vertical intercept corresponds to the instance in which they give all \$10 to their partners. The horizontal intercept corresponds to the instance in which they keep all \$10 to themselves. Chen has preferences such that he would like a 70%-30% split of the \$10 (his

$\alpha = 0.7$) and his highest indifference curve is tangent to the feasible frontier at a (7, 3) split shown by point **b'** in Figure 3.14 b.

Annette has preferences such that she would like a 100%-0% split of the $10 (her $\alpha = 1$, she places zero weight on Ben's payoff) and her highest indifference curve touches the feasible frontier at **b** in Figure 3.14 a at a (10,0) split (she keeps all the money). We can interpret the slope of her indifference curves as her maximum willingness to pay in order to give Ben a small positive payoff, and the ask: how much of her own payoffs would she be willing to give up to transfer a penny to Ben? The answer is that there is no amount, however small, that would motivate her to do this.

But what allocation does she choose? She chooses the highest utility that is within the feasible set. Her highest utility is where her vertical indifference curve u_2^A touches her highest feasible allocation to herself of $10. She keeps all the money. Her keeping all the money shouldn't surprise us because she gives no weight to Ben's payoff. In mathematics, a solution like this is called a *corner solution*. Notice that we couldn't use our standard requirement for finding the constrained utility maximum of $mrs = mrt$. $mrs(\pi^A, \pi^B)$ was undefined because her indifference curves were vertical. But the principle of constrained utility maximization, that Annette would find the point in the feasible set with the highest utility, still applied to our problem and we found the solution.

M-Note 3.10: The mrs for a self-regarding Dictator

Why are Annette's indifference curves vertical in Figure 3.14? To answer this question, we need to find her marginal rate of substitution. To find her mrs, we need the marginal utilities of the two arguments of her utility function: π^A and π^B the money payoffs that Annette and Ben respectively get.

Marginal utility to Annette of Annette's payoff:

$$u^A_{\pi^A} = \frac{\partial u^A}{\partial \pi^A} = 1 \cdot (\pi^A)^{(1-1)}(\pi^B)^0 = 1$$

Marginal utility to Annette of Ben's payoff:

$$u^A_{\pi^B} = \frac{\partial u^A}{\partial \pi^B} = 0 \cdot (\pi^A)^1(\pi^B)^{(0-1)} = 0$$

Therefore Annette's marginal rates of substitution is:

$$\begin{aligned} mrs(\pi^A, \pi^B) &= \frac{u_{\pi^A}}{u_{\pi^B}} \\ &= \frac{1}{0} = \text{undefined} \end{aligned} \tag{3.20}$$

Now, the result of Equation 3.20 should not surprise us because the slope of a vertical line is undefined. Annette's indifference curves endlessly *rise* and have *no run*, so the negative of an undefined number (the slope) remains an undefined number (the mrs). Her indifference map therefore represents a range of vertical lines where the horizontal intercepts correspond to the amount of money she keeps which is also the utility number associated with the particular indifference curve.

Now, we might ask ourselves, what is Annette's utility at her constrained utility maximum? Let's substitute in the values we have for $\pi^A = 10$ and $\pi^B = 0$.

$$\begin{aligned} u^A(\pi^A, \pi^B) &= (\pi^A)^1(\pi^B)^0 \\ &= (10)^1(0)^0 = 10 \end{aligned} \tag{3.21}$$

Annette has a utility that is equal to the amount of money she keeps for herself.

M-Note 3.11: An altruistic person splitting the pie

We will derive Chen's decision about splitting the pie between him and Diane. Using Equation 3.10, his marginal rate of substitution is (see his utility, Equation 3.19):

$$mrs(\pi^C, \pi^D) = \frac{7\pi^D}{3\pi^C}$$

Now, let's assume that the size of the pie is $z = 1$, therefore, the feasible allocations are represented by $\pi^D = 1 - \pi^C$, so his mrt (the negative of the slope of the feasible frontier) is

$$mrt = -\frac{d\pi^D}{d\pi^C} = 1$$

Equating the mrs with the mrt, we can obtain how much Chen allocates to himself and to Diane:

$$\begin{aligned} \frac{7\pi^D}{3\pi^C} &= 1 \\ \pi^D &= \frac{3}{7}\pi^C \\ \text{Using the feasible allocation set} \quad \frac{3}{7}\pi^C &= 1 - \pi^C \\ \frac{10}{7}\pi^C &= 1 \\ \therefore \quad \pi^C &= 0.7 \\ \text{and} \quad \pi^D &= 0.3 \end{aligned}$$

That is why Chen offers Diane $\$3$ of the total of $\$10$ that she is able to allocate.

Checkpoint 3.9: Chen's Choice and the $mrs = mrt$ rule.

1. What is the marginal rate of transformation in this in the game described in Figure Figure 3.14?
2. Why is the utility maximizing outcome bundle at point **b** in Panel b of Figure 3.14 an example of a case there the $mrt = mrs$ rule does not work. How does this case differ from the case shown in Panel b, where the rule does work?
3. Use the value of Chen's mrs at point **c'** in Figure 3.14 Panel b along with the value of the mrt to explain why for Chen the opportunity cost of giving more money to Diane is less than his willingness to pay (give up his own payoffs) so that Diane can have more.

3.11 Application: Environmental trade-offs

We think of environmental damage as something to be avoided, but stopping or slowing the damage – or "abating" the damage in the language of environmental science – is costly. Less damage means some combination of less consumption, changing our consumption patterns to be less damaging to the environment, or diverting our productive potential from producing goods that we can now consume to discovering and installing new technologies. We therefore face a trade-off between consuming goods and maintaining the quality of the environment. How much of these opportunity costs of improved environmental quality are we willing to pay?

The constrained utility maximization method we have developed provides a way of posing and answering these questions using the preferences, beliefs, and constraints approach.

HISTORY In the middle of the twentieth century, long before we worried about climate change and its unfolding calamities, Aldo Leopold, the American environmentalist raised an economic question: "Like winds and sunsets, wild things were taken for granted until progress began to do away with them. Now we face the question whether a still higher 'standard of living' is worth its cost in things natural, wild and free." (Leopold 2020 [1949], p. xxi).

Leopold was articulating a trade-off between, on the one hand, consuming goods and services – Leopold's higher "standard of living" – and on the other, the costs of environmental damage – the "cost in things natural, wild and free."

Feasible combinations of conventional goods and environmental quality

The opportunity cost of environmental quality is consumption of other (conventional) goods such as food, clothing, shelter, and transportation, which we must give up to secure a higher quality environment. There is a feasible frontier showing the combinations of environmental quality, x, and conventional goods, y, that are possible for a society. The feasible frontier in the case of environmental quality depends on the *abatement technology*, which represents how much consumption of conventional goods society has to give up to achieve a given level of environmental quality.

Figure 3.15 shows a feasible frontier between conventional goods (y) and environmental quality (x). We measure environmental quality on a numeric scale from 0 (the environment that we would have if no abatement done) to 20 (the environment resulting if we were to divert to abatement uses all of society's resources above some minimum level of consumption). We measure conventional consumption as billions of dollars.

The negative slope of the feasible frontier at any point is the marginal rate of transformation of reduced environmental quality into increased conventional consumption, or $-\frac{\Delta y}{\Delta x}$. The steeper the frontier, the greater is the increase in feasible consumption allowed by a reduction in environmental quality.

This is also the the *opportunity cost* of improved environmental quality. So a flatter frontier means a lower opportunity cost of abatement.

To see this, starting at no abatement expenditures ($y = \bar{y}$), the opportunity cost of improved environment is initially small (the frontier is nearly flat) and as Annette implements more abatement, the cost more abatement increases as the environmental quality increases. The shape of the feasible frontier reflects an increasing marginal rate of transformation, or an increasing marginal opportunity cost of environmental quality.

Put another way, if environmental quality is at its maximum at the intercept of the feasible frontier with the horizontal axis, society could consume a lot

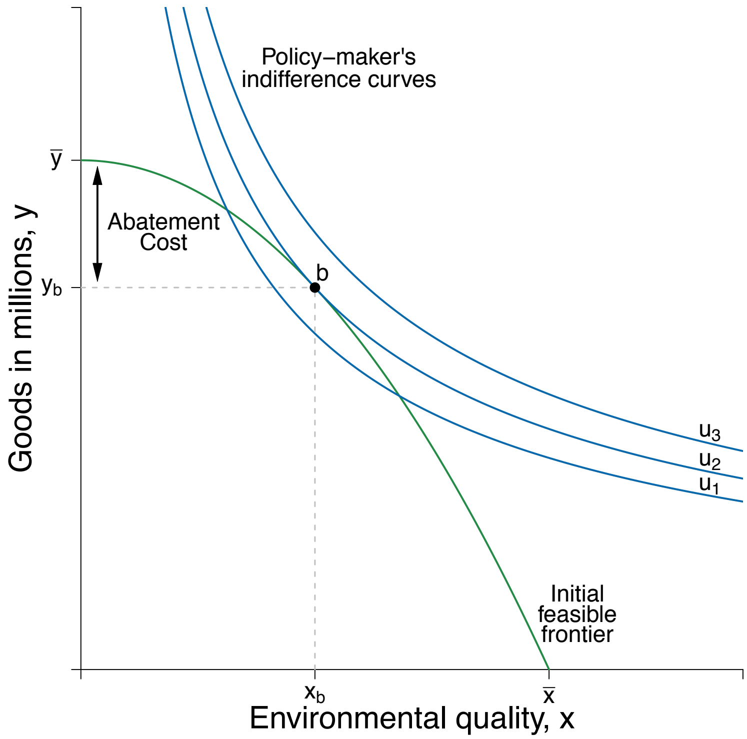


Figure 3.15: **Trade-off between consumption of conventional goods and environmental quality.** The constrained utility maximum is the point on the feasible frontier on the highest indifference curve u_2, shown as point **b** where the $mrs = mrt$ rule holds. The constrained maximum is at the point where the feasible frontier is tangent to the highest attainable indifference curve.

more conventional goods if it were willing to tolerate a small deterioration of environmental quality (the frontier is steep where it intercepts the horizontal axis). But the feasible increase in consumption of conventional goods allowed by a reduction in environmental quality falls as the level of environmental quality declines.

M-CHECK An example of the function representing the feasible frontier is shown below, where y is the goods available for consumption, $\bar{y}$ is the level of y that is feasible when environmental quality is at its minimum, and x is environmental quality.

$$y = 100 - \frac{1}{2}x^2 \qquad (3.22)$$

This is the equation for the feasible frontier is graphed in Figure 3.15.

Checkpoint 3.10: The mrt of the environmental feasible frontier

a. Refer to Equation 3.22 and find the marginal rate of transformation of the feasible frontier.
b. Practice sketching the feasible frontier and confirm the intercepts with the horizontal and vertical axes as shown in Figure 3.15.

3.12 Application: Optimal abatement of environmental damages

How much abatement is the right level, taking account of both preferences for conventional goods (consumption) and the quality of the environment along with the opportunity costs in lost consumption?

A citizen chooses a level of abatement of environmental damages

To begin with the simplest case, think of just one citizen, Annette, who might be representative of the attitudes of the whole society, trying to decide on the level of abatement that she would like to see implemented. She cares about both the quality of the environment, x, and the amount of conventional goods that will be available for people to consume

Annette's utility function has the following form: $u = u(x, y)$. Annette considers what she would like to see her society do about the environment (x), taking account of the effects on everyone. In other words, she is thinking from an other regarding like an ideal policy-maker.

Annette's indifference curves between environmental quality and conventional goods are downward sloping because she regards both environmental quality and conventional consumption as goods for which more is better. This means the marginal utility of both y and x are positive (e.g. $u_y > 0$ and $u_x > 0$). The negative slope of the indifference curves shown in Figure 3.15 at any point is Annette's marginal rate of substitution between more consumption of goods and a better environment. Her marginal rate of substitution shows the amount of goods she would be willing to give up for a small improvement in the environment. As before, Annette's indifference curves exhibit diminishing marginal utility of both environmental quality and consumption.

An example of a utility function that Annette might have is the Cobb-Douglas utility function:

$$u(x, y) = x^{\alpha} y^{(1-\alpha)} = x^{0.4} y^{0.6} \tag{3.23}$$

Figure 3.15 shows three indifference curves defined by equation 3.23: u_3^A is unattainable given the feasible frontier, u_1^A intersects the feasible frontier twice, and u_2^A is tangent to the frontier at point (x_b, y_b). Annette's constrained maximum allows her and her fellow citizens to consume 75 million units of conventional goods and enjoy environmental quality of about 7 (see M-Note 3.12 for the worked solution). If she were able to implement relevant environmental and fiscal policies, this point is the best society can do in Annette's opinion.

What is the total opportunity cost in foregone conventional consumption of a level of environmental quality of 7? The maximum feasible level of conventional consumption with no abatement is \$100 billion. The difference between the maximum feasible consumption of \$100 billion and Annette's preferred choice of conventional consumption of \$75 billion is the opportunity cost of an environmental quality of 7. In our example, the abatement costs are equal to $\$100$ billion $-$ $\$75$ billion $=$ $\$25$ billion in conventional goods. A citizen with Annette's preferences thinks that the sacrifice of $\$25$ billion consumption goods is more than worth paying to have an environmental quality of 7 instead of zero.

New technologies, and conflicts of interest

If, with a mind to the future, some of the abatement costs are devoted to research to improve abatement technologies, this would pivot the feasible frontier outwards, as shown in Figure 3.16. Remember this is very similar to the expansion of the feasible set shown in Figure Figure 8.3 when Keiko adopted improved studying so as to reduce the opportunity cost in reduced Living time associated with greater Learning.

FACTCHECK The pace of environment friendly innovation is astounding. Have a look at the reduction in costs of the photovoltaic cells used in solar panels dropping to one-onehundreth of there costs in 1975 in Figure 8.3.

As is shown in the figure, the shift of the feasible frontier would permit higher environmental quality of $x \approx 9.8$ at the same level of consumption of $\$75$ billion at the new $(x_r, y_r = y_b)$. But there would still be a trade-off: more conventional goods would require less environmental quality, or more environmental quality would require fewer conventional goods to stay on the feasible frontier.

We can also use Figure 3.15 to see why people often disagree about environmental policy.

- Preferences: peoples' *preferences* for conventional goods and the environment may differ
- Beliefs: people may disagree about the opportunity costs or the benefits of environmental quality
- Conflicts of interest: the costs and benefits of abatement fall on different people; those whose jobs or profits depend on carbon based energy, for example, stand to bear more of the costs of addressing climate change, while regions likely to be particularly hard-hit like Africa bear a larger share of the benefits.

M-Note 3.12: The trade-offs and opportunity costs of the environment

Let us work through the process that Annette the policy-maker would go through to identify the combination of goods in billions of dollars with environmental quality.

First, let us calculate her marginal rate of substitution from her utility function, $u^A(x,y) = (x^A)^{0.4}(y^A)^{0.6}$. From earlier in the chapter, we know that the $mrs(x,y)$ is the ratio of marginal utilities and we have already calculated this for $\alpha = 0.4$ and $(1-\alpha) = 0.6$ in Equation 3.13 in M-Note 3.7.

$$mrs(x,y) = \frac{2y}{3x} \tag{3.24}$$

Annette's feasible frontier, based on her beliefs and understanding of the existing science, is given by the equation $y = 100 - \frac{1}{2}x^2$, for which we can find her $mrt(x,y) = -\frac{dy}{dx}$:

$$\begin{aligned} \frac{dy}{dx} &= -x \\ \therefore -\frac{dy}{dx} &= x \end{aligned} \tag{3.25}$$

We now set the $mrt(x,y)$ given by Equation 3.25 equal to the $mrs(x,y)$ given by Equation

3.24 and we isolate one of the variables, y:

$$\begin{aligned} x &= \frac{2y}{3x} \\ \text{Multiply through by } 3x \quad 3x^2 &= 2y \\ \text{Divide through by 2} \quad y &= \frac{3}{2}x^2 \end{aligned} \tag{3.26}$$

We can now substitute Equation 3.26 into the feasible frontier to find x_b and y_b:

$$\begin{aligned} \frac{3}{2}x^2 &= 100 - \frac{1}{2}x^2 \\ 2x^2 &= 100 \\ x^2 &= 50 \\ \therefore x^b &= \sqrt{50} = 7.07 \end{aligned}$$

Having found x^b, we can substitute it back into 3.26 to find y^b:

$$\begin{aligned} y &= 100 - \frac{1}{2}(\sqrt{50})^2 \\ &= 100 - \frac{1}{2}(50) \\ y^b &= 75 \end{aligned}$$

So, as a result of Annette's policy-making utility function and feasible frontier, she would choose a combination of environmental quality, x, of value 7.07 with consumption of good and services of \$75 billion. \$75 billion is \$25 billion less than the maximum consumption of goods and services, $\bar{y} = 100$, so the cost of **abatement** is \$25 billion.

First, people may differ in their preferences over conventional goods and the environment. Another citizen, Brenda, may not worry as much as Annette about the climate and problems that future generations will inherit because she puts a lower weight on the welfare of others (in this case of future generations). Brenda would have different indifference curves. If she also had Cobb-Douglas utility, she would have less strong preferences for the environment than Annette, with a lower α and higher $(1-\alpha)$. Brenda would as a result choose a constrained maximum with lower social spending on abatement.

A second reason for disagreement is that Brenda thinks that the costs of abatement are much greater than Annette thinks they are.

If Brenda thinks that the actual costs of environmental quality are greater than Annette does, Brenda would work with a different feasible frontier *inside* the feasible frontier Annette considers. Brenda's feasible frontier would be steeper over its entire range, indicating that the opportunity cost of environmental quality is higher for any level of environmental quality than on Annette's feasible frontier. Brenda's constrained utility-maximizing bundle would be different from Annette's.

But there is a third reason for disagreement, not having to do with the preferences or beliefs of the citizens. Some members of the society may benefit from decisions that harm the environment. Brenda might be employed by or

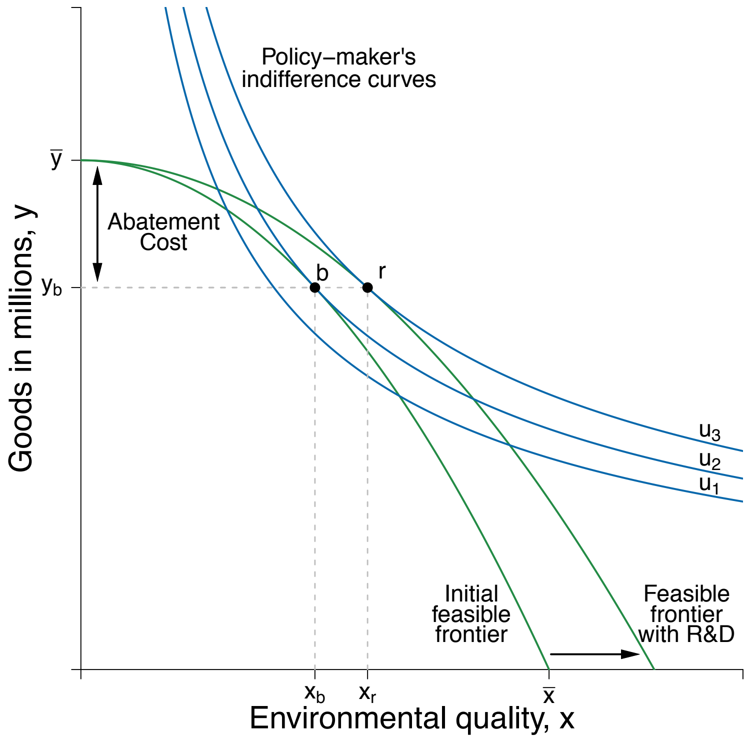


Figure 3.16: **Trade-off between consumption of conventional goods and environmental quality with R&D.** The choice between consumption of conventional goods and environmental quality with R&D pivoting the feasible frontier outwards leading to a new point of tangency on a higher indifference curve u_3.

an owner of a firm producing fossil fuels, for example. They will bear more than proportionally the costs of abatement and may prefer lesser levels of abatement for that self-regarding reason.

The fact that a policy of CO2 emissions abatement would affect Brenda adversely while benefiting Annette brings us back to the how we understand the term utility.

Checkpoint 3.11

a. Show that when Annette's utility function is defined by Equation 3.23 and the feasible frontier is defined by Equation 3.22 with $\bar{y} = 100$, that the utility-maximizing consumption bundle is \$75 billion with an environmental quality of 7.07.

b. Show that when Annette's utility function is defined by Equation 3.23 and the feasible frontier is defined by Equation 3.22 with a technological improvement where $y = 100 - \frac{1}{3}x^2$ that the constrained utility-maximizing consumption bundle is $(x_r = 8.66, y_r = \$75 \text{ billion})$.

c. Show that when Annette's utility function is defined by Equation 3.23 and the feasible frontier is defined by equation 3.22 with a technological improvement where $y = 100 - \frac{1}{4}x^2$, the constrained utility-maximizing consumption bundle is $(x_r = 10, y_r = \$75 \text{ billion})$. Draw a new feasible frontier tangent to a new indifference curve u_4 with accurate intercepts reflecting the new technology.

3.13 Cardinal inter-personally comparable utility: Evaluating policies to reduce inequality

Most policy choices involve conflicts of interest like that between Annette and Brenda about the abatement of environmental harms. Few policy choices are entirely win-win. Most policies – whether they concern taxation, immigration, health insurance, or the rate of inflation – result in benefits for some and losses for others.

Ordinal and cardinal utility in policy evaluation

How do we then evaluate competing policies? Don't think about this as a question about what would be a good outcome for *you* if you were a participant in the society. Instead, try to take the position of what Adam Smith called the Impartial Spectator who did not himself stand to gain or lose, but wanted instead to consider the gains and losses to society.

HISTORY Adam Smith in *The Theory of Moral Sentiments* conceived of the impartial spectator as follows, "We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it. If, upon placing ourselves in his situation, we thoroughly enter into all the passions and motives which influenced it, we approve of it, by sympathy with the approbation of this supposed equitable judge. If otherwise, we enter into his disapprobation, and condemn it."[3]

One answer you might give is just to count those who prefer each policy and select the most popular policy. All this requires is that people be able to rank the policies in question as better, worse, or indifferent. We could in this case treat utility as ordinal (that is, simply a ranking (or ordering) of outcomes). Something like this might occur in a majority rule democratic political system, especially if citizens could vote on policies as they do in many countries in referendums asking citizens to vote for or against a particular policy.

But this way of evaluating policies might result in evaluating positively those policies that confer minor gains to those in favor, and substantial losses to those preferring another policy. This does not seem like a sensible rule.

An alternative is to weigh the amount of the gains to the beneficiaries of each policy against the size of the costs incurred by those who would have done better under some other policy. This kind of comparison requires that we know not only *which* policies people prefer, but *how much* they prefer them.

HISTORY Lionel Robbins (1898-1984) was a leader in the "ordinal revolution" in economics. Economics, he wrote, does not need "to compare the satisfaction which I get from the spending of 6 pence on bread with the satisfaction which the Baker gets by receiving it. That comparison . . . is never needed in the theory . . . ." (123). Moreover, "There is no way of comparing the satisfactions of different people" (124).

To do this we treat utility as a cardinal measure for which utility is not just a ordinal ranking, but instead a number indicating how well off the person is under the option in question. Treating utility as cardinal allows us to say two very different things:

1. for Annette, the outcome (x', y') is twice as good as (x, y) because for example $u^A(x', y') = 2u^A(x, y)$
2. the sum of the Annette's and Brenda's utility is greater with outcome (x', y') than with outcome (x, y) because $u^A(x', y') + u^B(x', y') > u^A(x, y) + u^B(x, y)$

Both statements involve cardinal utilities, but they differ. The first statement compares how much *Annette* values two different states that she will experience. It does not compare her evaluation of a state that she will experience with someone else's evaluation of the state they will experience. The first statement is an example of the cardinal utility that we introduced in Chapter 2 as the basis of expected payoffs (or expected utility) and the analysis of decision-making in risky situations. The second statement compares Annette's utility *with Brenda's utility*. When utility is represented in this way it is called *inter-personally comparable cardinal utility* (or sometimes "cardinal full comparable utility"). If utility is cardinal in this inter-personally comparable sense, then we can compare how well off two or more people are, and how much better off or worse off a policy would make each of them. This provides a way to evaluate which policies should be implemented by asking whether the gains of those who benefit from a policy exceed the losses of those who do not.

HISTORY Philosopher-economist John Stuart Mill (1806-1873) referred to what we would now call the sum of the total utilities of a population as as "a good" that should be promoted: "the general happiness is desirable... each person's happiness is a good to that person, and the general happiness, therefore, a good to the aggregate of all persons."[4]

Why do these two methods of comparing utility matter? Remember that one of the problems with Pareto efficiency as a criterion for fair policy outcomes was *many* outcomes can be Pareto efficient. So Pareto efficiency does not provide an adequate basis for an Impartial Spectator preferring one outcome over the other. Using the second – stronger – conception of cardinal utility along with the judgement that one outcome is better than another if total utility is greater provides a rule for evaluating *which* Pareto-efficient outcome we might prefer as a society.

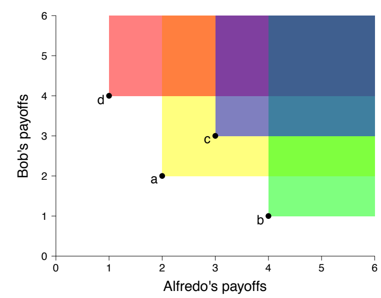


Figure 3.17: **Pareto comparisons in the Fisherman's Dilemma Game**. The Pareto criterion favors **c** over **a** (which it dominates) but cannot rank points **b, c** and **d** because are all Pareto efficient.

In the payoffs for the Fishermen's Dilemma in Figure 1.11 (shown here in the margin for easy reference) three of the four outcomes of the game are Pareto efficient. The Pareto criterion provides no way to choose among them. By contrast the rule — maximize total utility – selects the mutual cooperate outcome (point **c**) with total utility of 6.

Cardinal utility and the distribution of wealth

Suppose you are a policy-maker and you have to divide an amount of wealth between Annette and Brenda. The amount of wealth you have to divide is equal to 1, so each person can get a fraction of that wealth and, as long as the fractions sum to 1, then the outcome will be Pareto-efficient. Let the fraction that Annette gets be a with Brenda's fraction be $(1-a)$. Annette and Brenda have identical preferences for wealth given by the cardinal utility functions of how much wealth they get: $u^A(a)$ and $u^B(1-a)$. For both of them the marginal utility of wealth is diminishing with increased wealth.

EXAMPLE When you say "I'll do the shopping; it'll be less trouble for me than for you" you are representing utility (the trouble of shopping) as cardinal and making an interpersonal comparison of utility (less trouble for me than for you). In fact, much of our everyday ethical reasoning involves interpersonal comparisons of the benefits and costs that people experience.

The horizontal axis in Figure 3.18 shows all possible distributions of wealth between Annette and Brenda.

- Annette's share of wealth, a, varies from 0 to 1.
- At $a=0$, Annette gets nothing and Brenda gets everything.

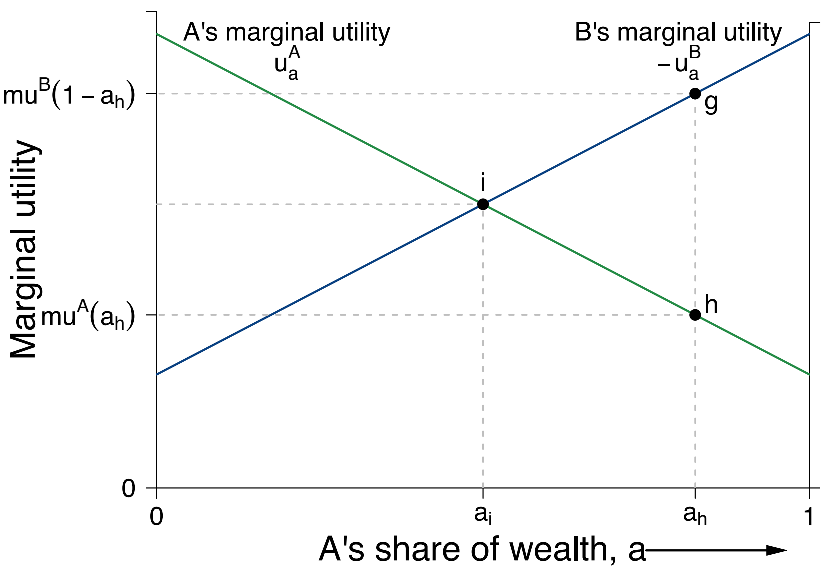


Figure 3.18: **Distribution of wealth, marginal utility, and total utility.** In the figure a is the proportion of wealth belonging to Annette (A). $1-a$ is the proportion of wealth belonging to Brenda (B). As a person's wealth increases, the marginal utility of wealth decreases. A's total wealth *increases* as you move along the bottom line from left to right, and as a result her marginal utility *decreases*. Because B's wealth increases as the division moves towards the left, B's marginal utility decreases from right to left.

- At $a = 1$, Brenda gets nothing and Annette gets everything.

Figure 3.18 shows the two marginal utility functions for wealth. Each of them has decreasing marginal utility in wealth. This means that the increment in utility associated with each additional unit of wealth they have is less when they have more wealth. Annette's marginal utility curve slopes downward as she gets more wealth (moving from left to right), and Brenda's marginal utility of wealth decreases as she gets more wealth (moving from right to left).

Suppose the status quo is a_h, a situation in which Annette is wealthy and Brenda is poor (Annette's has share of wealth a_h and Brenda's share of wealth is $1-a_h$). A policy that takes a small amount of wealth from Annette and transfers it to Brenda reduces Annette's utility by less than it increases Brenda's. We can see this by identifying that Annette's marginal utility at a_h, $u_a^A(a_h)$ is much lower than Brenda's utility at the same point $u_a^B(a_h)$. The vertical difference between points **g** and **h** shows the magnitude of the difference in their marginal utilities.

Redistributing wealth from Annette to Brenda therefore increases *total* utility (the sum of Annette and Brenda's utilities).

Applying this reasoning to other points in the diagram, we find that the distribution of wealth that maximizes total utility is a_i, where Annette's marginal utility of wealth equals Brenda's marginal utility of wealth. If Annette's and Brenda's utility functions are identical, the total utility maximizing point distributes wealth equally, $a_i = \frac{1}{2}$.

M-Note 3.13: Maximizing Total Utility

Consider a society of two people, Annette (A) and Brenda (B), in which maximizing total utility, $U = u^A + u^B$, is the goal of the policy-maker who we will assume is Adam Smith's Impartial Spectator . The Impartial Spectator selects a point, **i**, for their choice of policy. We assume Annette and Brenda assign the same utility numbers to some given level of wealth. They are identical in this respect. But they differ in their wealth. Annette's share of total wealth is $a > \frac{1}{2}$ and Brenda's is $1 - a < \frac{1}{2} < a$. Annette's utility is $u^A = u(a)$ and Brenda's utility, $u^B = u(1 - a)$, is also a function of a, but Brenda's utility *decreases* as a increases.

$$U(a) = u^A + u^B = u(a) + u(1-a) \tag{3.27}$$

To find the maximum total utility, we differentiate the total utility with respect to a and set the derivative equal to zero:

$$\begin{aligned} U_a(a) = \frac{dU}{da} &= u_a^A + u_a^B = u_a(a) - u_a(1-a) = 0 \\ \therefore u_a(a_i) &= u_a(1-a_i) \end{aligned}$$

The only way that $u_a(a_i)$ can be equal to $u_a(1 - a_i)$ is if a_i and $1 - a_i$ are the same number.

$$a_i = 1 - a_i = \frac{1}{2}$$

This maximization is depicted in Figure 3.18: Annette's marginal utility, $u_a^A(a)$, decreases as a increases (diminishing marginal utility) and Brenda's marginal utility $u_a^B(1 - a)$ increases as a increases (because her wealth, $1 - a$ decreases as a increases). The total utility maximizing choice occurs at a_i, where wealth is equal and as a result the marginal utilities are equal: $u^A(a_i) = u^B(1 - a_i)$.

3.14 Application: Cardinal utility and subjective well-being

A century ago economists thought that while ordinal comparisons like better or worse are possible empirical interpersonal comparisons expressed by a number indicating the degree of preferences were impossible to make. But today researchers are actively engaged in measuring individual happiness and life satisfaction, using techniques ranging from surveys and natural observation to the methods of experimental neuroscience. They are asking such questions as: "How important is income for happiness?" "Is being without a job a bigger source of unhappiness than being without a spouse?" These researchers refer to happiness or life satisfaction as *subjective well-being*.

EXAMPLE Daniel Kahneman, a psychologist and Nobel Laureate in economic, has advocated a *hedonistic* (meaning concerning pleasure and pain) theory of utility. Kahneman titled one of his papers "Back to Bentham?" to pay homage to the early 19th century philosopher economist Jeremy Bentham's utilitarian theory.[5]

To measure "pleasures and pains" in the lab, volunteers are exposed to an electrical shock and asked to report on their experience of that on a numerical scale. Others are asked to plunge their hands into extremely cold water for as long as they can stand it and immediately report their level of unhappiness having done so. Respondents in surveys are asked their "life satisfaction."

FACT CHECK The Satisfaction with Life survey is based on five questions each of which is rated on a 7-point scale from Strongly Disagree (1) to Strongly Agree (7). Here are the questions: In most ways my life is close to my ideal; The conditions of my life are excellent; I am satisfied with my life; So far I have gotten the important things I want in life; and If I could live my life over, I would change almost nothing.

This research has sought to understand the activities that make people most happy. Almost all people surveyed seem to like sex quite a lot, ranking "intimate relations" as having a high subjective well-being value. Ranked after sex, people like socializing, relaxing, sharing meals with friends, praying, and exercising. People don't like housework, childcare, commuting or working. People also report major changes in subjective well-being from painful events, like sudden loss of a job, a death in the family, or divorce, or from positive events like marriage, or the birth of a child.[6] But when you ask someone about their happiness over time the measures are surprisingly consistent: people are likely to report similar activities or outcomes as providing them with happiness when you ask them at different intervals.

FACT CHECK Non-laboratory measures of subjective well-being suggest that people with higher subjective well-being tend to be less likely to contract a cold virus and to recover more quickly when they do contract the cold. Similar evidence exists for people who have recovered from wounds and had baseline and subsequent subjective well-being measured: those who are happier recover more quickly.[7]

What are the take-home messages about subjective well-being, the choices about how we spend our time, and how we value effortful work?

First, people like a diverse array of activities and doing things they like provides them with happiness that can be measured in meaningful ways across different people and at different periods of time in our lives.

Second, people who report greater subjective well-being are also better off by physical biological measures. For example,they are less likely to be ill. Subjective well-being also manifests in hormone levels, brain patterns, and palm temperature.[8]

Third, while income matters for happiness (especially for people without much income) people value social relationships – marriage, a job, friendships – more than they value income.[9] Making the transition from unemployed to employed boosts a persons subjective well-being by much more than would be predicted simply by the increase in income. This is because having a job is a source of respect and dignity, especially as it provides a way for people to express autonomy over themselves, competence in their expression of their abilities, and relatedness to other co-workers and people around their work.

EXAMPLE The substantial subjective cost that people experience when they are out of work is one reason why employers (who have the power to terminate a person's job) have power over their employees. We shall return to this when we study the firm and the labor market.

Checkpoint 3.12: Joy or Misery?

Think about the kinds of activities that Kahneman and Krueger discuss above that provided people with joy (that they ranked highest in terms of providing them with subjective well-being).

1. Compare them with their opposites: those that result in disutility or even misery.
2. Come up with a list of activities that you engage in that provide you with joy which you try to prioritize.
3. Why do you spend the time that you do on these activities? Why do you not spend more?
4. Do you engage in activities that in the moment are unpleasurable but which you believe provide you with benefit nonetheless?

5. Do you think such activities appear in the models we've developed?

3.15 *Preferences, beliefs, and constraints: An assessment*

Many scholarly disciplines in addition to economics are devoted to understanding human behavior including psychology, sociology, anthropology and history, but also more distant endeavors including literature, philosophy, neuroscience, computer science and biology. The preferences, beliefs and constraints approach, while a standard set of tools in economics that is widely used in other fields, is just one of many approaches. People newly familiar with the approach often raise the following questions about it.

HISTORY In his 1953 work,*Essays in Positive Economics*, Nobel Laureate Milton Friedman (1912-2006) observed that "predicting the shots made by an expert billiard player" could be done on the basis of "the complicated mathematical formulas that would give the optimum directions" of the shots. But this prediction would not be "based on the belief that billiard players, even expert ones, can or do" actually make these calculations.[10]

- *Are people really all that selfish?* This concern is based on a misunderstanding of the model,which says nothing about whether people are seeking to help others, aggrandize themselves, or a little of both. Our treatment of altruism, reciprocity and fair-mindedness shows that the model – using indifference curves and feasible sets, for example – can apply to a variety of motives.

- *Do people consciously optimise*, for example, applying the $mrs = mrt$ rule when they shop? The model is not a description of how people actually think or their emotional states when they take a break from studying, or support a particular environmental policy. We model instead what people would do if they did the best that they could. The fact that the model often yields predictions similar to what we observe empirically (including by experiments, econometric and other quantitative methods) does not require that the model is an accurate representation of the process by which people come to take one course of action over another.

In some cases, people consciously optimize, going though mental calculations similar to the model. For example, a person buying a house or choosing between two job offers will weigh the pros and cons of the alternatives. But in other cases, the actions may not even appear to us as a decision, for example, what to eat for breakfast, what to wear today, or what our personal values should be. Without consciously trying to do so, people may arrive at something like the solution to these optimization problems by trial and error, or by observing others who seem to be successful or happy with their choices, or by following habits that will remain in place unless changed by some dramatically adverse consequences of following them.

Other concerns about the model are more serious.

- *What about emotions and visceral reactions, aren't they important?* This question points to a shortcoming of the approach; but it is not that the approach excludes emotions like fear, shame, and attraction. The shortcoming is that the preferences, beliefs, and constraints approach says almost nothing at all about motives, that is, it says nothing about the reasons *why* people rank some outcome as superior to another. Knowing more about motives like this would help us understand economic and other behavior.

- *Preferences and beliefs are not "just there" as facts of nature, they are products of environments we live in.* We have already seen an example of this as social norms concerning the kinds of work that are appropriate for women to do changed during the 20th century under the influence of new technologies for cooking cleaning and washing and the experience of women doing factory work producing armaments during the second world war. Our preferences and beliefs are to some degree "socially constructed". Our discussion of differing cultures around the world using experiments designed using the preferences, beliefs and constraints framework shows that the approach can help to clarify how cultures differ and how society shapes preferences and beliefs.
- *Commitments and consequences.* The framework is based on the idea that our behavior is based on our beliefs about the consequences our actions will bring about in the future. Don't we sometimes act to fulfill promises or other commitments made in the past, or just to "do the right thing" without regard to future consequences? Yes we do, and a shortcoming of the approach is that it does not address that kind of behavior.
- *Predicting behavior and evaluating outcomes*. Economists use the same concept "utility" in models designed to predict the actions that people will take and to provide the basis for evaluating economic outcomes and public policies to improve them. The idea is that whatever it is that motivates people to make the choices they do should also be the objective of public policy and form the basis for our preferring one societal outcome over another. But treating actual behavior as if it were the pursuit of a concept of well-being that should be the basis of our judgement of societal outcomes is a mistake. The reasons for our actions (that is, our preferences) include addictions, weakness of will, shortsightedness, and other well documented socially dysfunctional aspects of human behavior that in retrospect are often deeply regretted by those acting on them.

A sensible conclusion from reviewing these concerns about the preferences beliefs and constraints approach might be that the approach is better for answering some questions than others, and learning to distinguish which is which is an important learning objective. As we said at the beginning of the chapter: the map is not the territory. Good maps don't have all the information about the territory they depict and good economic models require us to leave some things out.

HISTORY: POSITIVE AND NORMATIVE ECONOMICS The distinction between the economics of "what is" called positive economics and "what ought to be" called normative economics was made by John Maynard Keynes in his 1893 *Scope and Method of Political Economy* and by Milton Friedman in his 1953 *The Methodology of Positive Economics.* The distinction is controversial in part due to differences about the appropriate role for "what ought to be" statements in economics.[11]

Checkpoint 3.13: Positive and normative uses of "utility"

> Consider the statement by J.S.Mill (in the above margin note): "each person's happiness is a good to that person, and the general happiness, therefore, a good to the aggregate of all persons." Explain how Mill is here using "happiness" both as a way of *predicting behavior* (sometimes called "positive economics") and as a way of *evaluating outcomes* from a societal standpoint ("normative economics.")

3.16 Conclusion

In this chapter we have studied the constrained optimization problems shown in Table 3.1. Though the problems concerned are quite different, the models and analytical tools we used to analyze them are very similar. In each case the analysis of the decision involves two kinds of trade-offs:

- The first trade-off that appeared in each of these situations is the actor's relative valuation of the things she cares about, measured by the negative of the slope of an indifference curve, that is, the marginal rate of substitution.
- The second trade-off is that at any point on the feasible frontier, the opportunity cost of having more of one good that the actor values is that she must have less of another good that she values. This opportunity cost trade-off is measured by the negative of the slope of the feasible frontier, that is, the marginal rate of transformation.

The result - the action taken doing the best she can under the constraints she faces - is determined in the same way in all the cases: by finding the point on the feasible frontier that is on the highest indifference curve. This will often be the bundle where the $mrs = mrt$ rule holds. The table demonstrates that many seemingly different kinds of action can be studied with a common model, one that we will use often.

In this chapter, we have focused on single actors and for the most part excluded from the model something important: other people. With the exception of the farmers of Palanpur, we have modeled the person facing a given situation defined by a feasible frontier and preferences represented by indifference curves. (We already explained that the second person in the Dictator Game is not really a player at all).

We now turn to a world populated by people interacting strategically, and we ask how economic institutions affect the outcomes of these interactions, and how these outcomes and institutions might be judged by the standards of Pareto-efficiency and fairness. We shall continue to employ the tools of constrained utility maximization: we will continue to understand people's trade-offs through their marginal rates of substitution and of transformation; and we shall need these to understand how one person might engage in exchange of

Actor	*Utility depends on*	*Action*	*Constraints*
Keiko	Learning, Living time	Time allocation	Learning-Living feasible frontier
Keiko	Learning, Study effort	Study effort	Study-Learning production function
Annette/Chen	Payoffs to two players	Payoffs to each	The total endowment
Annette/Brenda	Conventional goods, Environmental quality	Conventional goods	Consuming conventional goods degrades the environment

Table 3.1: The constrained optimization problems used in this chapter

goods and services with someone else.

Making connections

Strategic and non-strategic social interactions: In the previous chapters we considered strategic social interactions – like the fishers and the farmers from Palanpur. Here we look at simpler aspects of behavior when a person is attempting to do the best they can in situations that are not strategic because the choice of how hard to study, or how much fish to buy is not greatly affected by others choices.

Self-regarding and social preferences: In Chapter 2 we provided evidence that people can be self-regarding, altruistic, reciprocal, and fair-minded. These diverse behaviors can be modeled using the preferences, beliefs, and constraints approach using indifference curves and feasible frontiers, as we showed for the case of an altruist.

Opportunity costs and trade-offs: Regardless of whether a person's preferences are self-interested or social, people face trade-offs among the ends they wish to pursue and they face opportunity costs when trying to choose a course of action.

Public policy: Economics engaged: The idea of constrained utility maximization illustrated the trade-off between consuming more goods on the one hand or either consuming less and using some of the economy's resources to abate environmental damages, obtaining greater environmental quality. We also modeled the choices an altruistic person might make in sharing something of value thereby providing a model capable of analysing the kinds of result observed in the experiments reviewed in the previous chapter.

Evaluating outcomes: Treating utility as cardinal and inter-personally comparable rather than ordinal allows us to compare the benefits and burdens that a policy will impose on different people. This provides a basis (one of a number of alternatives) for saying that one policy or outcome might be preferred to another, as illustrated by the case of the distribution of wealth.

Important ideas

preference	constraints	beliefs
Homo economicus	altruist	reciprocator
ordinal utility	cardinal utility	utility function and value function
Cobb-Douglas utility	total utility	marginal utility
law of diminishing marginal utility	$mrs - mrt$ rule	slope
indifference curve	marginal rate of substitution	diminishing marginal rate of substitution
tradeoff	willingness to pay	iso-value curve
feasible/attainable	feasible frontier	production function
marginal product	marginal rate of transformation	increasing marginal rate of transformation
opportunity cost	increasing opportunity costs	utility-maximizing
point of tangency	price line	offer curve

Mathematical notation

Notation	**Definition**
$u()$	utility function
x	a good (or a "bad")
y	a good (or a "bad")
h	hours of work
α	Cobb-Douglas exponent of good x
$\bar{y}$	vertical intercept of the feasible frontier
$c()$	opportunity cost
a	A's share of wealth
p	price of a good
$\overline{u}$	constant utility along an indifference curve
z	endowment in the Dictator game
π	payoff in the Dictator game

Note on super- and subscripts: A, B, C, D: different people; CD: Cobb-Douglas utility function; Subscript b indicates where someone does the best they can; RD: feasible frontier with Research and Development.

Discussion questions

See supplementary materials.

Problems

See supplementary materials.

Selected Answers/Hints for Questions

See supplementary materials.

4

Property, Power & Exchange: Mutual Gains & Conflicts

> [T]he efforts of men are utilized in two different ways: they are directed to the production or transformation of economic goods, or else to the appropriation of goods produced by others.
>
> Vilfredo Pareto, *Manual of Political Economy* (1905) (Pareto (1971):341)

Ibn Battuta, the fourteenth century Moroccan scholar, reported that along the Volga River in what is now Russia, long distance trade took the following form: "Each traveler ... leaves the goods he has brought ... and they retire to their camping ground. Next day they go back to ... their goods and find opposite them skins of sable, miniver, and ermine. If the merchant is satisfied with the exchange he takes them, but if not he leaves them. The inhabitants then add more skins, but sometimes they take away their goods and leave the merchant's. This is their method of commerce. Those who go there do not know whom they are trading with or whether they be jinn [spirits] or men, for they never see anyone."

The Greek historian Herodotus describes similar exchanges between Carthaginian and Libyan groups in the 5th century B.C. After having left their goods, Herodotus reports, the Carthaginians withdraw and the Libyans "put some gold on the ground for the goods, and then pull back away from the goods. At that point the Carthaginians ... have a look, and if they think there is enough gold to pay for the cargo they take it and leave."

Herodotus describes how the process continues until an acceptable is price hit upon, remarking with surprise that "neither side cheats the other ... [the Carthaginians] do not touch the gold until it is equal in value to the cargo, and [the Libyans] do not touch the goods until the Carthaginians have taken the gold."

Alvise da Ca da Mosto, a fifteen century Venetian working for the Portuguese crown, reported a similar practice in the African kingdom of Mali, regarding it as "an ancient custom which seems strange and hard to believe."

DOING ECONOMICS

This chapter will enable you to do the following:

- Explain why, when people exchange goods, there are both mutual gains and also conflict over the distribution of these gains.
- Understand how an allocation of goods can be evaluated on grounds of Pareto-efficiency and fairness.
- Show how self-regarding as well as social preferences of parties to an exchange can affect the outcome, and how other-regarding social preferences may reduce the scope of conflicts over the distribution of the gains from exchange.
- Understand how both private property rights and the exercise of power by one of the parties to an exchange will affect the outcome of exchange.
- Use the ideas you have learned to explain how an employer and an employee might bargain over working hours and wages.

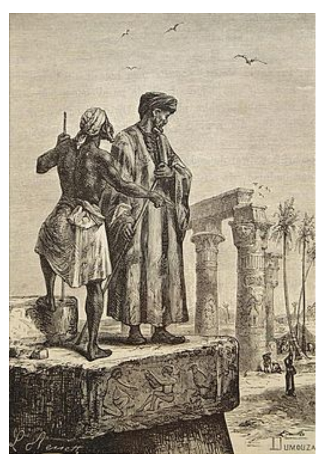


Figure 4.1: **A painting of Ibn Battuta (on the right) (1304-1369) Source: Wikimedia Commons.**

But is the so-called 'silent trade' really so odd? Transfers of goods among strangers can be dangerous. What one expected to be an exchange at mutually agreeable prices may end up as theft or an "offer you cannot refuse." But trade among strangers can also be highly profitable. The potential gains from trade are often greater the more distant geographically or socially the parties are to the exchange: the salt brought by Tuaregs from the Atlas Mountains in North Africa across the Sahara by camel to the Kingdom of Ghana was not available in West Africa. The gold and tropical nuts Tuaregs gained in silent trade with Ghanaians was not available north of the Sahara.

The silent trade – with its unusual etiquette in which parties interacted only at a distance – allowed both Tuaregs and Ghanaians to get some of what they lacked and wanted in return for giving up some of what they had in abundance and could readily part with.

They were exploiting the mutual gains that differences in geography, tastes, technologies, and skills allow. And the rules of the game for governing their exchange process – the institutions that we call "the silent trade" were a way of doing this and dividing the mutual gains without resorting to violent conflicts.

Other than these mutually advantageous exchanges, there are many other ways that goods change hands: from the use of violent coercion by private parties (i.e. theft), or by the use of one nation's military force to acquire the resources of another people. People have also been violently coerced into work through enslavement by private actors and states alike.

A key characteristic of these coerced transfers is that they are *not* motivated by mutual gain, but instead by the gain of one party facilitated by superior force and institutional power. These transfers of resources and lives have shaped the course of history and have had important economic consequences and enduring legacies.

But here we set aside the use of physical coercion and ask how societies organize the process of exchange motivated not by fear of physical harm but instead by the prospect of mutual gain. We also provide terms that allow us to evaluate some of these outcomes as better or worse than others.

Figure 4.2: **A statue of Herodotus.** Considered by many to be the first historian, Herodotus lived in the fifth century BCE. Source: Wikimedia Commons.

4.1 Mutual gains from trade: Conflict and coordination

In a modern economy we engage in indirect monetary exchange: selling some of our goods or some of our working time for money and using the money to purchase the goods we need rather than bartering directly as did the Libyans and Carthaginians. The principles of *barter exchange*, where goods are directly transferred among two parties without the use of money, however illustrate the fundamental considerations behind all types of exchange, including indirect monetary exchange.

REMINDER As in Chapter 3 a bundle is just a list of the quantity of the goods (or other thing of value) that a person has. We refer to the bundles held by all of those involved in an exchange as an *allocation*.

We will simplify by thinking about just two people who exchange goods directly with each other, thereby modifying the goods that they hold. To do this we will introduce two terms describing the *bundles* that each has before and after exchange:

- *The endowment bundle* or *endowment*, the quantities of goods a person has before exchanging goods.
- *The post-exchange bundle* the bundle a person has after exchanging goods with another person.

The bundles held by each of the people (either before or after exchange) is called an **allocation.**

ALLOCATION The bundles held by each of the people (either before or after exchange) is called an **allocation.**

Voluntary exchange: mutual gains and conflict over their distribution

VOLUNTARY EXCHANGE An exchange is **voluntary** if all parties to the exchange have the option to not engage in it but instead choose engage in the exchange. So each party must expect to be better off, or at least be no worse off, as a result of the exchange, which implies that each prefers (at least weakly) their post-exchange bundle to their endowment bundle

An exchange is **voluntary** if all parties to the exchange have the option to not engage in it but instead choose engage in the exchange. So each party must expect to be better off, or at least is no worse off, as a result of the exchange, which implies that each prefers (at least weakly) their post-exchange bundle to their endowment bundle.

REMINDER: An economic rent is the difference between a player's *fallback* payoff and the payoff (profit or utility) they obtain from participating in an interaction. The gains from exchange from an interaction is the sum of the economic rents of all participants.

Recalling the meaning of a Pareto-comparison, we can see that if an exchange is voluntary for both parties, the post-exchange allocation must be a Pareto-improvement over the endowment, otherwise one or both of the parties would have refused to participate in the exchange. The stipulation that the in order for an exchange to be called voluntary, the post exchange allocation must be a Pareto-improvement over the endowment bundle is termed the *voluntary transfer requirement*.

To make the idea of voluntary exchange concrete we often let the fallback position of the players be a bundle of goods that is their private property which they are free to dispose of in exchange or by gift to others, or to retain for themselves, excluding others.

PRIVATE PROPERTY Private property is the right to exclude others from the goods one owns, and to dispose of them by gift or sale to others who then become their owners.

VOLUNTARY TRANSFER REQUIREMENT The stipulation that in order for an exchange to be called voluntary, the post-exchange allocation must be a Pareto-improvement over the endowment bundle is termed the *voluntary transfer requirement*.

Let's review some of the terminology from earlier chapters and explain how they are used to study the process of exchange.

- A person's *fallback position* is what they experience in the absence of the exchange and the utility number they assign to that bundle (that is, the utility of their endowment bundle which is considered to be her *next best opportunity*.)
- The improvement in utility enjoyed by a party to an exchange is their *rent* resulting from the exchange, namely, the difference in utility associated with their post exchange bundle compared to their fallback position.

- The total rents received by parties to an exchange, also termed the *gains from trade* are the utilities of the exchanging parties at the outcome of the exchange minus the utilities at their fallback positions. .

PRIVATE PROPERTY Private property is the right to exclude others from the goods one owns, and to dispose of them by gift or sale to others who then become their owners.

The fact that an exchange is voluntary does not mean that is is fair. Some exchanges take place under conditions such that one party gains virtually all of the available rents. How the economic rents are divided between participants is the **distributional outcome** of the exchange. The rents may be captured by one party, leaving the other with a different set of goods than her endowment but no better off.

Or the rents may be split among the parties in a way that appears fair, or at least acceptable to both, as in the silent trade between the Carthaginians and the Libyans described by Herodotus. The division of the gains from exchange in the form of economic rents is parallel to the division of the pie in the Ultimatum Game of Chapter 2.

DISTRIBUTIONAL OUTCOME How the gains from exchange – the economic rents – are distributed between the people in an exchange; the share of the gains from exchange each player gets as a rent.

Exchange therefore has two aspects: *mutual benefit* and *conflict of interest*:

- *Mutual benefit* is possible because participants move from their endowment bundle to the post-exchange allocation where they share the gains from exchange and obtain an *economic rent*.

- A *conflict of interest* is present because the gains from exchange can be divided in many ways among the parties who find themselves in conflict over who gets the larger share.

Institutions and social norms govern the process of exchange that leads both to the re-allocation of goods, and to the distribution of the gains from trade. We will see that institutions and social norms have effects on:

- *Pareto efficiency,* facilitating or obstructing the realization every opportunity for mutual gain among the parties to an exchange, and

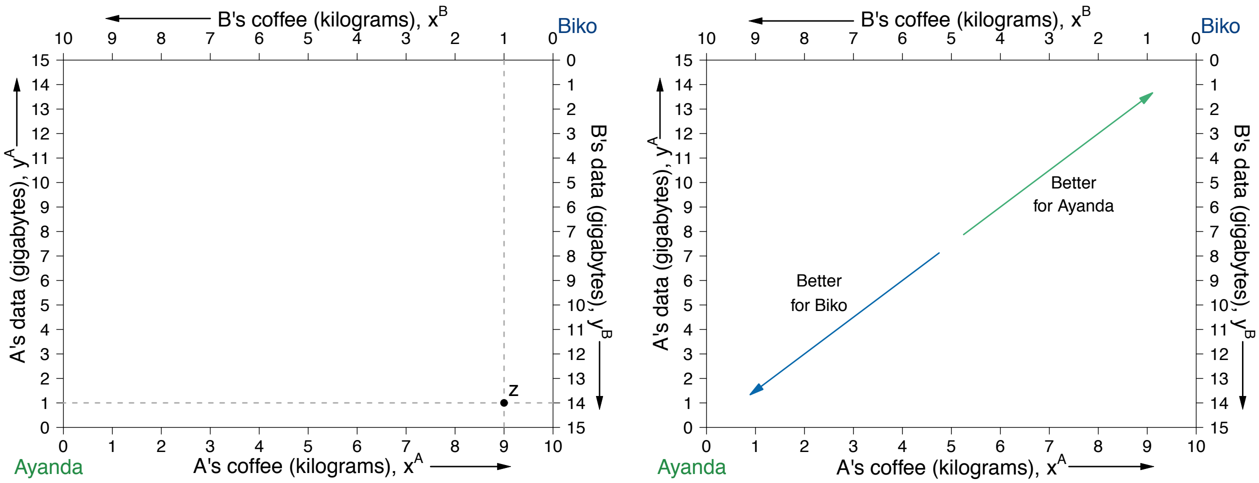
- *The fairness of the distributional outcome,* favoring one party or the other in the conflict of interest in the distribution of the economic rents.

A major institutional challenge today is to find *rules of the game* that will have as the Nash equilibrium allocations that are *both* Pareto efficient and fair. We will return to the interplay of these two objective frequently in the pages that follow.

4.2 Feasible allocations: The Edgeworth box

Lets consider a concrete setting in which two people might consider alternative possible distributions of two goods amongst themselves. Let's say that

HISTORY The Edgeworth box is named after the British economist Francis Ysidro Edgeworth (1845-1926) who is credited with having invented this clever way to represent exchange and bargaining.



(a) An allocation, **z**

(b) Indifference curves

Figure 4.3: **Feasible allocations that exhaust the supply of both goods.** Figure 4.3a shows an example of a feasible allocation at point **z**. Figure 4.3b shows the direction in which each person prefers to move to increase their utility. When indifference curves are plotted in this rectangle the graph is called an Edgeworth box.

Ayanda and Biko have to divide a total of 10 kilograms of coffee and 15 gigabytes of data between them. At the start, nobody owns the goods, the two quantities are simply amounts available to the two of them. Ayanda and Biko might now ask each other: what allocation of the coffee and data between the two of us would be the best?

We use the notation $\bar{x} = 10$ and $\bar{y} = 15$ to stand for the *total* amount of coffee (x) and data (y) available. We define x^A, and y^A as the quantity of goods x (coffee) and y (data) in Ayanda's bundle, and similarly x^B and y^B are the quantities in Biko's. The amount of the two goods in their respective bundles can be anywhere from zero to the entire amount available, namely, $\bar{x}$ and $\bar{y}$. Then, an allocation is a particular assignment of coffee and data to the two people that we can write as $(x^A, x^B; y^A, y^B)$. An allocation is *feasible* if the amounts of coffee and data it gives to Ayanda and Biko is no greater than the amount available:

$$x^A + x^B \leq \bar{x}$$
$$y^A + y^B \leq \bar{y}$$

Figure 4.3 (a) represents the *total supply* of the goods, with width and height equal to the total amount of coffee (x) and data (y) available. The box's *width* is the total amount of x, $\bar{x}$ (kilograms (kg) of coffee) and its *height* is the total amount of y, $\bar{y}$ (gigabytes (gb) of data). We measure A's allocation, (x^A, y^A) from the lower left-hand corner of the box, and B's allocation, (x^B, y^B) from the upper right-hand corner.

Any point in the box (or on its edges) is a bundle representing a *feasible* allocation of the two goods between the two parties, with the property that it fully *exhausts the total supply* of the two goods. You can see this because the width of the box is the total amount of x and the height of the box is the total amount of y. Allocation **z**, for example, gives Ayanda 9 kilograms of coffee and 1 gigabyte of data and Biko 1 kilogram of coffee and 14 gigabytes of data (exhausting the 10 units of x and the 15 units of y). There are also many feasible allocations of the two goods that are *not* shown in the box. For example, if Ayanda and Biko each got 1 kilogram of coffee and one gigabyte of data, that would be feasible given the total amounts, but it could not be shown in the Edgeworth box because the Edgeworth box only shows allocations where the two people divide up *all* of the goods so that they sum to $\bar{x}$ and $\bar{y}$.

As we move to the northeast in the box, Ayanda gets more of both goods, and as we move to the southwest in the box, Biko gets more of both goods. Because both are self-regarding we show this on the figure with the arrows labeled: "Better for Ayanda" and "Better for Biko" respectively .

REMINDER Recall that in Chapter 1, we used **z** to indicate the fallback position of people playing games in the general form of the Fishermen's Dilemma game with ranked outcomes. At **z**, the people experience their utilities, u_z^A and u_z^B, as their utility at the fallback position, that is, their endowments if they do not trade.

How can we evaluate whether some allocations are better than others? To do this we can represent the preferences of the two parties by plotting their indifference curves in the box. This allows us to say for both Ayanda and Biko that for any two allocations (points in the box) that the first is preferred to the second, the second is preferred to the first, or the person is indifferent between the two. To do this we need to know the utility functions of the two.

Both Ayanda and Biko enjoy consuming both coffee and data. Their utility functions are:

$$\begin{aligned} \text{Ayanda's utility function} \quad & u^A(x^A, y^A) \\ \text{Biko's utility function} \quad & u^B(x^B, y^B) \end{aligned}$$

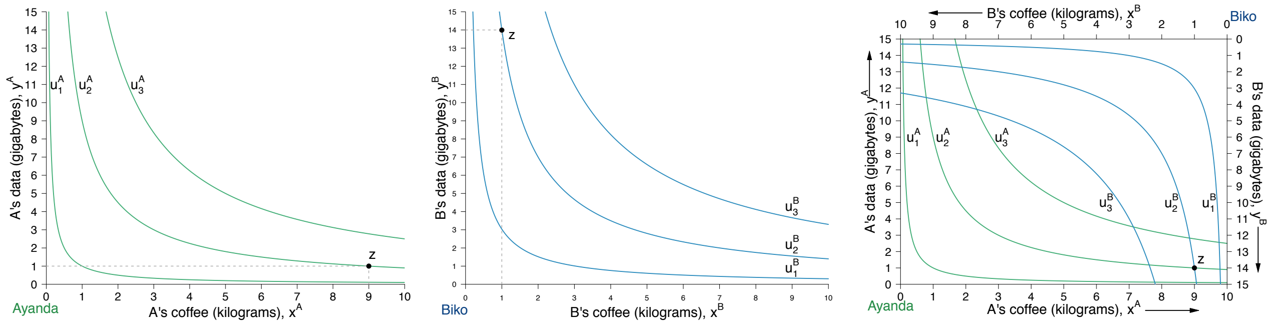
We assume that the indifference curves for both parties exhibit decreasing marginal utility for both goods. To provide a concrete example, we will assume that both Ayanda's and Biko's utility functions are Cobb-Douglas, but in some cases that follow, with different preferences for coffee and data:

$$\begin{aligned} \text{Ayanda's utility function} \quad & u^A(x^A, y^A) = (x^A)^{\alpha^A}(y^A)^{(1-\alpha^A)} \\ \text{Biko's utility function} \quad & u^B(x^B, y^B) = (x^B)^{\alpha^B}(y^B)^{(1-\alpha^B)} \end{aligned}$$

In numerical examples we will often contrast two cases:

- *Identical*: The two people have identical preferences for the two goods, such as $\alpha^A = \frac{1}{2}, \alpha^B = \frac{1}{2}$.
- *Different*: The two people have *different* preferences, for example, such that A's $\alpha^A = \frac{2}{3}$, whereas for B $\alpha^B = \frac{1}{3}$. So Ayanda has a stronger preference for coffee than Biko does.

We can visualize the allocation of coffee and data between Ayanda and Biko using an Edgeworth box. An Edgeworth box allows us to see both people's indifference curves in the same space to identify mutually beneficial trades.



(a) Ayanda's Indifference Curves (b) Biko's Indifference Curves (c) Indifference curves in an Edgeworth Box

Figure 4.4: **Indifference curves and an Edgeworth box.** In panels a. and b. we show three of Ayanda's and Biko's indifference curves respectively. In panel c., Biko's indifference curves have been flipped 180° so that the origin in the lower left of panel b. has become the origin of the Edgeworth box at the upper right.

Ayanda and Biko's indifference curves are shown separately in Figure 4.4 panels a and b. In panel c. we plot the same indifference curves together in the Edgeworth box. Ayanda evaluates the allocations from the point of view of the lower left-hand corner, and her indifference curves represent higher utility as we move to the northeast in the box.

Ayanda's indifference map looks exactly the same in the Edgeworth box as it does in the separate plot, because in both cases the origin from which we measure her allocation is in the lower left-hand corner. Biko evaluates the allocations in the box, however, from the point of view of the upper right-hand corner, and his indifference curves represent higher utility as we move to the southwest in the box. It may help you understand how we superimposed Biko's preferences on Ayanda's if you think about what we called their "point of view." In panel's a. and c., imagine Ayanda standing at the lower left origin and looking up her indifference map, as if the curves were contours of a mountain, the curves farther away being at higher altitudes. Now do the same with Biko, but for him when he looks to the north east in panel b., he is looking up his "utility mountain." But in panel c. he is standing at the upper left origin and the way up his utility map is to the south west.

REMINDER In Chapter 3 we defined the Cobb-Douglas (CD) family of utility functions as:

$$u(x,y) = x^{\alpha}y^{(1-\alpha)}$$

(with $0 \le \alpha \le 1$). The Cobb-Douglas utility function results in a marginal rate of substitution, $mrs(x,y) = \frac{\alpha}{(1-\alpha)}\frac{y}{x}$.

In the figures, at allocation **z** Ayanda and Biko have allocations $(x_z^A, y_z^A) = (9,1)$ and $(x_z^B, y_z^B) = (1,14)$. The indifference curves that go through allocation **z** provide Ayanda and Biko with utilities $u_z^A = u_2^A$ and $u_z^B = u_2^B$.

M-CHECK Biko's indifference map would look exactly the same as in Figure 4.4 b. if we rotated the Edgeworth box 180° to measure Biko's allocation from the lower left-hand corner.

In panels a. and c., $u_2^A = u_z^A$ is Ayanda's indifference curve through **z**. In panels b. and c., $u_2^B = u_z^B$ is Biko's indifference curve through **z**. The indifference maps for both Ayanda and Biko have indifference curves through every point in the box, but (following "the map is not the territory" principle) we show only three in the figure.

M-Note 4.1: Evaluating utilities at an allocation

Given the assumption of that Ayanda's utility function is a Cobb-Douglas with $\alpha^A = \frac{2}{3}$ and Biko's utility function is a Cobb-Douglas with $\alpha^B = \frac{1}{3}$, we can calculate their utilities at the allocation **z**. Remember that the exponents in the Cobb-Douglas utility function represent the person's intensity of preference for the good. In this example, Ayanda likes coffee more than Biko does.

Ayanda has a Cobb-Douglas utility function $u^A(x^A, y^A) = (x^A)^{\frac{2}{3}}(y^A)^{\frac{1}{3}}$:

- She has 9 kilograms (kgs) of coffee and 1 gigabyte (gb) of data.
- So her allocation at point **z** is $(x_z^A, y_z^A) = (9, 1)$
- At her allocation **z** her utility is $u^A(x_z^A, y_z^A)$.
- So for 9 kgs of coffee and 1 gb of data: $u^A(9, 1) = (9)^{\frac{2}{3}}(1)^{\frac{1}{3}} = 4.326749$.

Checkpoint 4.1: Biko's utility at allocation z

Using the method shown in M-Note 4.1, what is Biko' utility at the allocation given by point **z** in the Edgeworth box.

4.3 The Pareto-efficient set of feasible allocations

Which allocations in the Edgeworth box are Pareto-efficient?

It's easy to see that simply throwing away some of x or y cannot be efficient because allocating those portions to Ayanda and or Biko instead would have made at least one of them better off without making the other worse off. So Pareto-efficiency also requires that Equations 4.1 are satisfied as equalities, not as inequalities. By construction, any of the great many allocations in the Edgeworth box allocates all of the coffee and data to one or the other participant, and meets this criterion.

To narrow things down, Ayanda and Biko could agree that the final allocation chosen must be Pareto-efficient. In Figure 4.5 we show Ayanda and Biko's indifference curves through an arbitrary allocation **z** and three more indifference curves for each person: two indifference curves higher and one indifference curve lower than for allocation **z**.

The endowment allocation is not Pareto-efficient

Think about **z** as a hypothetical allocation, for example, if Biko said: "Ayanda, how about you have 9 kg of coffee and I get the 1 kg remaining, while I get 14 gb of the data, and you get the 1 gb remaining." We can see, however, that **z** in Figure 4.5 is not Pareto efficient. The reason is that at the allocations given by point **z**, Ayanda's and Biko's indifference curves:

- *intersect*, which means
- they have *different slopes*,
- indicating *different marginal rates of substitution*

REMINDER In games like the Ultimatum Game in Chapter 2 any allocation of the pie in which the entire endowment is allocated to one of the players or the other – in other words " no money left on the table" is Pareto-efficient. But the allocations resulting from the Ultimatum Game are frequently inefficient because when the Responder rejects the Proposers offer both players get zero, and *all* of the money is left on the table.

M-CHECK Even if for some reason we were not to allow the allocation to involve fractional quantities of the goods and require that allocations be integers, there 176 possible allocations to be exact (that's 11×16, in case you are wondering, because we would then have to include zeroes as possible allocations for the goods).

REMINDER The *marginal rate of substitution* is the negative of the slope of the indifference curve. It is also equal to the ratio of the marginal utilities of the two goods, x and y, i.e. $mrs^A(x, y) = \frac{u_x^A}{u_y^A}$. The marginal rate of substitution is also the willingness to pay for x in terms of y. The people's marginal rates of substitution have the dimensions data/coffee (data for coffee).

REMINDER For an outcome to be *Pareto-superior* to another, at least one participant must be made better off – get higher utility – and no participant can be made worse off – get lower utility.

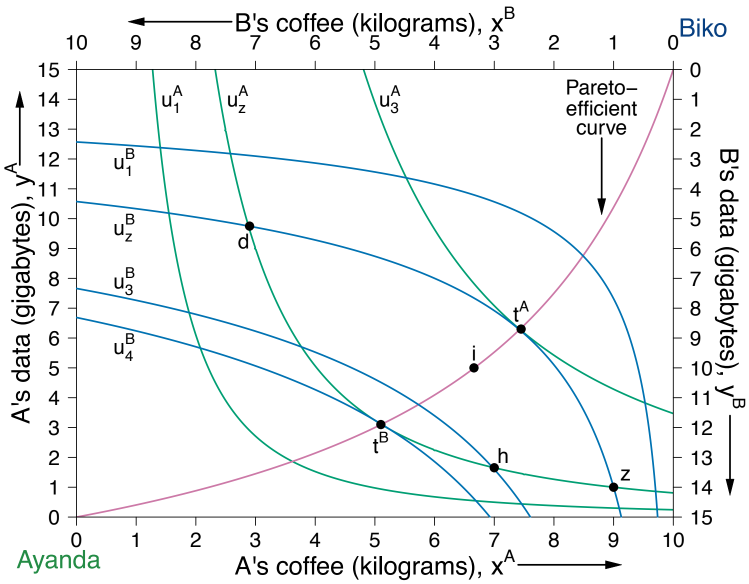


Figure 4.5: **Pareto-efficient allocations** To make this figure we let $u^A = (x^A)^{\frac{2}{3}}(y^A)^{\frac{1}{3}}$ and $u^B = (x^B)^{\frac{1}{3}}(y^B)^{\frac{2}{3}}$. So Ayanda prefers has a stronger preference for coffee, and Biko for data (they have asymmetrical preferences). Allocation **h** is Pareto-superior to allocation **z**, but it is not Pareto-efficient because an alternative point, e.g. allocation $\mathbf{t^B}$, is Pareto-superior to point **h** (Biko is better off without Ayanda being worse off). All points along the Pareto-efficient curve between **i** and $\mathbf{t^B}$ would be both Pareto-superior to **h** and **z** and Pareto-efficient.

- meaning that Ayanda and Biko have *different offer-prices* for the two goods, which means their willingness to pay to acquire more of one or the other good differ;
- and this means that there is *a feasible Pareto-improving exchange that has not been realized*,
- so these allocations are *not Pareto efficient*.

Figure 4.5 provides us with a numerical demonstration of the above logic. We have seen that the difference between the two people's marginal rates of substitution at the point **z** indicates that they can make a Pareto-improving trade – Ayanda giving up some of her coffee in return for some of Biko's data – at their endowment allocation **z**.

In M-Note 4.2 we show that Ayanda's mrs is $\frac{2}{9}$ and Biko's is 7 in Figure 4.5. This means that Ayanda is willing to pay at most a kilogram of coffee for $\frac{2}{9}$ of a gigabyte of data, while Biko is willing to trade at most 7 gb of data for one kg of coffee.

A mutually beneficial trade could therefore take place at any price of coffee between $\frac{2}{9}$ of a gb of data and 7 gb of data. The low price would benefit Biko, with Ayanda not improving her utility at all. Correspondingly, if they traded at the high price Ayanda would make all the gains.

We return to how the price might be determined later. For now, we can eliminate point **z** in Figure 4.5 as a candidate for being a Pareto-efficient allocation.

M-Note 4.2: mrs in the Edgeworth box

At allocation **z** $(9,1),(1,14)$ in Figure 4.5, we can calculate each person's marginal rate of substitution and compare them. We computed what a person's $mrs(x,y)$ is when she has Cobb-Douglas utility in M-Note 3.4 in Chapter 3. We obtain Biko's from the same reasoning. We shall assume for this example that the two have *asymmetrical* preferences as in Figure **??**.

Let's start with Ayanda, assuming $\alpha^A = \frac{2}{3}$:

- $mrs^A(x,y) = \frac{u_x^A}{u_y^A} = 2\frac{y^A}{x^A}$
- Substitute in A's allocation at **z**: $mrs_z^A(x_z^A, y_z^A) = 2\frac{1}{9} = \frac{2}{9}$

Ayanda is willing to pay $\frac{2}{9}$ of a gigabyte to get a kilogram of coffee, or to sell a kilogram of coffee for $\frac{2}{9}$ of a gigabyte of data.

Now for Biko, assuming $\alpha^B = \frac{1}{3}$:

- $mrs^B(x^B, y^B) = \frac{u_x^B}{u_y^B} = \frac{1}{2}\frac{y^B}{x^B}$
- Substitute in B's allocation at **z**: $mrs_e^B(x_e^B, y_e^B) = \frac{1}{2}\frac{14}{1} = 7$

Biko is willing to pay 7 gigabytes of data for a kilogram of coffee, or to sell kilogram of coffee for 7 gigabytes of data.

We can see that $mrs^A \quad < \quad mrs^B$ because $\frac{2}{9} \quad < \quad 7$. This shows up in Figure 4.5 where the slope of Ayanda's indifference curve is steeper than the slope of Biko's indifference curve at allocation **z**.

Which allocations are Pareto efficient? The $mrs^A = mrs^B$ rule

The same reasoning allows us to eliminate most of the other points too. Remember the demonstration that showed point **z** to be Pareto-inefficient started with "at the allocations given by these points Ayanda's and Biko's indifference curves intersect." We explain this further in M-Note 4.2. So *any* allocation at which the indifference curves intersect, like point **h** in Figure 4.5 cannot be Pareto efficient.

To find the Pareto-efficient allocations, we need to determine which allocations remain after we have eliminated all of those at which the indifference curves cross. To do this we can run the above reasoning in reverse.

If the two indifference curves (one of Ayanda's, one of Biko's) share a common point (that is, that represent the utilities at a particular allocation) but do not *intersect*, then the two indifference curves must be *tangent*. This tells us (reversing the logic above about indifference curves that intersect) that if Ayanda's and Biko's indifference curves:

- are *tangent*, this means that
- they have the *same slopes*, indicating

- *identical marginal rates of substitution*,
- meaning that Ayanda and Biko have *the same willingness to pay for the two goods*.
- This is the same as saying that *their maximum willingness to pay to acquire more of the other's good is not greater than the least price at which the other would part with their good*
- and this means that there is *no feasible Pareto-improving exchange* into which both would voluntarily enter
- so the *status quo allocation is Pareto efficient*.

This gives us the following rule for an allocation between two players, A and B, being Pareto efficient:

$$\text{The } mrs^A = mrs^B \text{ rule:} \qquad \text{mrs}^A(x^A, y^A) = \text{mrs}^B(x^B, y^B) \tag{4.1}$$

This rule differs from the seemingly similar $mrs = mrt$ rule for a single individual because this new rule applies to strategic interactions among two or more inter-dependent actors, of the kind that occur in markets for labor, credit, and many goods. The superscripts A and B are there to remind you that two (or more) players are involved in this rule.

The points $\mathbf{t^A}$, $\mathbf{t^B}$ and **i** lie on the purple Pareto-efficient curve in Figure 4.5. We will often abbreviate the Pareto-efficient curve to PEC. The Pareto-efficient curve consists of all Pareto-efficient allocations, including Ayanda getting all of both goods, or the reverse.

Confining allocations to the Pareto-efficient curve limits the choices that Ayanda and Biko need to make. But the question is still far from answered. Moving from one Pareto-efficient allocation to another must make one of the participants better off and the other worse off. The Pareto efficiency criterion is not going to help them decide which of the points on the Pareto-efficient curve they would consider to be the best.

So they face a problem and a conflict of interest.

- *The problem* is that there are still innumerable Pareto-efficient outcomes on the PEC and they need some way to decide which one to choose.
- *The conflict of interest* is that Ayanda prefers points on the PEC to the northeast in the Edgeworth box, while Biko prefers points to the southwest, so they will not agree on which Pareto-efficient division of the coffee and data to make.

M-Note 4.3: Computing the Pareto-efficient Curve

We will use $mrs^A = mrs^B$ rule to work out the equation for the Pareto-efficient curve.

REMINDER: THE $mrs = mrt$ RULE We derived a similar rule for single person interactions in Chapter 3 The $mrs = mrt$ rule (with a few exceptions) identifies the constrained optimal allocation for a single individual as the bundle at which the marginal rate of substitution (the person's willingness to pay for more of the y-good) is equal to the marginal rate of transformation (the opportunity cost of getting more of the y-good).

M-CHECK Like the $mrs = mrt$ rule, $mrs^A = mrs^B$ does not work in every case. The Pareto-efficient point may be a corner solution (not a tangency) at which one of the goods is not consumed at all by one of the players, and a tangency identified by the rule may be a minimum not a maximum. The reasons are the same as were explained for the $mrs = mrt$ rule.

PARETO-EFFICIENT CURVE The Pareto-efficient curve is all outcomes that are *Pareto-efficient*. At a Pareto-efficient outcome the marginal rates of substitution of the two parties are equal so the $mrs^A = mrs^B$ holds. The Pareto efficient curve is sometimes called the "contract curve", a term we do not use because there need not be any contract involved (e.g. when an outcome in our thought experiment was implemented by the Impartial Spectator.

To find the Pareto-efficient curve, we set Ayanda's marginal rate of substitution equal to Biko's marginal rate of substitution. We already know that $mrs^A(x^A, y^A) = 2\frac{y^A}{x^A}$ and $mrs^B(x^B, y^B) = \frac{1}{2}\frac{y^B}{x^B}$. We also know that $\bar{x} = x^A + x^B = 10$, so $x^B = \bar{x} - x^A$ and $\bar{y} = y^A + y^B = 15$, so $y^B = \bar{y} - y^A$. Solutions to these equations for x^A, y^A, x^B, y^B are Pareto-efficient allocations.

We set the marginal rates of substitution equal to each other and use these conditions to find the Pareto-efficient curve:

$$\begin{aligned}
mrs^A(x^A, y^A) &= mrs^B(x^B, y^B) \\
2\frac{y^A}{x^A} &= \frac{1}{2}\frac{\bar{y} - y^A}{\bar{x} - x^A} \\
4\frac{y^A}{x^A} &= \frac{15 - y^A}{10 - x^A} \\
4(10 - x^A)y^A &= x^A(15 - y^A) \\
40y^A - 4x^A y^A &= 15x^A - x^A y^A \\
(40 - 3x^A)y^A &= 15x^A \\
\text{Pareto-efficient Curve} \quad y^A &= \frac{15x^A}{40 - 3x^A}
\end{aligned}$$

The Pareto-efficient allocations lie on an upward sloping curve between the two origins.

Checkpoint 4.2: Conflict and symmetry of preferences on the Pareto-efficient curve

a. Using Figure 4.5 do the following:
   i. Explain Ayanda's and Biko's preference among the Pareto-efficient points $\mathbf{t^A}$, $\mathbf{t^B}$, and **i**.
   ii. Show that they rank these points in opposite order.
   iii. Explain why for any two points on the Pareto-efficient curve, Ayanda will prefer one point and Biko another point; they will never agree on which is preferable.

b. Work out the formula for the Pareto-efficient curve when the two people have identical Cobb-Douglas utility functions where $\alpha^A = \alpha^B = \frac{1}{2}$. (It's easier than the asymmetrical case.) The solution is that the Pareto-efficient curve is given by $y^A = \frac{3}{2}x^A$. But it's important for *you* to work out how to get the solution.

4.4 Adam Smith's Impartial Spectator suggests a fair outcome

Ayanda and Biko are going to have to figure out some way – other than each simply trying to get more – for picking an allocation. This means stepping back and looking at the problem without thinking about their own particular preferences. They would probably experiment with some simple rules. They could adopt:

- "*finders keepers*" rule and allocate the goods to whoever had first discovered the discarded coffee and data; but this might not seem fair.
- the *fifty-fifty* norm of the landlords and sharecroppers in Chapter 2, and

EXAMPLE To see how maximizing total utility might lead to unacceptable outcomes, think about two people, one who in order to minimize her carbon footprint or for other ethical reasons has cultivated a simple life style and is not much interested in increasing her material consumption and the other who has cultivated a taste for luxuries and will be miserable without them. Maximizing total utility would require giving most of the goods to the second person.

each take half the quantity of the two goods; but if they have different preferences (as is the case in panel b of Figure 4.5) splitting both goods equally would not even be Pareto-efficient (an equal split is not on the purple Pareto-efficient curve.)

- the *maximize total utility principle*; but this places no value on equality, and might result in selecting an allocation in which one person had most of the goods (and utility) and the other little of either.

To develop more satisfactory rules, they might consult an Impartial Spectator a fair and impartial spectator who can assist them (and us) in reasoning about what a good outcome might be. We use upper case letters for her name to remind you that she is an entirely made-up character, a thought experiment, and not a part of the game in which Ayanda and Biko are engaged. The Impartial Spectator is *not* a person, she is a thought experiment representing our conscience, allowing us to explore differing values and how they would lead us (and Ayanda and Biko) to select a particular allocation as the best.

We're going to follow the Impartial Spectator's thinking by looking at different criteria that she could adopt. For example, she could ask:

- Are the *procedures* that determined the allocation fair?
- Is the *outcome* itself fair?

The first criterion is referred to as a *procedural* judgement, and therefore she judges the outcome based on the *procedure* used to acquire the goods. She would ask for example if the original endowment bundles had been acquired fairly, for example through hard work, gifts, or exchanges in which both Ayanda and Biko had an equal opportunity to acquire the goods. She would go on to inquire if the process of trading had itself been fair: for example did either of them have unfair advantages in determining the price at which they would exchange.

The second criterion is called *substantive*: it asks about the *substance* of the resulting allocation, asking for example if it is fair (no matter how it came about).

Both criteria are important, but we will focus on the *substantive* judgements because it allows us to illustrate how the Impartial Spectator could select the "best" allocation by solving a constrained optimization problem. For the Impartial Spectator to make judgments among Pareto-efficient allocations that give Ayanda and Biko different levels of utility using the constrained optimization method, she needs to refer to two pieces of information:

- The set of all Pareto efficient combinations of utility levels that Ayanda and Biko could experience by allocating the goods in different ways;

- How she (the spectator herself) values each of these combinations of the utility levels of the two.

The utility possibilities frontier

Setting aside Pareto-dominated allocations, the Impartial spectator will concentrate on the boundary of the set of feasible utility pairs of the two. This is called the **utility possibilities frontier** (UPF) and it, shows all combinations of Ayanda and Biko's utilities associated with allocations on the Pareto-efficient curve.

In Figure 4.6 In panel a we show an Edgeworth box of the two player's allocation problem in which they have identical preferences in the way they each value coffee and data.

In panel b, we show the utility possibilities frontier for this case. For the moment, ignore the downward-sloping blue lines.

The utility possibilities frontier shows Ayanda's utility (u^A) on the horizontal axis and Biko's utility (u^B) on the vertical axis as we move from one extreme of the Pareto-efficient curve in figure 4.5 to the other.

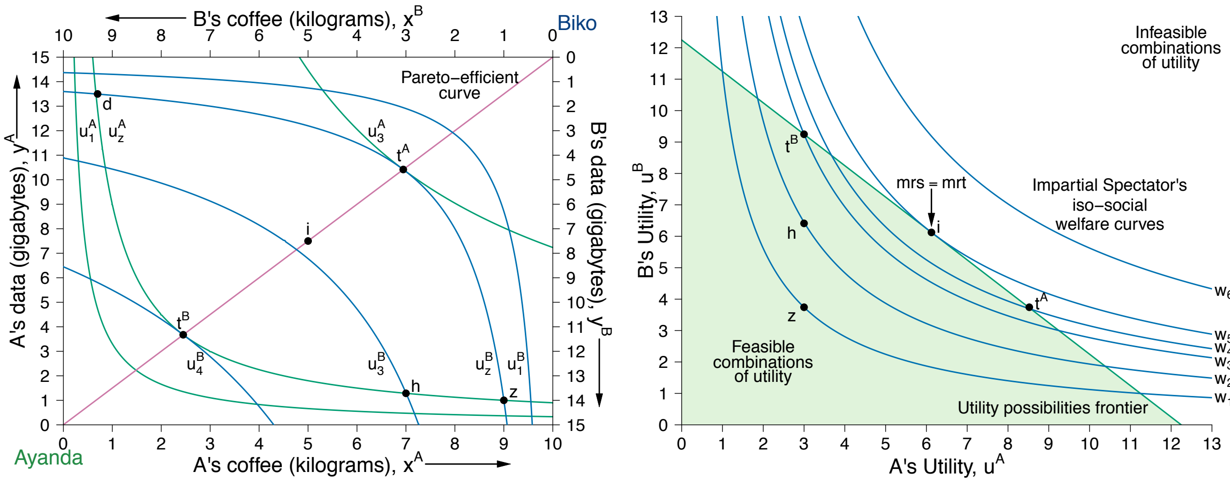
The UPF is downward-sloping because the participants are in *conflict* over who gets what share of the possible distributions of utility as the allocation changes on the Pareto-efficient curve in the Edgeworth box. The UPF is constructed from the Pareto-efficient curve by translating each Pareto-efficient allocation $(x^A, y^A; x^B, y^B)$ into a point $(u^A(x^A, y^A), u^B(x^B, y^B))$ that represents the utility levels of the two participants at that allocation. To construct it, take any point on the Pareto-efficient curve in Figure 4.5, say point $\mathbf{t^A}$, then read from the two indifference curves through $\mathbf{t^A}$ the two levels of utility of Ayanda and Biko at that allocation (namely 8.52 and 3.74 respectively), then go to Figure 4.6 where those two utility levels become the coordinates in the utility possibility graph of point $\mathbf{t^A}$ in the Edgeworth box graph.

UTILITY POSSIBILITIES FRONTIER (UPF) The utility possibilities frontier is a curve plotted with u^A on the horizontal axis and u^B on the vertical axis that shows the utility of the two participants at all Pareto-efficient outcomes.

Points $\mathbf{t^A}$, **i**, and $\mathbf{t^B}$ correspond to the same lettered points in Figure 4.5 and portray the *utilities* of each of the two at these Pareto-efficient allocations. In similar fashion, points **z** and **h** correspond to the same letters in Figure 4.5, but these allocations, being Pareto-inefficient are of no interest to the Impartial Spectator.

REMINDER The utility possibilities frontier is similar to what we did in Chapter 1 to understand the Pareto efficiency of different game outcomes. The UPF is another feasible frontier introduced in Chapter 3, since it shows the feasible combinations of utility possible given the available goods and the preferences of the participants.

As in the case of other feasible frontiers the negative of the slope of utility possibility frontier, $-\frac{\Delta u^B}{\Delta u^A}$, is the marginal rate of transformation of B's utility into A's utility by progressively giving A more of the goods and B less. This is also the opportunity cost of A having more utility in terms of the sacrifice in B's utility necessary to allow this. A steep utility possibility frontier means that for A to gain one unit of utility, B must sacrifice a lot.



(a) Identical Preferences Edgeworth box

(b) The utility possibilities frontier (UPF) and the Impartial Spectator's iso-social welfare curves (w)

Figure 4.6: **The utility possibilities frontier (UPF) and the Impartial Spectator's iso-social welfare curves (w).** The utility functions of the two players used to create panel a are identical, with in both cases $\alpha = 0.5$. Because they both value the two goods in the same way, they consume them in the same proportions at all points on the PEC. The only differences is which player has more. Each point in Panel b. corresponds to an allocation in the Edgeworth box shown in panel a. The downward-sloping curves in Figure b are the Impartial Spectator's iso-social welfare curves, corresponding to six levels in his judgement of social welfare w_1 through w_6. Social welfare is lower at points closer to the origin. The allocation given by point **i** is the social optimum determined by the $mrs = mrt$ rule.

Checkpoint 4.3: The UPF and the PEC

1. Explain why the utility possibilities frontier in Figure 4.6 is downward-sloping.
2. Explain why if the utility functions of the two differ, an even split of the two goods – half of each to Ayanda and half of each to Biko – could *not* be the Impartial Spectator's choice of the best allocation.

The Impartial Spectator's social welfare function

Which point on the UPF – in other words which allocation of the goods between Ayanda and Biko – the Impartial Spectator ranks as best will depend on her values. She has to compare how much she values the utility of Ayanda and Biko respectively and how this varies depending on the level of utility that each are experiencing.

To do this she has to be able to compare *how much* the levels of utility for the two for each of the allocations on the UPF. She knows that Ayanda prefers allocation $\mathbf{t^A}$ to $\mathbf{t^B}$ (and that Biko ranks these two allocations the other way around). She needs to treat the utility of each like ordinary numbers that measure the *size* not just the *rank* of something, in this case the *cardinal utility* of each.

A summary of the Impartial Spectator's evaluation of different utility distributions (u^A, u^B) is provided by her **social welfare function**, $W(u^A, u^B)$. This is similar to the utility function that expresses a person's preferences over bundles of goods, (x, y), but remember the Impartial Spectator is not a person, but a thought experiment. This is why it is called the social welfare function rather than the Impartial Spectator's utility function. A social welfare function provides a way of treating the utilities of the citizens as being cardinal

SOCIAL WELFARE FUNCTION A social welfare function is a representation of "the common good" based on a some weighting of the utilities (u^A, u^B, and so on) of the people making up the society. We can write a social welfare function in the form $W(u^A, u^B)$.

numbers that are comparable across people (like height or weight).

An example is a social welfare function that expresses *total welfare* as the *product* of the utility of the citizens, each utility raised to some exponent.

$$\text{Example Social Welfare Function:} \quad W(u^A, u^B) = (u^A)^{\lambda}(u^B)^{1-\lambda} \quad (4.2)$$

This social welfare function has the same form as a Cobb-Douglas utility function: the participants' levels of utility are the "goods" for the Impartial Spectator. When $\lambda = 0.5 = 1 - \lambda$, then the Impartial Spectator:

- weights the two peoples' utilities *equally*; and
- places diminishing marginal value on increases in the utility of either of Ayanda or Biko; the more they consume of the goods the greater is their utility and therefore the less they add to the Impartial Spectator's social welfare.

Because the Impartial Spectator values the two people's utilities equally, and (in the judgement of the Spectator) the marginal value of increased utility is diminishing, she will not rank highly any outcome in which one or the other gains most of both goods.

Just as we can use indifference curves to represent a person's utility function over goods, we can use **iso-social welfare curves** to represent the Impartial Spectator's social welfare function over the utility distribution between people. The level of social welfare is the same along an iso-social welfare curve, just as utility was the same along an indifference curve.

Given the Impartial Spectator's social welfare function, the problem of choosing the Pareto-efficient allocation of coffee and data becomes a constrained maximization problem similar to those we studied in Chapter 3. The feasible frontier for the Impartial Spectator is the utility possibility frontier, because it represents the levels of utility that are achievable given the amount of goods available and the preferences of the participants.

The iso-social welfare curves of the social welfare function are analogous to indifference curves for a single individual, but apply to the utilities of the two people not the two goods consumed by the single individual and express the valuations of the Impartial Spectator, not the preferences of the individual. Similar to the individual indifference curve, the negative of the slope of the iso-social welfare curve at any point (u^A, u^B) is the Impartial Spectator's marginal rate of substitution of Ayanda's utility in terms of Biko's utility. And we can use the $mrs = mrt$ rule to find the constrained social welfare-maximizing allocation. It is the point where the UPF is tangent to an iso-social welfare curve.

REMINDER Assigning cardinal utility numbers to bundles means that we can make statements like:

- for Annette, the outcome (x', y') is twice as good as (x, y) but also
- the sum of the utility experienced by Annette and Brenda is greater with outcome (x', y') than with outcome (x, y) because $u^A(x', y') + u^B(x', y') > u^A(x, y) + u^B(x, y)$

The sum of the utilities of the two – in the second statement – is an example of a social welfare function.

ISO-SOCIAL WELFARE CURVE Iso-social welfare curves show constant or equal ("iso") levels of welfare, $\bar{W}$, for different combinations of utility between A and B. The negative of the slope of the iso-social welfare curve is the Impartial Spectator's marginal rate of substitution ($mrs^{SW}(u^B, u^A)$) of Ayanda's utility for Biko's utility.

For example, suppose the Impartial Spectator's social welfare function is:

$$W(u^A, u^B) = (u^A)^{\frac{1}{2}}(u^B)^{\frac{1}{2}}$$

which puts a identical weight on the utilities of the two parties, then given that Ayanda and Biko have identical utility functions, the social welfare maximum shown in the Edgeworth box in Figure 4.5 Panel a is $x^A = 5, y^A = 7.5, x^B = 5, y^B = 7.5$ or a fifty-fifty split of each good. In the utility possibility frontier graph in Figure 4.6 this is point **i**. If their preferences *differed*, then the social optimum would result in each getting *different* amounts of x and y.

Different Impartial Spectators might have different social welfare functions to rank distributions of the utilities of the two parties, leading to the choice of different social welfare maximizing Pareto-efficient allocations. Societies do not have an Impartial Spectator to determine how to weight the competing interests of society's members in a social welfare function. Instead, in a democratic society we debate the question of distribution and sometimes come to a consensus (and sometimes to a deadlock). Controversy about the rights and wrongs of economic policies such as the tax rates paid by wealthy people and the provision of public services to all, are often implicitly about the weights (such as a, in Equation (4.2)) that policy-makers should place on the well-being of different people.

Here we see a sharp contrast between the Pareto-efficiency criterion and the maximization of social welfare. Preferring a particular Pareto-efficient allocation over an alternative allocating in which both are worse off cannot be a matter of conflict. Maximization of some particular social welfare function subject to the constrain of the utility possibility frontier – some gaining and some losing depending on the social welfare function used – and is certain to be controversial.

The imaginary Impartial Spectator helps us understand how values dictate what we think of as better or worse allocations. These outcomes, as we have seen in previous chapters and we will now see in greater detail, depend on the rules of the game. So the Impartial Spectator will have something to say about how we evaluate which are better or worse institutions by which organize the process of exchange.

4.5 Property rights and participation constraints

The scenario of Ayanda and Biko enjoying their coffee and data in their student residence and deciding how to allocate them helps us understand the abstract issues of Pareto efficiency and fairness. Very similar issues arise when instead we consider Ayanda and Biko to be total strangers, interacting in a market. But in this new setting the allocation will not be determined by some imaginary Impartial Spectator. Instead, the allocation will be determined by who initially owns which goods and the rules of the game that regulate how Ayanda and Biko might benefit by exchanging some of their goods with each other.

Market institutions: Property rights and participation constraints

Nobody actually owned the data and coffee that the Impartial Spectator allocated in our thought experiment and Ayanda and Biko were not really engaged in a game. This is not how markets work. Key aspects of the rules of the market game are:

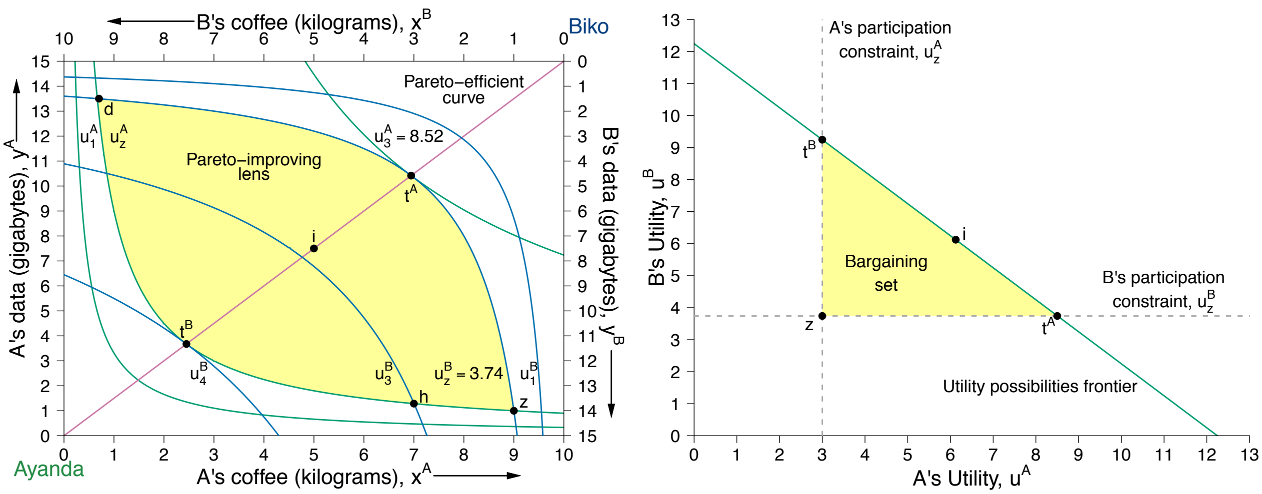
- The *rule of law* establishes that the institutions – the laws and other informal rules – governing the interaction are observed, and not violated by arbitrary acts (for example theft of the others goods by one of the traders or confiscating by a government official).
- *Private ownership*. At any moment in the game the goods are the private property of one or the other of the players, so a point in the Edgeworth box indicates a distribution of property between the two. The distribution at the start of the game is called each player's *endowment.*
- Private property and the rule of law mean that each player as *option to refuse offers* so any exchanges that a player will agree to participate in must be Pareto improvements over the endowment
- *Asymmetric bargaining power* will affect the nature of the exchanges that are executed, and who captures the greater share of the gains from exchange.

Private property does not distinguish between the two parties: each have *identical* rights to exclude the other from their bundle of goods. This would be true even if Biko initially owned all of the goods and Ayanda had none or the other way around. In this respect private property rights provide a level playing field because the right to exclude others from the use of your goods does not depend on how many goods you have, or on your identity.

HISTORY It has not always been true that one's property rights did not depend on your identity. In many societies, some people – such as women – did not have the right to own property, and some people – such as enslaved people – were treated as property.

The exchange process begins with the property people "start with," that is, their endowment allocation. These endowments exist *before the exchange we are considering happens*. But we are cutting into time at a particular moment. These endowments, which are the status quo of our game, are the result of similar games played in the past, and also other games in which who owns what goods may have been determined by *force* and not by *voluntary exchange*. This means that unlike the Impartial Spectator starting with a clean slate – any allocation in the Edgeworth box is up for consideration – and advising Ayanda and Biko on the division of a pile of goods they have tripped over in their student residence, market exchange starts from one particular point in the Edgeworth box the **endowment allocation**. The rules of the

ENDOWMENT ALLOCATION The ownership of goods at the start of the game is termed the endowment allocation. It is the starting point of the game, but in applications to real economies who owns what at any point in time is the outcome (not the starting point) of other interactions that have determine who owns what.



(a) Edgeworth box for self-regarding Ayanda and Biko

(b) The utility possibilities frontier

Figure 4.7: **Edgeworth box, the utility possibility frontier, and the bargaining set.** In Panel **a** u_z^A is Ayanda's utility at her endowment and is her *participation constraint* (shown by indifference curve u_z^A) and u_z^B is Biko's utility at his endowment and is his *participation constraint* (shown by indifference curve u_z^B).In Panel **b**) the coordinates of the x- and y-axis intercepts of the utility possibilities frontier give the utilities of the two when either Ayanda (x-axis intercept) or Biko (y-axis intercept) have all of the economic rents.

game then determine how the two can move to some other post exchange allocation. The endowment allocation is important for two reasons:

- it is the *starting point* of the process and
- and also, because the exchange is voluntary meaning they can refuse to trade, it is their fallback position, that is, the *worst they can do*.

The participation constraint (PC)

To see how the second bullet above will narrow down that the post-exchange allocation can be, starting at any given endowment allocation we introduce the following notation, along with panel a. of Figure 4.7 (we will explain panel b. below). The *endowment bundle* of person i is (x_z^i, y_z^i) where the superscript indicates who person i is ($i = A$ for Ayanda, $i = B$ for Biko). This allocation is point **z** in the figure. It is identical to point **z** in previous figures, but instead of being some hypothetical allocation that the Impartial Spectator was trying out it is now something entirely different: it is what Ayanda and Biko *own* at the start of the game.

From point **z** in the figure you can see that Ayanda's and Biko's endowments of coffee and data are:

- Ayanda's endowment: $(x_z^A, y_z^A) = (9, 1)$
- Biko's endowment: $(x_z^B, y_z^B) = (1, 14)$

Introducing history in the form of initially privately owned endowments, along with the voluntary transfer requirement, limits the possible allocations that can result from exchange.

REMINDER The participation constraint is also the *fallback* in the exchange scenario, the utility that a person can certainly secure if they choose not to participate in exchange at all.

Because they can refuse any deal and therefore experience the utility from their endowment bundle, they will not accept any post-exchange bundle that makes them worse off than their fallback utilities. The indifference curves, u_z^A and u_z^B, that include the endowment point are the post exchange bundles that yield a utility identical to their fallback position.

These two indifference curves are called their **participation constraints**. They are called participation constraints because Ayanda will not *participate* in (that is she will refuse) any offer that would give her a post exchange bundle below and to the left of u_z^A. Likewise Biko will not participate in any offer that would give him a post exchange bundle above and to the right of u_z^B (labeled as $u_z^B = 3.74$ in Figure 4.7).

The right to refuse an exchange that will make a player worse off – the basis of the participation constraints – reduces the possible set of post exchange allocations consistent with voluntary exchange, and starting from the endowment allocation indicated by point **z** in panel **a**.

The yellow colored space between the two constraints – the indifference curves, u_z^A and u_z^B – is the entire set of allocations that are Pareto superior to point **z** and which therefore could be the result voluntary modifications of the endowment allocation by means of exchange. This area is called the **Pareto-improving lens**.

PARETO IMPROVING LENSThe set of allocations that are Pareto superior to the fallback options of the players is the Pareto improving lens – shaded yellow in the figures to follow in the rest of the book.

4.6 Symmetrical exchange: Trading into the Pareto-improving lens

In this section we start with the assumption that the two traders have identical preferences. That is, that their Cob-Douglas utility functions have $\alpha^A = \alpha^B = 0.5$.

We used a hypothetical point **z** in Figure 4.5 to show that an allocation where the indifference curves cross cannot be Pareto-efficient. Our demonstration consisted of showing that at such an allocation both Ayanda and Biko could benefit from exchange.

We can now use the same reasoning to illustrate how starting at point **z**, now an endowment allocation – a real distribution of ownership of two bundles – the two could trade into the Pareto improving space, and eventually all the way to the Pareto-efficient curve. Each person has a willingness to pay for x in terms of y, their marginal rate of substitution at the endowment allocation **z**. Ayanda's maximum willingness to pay is her $mrs^A(9,1) = \frac{1}{9}$ and Biko's maximum willingness to pay is his $mrs^B(1,14) = 14$.

M-CHECK With $\alpha^A = \alpha^b = 0.5$, their marginal rates of substitution are $mrs^A = \frac{y^A}{x^A} = \frac{1}{9}$ and $mrs^B = \frac{y^B}{x^B} = 14$. This means they could make an exchange at a "price" between $\frac{1}{9}$ of a gigabyte for a kilogram of coffee (Ayanda's mrs) and 14 gigabytes of data for a kilogram of coffee (Biko's mrs).

The difference between Ayanda and Biko's willingness to pay (mrs) signals an opportunity for Ayanda to trade data with Biko at a rate of exchange between her own marginal rate of substitution and Biko's marginal rate of substitution. A small exchange on these terms would move them to a *post-exchange allocation* upward and to the left of the endowment.

To stress that the game is entirely symmetrical imagine that they have agreed on a set of rules to determine the price and the amounts to be exchanged. At any allocation at which the mrs of the two differs (meaning their indifference curves intersect), take the following steps:

1. Pick a "price" midway between the mrs of the two. (This means that at point **z** the price would be $14 + \frac{1}{9}$ divided by 2, or 7.06.)

2. Ask the amounts that each would like to transact at the price of 7.06 gb of data for a kilo of coffee, for example how much coffee Ayanda would like to 'sell' at this price, and how much coffee Biko would want to 'buy'(these desired amounts will differ between the two);

3. Because the transfer has to be voluntary (nobody can be forced to buy more than they wish), transfer the amounts desired by the person who wishes to transact least.

4. At the resulting post-exchange allocation determine if the indifference curves are intersecting. If so return to step one and continue. If not (that is, if the indifference curves are tangent) end the game with this final allocation.

We can see that by this process the two will have moved, step-by-step from the endowment allocation at point **z** to a final post-exchange allocation that will be on the Pareto-efficient curve. We know that they will get there for two reasons:

- *Trades are Pareto-improving:* each trade they take moves them in the direction of the Pareto-efficient curve because moving in the other direction could not be a Pareto-improvement and would violate the voluntary transaction requirement.

- *Trade concludes at a Pareto-efficient outcome*: by the rules of the game they have adopted they will keep on exchanging until they are at a place where their mrs's are identical, which must be on the Pareto-efficient curve.

They could have adopted a different set of rules or institution for exchange. For example, they could have said that for step 1 above there will be two alternative prices, one just a little less than Biko's willingess to pay, and the other just a little more than the lowest price at which Ayanda would part with her coffee; and then just flipped a coin to see which of these prices they would use in that transaction. Having made that transaction, do another coin flip to see whose preferred price will be used, and so on until they reached a Pareto efficient point at which no further trade was possible.

Other than knowing that they would *eventually* get to the Pareto-efficient curve, we do not know which specific point on the curve they would get to. If the coin flips went in favor of Ayanda, they could end up close to $\mathbf{t^A}$ with Biko sharing very little of the gains from exchange. Or it could have gone the other way, somewhere near point $\mathbf{t^B}$. They even could have ended up at point **i** the allocation chosen by the Impartial Spectator.

The utility possibilities frontier in Panel b. of Figure 4.7 translates these allocations and the transactions supporting them into the utilities of the two players. The Pareto-improving lens in panel a. corresponds to the **bargaining set** in panel b. The first panels shows all of the allocations – denominated in quantities of x and y allocated to the two – that are Pareto improvements over the endowment allocation. The second – the bargaining set – shows the utility levels associated with every allocation in the Pareto-improving set.

The yellow area in panel b. is called the bargaining set because it compares outcomes of their bargains relative to their fallback position (point **z**). The bargaining set shows all of the possible distributions of the rents (utility greater than their fallback positions) that might result from their bargaining, depending on the rules governing how they bargain. These rules will determine the extent of the bargaining power of the players.

Checkpoint 4.4: Pareto improvements, rents, and Pareto efficiency

If point **h** is the post exchange allocation based on the endowment allocation of point **z**, explain the following:

a. Did Biko benefit from the exchange?
b. Did Ayanda benefit from the exchange?
c. What is the rent that Ayanda receives as a result of this exchange?
d. Did the exchange result in a Pareto improvement?
e. Is the post exchange allocation (point **h**) Pareto efficient?

4.7 Bargaining power: Take-it-or-leave-it

BARGAINING POWER is the ability to gain a large share of the mutual gains from exchange (total rents) made possible from some interaction, as may be determined by the rules of the game governing the interaction and the skill of the players in securing a favorable agreement under these rules.

In the bargaining over the distribution of coffee and data above, the two examples of rules of the game were symmetrical. Neither "split the difference between the willingness to pay of the two" or "alternating coin flips to see whose preferred price will be used" gave any obvious advantage to either player

But many bargaining interactions are asymmetrical. One of the players has most of the bargaining power. Bargaining power is the ability to gain a large share of the mutual gains from exchange (total rents) made possible from some interaction, as may be determined by the rules of the game governing the interaction and the skill of the players in securing a favorable agreement under these rules.

An example is the Ultimatum Game in Chapter 2 (whose name already suggests the asymmetry). The Proposer makes a offer of some fraction of the "pie". The Responder's strategy set is simply: accept or reject, or "take it or leave it." Being in a position to make that kind of an ultimatum is called take it or leave it power, or TIOLI power for short.

TAKE-IT-OR-LEAVE-IT-POWER A player with *TIOLI power* in a two-person bargaining game can specify the entire terms of the exchange – for example, both the quantity to be exchanged and the price – in an offer, to which the other player responds by accepting or rejecting.

In the coffee for data bargaining game if Ayanda had TIOLI power, she could have said to Biko: "I'll give you 2 kilograms of coffee and you give me 9 gigabytes of data. If you refuse, I will not agree to any other trade you might propose." In other words, "either accept the allocation I impose, or we both stay at our endowment, **z**." The assumption that Ayanda's offer is *credible* is important: if Biko suspects that he could refuse and Ayanda *would* listen to a counter-offer, the threat in the TIOLI offer would be empty.

REMINDER The Ultimatum Game discussed in Chapter 2 has this TIOLI structure including returning to the endowment point the the Responder rejects – both getting a payoff of zero, namely what they would have received had they not interacted. That is why it is called the Ultimatum Game, as the Proposer's offer is an ultimatum.

A bargainer with TIOLI power can often capture most or even all of the total rents that an economic interaction provides. This is because TIOLI power allows a bargainer to specify both:

- the *price* at which the goods will be exchanged and
- the *amount* of goods that will be exchanged

This means that the person with TIOLI power can just pick some preferred allocation – a point in the Edgeworth box different from the endowment point – and make that the TIOLI offer.

What take-it-or-leave-it offer will Ayanda make to Biko?

Ayanda does not care about Biko's utility, but she does care about how he will respond to her offer. If he rejects, then she gets her fallback option. She will realize that she must offer Biko a deal that Biko regards as better – or at least not worse – than the endowment. In other words, Ayanda has to take Biko's participation constraint as a limit on the kind of offer she will make. This is an example of the backward induction method that you learned in Chapter 2: Ayanda has to reason backwards from her understanding of what Biko will do *after* she has made her offer to what offer she should make **now**.

So Ayanda has the following constrained maximization problem: find a final allocation (different from the endowments) to propose at which Biko is no worse off than at the endowment and Ayanda is as well off as she can be. Ayanda knows that the solution to this problem must have two characteristics: It must:

- *satisfy Biko's participation constraint*, that is, be in (or on the boundary of) the Pareto-improving lens in Figure 4.7.
- *be Pareto-efficient*, but this is not because Ayanda cares any more about efficiency than she does about Biko: if she offered an allocation that satisfied Biko's participation constraint and was *not* Pareto efficient then there

would be some *other* allocation at which *she* could be better off and Biko not worse off.

Ayanda would probably offer Biko something just a tiny bit better than Biko's fallback utility to make sure he accepts. But to avoid having to keep track of that tiny amount in our thinking, here and in the rest of the book, we will assume that Biko will accept an allocation that *just meets* his participation constraint.

M-CHECK Remember that in Chapter 3, a utility maximizer is often constrained by a feasible frontier. Even with TIOLI power, Ayanda is constrained by Biko's *participation constraint*, that is, $u^B(x^B, y^B) \geq u^B_z$.

That solves the problem for Ayanda: to meet the two requirements bulleted above , she must find the intersection of the Pareto-efficient curve and Biko's participation constraint u^B_z. Therefore, Ayanda offers an exchange that implements point $\mathbf{t^A}$ at the extreme end of the Pareto-efficient curve in Figure 4.7 a. The same result is shown in Figure 4.7 b, where $\mathbf{t^A}$ represents the distribution of utilities resulting from the TIOLI allocation that Ayanda offered and Biko (barely and grudgingly) accepted. Point $\mathbf{t^B}$ at the other extreme of the Pareto-efficient curve in the figure in the Edgeworth box corresponds to the allocation where Biko has TIOLI power and point $\mathbf{t^B}$ on the utility possibilities frontier is the corresponding distribution of utilities.

REMINDER An outcome is socially efficient when it maximizes a social welfare function; what is deemed socially efficient depends on how the utility of each member of the population is weighted in the social welfare function.

We can see that the TIOLI allocation does not weight the two utilities identically (as did the social welfare function of the Impartial Spectator, which led to point $\mathbf{i}$). This why we say that allocation $\mathbf{t^A}$ is Pareto-efficient but not socially efficient, where the latter term is whatever the Impartial Spectator selected based on maximizing an equally-weighted social welfare function.

Two features of the TIOLI outcome where the participation constraint holds are important because they arise in many social coordination problems that involve a participation constraint:

1. *Pareto efficiency:* The PC-constrained outcome is Pareto-efficient.
2. *Inequality*: At PC-constrained outcomes the bargainer with TIOLI power gets all of the economic rent.

M-Note 4.4: Finding Ayanda's TIOLI Offer

We need two pieces of information to find Ayanda's TIOLI offer:

- The equation for the Pareto efficient curve (because we know that the resulting allocation will be Pareto efficient) and
- The equation for Biko's participation constraint (because we know that Ayanda will not offer him anything better than his utility at his endowment bundle).

The Pareto efficient curve: At Checkpoint 4.3 we asked you to find the Pareto-efficient curve for Ayanda and Biko when they have identical Cobb-Douglas utility functions with $\alpha = 0.5$. The solution is that the Edgeworth box has the following Pareto-efficient curve defined over the two people's allocations of x and y:

$$y^A = \left(\frac{3}{2}\right)x^A \tag{4.3}$$

$$\tag{4.4}$$

We can re-write Equation 4.5 in terms of x^B and y^B by substituting $x^A = \bar{x} - x^B$ and $y^A = \bar{y} - y^B$ in the equation to find:

$$y^B = \left(\frac{3}{2}\right)x^B \tag{4.5}$$

As you can see, the Pareto-efficient curve is a line from the one corner of the Edgeworth Box to the other. The utility functions, endowments and TIOLI offers calculated in this M-Note are the basis for Figure 4.7.

To find the TIOLI offers, we need the players' participation constraints because the player with TIOLI power wishes to maximize their utility subject to the participation constraint of the other player.

Biko's participation constraint:

B's fallback utility (his participation constraint (PC)) at his endowment $x_z^B = 1, y_z^B = 14$ is:

$$u_z^B(1,14) = (1)^{0.5}(14)^{0.5} = 3.74,$$

So we need to find the point on the PEC at which Biko has this level of utility.

A's TIOLI Offer: We substitute the Pareto-efficient curve's value for x^A into B's utility function that equal to B's fallback utility:

$$\begin{aligned}
u^B = (x^B)^{0.5}\underbrace{\left(\frac{3}{2}x^B\right)^{0.5}}_{\text{PEC}} &= \underbrace{u_z^B = 3.74}_{\text{PC}} \\
\Rightarrow (\frac{3}{2})^{0.5}x^B &= 3.74 \\
x_{TA}^B &= 3.74/(\frac{3}{2})^{0.5} = 3.05 \approx 3 \\
\therefore y_{TA}^B &= \frac{3}{2}x^B = \frac{3}{2}(3) = \frac{9}{2} = 4.5 \\
\therefore x_{TA}^A &= \bar{x} - x^B = 7 \\
\therefore y_{TA}^A &= \bar{y} - y^B = 10.5
\end{aligned}$$

So where "TA" means A had TIOLI power, the post-exchange allocation will be $x_{TA}^A = 7, y_{TA}^A = 11.5, x_{TA}^B = 3, y_{TA}^B = 4.5$. The post-exchange allocations imply that A made a TIOLI offer to B of 2 units of x ($x_{TA}^B - x_z^B = 3 - 1 = 2$) in exchange for 9.5 units of y ($y_{TA}^A - y_z^A = 14 - 4.5 = 9.5$). A's utility is $u_{TA}^A = 8.97$ and B remains on his participation constraint at $u_z^B = 3.74$.

Checkpoint 4.5: Biko's TIOLI offer to Ayanda

Given the TIOLI offer you just saw for Ayanda in M-Note 4.4, what would happen if the players' positions were reversed and Biko had TIOLI power over Ayanda? To answer this here is A's fallback utility at her endowment $x_z^A = 9, y_z^A = 1$:

$$u_z^A(9,1) = (9)^{0.5}(1)^{0.5} = 3,$$

a. What offer would Biko make?

b. What would the post-exchange allocations be? Explain.
c. Use Ayanda's utility at her fallback position that we found in M-Note 4.4.

4.8 Application: Bargaining over wages and hours

We illustrate TIOLI power below by a case in which the two bargainers drop their student personas to take on familiar roles in what is arguably the most important market in a modern economy: Ayanda is the owner of a company interacting with Biko, a prospective employee. In labor market bargaining over wages and working conditions the employer almost always has TIOLI power, stating the wage, the job and the hours. The worker accepts or not. We postpone until Chapter 15 the question: why might Ayanda get to have this power and not Biko?

So, leaving the world of coffee and data behind us, we will see that the sum of the mutual gains enjoyed by the two and how these are divided between them will depend on both their preferences and the rules of the game:

- *Power:* Do the two bargain symmetrically with neither one nor the other of them having first mover advantage? Is one of them first mover with take-it-or-leave-it power (TIOLI power)?
- *Fallback:* What is each person's fallback position? how well off are they if they do not exchange at all? Does Biko have other options than being employed by Ayanda? If Ayanda does not employ Biko, are there others she she could employ?

To fill in some answers to those questions, our two actors are now:

- *Ayanda, an employer*: whose endowment bundle is a sum of money only (no employees), and who in the absence of any exchange with Biko has nobody work for her; she will make Biko a take it or leave it offer of a sum of money in return for some number of hours of work for her, and
- *Biko, a worker*: who is applying to work in Ayanda's company, whose endowment bundle is free time only (no money); he has a maximum of 16 hours of (non-sleeping) time to spend, possibly working for Ayanda.

We introduce a more complete model of the labor market with competition among firms for workers and customers and among workers for jobs in Chapter 11 including the ways that unemployment benefits, competition among firms, and antitrust could affect these outcomes.

Quasi-linear preferences for money and time

QUASI-LINEAR FUNCTION A quasi-linear function depends *linearly* on one variable, e.g. y, and *non-linearly* on another variable, e.g. x, and has the form $u(x,y) = y + h(x)$, where $h(x)$ is a non linear function.

To represent the preferences of Ayanda and Biko we will introduce a new form of utility function, one that will simplify our analysis while still conveying

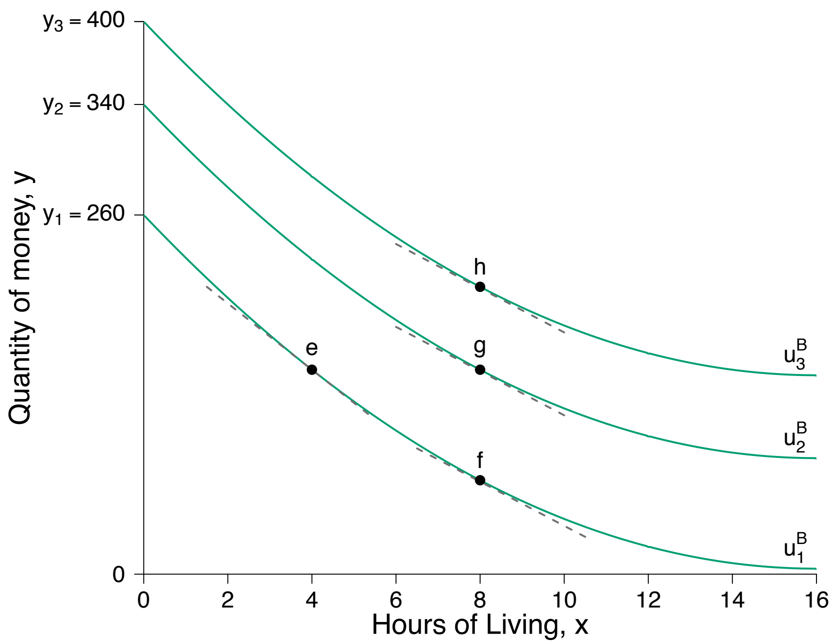


Figure 4.8: **Marginal rates of substitution with quasi-linear preferences.** With quasi-linear preferences, marginal rates of substitution depend only on the amount of the good x (here, Hours of Living for Biko), and not at all on the amount of money left over to buy other goods, y. As a result, indifference curves with different levels of utility are vertical displacements of a single curve – you can add or subtract an amount of y from the indifference curve to move it up or down.

the main insights. The function is called quasi-linear because utility is partly (“quasi") proportional to one of the arguments of the function, while being non-linear in the other arguments. The Cobb-Douglas utility function is not quasi linear because it is non linear with respect to both x and y.

As in the case of Harriet deciding how much fish to buy from Alfredo or Bob in Chapter 3, we will consider the second good as “money left over" after the exchange. This may seem odd because money is not something you value for itself. But money can buy you other goods which you do value: The utility of "money left over" is the utility of the goods which the person can purchase as a result.

We now illustrate a case where one person starts off with all of one good and none of the second, so the other person all of the second good,but none of the first. This could model you walking into the supermarket with money in your pocket (or more likely a credit card) and nothing in your shopping bags, and planning to walk out with less in your credit card and some groceries in our shopping bag. So it is a model of any kind of exchange. But here illustrate it by Ayanda (possibly) employing Biko.

The marginal rate of substitution for a person with quasi-linear preferences that are linear in "money" (y) depends only on the amount she has of the good or service for which her preferences are non-linear, not on the amount of money.

The reason why this is true is because:

- The marginal rate of substitution is the ratio of the marginal utility of x to the marginal utility of y
- The person's marginal utility for y is always a constant and it does not decline as she gets more y, so
- The marginal rate of substitution depends only on the marginal utility of x which varies with the quantity of x consumed because the function is non linear in this variable.

You can see this in Figure 4.8 by noticing that for a given amount of x the slope of the indifference curve (shown by the dashed tangent lines) is the same no matter how much y the person has, such as at $x = 8$ hours of living, as shown by points **f**, **g**, and **h**. This is because, given the quasi-linear utility function that we used to draw the figure, the willingness to pay for an additional hour of living (the marginal rate of substitution, that is, the negative of the slope of the indifference curve) does not depend on the amount of money lef over that the person has; it depends only on how many hours of living they have. This means that the indifference curves u_1, u_2, and u_3 in the figure are just shifted up replicas (you can see the amounts by which they are shifted up by comparing the vertical axis intercepts).

We can also compare points **e** and **g** at the same level of y: Biko likes to have more living time ($u_1^B < u_2^B$) and his willingness to pay for additional hours of Living declines the more he has (the indifference curve is less steep at **g** than at **e)**.

It is of course unrealistic to think that anyone would have truly linear preferences in any amount imaginable of money left over, for this would require that the person did not have diminishing marginal utility in the things that money can buy. But because "money" can be considered as generalized purchasing power that can be spent on a vast array of things, and because we do not consider changes in people's bundles of money making them either billionaires or paupers, it is a useful simplifying assumption.

M-Note 4.5: A quasi-linear utility function

Quasi-linear utility functions have the form:

$$u(x,y) = y + h(x) \tag{4.6}$$

Equation 7.28 is quasi-linear because it is *linear* in one variable, y, and non-linear in the other x as in the non linear function $h(x)$. For example, $h(x)$ could be quadratic in x or could include the natural log of x, $ln(x)$.

A quasi-linear utility function depends *linearly* on one variable, e.g. y, and *non-linearly* on another variable, e.g. x, and has the form $u(x,y) = ay + h(x)$, where a is a constant. Hence it is *quasi* or 'partly' linear. We often set $a = 1$. Two examples of such functions include, $u(x,y) = y + 20x - x^2$ and $u(x,y) = y + 10\ln(1+x)$.

Allocating money and time

For simplicity, we assume that Ayanda, the employer, and Biko, the worker, have quasi-linear utility functions. Both place a *constant value* per monetary unit on the money they have after the transaction (to purchase other things, for example). That is, the marginal utility of money is constant. So we can measure the utility of each in whatever monetary units they are using, which since their names are from South Africa, might as well be the South African Rand.

FACT CHECK At the time of writing this 1 Euro was equal in value to about 16 South African Rand (ZAR), 1 Pound Sterling was equal to about 21 South African Rand. In 2020, the hourly minimum wage in South Africa was ZAR 20.76.

Biko values his Living (that is his 16 waking hours, minus the time he "hires out" of himself to work for Ayanda). But the marginal utility of free time is decreases as the amount of free time he has increases, just as was the case for Aisha in Chapter 3. His first hour of free time is high, but his free time gets less and less valuable the more free time he has.

Ayanda places a value, too, on Biko's free time, but it is the opposite of Biko's value: she benefits by Biko having *less* free time and her having *more* of Biko's time working for her. The positive value she places on Biko's labor – like the positive value he places on his free time – depends on how much of it she gets. The marginal utility to Ayanda of Biko's labor decreases as she hires more of his time: the value of Biko's work is high the first hour Ayanda hires, less valuable the second hour, less valuable the third, and so on.

This is because if she has just an hour of his time, she assigns him to really important tasks, but the tasks he does in later hours are less essential to Ayanda. (This is similar to why the marginal productivity of time studying diminishes as the amount of time studying increases). Not accounting for the wage she must pay him, each hour of Biko's work gives Ayanda utility, but at a decreasing rate.

Figure 4.9 shows the setting for this interaction as an Edgeworth box, with the quantities interpreted amounts per day. The endowment point **z** is in the upper left corner of the box showing that Biko has 16 hours of Living time and no money. Ayanda has $400 but no Labor from Biko to work in her company. As before, like **z** every point in the box represents an allocation that is feasible given the amount of money that Ayanda has in her endowment bundle and the amount of free time that Biko has in his.

Three of Biko's indifference curves and three of Ayanda's are shown in Figure 4.9. For both Ayanda and Biko, their reservation indifference curve (their participation constraint) goes through the endowment point where Biko has 16 hours of Living and Ayanda has $400 per day to pay workers, meaning that the indifference curves include the endowment point **z**.

Also shown is one of Biko's indifference curves labeled u_3^B, which is tangent to Ayanda's participation constraint (u_z^A) at point $\mathbf{t^B}$. The allocation given by the

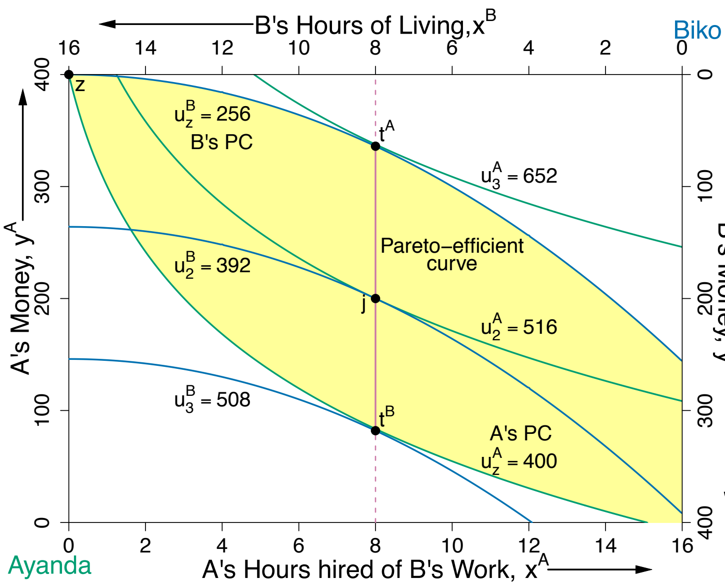


Figure 4.9: **Bargaining over hours and wages**. Shown are three each of Ayanda's and Biko's indifference curves and the utility that they experience at any of the allocations indicated by the points making up these curves. Point **z** is the endowment allocation which is a point on the participation constraints of each of the two. Points $\mathbf{t^A}$ and $\mathbf{t^B}$ respectively are the allocations resulting when Ayanda or Biko are first mover with TIOLI power. The yellow shaded area is the Pareto-improving lens. The vertical line (including its dashed portions) is the Pareto-efficient curve made up of all points of tangency between the indifference curves of the two such as **j**, $\mathbf{t^A}$ and $\mathbf{t^B}$.

that tangency is a Pareto-efficient allocation (because the marginal rates of substitution of the two are equal). We also show a third indifference curve for Ayanda, labeled u_3^A, which is tangent to Biko's participation constraint (u_z^B) at point $\mathbf{t^A}$. These two tangencies are points on the Pareto-efficient curve, which is a vertical line through these points all the potential tangencies above each person's fallback.

The reason why the Pareto-efficient curve is *vertical* here (remember it was a diagonal line in the previous Edgeworth boxes) is that Ayanda and Biko have *quasi-linear* utility functions. With quasi-linear utility, the marginal utility of hours depends only on the quantity of hours and not on the amount of money they have. If the two curves are tangent at 8 hours when Ayanda has most of the money and Biko little, they will also be tangent at 8 hours when Biko has most of the money and Ayanda has little.

4.9 Application. The rules of the game determine hours and wages

The Edgeworth box and the indifference curves by themselves do not determine the outcome of the interaction. Without knowing more, any point in the box is a possible outcome. Knowing the endowment allocation **z** narrows down the possible post exchange allocations but not by very much.

Employment in most modern economies is *voluntary* (but see the Fact Check), so we will assume that the outcomes are limited to those that are at least as good for each participant as their fallback position given by point **z**. As a result, outcomes of bargaining between the employer and the worker must be in the yellow shaded Pareto-improving lens in Figure 4.9.

FACT CHECK In the past slavery has meant the ownership of one person by another, including the right of sale of the slave to another owner. The term *modern slavery* refers to any situation in which, like historical slavery, the services or goods that one party provides for another are not voluntarily offered but are motivated by fear of severe harm. Ownership of one person by another need not be part of modern slavery. Prisoners, immigrants without legal rights of residence, residents of undemocratic countries, “sex slaves," and children are over represented among contemporary 'modern slaves.'

We illustrate the importance of institutions by showing the allocations will result under four different rules of the game. Each set of rules is a specific account four different ways that an employer and worker might interact. :

- The *employer can make a take-it-or-leave-it offer* of both the wage and the hours worked.
- As members of *a trade union*, the employees (we will take Biko as a representative worker) *can make a take it or leave it offer* specifying both the wage rate and the length of the working day (hours).
- *Legislation* is passed limiting working hours per day to no more than 5 hours and the total pay or this period to not less than 254 Rand or 50.80 Rand per hour.
- The above legislation is passed, but it has a proviso that if the *two parties can agree on an alternative allocation* their agreement can be implemented.

Employer has TIOLI power

Imagine that, like most employers, Ayanda can offer Biko a job description: work a given amount of hours for a given amount of pay (and therefore for a particular hourly wage). Biko's only choice is to accept or reject, so Ayanda has take-it-or-leave-it power. For Biko to accept, Ayanda knows the offer must be at least as good as Biko's reservation option, so the relevant constraint for her is Biko's participation constraint (as was the case for the coffee and data bargaining).

She will choose the point that she values most along this indifference curve, and therefore implement an offer indicated by point $\mathbf{t^A}$. Having TIOLI power, the employer has gotten all of the economic rent, leaving Biko indifferent between taking the job and refusing it (as before in cases like this we just assume he takes the job).

What is Ayanda's rent from this transaction, meaning the excess of her utility at point $\mathbf{t^A}$ compared to at point **z**, the endowment allocation at which no trade has occurred? Reading the utility numbers from her indifference curve at point $\mathbf{t^A}$ and her reservation indifference curve through point **z** we can see that her rent is $u_3^A = 652$ minus $u_z^A = 400$ or 252. Because utility is measured vertically in terms of money this is the same thing as the vertical distance between points $\mathbf{t^A}$ and $\mathbf{t^B}$ in the graph.

EXAMPLE Put yourself in Biko's shoes if the allocation is point $\mathbf{t^A}$. How do you think he feels about his employer and his job? Would he be motivated to work hard, not to steal from his employer, and otherwise contribute to the profitable operation of her firm? These are serious problems and a reason why extreme allocations – like Ayanda getting all of the economic rent from the interaction and Biko being indifferent between his job and being fired – are not commonly observed. If Ayanda has an interest in Biko's good will and hard work, she may have to share at least a bit of the gains from exchange with Biko so that he receives a rent. This fact will become important when we consider the labor market.

Employees and their trade union have TIOLI power

Turning to the opposite case Biko, through his trade union, is now first-mover with TIOLI power. The offer he will make (and she will accept) is the opposite

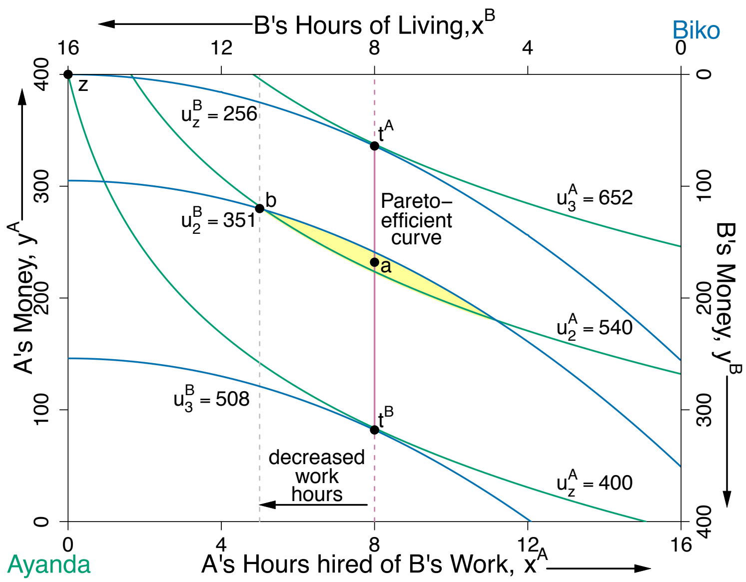


Figure 4.10: **Allocations with legislation and bargaining**. The legislation stipulating hours and pay results in the allocation indicated by point **b**. Because **b** is preferred to the no exchange option **z** by both of them, they will definitely make an exchange. But then can both do better than at **b**. Taking the allocation at **b** as their new fallback position, they could bargain to point **a** or any other allocation in the yellow Pareto improving lens.

of $\mathbf{t^A}$ the allocation resulting when Ayanda had TIOLI power. Biko will recognize Ayanda's participation constraint – he as to make her an offer she will not refuse. And he will choose the allocation indicated by point $\mathbf{t^B}$ in which his post exchange bundle gives him all of the economic rents of 252.

This is the most that Biko could demand without Ayanda simply going out of business. This constraint on the demands that workers can make on employers in a market and profit based economy will be a major theme in the chapters to come.

Legislation imposes hours and pay limitations.

As described above the legislation imposes on both Ayanda and Biko the allocation at point **b**. But the allocation imposed by the legislation is Pareto inefficient. But it does set a new status quo, a fallback position that, if they cannot come to some agreement will be the post exchange allocation.

Both Ayanda and Biko can see that at **b** they could both do better by agreeing that Biko should work more than 5 hours, and Ayanda should pay him more. The yellow Pareto-improving lens shows the space for their possible bargains.

Bargaining to override the legislation: more work and more pay.

They could bargain to agree upon any point in the yellow Pareto improving lens, possibly agreeing on the Pareto efficient allocation at point **a**. Where they ended up in or on the boundary of the Pareto improving lens would depend on the rules of the game governing that bargaining process. They might even fail to agree on any bargain – as is often the case with players in the Ultimatum Game – and remain at point b.

Figure 4.11 shows how Ayanda and Biko do under these differing rules of the game as indicated by the rents they enjoy in the Nash equilibrium of each game, that is the utility associated with their fallback options 400 and 256 respectively.

Introducing a historically realistic set of rules of the game – making the employer the first mover with TIOLI power – has two effects: it generates 252 units of utility in gains from trade, and it makes the final allocation more unequal than the endowment allocation (because the employer captures all of the mutual gains made possible by exchange). Biko's share of the total utility (not shown) falls from two fifths to less than a third.

In many countries during the 20th century the response to the unequal allocations implemented when the employer has TIOLI power was the formation of trade unions. And you can see from the figure that if the union is powerful enough for it to have TIIOLI power, then Biko (and his trade union colleagues) capture the entire rent, Ayanda getting nothing more than her reservation utility. Biko's share of the total utility jumps from two fifths at the endowment allocation to well over half.

Even before workers had the right to vote and before trade unions were legal, political movements mobilized to pressure governments to regulate working conditions. In the model the introduction of hours and wage regulations implemented an outcome in which both Biko and Ayanda captured some of the gains from trade. The reforms implemented a Pareto inefficient allocation but the shortfall from the maximum possible joint rents was minor (from 252 to 235).

The final case – bargaining up from the regulated hours and wages – describes labor markets in many countries today. Government regulations establish a fallback position, and then employers and workers (either individually or in trade unions) seek bargains that improve on that allocation.

Though they differ radically in their distributional aspect, all of the scenarios are Pareto superior to the endowment allocation. We can also see that the negotiated allocation after legislation is Pareto superior to the allocation implemented by the legislation.

We cannot say which of the three Pareto efficient allocations is preferred from a fairness standpoint without knowing more about Ayanda and Biko's other wealth, their needs, and other aspects that might affect their ethical claims on the benefits of their interaction.

Checkpoint 4.6: Bargaining over hours and wages

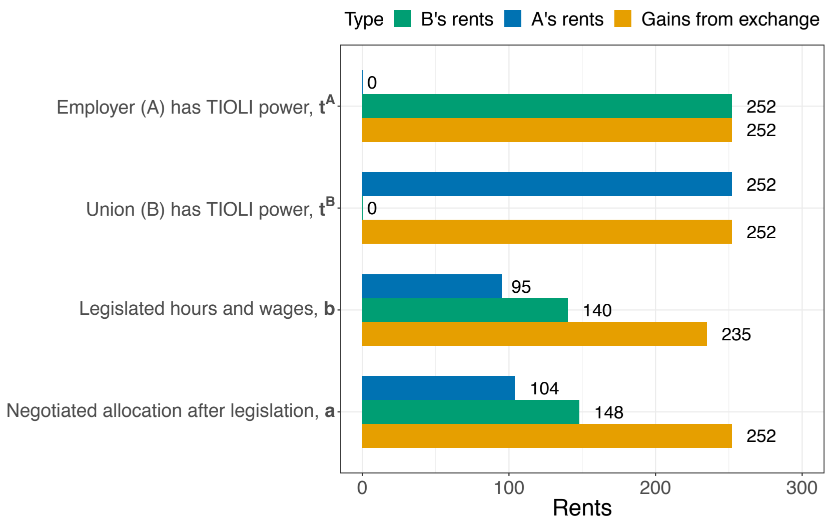


Figure 4.11: **Rents under differing rules of the game, with Ayanda as employer and Biko as worker** The rents and gains from exchange of each set of rules are shown in the figure. That is, the figure shows each player's utility under each set of rules minus that player's fallback option ($u_z^A = 400$ and $u_z^B = 256$ respectively). The gains from exchange are the sum of the rents received by Ayanda and Biko. Source: Authors calculations described in the text.

Using the figure, explain how the following two things (taken separately) would affect the outcome under the four different rules of the game above (start by explaining how the endowment point **z** would be affected):

- If Biko does *not* exchange their time with Ayanda and is unemployed, he receives what is called an unemployment benefit, that is a payment from the government equal to $\$100$, and this is financed by a tax on Ayanda equal to $\$100$.
- Ayanda now has free access to a robot that will at no cost do work equivalent to two hours of Biko's time.

4.10 First-mover advantage: Price-setting power

Returning to Ayanda and Biko with their former personas as students exchanging coffee and data, we will now see that while first movers typically have advantages, these advantages may not be due to TIOLI power. Ayanda may be first-mover but be unable to *commit* to take-it-or-leave-it offer that stipulates an exchange of a specific amount of coffee for a specific amount of data.

FIRST-MOVER ADVANTAGE A player has a first-mover advantage when the institutions, history, or power structures of a game give the player the opportunity to make an offer or move *before* the other players in the game can take action. The opportunity to move first can confer an advantage that results in *higher utility* or a greater share of *economic rents* in the outcome of an interaction.

PRICE-SETTING POWER A first-mover with price-setting power (PS power), can commit to a price – the ratio in which goods will be exchanged – but not the quantity that will be transacted at that price.

Price-setting power

She may have what is called **price-setting power** (PS power) if she can specify a price – either a monetary price or the ratio at which the two will exchange goods – but not how much (the quantity) of her good Biko will buy.

Ayanda might say, for example: “I will give you one kilogram of coffee for every three gigabytes of data you give me. You can decide how much data you would like to exchange for coffee at that ratio, but the ratio is not going to change. Of course you are free to buy nothing."

We saw that owners of companies typically have TIOLI power when hiring employees; but in their interactions with their customers they typically have price-setting power. They set a price at which they will sell their product, and sell as much to each customer at that price as the customer wants to buy.

The incentive compatibility constraint (ICC).

If Ayanda has price setting power she must find a way to determine the price when it is the price alone that makes up her offer. So her constrained optimization problem is not the same as it was when she had TIOLI power.

When Ayanda had TIOLI power she had only to satisfy Biko's participation constraint: if her take-it-or-leave-it offer were a post exchange bundle that would make Biko worse off than at his endowment bundle, Biko would leave it rather than taking it! Of course whether she has TIOLI power or just price setting power, if Ayanda wants to exchange with Biko, she will have to satisfy his participation constraint.

But there is now a second constraint she must satisfy called the **incentive compatibility constraint**. What this means is whatever post-exchange bundle Ayanda would like to implement, she must provide Biko with incentives so that his best response will be to exchange the amount that will allow her to "move" from her endowment bundle to her desired post-exchange bundle.

This is called the incentive compatibility constraint because she must provide Biko with *incentives* that motivate Biko to act in a way that is compatible with (meaning, that implements) her desired outcome. The incentive compatibility constraint is based on Biko's best response – the amount of coffee he is willing to buy – to the price Ayanda offers.

M-CHECK A continuous variable can take on any value over some interval. So, a variable that can take the value of any number between 0 and 5 is a continuous variable; a variable that is restricted to the integers between 0 and 5, namely, 1, 2, 3 or 4 is discrete. The number of your sisters or brothers is discrete, the height any one of them is continuous.

INCENTIVE COMPATIBILITY CONSTRAINT The incentive compatibility constraint, ICC, requires that first mover provide incentives that make the second mover's best response be to act in ways that implement the post exchange allocation which the first mover prefers.

You have encountered best responses in Chapter 1. There the the strategy sets were particular actions and therefore best responses were limited to actions like "Plant Late," or "Fish 12 hours." Options like "Plant a little earlier" or "Fish 10 hours and 15 minutes" were not possible. Sometimes discrete strategy sets and best responses like this make sense (think: "Drive on the the left if you are in the U.K., or Japan").

But sometimes the strategy sets for players are continuous, as for example in setting a price for a good or when choosing the amount of time for an activity, like fishing. When this is the case – as with Ayanda's decision to set a price – we consider the players' best responses as *continuous variables* and describe them by *best-response functions*.

As was the case when she had TIOLI power, Ayanda will reason backwards from her understanding of what Biko how Biko will respond to each of her possible offers and how that will affect her utility. That is, she will use backward induction.

To determine how Ayanda can maximize her utility subject to Biko's incentive compatibility constraint (the price-setting case) is a somewhat more complex problem than maximizing her utility subject only to Biko's participation constraint (the TIOLI power case). The reason is that in the TIOLI case there are just two things that Biko can do: accept or reject her offer. But when Ayanda has price-setting power only, Biko can choose from the entire range of possible amounts that he might be willing to exchange with her, depending on the price.

As a result, Ayanda has to think in two stages when choosing a price ratio.

First stage: *What will Biko do*? How much coffee will Biko buy at each price ratio Ayanda offers? This is Biko's *price-offer curve*, which is his best response.

Second stage: *What should I do, given what he will do?* Given her estimate of Biko's best response, which price ratio maximizes Ayanda's utility? That is, which price ratio takes Ayanda to her highest indifference curve, given the constraint of Biko's price-offer curve?

Best response and incentive compatibility

For the first stage, that is, determining how Biko will respond to each price she might offer, Ayanda uses whatever information she might have, such as her experience in the past with Biko's response to offers, her best guess as to Biko's utility function, or her experience with other people she thinks are similar to Biko.

REMINDER The method is identical to how we derived Keiko's price-offer curve – offering money in return for fish – in Chapter 3, except that here Biko is not 'buying' coffee using money, he is exchanging data for coffee. As a result the "price" is not in terms of dollars per kilogram of coffee, but *gigabytes of data* per kilogram of coffee.

Just as in Chapter 3 there is a budget constraint limiting the exchanges he can undertake, but this is now a line giving feasible combinations of data and coffee available to him through exchange at some given price.

If the price p – the number of gigabytes of data per kilogram of coffee – and his post-exchange bundle is denoted as (x^B, y^B) then Biko's budget constraint requires that the value of his post-exchange bundle must be the same as the value of his endowment bundle, or:

$$px^B + y^B = p\bar{x}^B + \bar{y}^B \tag{4.7}$$

$$\text{or } p(x^B - \bar{x}^B) = -(y^B - \bar{y}^B) \tag{4.8}$$

The second version of the budget constraint means that the value of the coffee that he acquires (at the price p) or $x^B - \bar{x}^B$ must be equal to the value of the data that he gives up $\bar{y}^B - y^B$.

We can rearrange Biko's budget constraint another way to show that the price p must be equal to the ratio of the amount of data he gives up to the amount of coffee he gets

$$p = \frac{\bar{y}^B - y^B}{x^B - \bar{x}^B} \tag{4.9}$$

We show the derivation of Biko's best-response function in Figure 4.12. We start, in panel **a** by showing Biko's best response to *one particular price*. Then, in panel **b** we repeat the same reasoning for *many prices*, showing how his best response to *any* price can be determined.

We know that given the price p_4 Biko will choose how much data to transfer to Ayanda in return for her coffee in order to maximize his utility subject to his budget constraint. In panel **a** we show his feasible set with his budget constraint for that particular price p_4 its feasible frontier. The budget constraint includes the point **z** because one of the feasible choices he could make while respecting the budget constraint is to exchange nothing.

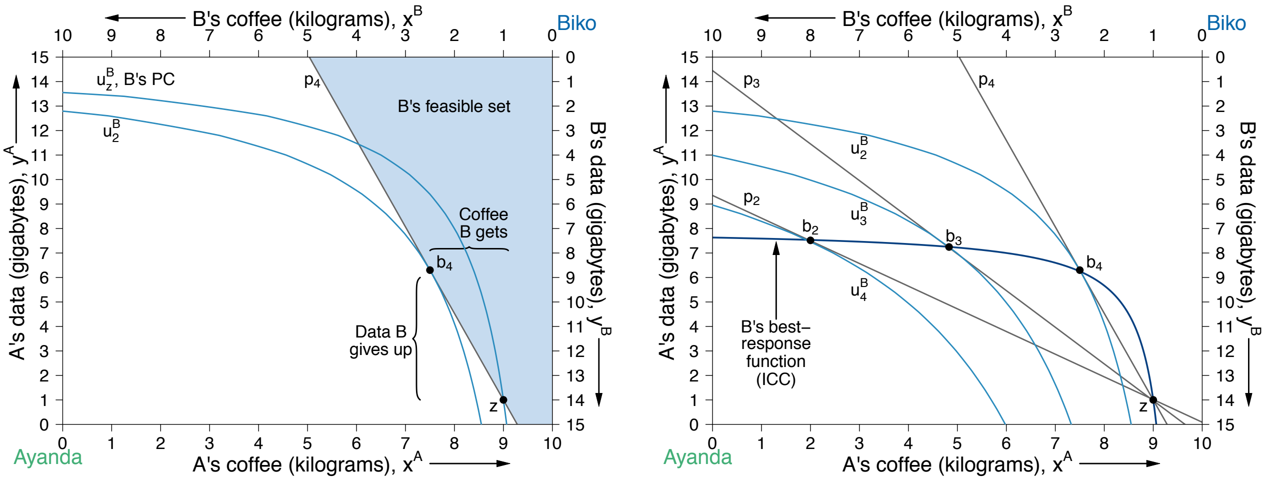
In Figure 4.12 the slope of the p_4 line is the amount of data that Biko gives up (Δy^B) divided by the amount of coffee that he gets (Δx^B) , when the price is p. So:

$$\begin{aligned} p &= -\frac{\Delta y^B}{\Delta x^B} = \frac{\bar{y}^B - y^B}{x^B - \bar{x}^B} \\ \text{marginal rate of transformation } (\textit{mrt}) &= -\text{slope of the price line} \end{aligned}$$

For any given price this is the kind of individual utility maximization problem that you studied in Chapter 3 in which the solution is to find the allocation at which the $mrs = mrt$ rule holds. You can see in panel a that the highest indifference curve that Biko can reach, consistent with his budget constraint (labeled u_2^B) is tangent to his budget constraint at point $\mathbf{b_4}$. This result expresses the principle of constrained optimization that you have already learned. It is a point equating:

- The slope of his indifference curve, which is the negative of the *marginal rate of substitution* and
- The slope of the feasibility frontier – in this case the budget constraint – which is the negative of the *marginal rate of transformation* of coffee into data.

The mrt is the price p set by Ayanda, that tells Biko how many gb of data he has to give up to get a kilo of coffee. Biko's best response is to choose a



(a) B's best response to a price = p_4

(b) B's best-response function (ICC)

Figure 4.12: **Constructing B's best-response function (ICC).** In panel a, B's feasible set is in the upper right corner of the Edgeworth box because, as we explained in Figure 4.4, the upper left corner of the box is the origin for him (indicating zero of both goods). In panel a, when the price p_4 is equal to 3.53 Biko reaches his highest feasible indifference curve (u_2^B) by giving up 5.3 gb of data in return for 1.5 kg of coffee. In panel b he chooses post-exchange bundles indicated by points $\mathbf{b_3}$ and $\mathbf{b_2}$ in response to prices $p_3 < p_4$ and $p_2 < p_3$. B's best-response function (ICC) connects these and similar points all of them B's utility-maximizing bundle, for different prices.

post-exchange bundle that satisfies the two conditions:

$$\textit{mrs} = \textit{mrt} \text{ tangency: } mrs^B(x^B, y^B) = mrt = p \tag{4.10}$$

$$\text{and, budget constraint: } px^B + y^B = p\bar{x}^B + \bar{y}^B abbud \tag{4.11}$$

Equation 4.10 expresses the optimizing part of Biko's choice, while Equation **??** expresses the constraint. The utility Biko enjoys at $\mathbf{b_4}$ in the figure is greater than the utility of his endowment bundle ($u_2^B > u_z^B$). From this we conclude that if the price is p_4 Biko will choose the post-exchange bundle given by point $\mathbf{b_4}$. This gives us one point on Biko's best-response function.

In panel b we construct Biko's best response function, by repeating the analysis in panel a but for differing prices tracing out a curve in the (x, y) coordinates. This is his *best-response function* because, by construction, points on the curve show for each the value of p the post-exchange allocation that maximizes his utility if could buy any amount of Ayanda's coffee at the price p. Ayanda now has all the information she needs to set the price.

M-Note 4.6: The incentive compatibility constraint

Here we show the derivation of the incentive compatibility constraint for Ayanda's utilitiy choice of a utility maximizing price to offer Biko. This equation will show, for every price that Ayanda could offer, the amount of goods that Biko will be willing to exchange.

To do this we use Equations **??** and 4.10, the two conditions that Bikos response must satisfy.

Given the price p offered by Ayanda, Biko's budget constraint is

$$px^B + y^B = p\bar{x}^B + \bar{y}^B \tag{4.12}$$

That is

$$y^B(x^B) = -px^B + p\bar{x}^B + \bar{y}^B$$

To maximize his utility $u^B(x^B, y^B(x^B))$, Biko will choose the bundle $(x^B, y^B(x^B))$ satisfies the first-order condition,

$$u_x^B + u_y^B \frac{dy^B}{dx^B} = u_x^B - u_y^B p = 0$$

That is

$$mrs^B(x^B, y^B) = -\frac{u_x^B}{u_y^B} = -p = mrt \tag{4.13}$$

Suppose that $u^B = (x^B)^{\frac{1}{3}}(y^B)^{\frac{2}{3}}$, we can derive the incentive compatibility constraint using Equations 4.12 and 4.13. From M-Note 4.2, we have

$$mrs^B(x^B, y^B) = \frac{1}{2}\frac{y^B}{x^B}$$

Moreover, the budget constraint can be rewritten as Equation 4.9, i.e.,

$$p = \frac{\bar{y}^B - y^B}{x^B - \bar{x}^B}$$

Therefore, we have

$$\frac{1}{2}\frac{y^B}{x^B} = -\frac{\bar{y}^B - y^B}{x^B - \bar{x}^B} \tag{4.14}$$

which defines the incentive compatibility constraint shown in the Edgeworth box.

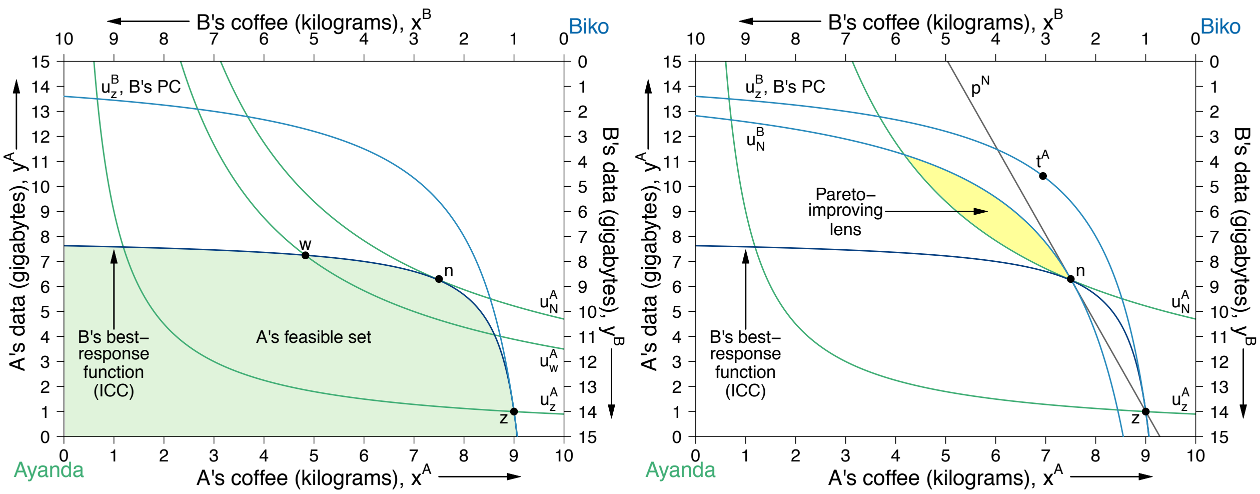
4.11 Setting the price subject to an incentive compatibility constraint

Biko's best-response function is the incentive compatibility constraint for Ayanda's optimizing problem, shown in Figure 4.13. Notice that the incentive compatibility constraint is more limiting to Ayanda than is Biko's participation constraint labeled $u_z^B, B'sPC$. This means that there are some allocations (between the participation constraint and the incentive compatibility constraint) which would make Biko better off than at his endowment bundle, and which Ayanda would prefer to any point in her feasible set, but which Ayanda could not implement when she has price-setting power but not take-it-or-leave-it power.

Because Ayanda always has the option of simply discarding some of the data she gets from Biko, we can think about the green shaded area under Biko's best response function as her *feasible set*. The slope of Biko's best-response function is (from Ayanda's viewpoint) the marginal rate of transformation of coffee into data, given how Biko responds to each of the prices she could set. You can see that starting at the endowment allocation, the best-response function is initially steep, so a modest amount of coffee that she gives up is transformed – through exchange – into a substantial amount of data. But the more data she wishes to acquire – moving up on the best-response function – the less favorable to her the mrt becomes.

Ayanda's choice of what price to set is a familiar constrained optimization problem. It proceeds in two steps:

1. Determine the *final allocation* she would like to implement by finding the point in the feasible set that is associated with the higher utility. To do this she uses the $mrs = mrt$ rule and selects point **n** in the figure, with its associated utility u_N^A (which exceeds that associated with point **w**, namely



(a) Maximizing utility subject to the incentive compatibility constraint

(b) Pareto-superior alternatives to Ayanda's choice of **n**

Figure 4.13: **A sets the price subject to B's best-response function (ICC)** Ayanda's utility-maximizing post-exchange bundle is indicated by point **n** where her indifference curve is *tangent* to Biko's best-response function (his price-offer curve). The negative of the slope of the solid gray line through both **n** and the endowment point **z** is equal to the price Ayanda chooses, p^N. Biko's budget constraint given by Ayanda's choice of p^N is tangent to Biko's indifference curve through **n** by construction, that is, because **n** is on Biko's best-response function. To interpret the lower shaded area as a feasible set, it must be the case that A could choose *not* to consume the data or coffee she has in that area (that is, some of it could be thrown away).

u_w^A, which was also feasible. This is where her indifference curve is tangent to Biko's best best response function. This is shown in Panel a of Figure 4.13

2. Determine the *price* that will implement this outcome. Every allocation on the best response function corresponds to some particular price that will implement it. Price p^N shown in Panel b of Figure 4.13 implements point **n**

We have given the price that Ayanda sets a superscript N because the allocation that it implements is a Nash equilibrium. To confirm that this is the case we ask two questions:

- Given the strategy that Ayanda has adopted – that is, setting the price p^N – is there any way that Biko do better than he does by trading with her so as to implement her chosen allocation (point **n**)? The answer is no, because **n** is a point on his best response function, which tells us that if she offers the price p^N the best he can do is to trade with her so as to implement her desired point.
- Given the strategy that Biko has adopted – his best-response function – is there any way that Ayanda could do better than she does by setting the price p^N? The answer is no, because she found point **n** exactly by doing the best she could given his best-response function.

There are two important aspects of the Nash equilibrium (**n**) of this game.

First, the Nash equilibrium is *not Pareto efficient*. Ayanda's and Biko's indifference curves are **not** tangent at **n**, they intersect, and you know from the $mrs^A = mrs^B$ rule for a Pareto efficient outcome that any allocation at which the indifference curves intersect is not Pareto efficient (because then the rule is violated). The reason why Ayanda implemented an Pareto inefficient allocation is that the the constraint she faced was not Biko's PC (the slope of which is mrs^B) but instead his best response function (the slope of which is the mrt). So she implemented $mrs^A = mrt \neq mrs^B$ violating the Pareto efficiency rule. The allocations that are Pareto-superior to **n** are shown by the yellow lens between the indifference curves through **n**.

Second, the person who is *not* the first mover (Biko) receives a rent in the Nash equilibrium: as you can see from Figure 4.13 at **n** he is better off (on a higher indifference curved) than with his endowment bundle (which is his fallback option, namely no trade) indicated by the indifference curve labeled u_z^B, B's PC.

These two results of the first mover with price-setting power only case contrast with the case of the TIOLI power. There is an important lesson here: when one of the two parties has price-setting power, but not TIOLI power, she may use that advantage to advance her distributional interests in a way that implements an inefficient outcome. For example, Ayanda could have implemented a Pareto-efficient outcome, like point **w** shown in the figure. This point is on the Pareto-efficient curve (not shown in the figure), and had she offered Biko the lower price given by the slope of a budget constraint from point **z** to point **w**, he would have purchased just the amount of coffee that would have implemented point **w**. But her utility is higher at point **n** which she can implement by charging a higher price (steeper budget constraint for Biko).

This explains why it is the case that When Ayanda has price-setting power only she uses it to get a larger piece of a smaller pie. When she had TIOLI power she knew that she would get the maximum economic rent (because the only constraint she faced was Biko's participation constraint). Subject only to the participation constraint she could dictate the entire outcome, so she had no reason to adopt any allocation that was not Pareto efficient.

We will see that this feature of the price-setting case also reappears in other economic interactions – including credit markets, labor markets, and markets for goods with limited competition.

Checkpoint 4.7: PSP vs. TIOLI

a. Using Figure 4.13, by reading the relevant points on the x and y axis, say what the post-exchange allocations for Ayanda and Biko (how much coffee for each, how much data for each). Compare this to the post-exchange allocations when Ayanda has TIOLI power, calculated in M-Note 4.4.
b. Test your understanding of the first-mover case by explaining the outcome when Biko is the first-mover. Draw a new version of Figure 4.13.

4.12 Application. Other-regarding preferences: Allocations among friends

Ayanda and Biko are about to experience one final change in their identities, along with a personality transplant: they have become friends and they care about each other. Both are altruistic: they place some positive weight on the well being of the other. This means, as you will recall from Chapter 2 that they are other regarding, when evaluating an allocation they take account not only of the utility they will experience from their bundle but also the utility the other will experience from their bundle.

They still have a decision to make: how to divide up their coffee (still 10 kilos of it) and the data (15 gigabytes of it as before). But we will assume now that neither of them own any portion of either good – so there is no endowment allocation like our interpretation of point **z** so far.

The see how the Edgeworth box helps us to understand their decision problem and because this involves some unusual indifference curves, we first treat a hypothetical case in which Ayanda is completely altruistic and Biko is as before entirely self-regarding. (We do not imagine that Ayanda would put up with this, it is just a first step along the way to seeing how two other regarding friends would look at the problem).

An altruistic utility function

Altruistic Ayanda cares not only about her bundle at an allocation, but also what Biko gets. Ayanda's utility therefore depends not only on x^A and y^A but also on x^B and y^B. We measure how much she cares about what Biko gets – her *degree of altruism* – by λ ("lambda") a number that varies from 0, if she is entirely self regarding, to one-half if she places as much weight on what Biko gets as what she herself gets, in which case she would be called a *perfect altruist*.

M-Note 4.7: An altruistic utility function

Remember if Biko did not exist so that Ayanda were making her choice of an allocation in isolation, her utility would be

$$u^A(x^A, y^A) \quad = \quad x_A^{\alpha} y_A^{(1-\alpha)} \tag{4.15}$$

But interacting with Biko and dividing goods with him, for $\lambda \quad > \quad 0$ we have Ayanda's utility function as an altruist:

$$u^A(x^A, y^A, x^B, y^B) \quad = \quad \left(x_A^{\alpha} y_A^{(1-\alpha)}\right)^{(1-\lambda)} \left(x_B^{\alpha} y_B^{(1-\alpha)}\right)^{\lambda} \tag{4.16}$$

To see why we say that λ is a measure of how much Ayanda cares about what Biko gets we can take the natural logarithm of equation 4.16

$$ln(u^A) = (1-\lambda)ln\left(x_A^{\alpha}y_A^{(1-\alpha)}\right) + \lambda ln\left(x_B^{\alpha}y_B^{(1-\alpha)}\right) \tag{4.17}$$

Equation 4.17 says that the natural logarithm of A's utility is $(1 - \lambda)$ times the natural logarithm of her valuation (if made in isolation) of her own bundle plus λ times the natural logarithm of B's evaluation (if made in isolation) of his bundle.

Biko's utility function a has the same structure as Ayanda's but the interpretation of λ is the opposite. In Biko's utility function λ is the exponent of Ayanda's bundle , and $(1 - \lambda)$ is the exponent on his own bundle, the opposite of where these terms appear in Ayanda's utility function. The totally self-regarding person, Biko in this case, places no weight on the bundle of the other person; his degree of altruism, $\lambda = 0$. So self-regarding B's utility function is:

$$\begin{aligned} u^B(x^A, y^A, x^B, y^B) &= \left(x_A^{\alpha}y_A^{(1-\alpha)}\right)^0 \left(x_B^{\alpha}y_B^{(1-\alpha)}\right)^1 \\ &= x_B^{\alpha}y_B^{(1-\alpha)} \end{aligned} \tag{4.18}$$

which is just his previous utility function before we introduced λ. The rearrangement of the equation in the second line is true because any term raised to a zero exponent (as in Biko's utility function) has a value of 1.

Checkpoint 4.8: Spite and love

.

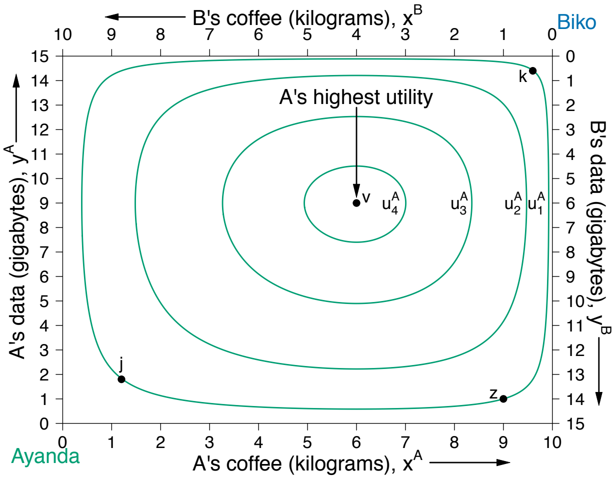
a. What would it mean in the utility function 4.16 if we had λ < 0? Can you give an example of someone acting as if they had preferences like this?
b. Can you imagine a person having a value of λ greater than one half, what would this mean? Can you think of situations in which people have acted on preferences of this type?

An altruistic indifference map

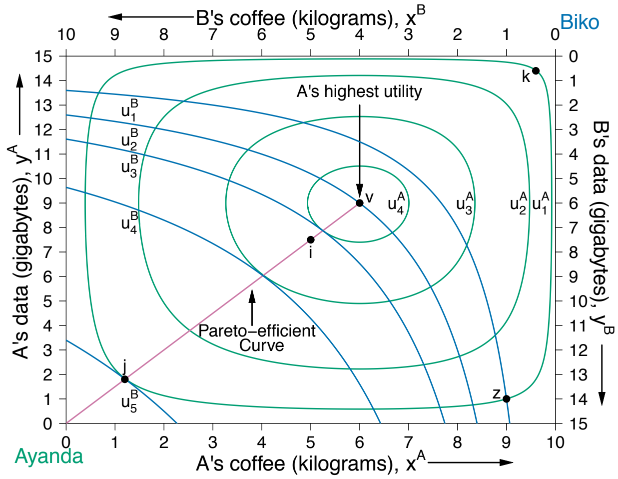
To draw her indifference map, we will give Ayanda some particular value of λ. Figure 4.14 shows an Edgeworth box representing a-not-perfectly-altruistic Ayanda with $\lambda = 0.4$.

Ayanda's indifference curves look like the contours on a topographic map of a mountain. We described the constrained optimization process in Chapter 3 as a kind of hill climbing. where both elements in the bundle were a "good" and over the entire map, the mountain rose to higher levels if you moved in the "north east" direction, that is more of both goods. In those figures you never saw the top of the mountain, because there was not any top. There was no such thing as "too much" of either good.

But Ayanda's indifference map has a definite peak at the allocation indicated by point **v**. The reason is that from her other-regarding perspective she can have "too much" of a good when that means that Biko (who she cares about) too little. This is why Ayanda's indifference curves oval shaped, just like the description of a mountain and its peak on a contour map.



(a) Altruistic Ayanda's indifference curves



(b) Altruistic Ayanda and self-regarding Biko

Figure 4.14: **Allocation and distribution with one altruistic person and one self-regarding person.** In panel **a** the green oval shaped curves labeled u^A are the indifference curves based on Ayanda's utility function. In both panels, points **z** and **i** are the same allocations here as in Figure 4.6. Notice that in panel **a** because Ayanda values what Biko gets she regards the **j** as equivalent to the endowment **k**, despite the fact that she receives less of both goods at **j** than she does at **k**. For the same reason, Ayanda's utility reaches a maximum at the allocation **v** indicated in the figure. The Pareto-efficient curve now does not include **k**, because Biko is so deprived of both goods at that point that Ayanda prefers **v** to **k**.

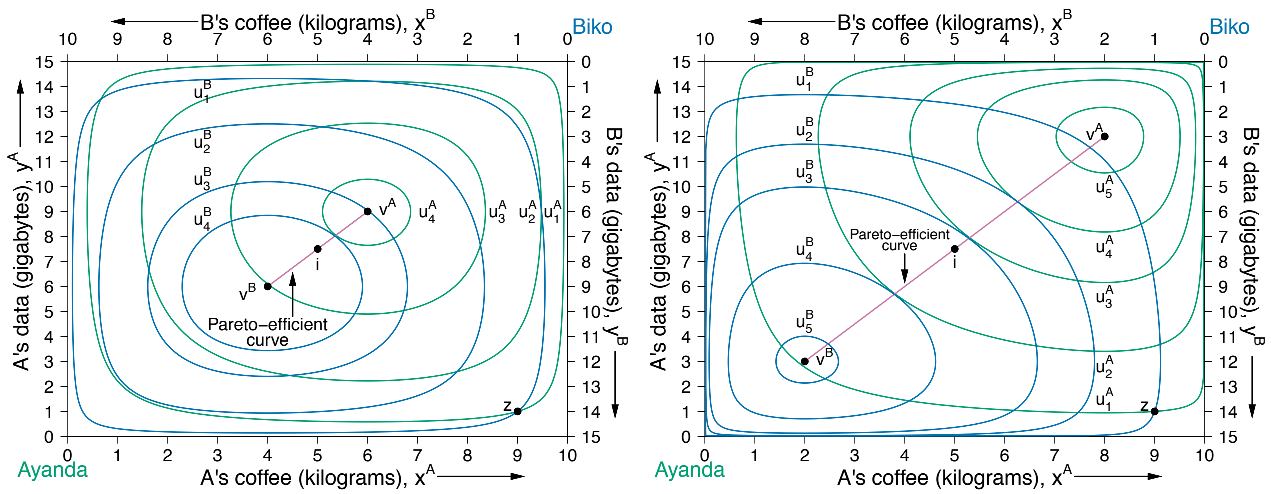
Notice that when she has little of either good (close to her origin in the lower left of the box) her indifference curves look as you have seen before. In this situation both coffee and gigabytes are "goods" so more of each is better, and the indifference curves slope downward, as you would expect. Moving up or to the right brings you to a higher indifference curve. In this part of the figure "more is up."

But beyond a certain point "more" for Ayanda is no longer "up". If she has most of both goods then getting even more is not something she values, so moving up and to the right leads her to lower not higher indifference curves.

To understand the upward-sloping parts of Ayanda's indifference curves, remember that if one of the axes represents a good and the other a bad, then the indifference curve slopes *upwards*, as in the case of study time (a bad) and expected grades (a good). In the upper right of the box for example near point **k** where she has *most* of both goods and Biko has *little* of either the indifference curves slope downward because for Ayanda having more of either good (and Biko having less) *reduces* her utility: both her coffee and her gigabytes are "bads" not goods.

In panel **b** of Figure **??** we add Biko's conventional (self-regarding) indifference curve, so we now know how both of them evaluate every feasible allocation given by the dimensions of the box. To do this we use Biko's self-regarding utility function with the value he places on Ayanda's utility being zero that is $\lambda = 0$ because he is entirely self-regarding (that is, zero altruism).

The Pareto-efficient curve is, as before, made of points of tangency between Ayanda's and Biko's indifference curves. But now we *exclude* tangencies at allocations for which Ayanda places a *negative* value on having more of one or



(a) More altruism ($\lambda = 0.4$)

(b) Less altruism ($\lambda = 0.2$)

Figure 4.15: **Altruistic indifference maps.** The two panels depict two different levels of altruism: high ($\lambda = .4$) in panel **a** and low ($\lambda = .1$) in panel **b**. The allocations indicated by the points v^A and v^B are respectively A's and B's preferred allocation. The Pareto-efficient curve is composed of all allocations at which both own coffee and own data are "goods" rather than "bads" to both A and B, and where their marginal rates of substitution are equal, that is, their indifference curves are tangent.

both of the goods, above and to the right of her "utility peak" at **v**. As a result the Pareto-efficient curve in Figure 4.14 looks different from the one in Figure 4.6 as it does not extend upwards and to the right beyond Ayanda's maximum **v**. Ayanda does not want more of either good than she gets at her maximum **v**, while Biko prefers **j** to any allocation in which she gets less of either or both of the goods.

Checkpoint 4.9: Altruistic comparisons

Consider Figure 4.14

a. Where is Biko's utility peak in the figure (analogous to Ayanda's allocation at point **v**?
b. where would point **v** be if $\lambda = \frac{1}{2}$ (or as close to $\lambda = \frac{1}{2}$ as possible?
c. What happens if Ayanda is self-regarding and Biko is an altruist? How would the Edgeworth box change?

Efficiency and fairness among altruists

With these analytical tools we can now look at the decision problem faced by the friends Ayanda and Biko both with other-regarding social preferences. Figure 4.15 hows for the same Edgeworth box, the indifference maps of the two.

Unlike the case of one altruistic actor, now both participants have preferred allocations in the interior of the Edgeworth box. They both would like to avoid "too much of a good thing."

Both of them dislike extreme allocations giving most to one or the other. This would not be the case were they evaluating bundles in isolation, that is if the other person did not exist. The reason why they place a negative value on getting more when they already have a lot is not due to diminishing marginal utility, it is because getting more for yourself means getting less for the other.

Each of their preferred allocations are shown in the figures by the allocations, v^A for Ayanda and v^B for Biko. Around each person's preferred allocation, their iso-social welfare curves move outwards and downwards in all directions, corresponding to lower and lower levels of utility.

As you can see from panel **a** of Figure 4.15 the Pareo efficiency curve is a line between their two preferred "utility peaks" v^A and v^B. By comparing panels **a** and **b** depicting greater and lesser degrees of altruism, you can see that the more altruistic they are, the shorter the Pareto-efficient curve is, because greater altruism eliminates more of the extremely unequal allocations.

There is still a conflict of interest, however. At Ayanda's preferred allocation Biko has a level of utility less that than the utility he enjoys at this own preferred allocation. The same is true of Ayanda: she does much better at her preferred allocation than at Biko's

Along the Pareto efficient curve movements in one direction or the other necessarily involve one gaining and the other losing. As always the Pareto efficient curve is a conflict region even among altruists. The fact that the 'utility peaks' are closer together in panel **a** illustrating a greater degree of altruism means that the conflict of interest between them is lesser the more altruistic they are.

How might they resolve their remaining conflicts of interest? Here, to make a decision, they need to go beyond their own utilities (even taking account of their altruistic nature) to bring in some additional way of making a judgement. They might adopt:

- a *social norm* that they both share, for example if one of the two found the coffee and the data they could go by "finders keepers"; in this case whichever of them who found the goods could make the decision, presumably implementing his or her preferred allocation.
- a *procedural rule of justice*, for example flipping a coin to see whose preferred allocation v^A or v^B would be implemented; or
- *a substantive rule of justice*, for implementing the allocation recommended by the Impartial Spectator.

Point **i** in the figures is a reference point showing the allocation that the Impartial Spectator (who weights Ayanda's and Biko's utilities equally) would implement. This is the same allocation that they would have *both* preferred had they been perfect altruists.

Checkpoint 4.10: Altruism and Rents

Why does altruism reduce the conflict over which allocation to implement?

4.13 The rules of the game and the problem of limited information

Case	*Constraints implied by the rules of the game*	*Objectives of the actor(s)*	*Characteristics of the resulting allocation*
Impartial spectator	The available goods (dimensions of the Edgeworth box)	Social welfare equally weighting the utility of each	Pareto efficient and fair (by the standard of the social welfare function).
Symmetrical bargaining with no first-mover advantage	Endowment allocation (private property). Each player's participation constraint (PC) at each stage of the bargaining	Utility of the two traders	Pareto-efficient if no impediments to bargaining, otherwise possible Pareto improvements over the endowment allocation
Take-it-or-leave-it power	Endowment allocation (private property) Second mover's PC	Utility of the first and second movers	Pareto efficient, first mover's rent is all of the gains from exchange
Price-setting Power	Endowment allocation (private property) Second mover's incentive compatibility constraint (ICC)	Utility of the first and second movers	Pareto inefficient; first mover gets most of the gains from exchange but 2nd mover gets some
Legislation	The available goods (money and time)	Whatever the legislators were seeking to (possibly the social welfare optimum)	Pareto inefficient, could be improved upon by private bargaining
Bargaining away from legislated hours and wages	The new participation constraints given the fallback position implemented by the legislation	Utilities of the two players	Pareto-efficient if no impediments to bargaining, otherwise possible Pareto improvements over the new fallback
Altruism	The available goods (dimensions of the Edgeworth box)	Utilities of both (taking account of how much they value the other's bundle); fairness	Pareto efficient and (if they can agree on a fairness principle) fair.

Table 4.1: **The rules of the game: Objectives, constraints and the characteristics of the resulting allocations.**

We have examined several institutional approaches to resolving the conflict between Ayanda and Biko over allocations of available goods. They all illustrate the dilemma posed in social interactions between:

- The goal of reaching an allocation that is Pareto-superior to the endowment and possibly even Pareto-efficient.
- The goal of resolving the conflict over the distribution of the resulting eco-

nomic rents in a way that is fair.

Table 4.1 summarizes some of the key aspects of the cases we have discussed. Which of the scenarios in the table are relevant in any particular case depends on the rules of the game for the society of which the players are a part. How well the rules work depends in important part on whether the actors have the information that we have attributed to them.

- *The Impartial Spectator* needs to know a lot about Biko and Ayanda to implement his preferred outcome, in particular their preferences.
- *The symmetrical traders* need little information other than their own preferences; they simply continue accepting exchanges as long as the post exchange bundle is preferable to the pre-exchange bundle.
- *The person with TIOLI power* needs to know the second-mover's participation constraint (a single indifference curve), which is less information than the Impartial Spectator requires.
- *The person with price-setting power* needs to to estimate the second-mover's best-response incentive compatibility constraint, which requires more information than the participation constraint, but less information than the Impartial Spectator. If the legislator (who imposed the hours and wages law) was intending to implement an efficient and fair outcome such as the one recommended by the Impartial Spectator, then he (or they) would have to know as much as was required of the Impartial Spectator, namely the entire preference maps of the two.
- *The two altruists* need to know both their own and the others preferences (without knowing what the other cares about it is impossible to care about the other). This is as demanding as the information required of the Impartial Spectator.

A basic fact of economic life is that information is scarce and local. For example in their altruistic friends scenario Ayanda and Biko probably know a lot about each others preferences, but this is unlikely when Ayanda is the employer and Biko her prospective worker. This will have important ramifications in the chapters to come especially when we study the labor market, the credit market, and other exchanges where limited information makes it impossible to implement Pareto eficient allocations.

4.14 Conclusion

From the silent trade that Ibn Battuta and Herodotus described centuries ago to eBay, Amazon and Alibaba today, people have exchanged goods to their mutual advantage and engaged in conflicts over who would get the lion's share of the the gains from exchange. The four scenarios we have

HISTORY For Friedrich Hayek, an important 20th century economist and philosopher, the fact that information is scarce and local was the basis of his criticism of centrally planned economies – such as the Soviet Union at the time – and his advocacy of private property and markets. See his thoughts on this in the headquote for Part IV of this book and in Chapter 14.

introduced have made it clear that the outcomes of these exchanges and conflicts depend on the institutions under which they take place, and the preferences of the people involved. We will see in later chapters that it is quite common that one of players in an economic interaction has price-setting power or its equivalents: the power to set wages, interest rates, and other terms of an exchange.

We have also seen some of the scenarios require that an actor knows a lot about the other person which is not very realistic even in the two person case we have used as a simplification of societal interactions. The fact that information is both scarce and local will play an important role in our analysis of capitalism as an economic system in later chapters.

The abstract scenarios we have introduced here do not capture the often step-by-step dynamics by which people move from their endowments to eventually reach an allocation through a series mutually beneficial trades. We will see that exactly how an economy moves from an out-of-equilibrium endowment to an equilibrium final allocation makes a difference for the fairness of the outcomes, but to economists it remains a vexing and far from settled problem.

Making connections

Constrained optimization in strategic interactions: The constrained optimization techniques developed in Chapter 3) are used to better understand strategic interactions introduced in Chapters 2 and 1.

Optimization rules. In addition to the $mrs = mrt$ rule which we developed in Chapter 3 for individual optimization we also have the $mrs^A = mrs^B$ rule defining a Pareto efficient outcome, both of which are used in strategic interactions.

Mutual gains from trade: If the endowment allocation (status quo) is not Pareto-efficient, then mutual gains are possible by implementing some different allocation of the goods which people may be able to agree to voluntarily.

Rents and conflicts: These improvements over the fallback option accruing to the players are *rents*, made possible by the gains from exchange.

Institutions (rules of the game) and bargaining power: The distribution of these rents in the Nash equilibrium allocation depend on the players preferences and the initial endowment as well as on the property rights in force, other institutions and the forms of bargaining power that each participant can exercise.

Pareto efficiency, institutions: If players have sufficient information some institutions will result in Pareto-efficient outcomes. Examples are the allocation implemented by the imaginary Impartial Spectator, and the situation in which one person has take-it-or-leave-it power. Price-setting power by one person, however, results in a Pareto-inefficient outcome even with unlimited information.

Self-regarding and social preferences: Among the set of Pareto-efficient allocations there will generally be conflict of interest among the participants. But, the extent of these conflicts may be reduced by social preferences such as altruism.

Important Ideas

utility function	marginal rate of substitution	Cobb-Douglas Utility
Edgeworth box	pareto-criterion	pareto-improving lens
pareto-efficiency	pareto-efficient curve	utility possibilities frontier
endowment	post-exchange allocation	impartial Spectator
social welfare function	$mrs^A = mrs^B$ rule	iso-welfare curve
altruism	private property	first-mover advantage
take-it-or-leave-it power	price-setting power	participation constraint
incentive compatibility constraint	price-offer curve	institutions
gains from trade	economic rent	

Mathematical Notation

Notation	**Definition**
α	exponent of good x in the Cobb-Douglas utility function
$u()$	utility function
$\bar{x}, \bar{y}$	total amounts of x and y available for trade
p	price of coffee in terms of data
W	Social Welfare function
λ	extent of altruism
$h()$	non-linear term of quasi-linear utility function
a	parameter in the linear term of quasi-linear utility function

Note on super- and subscripts: A, B and i: people; z: endowment point; $\mathbf{t^i}$: outcome with a take-it-or-leave-it offer.

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

5

Coordination Failures & Institutional Responses

> Right now, my only incentive is to go out and kill as many fish as I can...any fish I leave is just going to be picked by the next guy.
>
> John Sorlien. Rhode Island (USA) lobsterman[2]

Don't get him wrong: John Sorlien, the lobsterman, is not the kind of self-interested and amoral *Homo economicus* you might find in an economics textbook. He is actually an environmentalist of sorts, and as President of the Rhode Island Lobstermen's Association he was up against a serious problem of incentives, not a shortcoming of human nature. When he started lobstering at the age of 22, he set his traps right outside the harbor at Point Judith, within a few miles of the beach, and made a good living. But the inshore fisheries have long since been depleted, and now his traps lie 70 miles offshore. He and his fellow lobstermen are struggling to make ends meet.

Across the world in Port Lincoln on Australia's south coast, Daryl Spencer, who dropped out of school when he was 15 and eventually drifted into lobstering, has done much better. During the 1960's the Australian government assigned licenses – one per trap – to lobstermen working at the time, and from that time on, any newcomer seeking to make a living trapping lobsters off of Port Lincoln had to purchase licenses.

Spencer purchased his start up licenses for a modest sum and by 2000 his licenses were worth more than a million U.S. dollars (in 2000 prices); considerably more valuable than his boat. More than giving Spencer a valuable asset, the policy has limited the Australian lobstermen' s work: Spencer has 60 traps, the maximum allowed; in the Atlantic off of Point Judith John Sorlien pulls 800 traps and makes a lot less money.

Regulating the amount of lobsters trapped is a coordination problem. Point Judith and Port Lincoln represent extremes along a continuum of failure and success; with the lobstermen of Port Lincoln reaping the mutual gains made

DOING ECONOMICS

This chapter will enable you to do the following:

- Understand how the external effects of our actions on others that are not taken into account when people make choices lead to coordination failures.
- Represent social interactions with graphical and algebraic indifference curves, feasible sets and best responses functions.
- Use best response functions to see how the fairness and Pareto efficiency of the resulting allocations will depend on the rules of the game.
- Understand how improvements in property rights, government policies such as taxes or direct regulation, the exercise of power by private individuals and social preferences can all result in Nash equilibria that are Pareto-superior to the Nash equilibrium that would result in their absence.
- See that the Pareto-improvement made possible in each of these cases occurs because (in very different ways) they induce actors to internalize the external effects that their actions have on others.

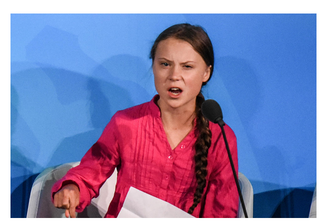


Figure 5.1: **Sounding the alarm on climate change, a coordination problem.** Greta Thunberg, then 16 years old, speaking at the United Nations in 2019 about what is probably the most serious coordination problem that humanity has ever faced. She said: "We are in the beginning of a mass extinction, and all you can talk about is money and fairy tales of eternal economic growth. How dare you!"[1]

possible by a joint decision to limit the number of traps. One may wonder why the Point Judith fishermen do not simply emulate the Australians. This is especially surprising since one of Sorlien's friends and a fellow Point Judith lobsterman visited Port Lincoln, returning with tales of millionaire fishermen living in mansions. But getting the rules right is a lot more difficult than the Port Lincoln story may suggest, and good rules often do not travel well.

One of the common obstacles to successful coordination is that the rules that address the coordination problem also implement a division of the gains to cooperation. In Port Lincoln, those who were awarded the licences benefited; others did not. Had the young Daryl Spencer not agreed one day to help out a lobsterman friend and then decided to become a lobsterman himself, someone else would be a millionaire, and Spencer might still be painting houses and complaining about the high price of lobsters.

Even if policies to address coordination failures could result in benefits for everyone affected, how a group coordinates, and what policies they coordinate on will affect how these benefits will be distributed. And this makes it difficult to agree on a policy. An example is the Ultimatum Game experiment, in which conflicts over the size of the Proposers and Respondent's "slice of the pie" sometimes result in neither getting any piece of the pie at all.

REMINDER A coordination problem is a situation in which people could all be better off (or at least some be better of and none be worse off) if they jointly decide how to act – that is, if they coordinate their actions – than if they act independently.

Conflicts over the distribution of the gains to cooperation have sunk many otherwise viable agreements to limit the depletion of fishing stocks. To give an example, a confederation of tribes of North-west U.S. Native American salmon fishermen seeking to limit their catch decided to allocate shares of a given maximum catch to each tribe.

In the course of months of debate and bargaining the following principles of division were proposed, with each proposal more or less transparently benefiting one or another tribe or class of people:[3]

PERMIT A permit allows a firm or person to engage in an activity: it gives them permission. A permit gives the holder a property right to a certain amount of a good or output. For example, a fishing permit would allow a certain number of fish to be caught or a carbon emission permit would allow a certain amount of carbon dioxide to be emitted during production. When permits are *transferable*, firms and people can buy and sell permits at a price.

- One tribe one share.
- Shares allocated in proportion to a tribe's number of members.
- Shares to each tribe based on the tribe's expenditure on lobbying (seeking to influence) the U.S. Federal Government to adopt policies more favorable to the tribes, and finally.
- Shares to each tribe in proportion to the relative quantities of fish taken at the time of the initial treaty with the U.S. Government.

Neither unrestricted competition nor marketable **permits** to catch specified amounts (similar to Australia's lobstering licenses) was proposed. The variety of proposals and their different effects on the distribution of income among the tribes suggest how challenging it may be to agree on a rule for sharing the gains to cooperation.

Depleting a fishing stock is little different in the structure of its incentives and its consequences from many other social interactions. In Chapter 9, for example, using exactly the model we develop here of the coordination problem that fishers face "harvesting fish," we will study how firms compete on markets "harvesting customers" by attempting to charge lower prices than their competitors.

What do firms competing on markets have in common with fishing people depleting the basis of their livelihood? The common idea is *over*-harvesting – whether it is fish or customers – that could be prevented if the firms or fishermen coordinated their actions rather than acting singly. Just as the Port Lincoln lobstermen discovered that they could benefit by making a common decision to limit the number of traps they set, so too will firms discover that they could make higher profits if they were able to agree on a price at which to sell, rather than competing.

In Chapter 9 we will return to the fact that coordination among the firms to set a common price – which is illegal in many countries – raises profits but harms buyers and as a result increases inequality.

The fact that coordination problems take so many familiar forms explains both the continuing interest in Hardin's "Tragedy of the Commons" introduced in Chapter 1 as well as the impressive amount of human ingenuity that has been invested in finding ways to avoid or mitigate the costly consequences of uncoordinated individual optimization in these situations.

In this chapter we develop tools to understand the nature of coordination problems like the Tragedy of the Commons. We use these tools to analyse some of the policies (changes in the rules of the game) that improve the Nash equilibrium outcome when external effects are present.

5.1 Common property resources, public goods, and club goods

Coordination problems are common because when we interact with others we affect their well being – positively or negatively – and these external effects are not taken into account when we decide on a course of action. The nature of these external effects differs depending on the type of interaction in question. In the case of over fishing or 'over-harvesting' consumers, when one person fishes more, or a firm cuts prices, the external effects – on the catch of the other fishermen or the profits of other firms – are negative.

A taxonomy of goods

To better understand the kinds of coordination problem that we face and how we might design effective remedies, we classify goods according to their the kinds of external effects associated with them and the reason why these are a problem. To do this we ask two questions, introducing two new terms:

- Is the good *rival* or *non-rival*?
- Is the good *excludable* or *non-excludable*?

When a good is *rival*, the benefits of its use are limited: more people using the good reduces the benefit available to others. Your phone is a rival good (our using it precludes others using it at the same time) while information typically is non-rival (the fact that I know what time it is and share this information with you does not preclude your benefiting from the same information).

The distinction between rival and non rival goods can be dramatized by considering how different the reaction would be if you met someone in the street and politely asked:

- "Excuse me could you give me the time of day?" or
- "Excuse me, could you give me your phone?

When a good is *excludable* a potential user may be denied access to the good (or excluded from its usage) at low cost. Your home is an excludable good. The music from an outdoor concert in a park is not excludable.

We make use of these distinctions to provide the taxonomy shown in Table 5.1. The four categories shown there are "pure cases" introduced to clarify distinctions. In reality many goods or resources have some aspects of a public good (they may be a little bit rival and a little bit excludable). The same is true of the other three categories.

Non-excludability and external effects

But if we just think about the pure cases for now, we have the following:

Common property resources are *rival* and *non-excludable*, like in the Fishermen's Dilemma in Chapter 1. As was the case for the lobstermen above, the more one fished, the less others caught; but in the absence of a permit system like they adopted in Australia, no fishermen could be stopped from fishing, so the common property or pool (the lake or the ocean) was non-excludable.

Examples of *common property* resources and their associated coordination problems include congestion in transportation and communications networks, overuse of open access forests, fisheries, water resources. Status is another common property resource, not everyone can be high status (there is a limited amount to go around) so it is rival. But nobody can be *excluded* from acquiring status symbols and engaging in other social climbing activities.

	Excludable	*Non-excludable*
Rival	Private good (clothing, food)	Common Property (Pool) Resource (fishing stocks, potential buyers),
Non-rival	Club Good (streaming music, online movies)	Public Good (global climate, rules of calculus)

Table 5.1: **Public, private, common property and club goods.** In parentheses are examples of the kinds of goods.

A public good is both non-rival and non-excludable. A private good is neither: it is both rival and excludable.

A slice of pizza is a private good: it is *rival* because if you eat it nobody else can enjoy it. It is *excludable* because the pizza seller can exclude you from eating it if you do not pay for it. By contrast, weather forecasts (on your phone, website, or the radio) are a public good. As more people use the weather forecasts the benefits that those already using the forecasts receive do *not* decrease, the benefits of the weather forecasts are *non-rival*. No person can be excluded from access to the information about the weather, therefore the benefits are *non-excludable*.

When a person contributes to a public good – for example by producing some new information of value to everyone – she is contributing benefits to others, so she confers external benefits on others. The problem here is that the person does not benefit from the positive external effects that her actions convey on others. So unless the actor values the well being of others as much as hers own (very unlikely) the public good will be under-provided. Common pool resources, as the lobsterman John Sorlien explained, will be *over-exploited*.

In contrast with public goods and common property resources, there are "club goods." Club goods are non-rival, but people can be excluded from their consumption. Common examples include collecting a toll on a little used highway, charging admission to an uncrowded museum, or making people pay for streaming video and music.

Intellectual property rights such as patents and copyrights are club goods: allow people to be excluded from the use of information, which in the absence of the intellectual property rights would be a public good. This makes it clear that how some good or resource is classified in our two-by-two taxonomy depends not only the nature of the good itself, but also on the rules of the game that determine whether it is excludable or not.

In this chapter we illustrate how coordination failures occur and how policies might address them with the example of common property resource problems (or common pool resource problems). The "common property" or "common pool" is the stock of fish available for catching or the pool of customers who might purchase the goods sold by the firms. Because common property resources are non-excludable and rival, people who use them impose *external*

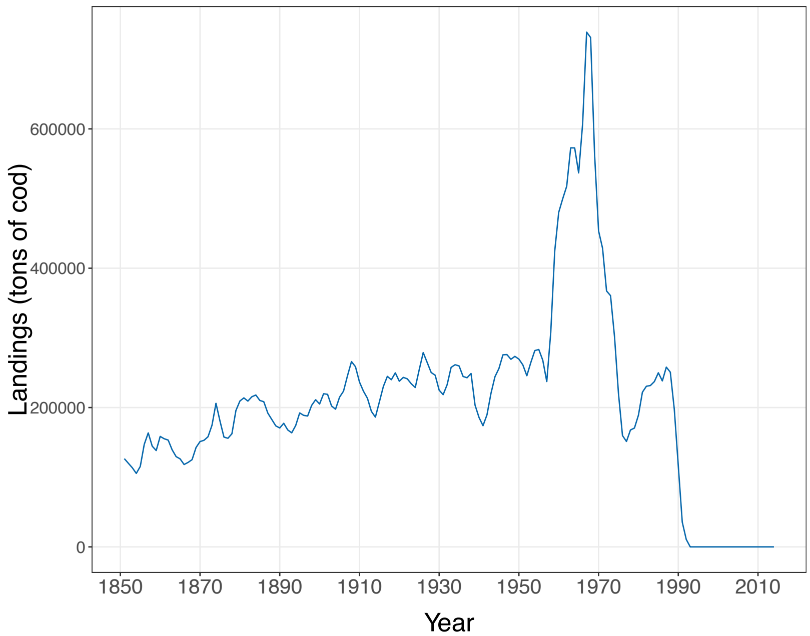


Figure 5.2: **he Grand Banks (North Atlantic) fisheries: cod landings in tons (1851-2014)**. In the 1960s new fishing technologies allowed a dramatic increase in cod fish caught ("landings") far outpacing the capacity of the fish to reproduce. This led to a partial collapse of the fishery in the 1970s and a total collapse in 1992 when the Canadian government banned fishing entirely. The fishing stocks to sustainable levels of the past by the 2030s. Source: Ecosystems and Human Well-being: Synthesis. A Report of the Millennium Ecosystem Assessment (2005)[4]

costs on each other. The "problem" is that self-regarding people will *over-exploit* the resource because they will not place any value on the negative external effects of their actions on others. Just such a pattern of exploitation is shown in Figure 5.2, which displays the catches of cod fish in the North Atlantic fisheries.

Checkpoint 5.1: A taxonomy of types of goods

Look again at Table 5.1 think of at least two further examples for each of the four categories of goods.

5.2 *A common property resources problem: Preferences*

Let's consider a specific example of a common property resource problem: the over-exploitation of an environmental resource. It could be the oceans, or forests, or a livable planet, but we'll stick to the problem of over-harvesting fish.

Our questions look at the ways that the rules of the game and the preferences of the actors determine what we should expect to happen in these situations.

Preferences over fishing time and fish consumed

We turn now to the the problem confronted by two fishermen, called Abdul (A) and Bridget (B). We model just two fishermen as a way of representing

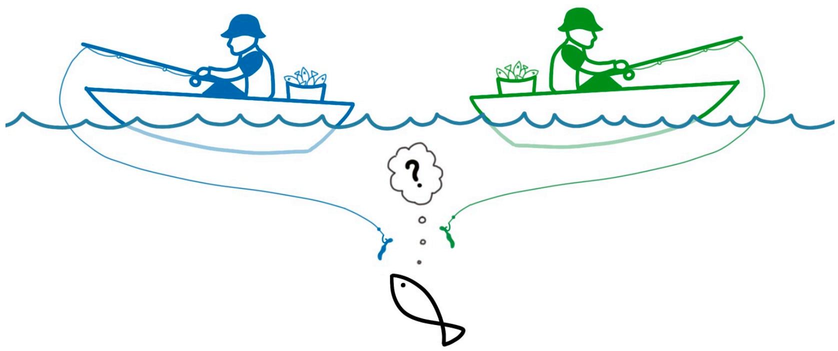


Figure 5.3: Abdul and Bridget trying to catch the same fish. The lake is a common pool resource, so the benefits are rival and each imposes a negative external effect on the other.

how a large number of them might interact. They fish in the same lake, using their labor and their nets. To start, we assume they consume the fish they catch (what we call their "catch") and do not engage in any kind of exchange. We will begin by assuming that they do not make any agreements about how to pursue their economic activities. (Recall that this means that they are engaged in a *non-cooperative game*.)

DISUTILITY OF WORK Working doesn't only take up time, it is also costly to people because of the effort that they need to exert. Manual labor is physically tiring and often, with activities like construction and mining, can be dangerous as well as complex and challenging mentally. Working as a waitress burns as many calories in an hour as doing construction work. Office work, too, requires effort, requiring concentration and attention. Exerting this effort often isn't pleasant and therefore results in disutility or a cost of utility to exert.

Each derives well-being from eating fish and experiences a loss of well-being (disutility) with additional fishing time. We represent their preferences when they are engaged in some amount of fishing with the following quasi-linear utility functions:

$$\begin{array}{lrcll} & \text{Fisherman's utility} & = & \text{Fish consumption} - \text{Disutility of fishing} & \\ \text{Abdul's utility} & u^A(h^A, y^A) & = & y^A - \frac{1}{2}(h^A)^2 & (5.1) \\ \text{Bridget's utility} & u^B(h^B, y^B) & = & y^B - \frac{1}{2}(h^B)^2 & (5.2) \end{array}$$

The utility function given by Equation 5.1 tells us three things about Abdul's preferences:

- *Consumption* (y^A) measured in pounds of fish is a "good;" Abdul derives utility from obtaining more consumption (consuming more fish) which is why y^A has a positive sign.
- *Time spent fishing* (h^A) measured in hours is a "bad": the second term has a negative sign.
- *Utility* (u^A) is increased by one unit if he is able to consume one more pound of fish, so the units in which we can measure utility are pounds of fish.
- *Marginal utility* is not diminishing but instead is a constant (equal to 1,

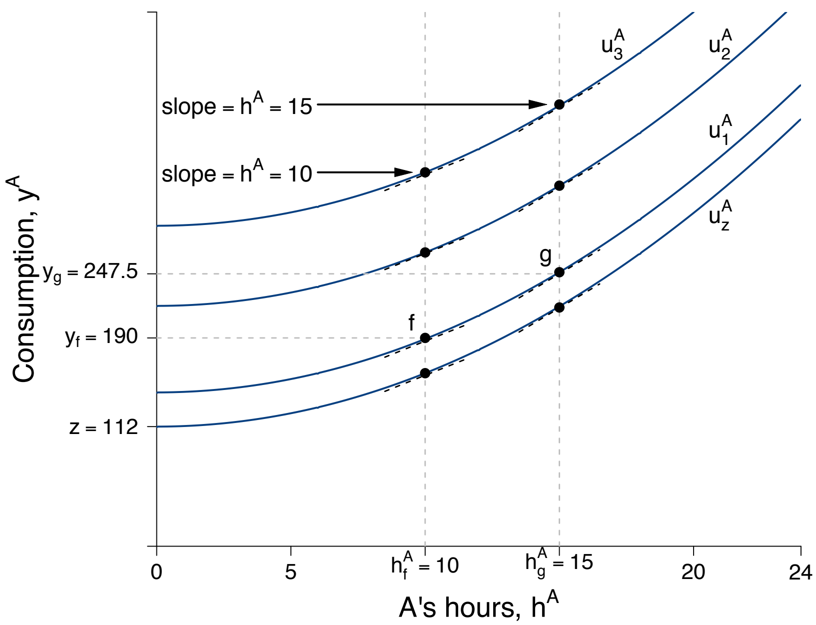


Figure 5.4: **Abdul's indifference curves over output (y^A) and fishing time measured in hours (h^A).** Output (fish) (y^A) is a "good" and provides Abdul with positive utility, whereas fishing time (h^A) is a "bad". Notice that Abdul's indifference curves in fishing hours and output are *upward-sloping*, similar to the indifference curves over money (income, a good) and working time (a bad) in Chapter 4.

because an additional pound of fish provides him with a one unit increase in utility).

Bridget's utility function Equation 5.2 is interpreted in the same way as Abdul's. Both of them refer to some given time period, such as a week. So *output and consumption* are pounds of fish caught and eaten in a week, while *time spent fishing* is hours fished over the course of a week.

If for some reason they do not fish at all, they receive a small amount of fish y_z from neighbors or the government, labeled with the subscript z because this is their fallback position (as the endowment allocation was in Chapter 10)

To decide how much time to fish, people like Abdul have to balance their disutility of hours of work with the utility of consumption that they get from consuming the fruits – or the fish – of their work time. To understand this process, we look at Abdul's indifference curves.

REMINDER The indifference map provides information on how he evaluates all of the imaginable combinations of fishing time and fish caught. It says nothing about the actions and outcomes that are *feasible* for him.

Four indifference curves derived from Abdul's utility function, equation 5.1, are presented in Figure 5.4. Notice that:

- the higher numbered (meaning more preferred) indifference curves are above (more fish) and to the left (less work);
- the curves slope upwards because fish is a good and fishing time is a bad, so comparing points **f** and **g** he is indifferent between fishing less and consuming less (point **f**) and fishing more and consuming more (point **g**)

- the lowest indifference curve is labeled u_z^A and its vertical axis intercept is point **z** or the level of utility measured in fish per week, y_z^A that he will receive if he does not fish at all; and finally
- for any given level of y^A the indifference curve is steeper the more hours Abdul works: the more he works the greater is his dislike of working more compared to how much he likes eating more fish.

The negative of the slope of his indifference curve is Abdul's marginal rate of substitution between fish (y^A) and fishing time (h^A), the ratio of his marginal utility of fishing time to his marginal utility of fish. This quantity takes a particularly simple form in this case because (as is shown in M-Note 5.1). Abdul's marginal utility of fish is 1 and his marginal utility of fishing time is $-h^A$. So, the marginal rate of substitution of fish consumption for fishing time is:

$$mrs^A(h^A, y^A) \quad = \quad -h^A \tag{5.3}$$

M-CHECK Abdul's utility function in fish and fishing time is *quasi-linear*: since it is linear in fish – he derives a positive and constant marginal utility from consuming fish – but is *negative* (it is a *dis*utility) and non-linear in fishing time. His marginal disutility of fishing time it is not constant, it is greater the more time he spends fishing.

Abdul's marginal rate of substitution of fish consumption for fishing time is $-h^A$, and this is also his marginal disutility of fishing time, which is negative, because he regards fishing time as a "bad."

If he were already working 12 hours, then his disutility of hours of fishing (which is h^A itself) is the *greatest* amount of fish he would be willing to give up in order to be able to work an hour *less*. This is his willingness to pay (in fish) to have more free time. Or if Abdul were employed where he is already working h^A and paid a wage, then the quantity h^A is the lowest wage (paid in fish) in return for agreeing to work an extra hour, that he would *accept*.

M-Note 5.1: The $mrs(h, y)$ with quasi-linear preferences

When Abdul's utility is given by Equation 5.1, we have.

$$\text{Marginal utility of consuming fish} \quad = \quad \frac{\partial u^A(h^A, y^A)}{\partial y^A} = 1 \tag{5.4}$$

$$\text{Marginal disutility of fishing time} \quad = \quad \frac{\partial u^A(h^A, y^A)}{\partial h^A} = h^A \tag{5.5}$$

The marginal *utility* of fishing time is *negative* (it *reduces* Abdul's utility and is equal to $-h^A$) and we use the term marginal *dis*utility of fishing time for the same quantity but with a positive sign (it *increases* Abdul's *dis*utility).

The marginal rate of substitution of output for hours of work ($mrs^A(h^A, y^A)$) is the negative of the slope of the indifference curve, which is ratio of the marginal utilities:

$$mrs^A(h^A, y^A) \quad = \quad \frac{-h^A}{1} = -h^A \tag{5.6}$$

So the slope of the indifference curve is the marginal disutility of working time, or just h^A the amount of working time itself.

Furthermore, along any of the indifference curves, u_1^A, u_2^A and u_3^A, the vertical intercept is the amount of utility in fish if they were not working at all, that would be the same as the utility at every point on that indifference curve.

The marginal rate of substitution of Abudl's fish consumption for Abdul fishing time

$mrs(h^A, y^A)$ is an entirely different quantity than the marginal rate of substitution given by indifference curves for the fishing times of the fishermen $mrs(y^A, y^B)$ that we introduce later.

Checkpoint 5.2: The lake as a common property resource

a. Explain why the lake that Abdul and Bridget are fishing is a common property resource. What are its characteristics? Explain.

b. Return to Chapter 1 and the choice of strategies that the fishermen had in the Fishermen's Dilemma to Fish 10 hours or Fish 12 hours. Substitute these values into the utility functions to see what the payoffs in the corresponding game table would be if the fishermen could only choose these two strategies. Find the Nash equilibrium of the game.

5.3 *Technology and environmental limits: The source of a coordination failure*

TECHNOLOGY A technology is a description of the relationship between inputs –including work, machinery, and raw materials – and outputs.

A coordination problem arises because Abdul or Bridget fishing more reduces the amount of fish the other catches in an hour of fishing. This negative external effect that each has on the catch of the other is the source of the coordination problem.

These external effects are part of the **technology** of fishing. A technology is a description of the relationship between inputs – such as fishing time, equipment, and fish in the wild – and outputs – in this case caught fish. A technology is often described mathematically in a production function. You already used a production function in Figure 3.9 where the input was time spent studying and the output as learning. (We postpone a detailed discussion of production functions until the next chapter).

PRODUCTION FUNCTION A technology is a way of transforming inputs into outputs, described mathematically as a production function.

Here are the production functions for Abdul and Bridget, where x^A for Abdul and x^B for Bridget represents the number of fish caught by each of them in a week and h^A and h^B are the hours of fishing time they work during the week. The production functions translate the actions taken by the two – their fishing hours (h^A and h^B) – into the amount that each catches (x^A and x^B) and consumes (y^A and y^B).

$$\text{Abdul's catch:} \quad x^A(h^A, h^B) = h^A(\alpha - \beta(h^A + h^B)) \tag{5.7}$$

$$\text{Bridget's catch:} \quad x^B(h^A, h^B) = h^B(\alpha - \beta(h^A + h^B)) \tag{5.8}$$

The two parameters of the production function are:

- α (Greek *alpha*) is the the fisherman's *maximum average productivity*, that is, total catch divided by time spent fishing which would occur if one of them fished some small amount of time and the other did not fish at all. We assume that $\alpha > 0$ otherwise they could not ever catch any fish.

- β (Greek *beta*) measures the decrease in average productivity for each hour fished in total by the two. We assume that $\beta > 0$ to reflect the fact their interdependence and the negative external effect that each fishing has on the other's catch.

Were we to be applying this model to real fishermen, we would find that β and α differ. For example, one may have a larger boat and for this reason may catch more fish in an hour and also have a larger effect on the fishing productivity of the other. However, because our interest here is not in the effects of differing sizes of their boats, but instead on differing amount of time fished, we assume that β and α are the same for the two fishermen.

M-CHECK We adopt parameters for the production functions so that Bridget and Abdul cannot work so many hours that their average productivity becomes negative, so that fishing more would reduce their total catch. This is why we do *not* extend the lower of the two production function curves in Figure 5.5 beyond 24 hours, the point after which the function turns downwards.

From Equation 5.7 you can see that Abdul's total catch is his hours of fishing multiplied by his total catch per hour of fishing, termed the *average productivity of* his *fishing time*. Dividing both sides of Equation 5.7 by h^A we have his average productivity of fishing time:

$$\frac{x^A}{h^A} = \alpha - \beta(h^A + h^B) \tag{5.9}$$

What this means is that:

Average productivity $=$ Maximum $-$ Reductions due to own and other's fishing time

M-CHECK The marginal product of Abdul's fishing time is found by partial differentiating his total catch (x^A) given by the production function (Equation 5.7) with respect to his working time (h^A), which gives us Equation 5.10 We study the mathematical and conceptual properties of production functions and marginal products in the next chapter.

We are also interested in what is termed the marginal productivity of Abdul's fishing time. This is the effect of fishing a little more on the size of his total catch. In the M-check we show that Abdul's marginal productivity of fish time is:

$$mp_{h^A} = \alpha - \beta(h^A + h^B) - \beta h^A \tag{5.10}$$

This equation can be read as:

Marginal productivity $=$ Average productivity $-$ Reduction due to own fishing time

We call α the maximum productivity of fishing because it is the amount that would be produced per unit of fishing time when there is *no* fishing being done.

The parameter β expresses three important aspects of the technology:

- *Decreasing average productivity:* If Abdul spends more time fishing, his catch will be larger, but his average productivity – the size of the catch per hour fished – decreases.

- *Decreasing marginal product of work time :* If Abdul already fishes a lot, then the additional amount of fish that he catches were he to fish a little more will be less than if he were initially fishing a lesser amount.

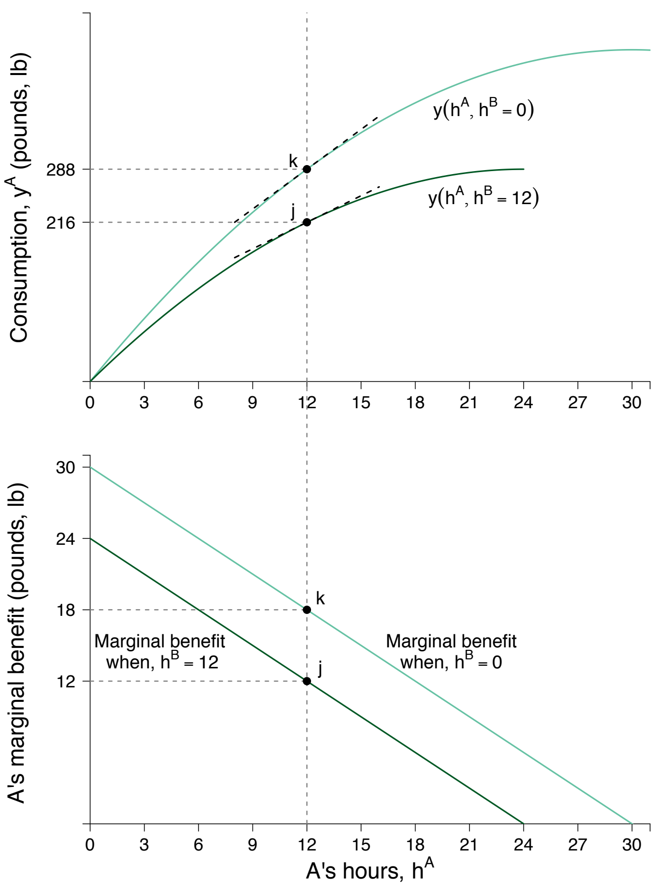


Figure 5.5: **Abdul's production of fish with hours of fishing and marginal benefit of hours spent fishing.** In the top panel, Abdul's light green total product line corresponds to when Bridget does not fish ($h^B = 0$) and Abdul's light green total product line corresponds to when Bridget fishes 12 hours ($h^B = 12$). Similarly, in the lower panel, Abdul's light green marginal benefit line corresponds to when Bridget does not fish ($h^B = 0$) and Abdul's light green marginal benefit line corresponds to when Bridget fishes 12 hours ($h^B = 12$).

- *Interdependence:* The fact that h^A appears in Bridget's production function and h^B in Abdul's represents the *external effects* and therefore the *interdependence* between the fishermen. The fact that the sign of these terms is *negative* means that the external effect is *negative*.

We depict Abdul's production function in the top panel of Figure 5.5. The higher of the two green curves represents the relationship between his labor input and his fish output when Bridget is not fishing at all, that is: $h^B = 0$. (We will explain the lower curve in a moment.)

Abdul's production function is increasing but becomes flatter the more time Abdul fishes. The slope of this curve is the marginal product of time fishing, indicating for each level of h^A the increase in the amount of his catch that would result if he increased his fishing time a little. We also call this the marginal benefit of fishing time because it indicates how much he benefits if he fishes a little more (how much the larger catch from additional fishing time raises his utility).

The second, lower, dark green curve in the top panel shows how Bridget's fishing for 12 hours reduces the amount of fish Abdul will catch for each level of Abdul's fishing time. Like the top curve, it is rising: as Abdul spends more time fishing he catches more fish. But two things about it are important:

- it is *lower* than when Bridget *does not fish*, and
- its slope is also lower (it is flatter for each hour that Abdul spends fishing).

Both are the result of the negative external effect that Bridget's fishing inflicts on Abdul.

In the lower panel of Figure 5.5 we show the marginal product of an hour of fishing based on the production function shown in the top panel, labeled the *marginal benefit* of hours of fishing.

In Figure 5.5 when Abdul fishes 12 hours a week (and Bridget does not fish), his catch is 288 but when she also fishes 12 hours (the lower green curve) his catch is just 216 lbs. Equally important, when Bridget is not fishing, and Abdul is fishing 12 hours, his marginal product is 18. The fact that the marginal benefit curve shifts downward when Bridget fishes 12 hours reflects the fact that in the top figure for any given amount of fishing time by Abdul, the curve is flatter.

M-Note 5.2: Numerical Examples for Productivity and External Effects

Throughout the chapter, we'll cover a worked example where Abdul and Bridget have the same level of productivity and external effect on each other. We shall assume that $\alpha = 30$ and that $\beta = \frac{1}{2}$.

Abdul and Bridget's utility functions therefore become the following:

$$\text{Abdul's utility:} \quad u^A(h^A, h^B) = h^A(30 - \frac{1}{2}(h^B + h^A)) - \frac{1}{2}(h^A)^2 \tag{5.11}$$

$$\text{Bridget's utility:} \quad u^B(h^A, h^B) = h^B(30 - \frac{1}{2}(h^A + h^B)) - \frac{1}{2}(h^B)^2 \tag{5.12}$$

In the case where the fishermen fished *alone*, that is the other fishermen had zero hours fishing, Abdul's utility would therefore be: $u^A = 30h^A - \frac{1}{2}(h^A)^2 - \frac{1}{2}(h^A)^2 = 30 - (h^A)^2$.

When Abdul and Bridget both spend time fishing, the external effect reduces Abdul's utility, therefore he would have $u^A = 30h^A - \frac{1}{2}h^A h^B - \frac{1}{2}(h^A)^2 - \frac{1}{2}(h^A)^2 = 30h^A - \frac{1}{2}h^A h^B - (h^A)^2$.

5.4 A best response: Another constrained optimization problem

REMINDER A player's best-response function gives, for every possible set of strategy chosen other players player, the strategy that maximizes the player's utility. A strategy profile in which all players are playing a best response, is a Nash equilibrium.

To understand the Nash equilibrium of the interaction between Abdul and Bridget we will need to know how each will best respond to any of the possible levels of fishing chosen by the other. This is because a Nash equilibrium is a mutual best response. To do this we will derive the *best response function* of each. But to do this we begin, as we did in Chapter 4, with a simpler problem:

showing how one of them, Abdul, will choose how many hours to fish, when Bridget is fishing at some given number of hours.

Abdul, choosing a level of fishing time

As a first step we bring together the information we have from the previous two sections on their preferences and their technology in a single equation expressing the benefits and costs of fishing:

Utility = Total benefit (fish caught and consumed) – Total cost (disutility of fishing time)

So, for each of them, we substitute the production functions (Equations 5.7 and 5.8) for $y^A = x^A$ and $y^B = x^B$ into their utility functions for their total consumption (Equations 5.1 and 5.2). Doing this we obtain for each of them a single function showing how their utility depends on their own and the other's fishing times, which is why we write their utility functions as $u^A(h^A, h^B)$ and $u^B(h^A, h^B)$

$$\text{Abdul's utility:} \quad u^A(h^A,h^B) = h^A(\alpha - \beta(h^A + h^B)) - \frac{1}{2}(h^A)^2 \tag{5.13}$$

$$\text{Bridget's utility:} \quad u^B(h^A,h^B) = h^B(\alpha - \beta(h^A + h^B)) - \frac{1}{2}(h^B)^2 \tag{5.14}$$

For concreteness let's suppose that Bridget is not fishing at all: she is a farmer and does not interact with Abdul in any way. We can therefore substitute $h^B = 0$ into Abdul's utility function, Equation 5.13. Then Abdul's constrained optimization problem is to maximize his utility subject to the constraint given by how productive his fishing time is when Bridget is *not* fishing.

This problem is set out in Figure 5.6 which combines Abdul's indifference curves from Figure 5.4 with his production function (when $h^B = 0$) from 5.5.

Abdul might first consider fishing six hours, with results indicated by points **f**, **g**, and **h** in Figure 5.6. To determine if he should fish 6 hours he would compare:

- *the marginal cost of working more*: namely the marginal disutility of working time, which is the slope of the indifference curve at **f** shown as point **h** in the lower panel with
- *marginal benefit of working more*: namely, the marginal productivity of his fishing time, which is the slope of the production function at **f** shown as point **g** in the lower panel.

From either the two slopes at point **f** ($mrs \neq mrt$) in the top panel or their representation by points **g** and **h** in the bottom panel ($mb > mc$) Abdul would see he would increase his utility by working more than 6 hours.

How much more? He will adopt the following method. He will compare:

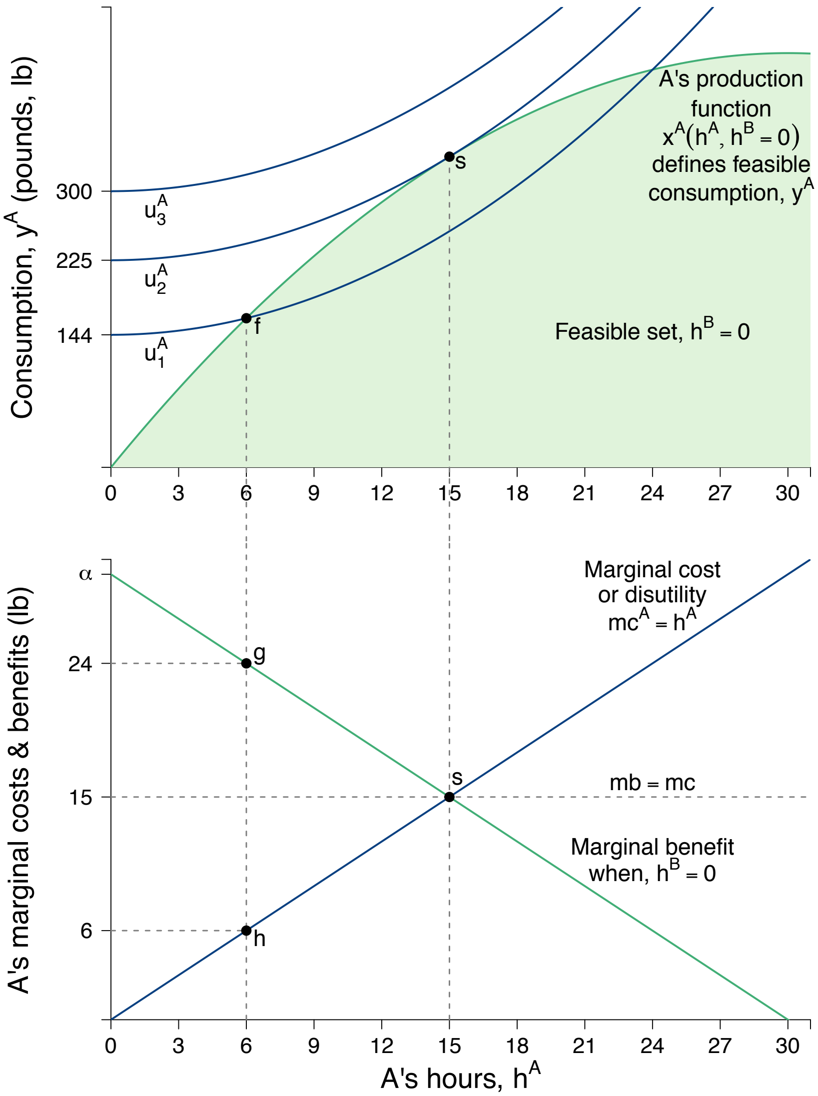


Figure 5.6: **Abdul maximizes his utility subject to the constraint of his production function when Bridget does not fish.**

Marginal benefit (mb): The additional fish he would catch is the marginal benefit of fishing more.

Marginal cost (mc): The additional effort he exerts or disutility it costs him is the marginal cost of fishing more.

Abdul will best respond if he follows some simple rules:

$mb > mc$ If the marginal benefit exceeds the marginal cost as at point **f**, then fish more.

$mb < mc$ If the marginal cost exceeds the marginal benefit, then fish less.

$mb = mc$ *If the marginal cost equals the marginal benefit, do not change how* *much you fish.*

5.5 A best-response function: Interdependence recognized

You can confirm from the figure that following the rule in italics just above, Abdul will fish 15 hours if Bridget is not fishing, indicated by point **s** in the figure (**s** for "solo" because Bridget is not fishing). This gives us just one point on his best response function $h^A(h^B = 0) = 15$ hours.

What about when Bridget *is* fishing, for example, fishing 12 hours? Abdul's reasoning is identical to the above rule. This case is illustrated in Figure 5.7 where the new feasible set constraining Abdul is smaller, because his catch for any amount of time that he spends fishing is *reduced* by Bridget also fishing.

Abdul knows that the level of fishing that will maximize his utility under these new conditions is that which equates:

- the slopes of an indifference curve and his production function so that the two are tangent in the top panel
- or, to put it another way, the marginal benefit and the marginal cost of more fishing in the bottom panel.

This gives us a second point on Abdul's best response function, $h^A(h^B = 12) = 12$. Abdul fishes less when Bridget fishes more. This occurs because Bridget's fishing more reduces the marginal benefit to Abdul's fishing.

What about Abdul's response to Bridget fishing different hours. We do not have to go through the above process, tediously making a separate figure for each level of fishing time she might choose.

Instead we can use mathematical expressions for the marginal costs and benefits of fishing to determine Abdul's best response not as a discrete point, but as a continuous function, giving us his fishing time for any level of fishing Bridget might do.

Using the rule that the best response is the number of hours that equates marginal benefits to marginal costs we have a general rule that can be expressed mathematically and which allows us to isolate h^A as a function of h^B and the parameters α and β. Here is the rule: a best response is a value of h^A that satisfies the following rule:

$$\begin{aligned} \text{Marginal benefit} &= \text{Marginal costs} \\ \alpha - \beta(2h^A + h^B) &= h^A \end{aligned} \tag{5.15}$$

Re-arranging Equation 5.36 to isolate h^A and to express his optimal fishing

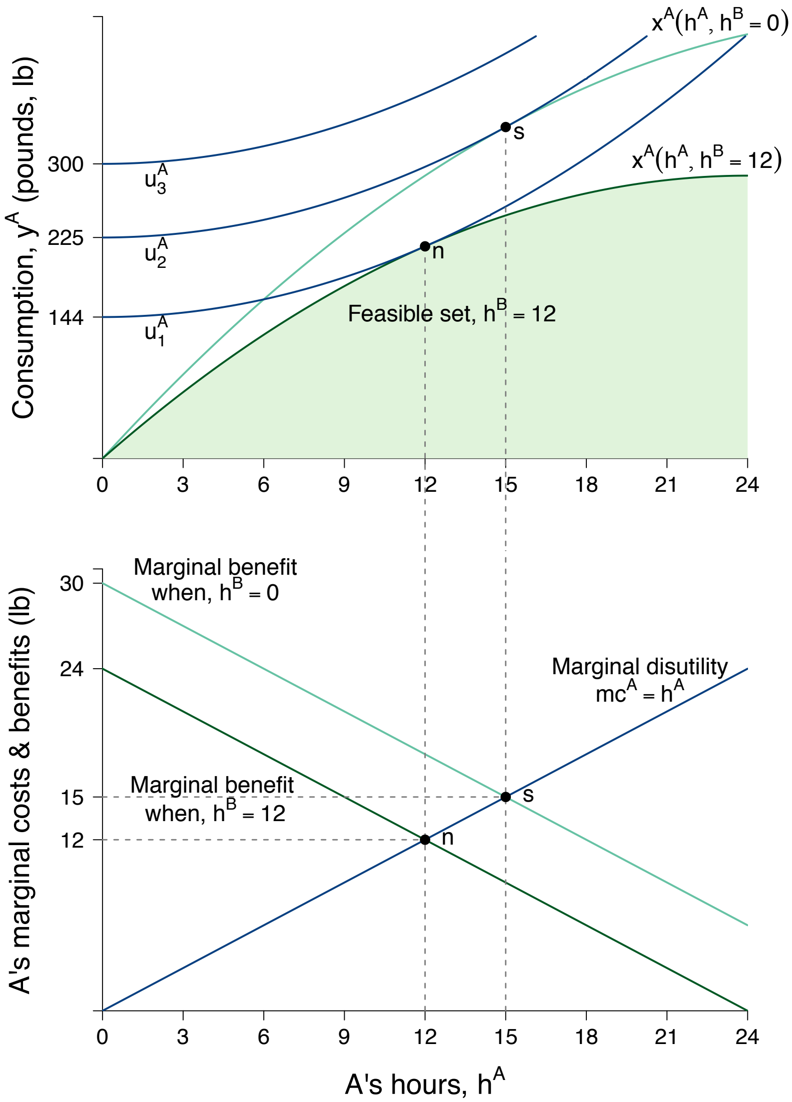


Figure 5.7: **Abdul maximizes his utility subject to the constraint of the production function when Bridget spends 12 hours fishing.** The feasible set is now smaller because of the negative external effect that her fishing imposes on Abdul. In the top panel, at point **n** his indifference curve labeled u_1^A is tangent to his production function, meaning in the lower panel, that the marginal disutility of fishing time is equal to the marginal productivity of fishing time, or the marginal cost of fishing more is equal to the marginal benefit.

hours as a function of Bridget's hours $h^A(h^B)$, we have

$$\text{Abdul's best-response function:} \qquad h^A(h^B) = \frac{\alpha - \beta h^B}{1 + 2\beta} \qquad (5.16)$$

How does Abdul's fishing time h^A change when the variable (h^B) and parameters (α and β) change?

- *Change in Bridget's fishing time* (h^B): If Bridget *decreases* her fishing time, Abdul's marginal benefit curve shifts *up*, and Abdul's best response is to *increase* his fishing time to balance his marginal cost with the higher marginal benefit. Abdul's best-response function does not shift, he chooses a different level of fishing due to the change in Bridget's fishing time.

- *Change in maximum productivity* (α): If Abdul's basic productivity *increases*, and nothing else changes, this shifts his marginal benefit curve *up* and *independently of any change in Bridget's fishing time,* he will *increase his fishing time* to balance his marginal cost with the higher marginal benefit. This is a *shift* in Abdul's best-response function itself, not just a movement from one point on it to another as in the bullet above.
- *Change in the over-fishing effect* (β): If the external effect *increases*, Abdul's marginal benefit curve pivots *downward* with a corresponding decrease in fishing time (β changes the *slope* of his marginal benefit curve, as can be seen from Equation 5.36). Like the increase in α, in this case Abdul changes his fishing time due to a *shift* in this best-response function.

The best-response function for Bridget can be derived in the same way we derived Abdul's. Therefore her best-response function is:

$$\text{Bridget's BRF :} \qquad h^B(h^A) = \frac{\alpha - \beta h^A}{1+2\beta} \tag{5.17}$$

M-Note 5.3: Marginal benefits, marginal costs, and finding the best responses

In M-Note 5.2, we used the example of $\alpha = 30$ and $\beta = \frac{1}{2}$ to provide utility functions for Abdul and Bridget, as represented in equations 5.11 and 5.12. We now use those parameters to identify the first-order condition for Abdul's utility maximization where his marginal benefits equal his marginal costs and therefore to provide a best-response function.

$$\begin{aligned} u^A(h^A, h^B) &= h^A(30 - \frac{1}{2}(h^B + h^A)) - \frac{1}{2}(h^A)^2 \\ u^A_{h^A} = \frac{\partial u^A}{\partial h^A} &= \underbrace{(30 - \frac{1}{2}h^B - h^A)}_{\text{Marginal benefit}} - \underbrace{h^A}_{\text{Marginal cost}} = 0 \end{aligned}$$

We can isolate Abdul's hours of work, h^A to find his best response to Bridget's hours of work:

$$\text{Abdul's BRF:} \quad h^A(h^B) = \frac{30 - \frac{1}{2}h^B}{2} = 15 - \frac{1}{4}h^B \tag{5.18}$$

$$\text{Bridget's BRF:} \quad h^B(h^A) = \frac{30 - \frac{1}{2}h^A}{2} = 15 - \frac{1}{4}h^A \tag{5.19}$$

Each of them therefore has a best-response function that is a function of the other person's time spent fishing: $h^A(h^B)$ for Abdul and $h^B(h^A)$ for Bridget.

M-Note 5.4: Mathematics of the best-response function

To understand each player's response to the other, it is useful to understand their marginal utilities of hours of fishing. We do this for Abdul, in the understanding that the Bridget will have symmetrical results. We will therefore find $u^A_{h^A}$, Abdul's marginal utility of his own hours of fishing, $u^A_{h^B}$, marginal utility of Bridget's hours of fishing, and $h^A(h^B)$, Abdul's best-response to Bridget's choice of hours. We start with Abdul's utility function:

$$u^A(h^A, h^B) = h^A(\alpha - \beta(h^A + h^B)) - \frac{1}{2}(h^A)^2$$

We can differentiate Abdul's utility function with respect to his own hours (h^A) to find his marginal utility of his own hours of work. We also differentiate his utility function with respect to Bridget's hours of work to find how his utility changes when Bridget changes her hours (h^B).

$$\begin{aligned} \text{A's marginal utility of } h^A \quad u^A_{h^A} = \frac{\partial u^A}{\partial h^A} &= \alpha - \beta h^B - 2\beta h^A - h^A \quad (5.20) \\ &= \alpha - \beta h^B - h^A(1 + 2\beta) \quad (5.21) \\ \text{Marginal effect on A's utility of } h^B \quad u^A_{h^B} = \frac{\partial u^A}{\partial h^B} &= -\beta h^A \quad (5.22) \end{aligned}$$

If we set Abdul's marginal utility $u^A_{h^A} = \frac{\partial u^A}{\partial h^A} = 0$, then we can find his best response to Bridget's hours of work:

$$\begin{aligned} u^A_{h^A} = \frac{\partial u^A}{\partial h^A} &= \alpha - \beta h^B - h^A(1 + 2\beta) = 0 \\ \text{Isolate } h^A \text{ term} \quad h^A(1 + 2\beta) &= \alpha - \beta h^B \\ \text{A s BRF:} \quad h^A &= \frac{\alpha - \beta h^B}{(1 + 2\beta)} \end{aligned}$$

Which is what we found from setting marginal benefit equal to marginal cost to find Abdul's best-response function in Equation 5.16.

We can use the above to define Abdul's marginal rate of substitution:

$$\begin{aligned} mrs^A(h^A, h^B) &= \frac{u^A_{h^A}}{u^A_{h^B}} \\ &= \frac{\alpha - \beta h^B - (1 + 2\beta)h^A}{-\beta h^A} \quad (5.23) \end{aligned}$$

Checkpoint 5.3: How the BRFs change

Consider how the parameters would change the best response functions in M-Note 5.2.

a. What would happen if $\alpha = 30$ changed to $\alpha = 24$ and $\beta = \frac{1}{2}$ changed to $\beta = \frac{1}{3}$?
b. What would the vertical and horizontal intercepts be for each of Abdul and Bridget? Sketch the best response functions you found in a.

5.6 How will the game be played? A symmetric Nash equilibrium

We do not have enough information to answer the question in the section title. To do this we need answers to other questions. Is one of them powerful enough determine the allocation unilaterally, stating: I fish 15 hours, and you are excluded from fishing? Is there a government that can place a tax on fishing to discourage over-harvesting the stock? Can Abdul and Bridget agree to fish less? If they did, can they count on their agreement being enforced? In other words, we need to know more about the *rules of the game*.

REMINDER We began our analysis of Ayanda and Biko trading data and coffee in a similar way, with the two being symmetrical traders with neither of them having any particular advantage in the bargaining process.

One possibility is that the two are independent (they do not make agreements with each other), self-regarding, and symmetrical, in that neither has any particular advantage in their interaction. So they simply try to do the best that they can, given what the other is doing and given the information they have. We will investigate other rules of the game later.

A stationary allocation among symmetric players

To study this case, we graph the two best-response functions in Figure 5.8. This gives us all the information we need to determine the Nash equilibrium of their interact on.

A Nash equilibrium is a mutual best response, so Abdul's choice of fishing hours must be a best response to Bridget's choice of fishing hours, which must in turn be a best response to Abdul's choice of fishing hours. This sounds complicated but with a little help from the mathematics we have already done, it is not: A Nash equilibrium is a point that is on both of the players' best-response functions. We label the point **n** and define the hours that they work at the Nash equilibrium as (h^{AN}, h^{BN}), where point **n** is the Nash equilibrium in the figure and the superscript N indicates each player's Nash equilibrium hours. A Nash equilibrium is a pair of fishing times (h^{AN}, h^{BN}) that satisfy each fisherman's best-response function.

When both players act according to their best-response functions, the outcome is a Nash equilibrium. In Figure 5.8 we plot the two best-response functions. The Nash equilibrium is the point where the two best response functions intersect the only point that the two lines have in common, shown by h^{AN} and h^{BN}.

We show in the M-Note 5.5 how to find the Nash equilibrium hours of fishing for each person. At the Nash equilibrium, the two fishermen will spend the same amount of time fishing.

$$h^{AN} = h^{BN} = \frac{\alpha}{1+3\beta} \tag{5.24}$$

Equation 5.24 shows that each fisherman's hours spent fishing is defined by the parameters α and β, capturing the effects on their best-response of their average productivity, their decreasing marginal productivity, and the negative external effect each has on the other.

The Nash equilibrium fishing hours, h^{AN} and h^{BN}, are equal because the Abdul and Bridget have identical utility functions (other than reversing the superscripts), and they are determined by the parameters α and β. The greater is the maximum average productivity, α, the greater will be their equilibrium hours of fishing. The larger is the negative external effect each has on their own productivity and on the other person's productivity, β, the lower their equilibrium hours will be.

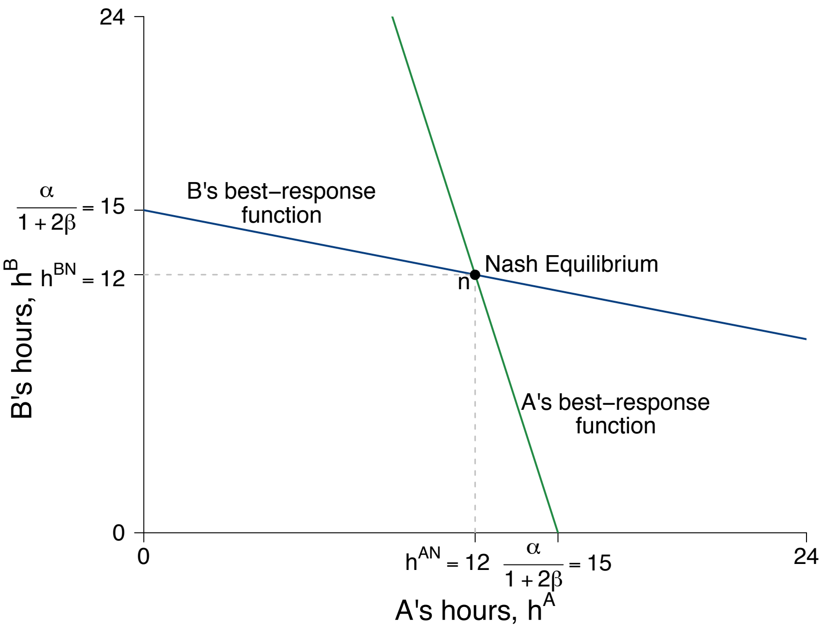


Figure 5.8: **Nash equilibrium: mutual best responses for Bridget and Abdul.** The equations for the best-response functions are:

$$h^B(h^A) = \frac{\alpha - \beta h^A}{1+2\beta}$$

$$h^A(h^B) = \frac{\alpha - \beta h^B}{1+2\beta}$$

If $\alpha = 30$ and $\beta = 0.5$, the parameters we used in the previous figures, then we can see that when Bridget does not fish (the intercept of Abdul's best response function with the horizontal axis) he fishes 15 hours. The point at which their best-response functions intersect is the Nash equilibrium of the interaction. Using these same parameters, we can see that the Nash equilibrium given by Equation 5.24 is that they both fish 12 hours.

M-Note 5.5: Finding Nash Equilibrium fishing time

By definition of the Nash equilibrium, Abdul's Nash equilibrium fishing time must be a best response to Bridget's Nash equilibrium fishing time, and Bridget's Nash equilibrium fishing time must be a best-response to Abdul's Nash equilibrium fishing time. A Nash equilibrium is therefore a pair of fishing times (h^{AN}, h^{BN}) that satisfy the following equations:

$$h^{AN} = h^A(h^{BN}) = \frac{(\alpha - \beta h^{BN})}{(1+2\beta)} \tag{5.25}$$

$$h^{BN} = h^B(h^{AN}) = \frac{(\alpha - \beta h^{AN})}{(1+2\beta)} \tag{5.26}$$

Equations 5.25 and are two linear equations in two unknowns. We can solve the equations for the unknowns, which are the fishing times at the Nash equilibrium.

There is a particularly simple way to do this in our case because:

1. The two fishermen have identical utility functions (they are mirror images of each other); so
2. we know that it must be that $h^{AN} = h^{BN}$, and
3. we can therefore set the Nash equilibrium level of fishing of the one equal to the best-response function of the other.

So substituting $h^B = h^A$, into Abdul's best-response function is:

$$h^A(h^B) \quad = \quad \frac{\alpha - \beta h^A}{1+2\beta}$$

Multiplying out and isolating h^A :

$$\begin{aligned} h^A + 2\beta h^A &= \alpha - \beta h^A \\ h^A + 3\beta h^A &= \alpha \\ h^A(1+3\beta) &= \alpha \\ h^{AN} &= \frac{\alpha}{1+3\beta} = h^{BN} \end{aligned}$$

5.7 How would the players get to the Nash equilibrium? A dynamic analysis

When we used the equation for the Nash equilibrium level of hours of fishing (Equation 5.24) to say what the effect of a change in α or β would be, we used what is called **comparative static** analysis.

COMPARATIVE STATICS When using comparative statics we compare the status quo outcome or the Nash equilibrium *before* the change with the outcome or Nash equilibrium *after* the change.

When using comparative statics we compare the status quo outcome or the Nash equilibrium *before* the change with the outcome or Nash equilibrium *after* the change.

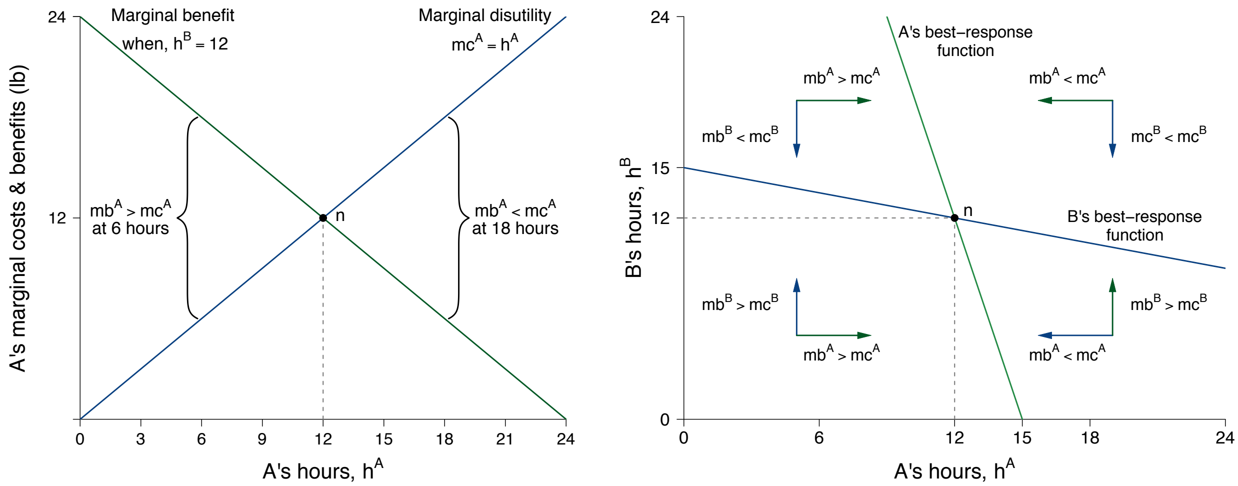
- The word *static* refers to the Nash equilibrium because at a Nash equilibrium there are no reasons for the actors to change what they are doing.
- The process is *comparative* because we compare two or more states before and after a change.

REMINDER When we say "other things equal" we are using the *ceteris paribus* assumption which allows us to compare what happens when *one* variable of interest changes.

We did the following:

- We started by assuming that Aram and Bina were at the Nash equilibrium, each fishing 12 hours.
- We then assumed that other things (like the weather) that might affect their fishing time are held constant (this is called the *ceteris paribus* assumption or 'other things equal').
- Then we compared the two Nash equilibria, one before the change and the other after the change, and.
- We assumed that after the change Aram and Bina would be at the new Nash equilibrium, working some different number of work hours.
- Finally we considered the difference in work hours between the two Nash equilibria to be the effect of the change on work hours.

This type of analysis gets the name *comparative static* because it *compares* two *static* (unchanging) situations *without* looking at *how* the change takes place. This is an essential method of economic analysis, simplifying the matter by a shortcut. The shortcut is that we *did not* explore the *process* by which the move from the first to the second Nash equilibrium occurs, namely who did what to implement the move. This is called a *dynamic analysis* because it is



(a) Marginal benefits and costs when B fishes 12 hours

(b) Dynamic analysis of A and B's choices

Figure 5.9: **How players can get to the Nash equilibrium: A dynamic analysis**. Panel a. shows the marginal costs and benefits of Abdul's fishing if Bridget fishes 12 hours. In panel b. shows the dynamics of the choices in terms of the fishermen's marginal benefits and marginal costs. The horizontal arrows show the direction Abdul will move if he is initially at the base of the arrow. The vertical arrows show the same for Bridget. The inequalities involving marginal benefits and costs (mb, mc) are the reason for the movement shown in the arrows (which are called "vectors").

based on the *process of change*. (The term dynamic refers to change, it is the opposite of static.)

We did not even explain why Abdul and Bridget would have been at the original Nash equilibrium in the first place. Fortunately, the way we have derived our best responses provides a way to fill in the necessary dynamic analysis.

Remember, when Abdul was selecting a best response he adopted a simple check list based on the marginal *benefits* of fishing more (mb) and the marginal *costs* of fishing more (mc):

- *if* $mb > mc$, then *fish more*
- *if* $mb < mc$, then *fish less*
- if $mc = mc$ don't change how much you are fishing.

When we introduced this check list we focused on the last line, because that is the equality that determines the utility-maximizing level of fishing for Abdul, that is, it is a point on his best-response function.

The top two lines of the checklist tell Abdul what to *change* when he is *not* fishing the optimal amount given what Bridget is doing, that is when he is 'off' his best-response function.

As Figure 5.9 shows, these first two lines of Abdul's checklist tell us that starting at any allocation (that is any combination of fishing hours of each of them) in which direction he should move, shown by the arrows. The dynamic analysis gives the following simple instruction: if you are not on your best-response function, move toward it.

Abdul's arrows are green and horizontal (when he changes his fishing hours

M-CHECK Abdul might adopt the instruction: close half of the difference between the hours I am now working and the hours indicated by my best-response function, given how many hours Bridget is now working. For example, if Abdul were fishing 6 hours while Bridget fished 12 hours, he would increase his hours by $(12-6)/2 = 3$ hours.

he moves left or right). The same reasoning allows us to show the dynamic arrows for Bridget, they are blue and horizontal, because when she changes her hours that moves the allocation point up or down.

For example, in Figure 5.9 a, if Abdul is fishing 6 hours the marginal benefits of fishing more exceed the costs (the bracketed term on the left). So in Figure 5.9 b, the arrows show that he will fish more. Similar reasoning (in reverse) applies to the case where he is fishing, for example, 18 hours. The extent by which the benefits differ from the costs depend on how much fishing Bridget is doing. Figure a shows the case for when she is fishing 12 hours. You can also work out how Bridget will adjust her hours if she is fishing more or less than the amount indicated by her best response function.

You can see from the figure that unless the allocation is at point **n** one or both of them will have an incentive to move (horizontally for Abdul, vertically for Bridget) in ways that will lead them to the Nash equilibrium.

This explains why we would expect both Bridget and Abdul to be at (or very close to) the Nash equilibrium. It also explains, if the Nash equilibrium shifted because of some change in either α or β, why we would *expect* the two to alter their fishing hours to move towards the *new* Nash equilibrium.

We now introduce a way that we can evaluate all of the possible equilibria of this game by the standards of Pareto efficiency and fairness.

M-Note 5.6: Numerical Nash Equilibrium

In M-Note 5.3, we found the best responses for Abdul and Bridget given by Equations 5.18 and 5.19. Using the method we outlined above, we set Abdul's Nash equilibrium hours of fishing equal to Bridget's BRF to find the Nash equilibrium level of fishing time:

$$
\begin{aligned}
\text{Bridget's BRF :} \quad h^{AN} &= 15 - \frac{1}{4}h^{AN} \\
\text{Collect terms} \quad h^{A} + \frac{1}{4}h^{A} &= 15 \\
\left(\frac{5}{4}\right)h^{A} &= 15 \\
\text{Multiply by } \frac{4}{5} \quad h^{AN} &= \left(\frac{4}{5}\right)15 = 12 = h^{BN}
\end{aligned}
\tag{5.27}
$$

As a result, we see that each will fish 12 hours at the Nash equilibrium. Therefore they each obtain the following Nash-equilibrium utility (by substituting h^{AN} and h^{BN} into their utility functions):

$$
\begin{aligned}
u^{AN}(h^{AN}, h^{BN}) &= h^{AN}(30 - \frac{1}{2}(h^{BN} + h^{AN}) - \frac{1}{2}(h^{AN})^2 \\
&= 12(30 - \frac{1}{2}(12+12)) - \frac{1}{2}(12)^2 \\
&= 216 - 72 = 144 = u^{BN}
\end{aligned}
$$

Each of them has a utility of 144 at the Nash equilibrium and the total welfare (sum of utilities) is $W^N = u^{AN} + u^{BN} = 288$.

Checkpoint 5.4: Storms and sustainability

Imagine that the external effect increased, as it would, for example, if greater climate volatility produced storms that caused the two fishers fish in the same limited part of the lake.

- Use the equation for the best responses of the two to redraw the figure. Why do the fishermen best respond by fishing less?
- Use the equation for h^{AN} and h^{BN} to show that the Nash equilibrium level of fishing will decline.
- Use what you have learned to explain how the best-response functions and the Nash equilibrium would change if the fishermen jointly adopt a strategy to let go of young fish to make the fish population more sustainable and reduce the external effect they have on the other fisherman.

5.8 Evaluating outcomes: Participation constraints, Pareto improvements and Pareto-efficiency

Because the symmetrical interaction is just one of many possible rules of the game that Bridget and Abdul might engage in, we need to go beyond the Nash equilibrium for that game and think find a way to evaluate all of the possible allocations that they might experience.

To do this, as in Chapter 4 we use the indifference maps of the two players superimposed on the same set of outcomes. Recall that in the previous chapter every point in the Edgeworth box indicated an allocation composed of a bundle of goods for Ayanda and another bundle of goods for Biko. We will see that the same is true in this case if we plot an allocation as the pair of fishing hours of the two, h^A, h^B. We start with Abdul's preferences.

Because the utility of each depends on their own fishing time and the fishing time of the other, that is because

$$\text{Abdul's utility:} \quad u^A(h^A, h^B) \tag{5.28}$$

$$\text{Bridget's utility:} \quad u^B(h^A, h^B) \tag{5.29}$$

we can plot indifference curves with fishing time on each axis: Bridget's fishing time (h^B) on the vertical axis and Abdul's fishing time (h^A) on the horizontal axis.

We do this in panel a of Figure 5.10, where every point in the figure is a particular allocation of fishing times (h^A, h^B). Using these allocations we can calculate the utility that Abdul would experience were that allocation to occur. On this basis we can calculate Abdul's indifference curves based on in his hours of fishing and Bridget's hours of fishing. Abdul prefers curves labeled with higher numbers, $u_3 > u_2 > u_1$. Notice two things about the indifference curves:

- *The vertical dimension, or the effect of Bridget fishing more*. Abdul's preferred indifference curves are lower. This is because the less Bridget fishes the better it is for Abdul.

- *The horizontal dimension, or the effect of Abdul fishing more*. If Bridget is fishing at the "low" level indicated in the figure and supposing that Abdul initially does not fish at all but considers fishing a little, he will start by finding himself at successively higher indifference curves as he fishes more, crossing the indifference curves labeled u_1^A, and then u_2^A and up to u_3^A. But if he spends too much time fishing he will then cross from u_3^A back down to indifference curve u_2^A and again go back down to u_1^A.

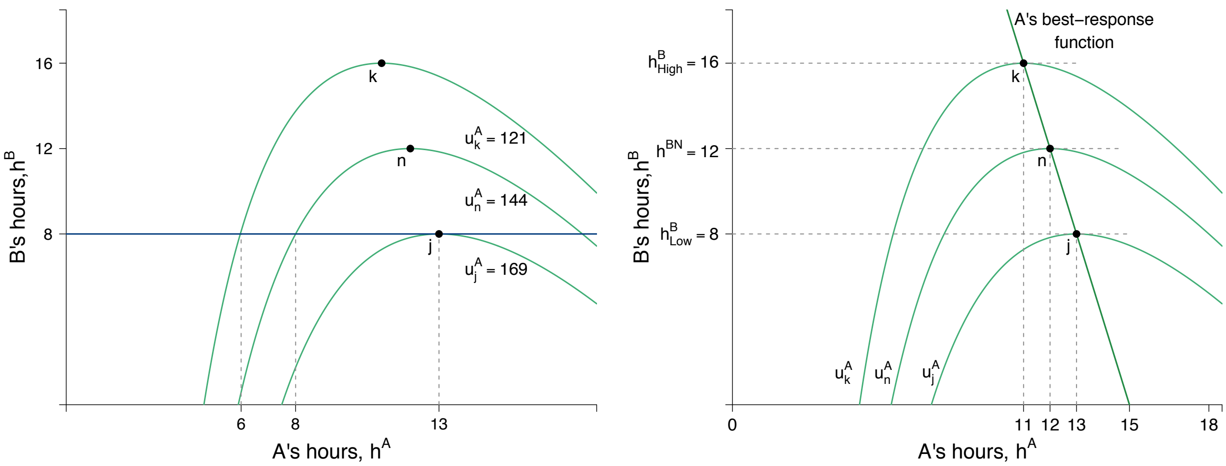
Another perspective on a best-response function

We can use the horizontal dimension of the figure to identify a point on Abdul's best-response function, associated with Bridget hypothetically fishing just 8 hours. We take this thought experiment as a constraint on Abdul's utility maximization. In the figure the green shaded area is his feasible set, and the horizontal line is its frontier. Remember Abdul prefers indifference curves that are lower down (indicating Bridget fishing less). The most preferred indifference curve that is feasible is the one tangent to the constraint, at point **j**, which is therefore a point in Abdul's best-response function.

It may help to think of his indifference map as showing the contours of the shoulder of a hill, and Abdul as walking along the frontier of the feasible set towards point **j** trying out different amounts of time he might devote to fishing. This is exactly what he did in Figure 5.6, comparing the marginal benefit and marginal cost of fishing more. At first he is climbing – crossing contours indicating ever-higher altitudes - higher utility. When he fishes 6 hours, he achieves utility $u_k^A = 121$, proceeding on to fish 8 hours, he achieves $u_n^A = 144$, and finally fishing 13 hours, he achieves $u_j^A = 169$. At point **j** his path levels off and if he continues to increase his fishing time he will descend to lower altitudes – lower utility – once more.

Panel b of Figure 5.10 illustrates how two additional points on Abdul's best response-function are derived. The best-response function is constructed by considering all of the possible levels of fishing that Bridget could hypothetically do, and then reason as we did for point **j**.

Notice that Abdul's best-response function intersects the indifference curves where the indifference curves are flat. If the indifference curve is flat then the mrs must be zero. In M-Note **??** we show why this must be true. In the right panel you can see that as Bridget's fishing time increases from 8 to 16 hours, Abdul's fishing time declines from 13 to 11 hours. He identifies his best-response hours of fishing by finding the point on his best-response function that corresponds to the number of hours Bridget fishes.



(a) Abdul making a choice constrained by Bridget's hours

(b) Abdul's best-response function

Figure 5.10: **A new look at Abdul's constrained optimization problem for selecting his fishing time depending on Bridget's fishing time.** .
To illustrate the construction of Abdul's best-response function, in Panel a we consider Abdul's decision about how many hours to work, given that Bridget has (hypothetically) decided to work 8 hours. The horizontal blue line is the constraint on Abdul's utility-maximizing process. In panel b, we consider three hypothetical levels of Bridget's fishing time. The horizontal lines represent Bridget's fishing time at each of these levels, and are the constraint on Abdul's maximization process. One of these horizontal lines is tangent to each of Abdul's indifference curves $h^B = 8$ tangent to u^A_j at point **j**, $h^B = 16$ tangent to u^A_k at point **k**, and $h^{BN} = 12$ tangent to u^A_n at point **n**. Abdul's entire best response function is made of of points like **j**, **n**, and **k**, for each of Bridget's possible levels of fishing hours.

M-Note 5.7: Why is the best-response function made up of points where the indifference curves are flat?

The negative of the slope of the indifference curve is:

$$mrs^A(h^A, y^A) \quad = \quad \frac{u^A_{h^A}}{u^A_{y^A}} \tag{5.30}$$

Abdul's best-response function gives the values of h^A and h^B for which the derivative of Abdul's utility with respect to his fishing time is equal to zero or:

$$u^A_{h^A} \quad = \quad \alpha - \beta h^B - (1+2\beta)h^A = 0$$

If $u^A_{h^A} \quad = \quad 0$, then the numerator of Equation 5.30 is zero, so the slope of the indifference curve is equal to zero, which means that it is flat.

Checkpoint 5.5: The marginal rate of substitution

Although each of the fishermen doesn't "control" the hours of work that the other does, we can still think in sensible ways about the marginal rate of substitution.

a. Consulting Figure **??**, what is the sign of the marginal rate of substitution going from left to right along the indifference curve leading up to point **n** on u^A_2 (on the left-hand side of point **n)**? Why?

b. Continuing to consult Figure **??**, what is the sign of the marginal rate of substitution going from left to right along the indifference curve on the right-hand side of point **n** on u^A_2? Why?

Fallback positions and the Pareto improving lens

We have said that rules of the game other than symmetrical interaction will lead to different Nash equilibrium allocations. As long as the interaction among the two is voluntary – there are no "offers you cannot refuse" – we can limit the possible outcomes by thinking about the alternatives that the two have should they decide *not* to fish at all. Any allocation in which they both fish and that makes either of them (or both) worse off than how they would do if they did not fish at all will not occur for the simple reason that they will not fish if they could do better by not.

We have already introduced the idea that Abdul, if he does not fish at all, will receive an income of y_z possibly from family, friends or the government. Suppose the same opportunity applies to Bridget. If they do not fish and receive y_z then their utility is just y_z. This is their fallback position (like the allocation **z** in the Edgeworth box of the previous chapter). But they only receive their fallback if they *do not* fish, so the opportunity cost of fishing – what they cannot have if they fish – is y_z. In Figure 5.11 we show both Abdul's and Bridget's indifference maps. We see from the numbering of the utility labels on the curves, that Bridget's indifference curves give greater values the closer they are to the vertical axis (as Abdul's did with the horizontal axis).

One of their indifference curves is particularly important: it is labeled u_z^A and u_z^B. These two curves show all of the allocations h^A, h^B that yield, for Abdul and Bridget respectively a level of utility equal to the utility of their fallback position namely $u_z = y_z$. This is the *participation constraint* for each of them: they will not participate in fishing unless they can do at least this well. Any point between these indifference curves is a Pareto-improvement over their fallback position: both are better off than their fallback option.

Any point between these indifference curves is a Pareto-improvement. The Pareto-improvements are shown by the Pareto-improving lens shaded in yellow.

The Pareto-efficient curve

There is another important curve in Figure 5.11: the purple solid and dashed Pareto-efficient curve. We know that Pareto-efficiency requires that the fishermen's indifference curves be *tangent*, that is, for their marginal rates of substitution to equal. You can see two of these tangencies in the interior of the Pareto improving lens. The other tangencies defining the Pareto-efficient curve are not shown. The Pareto-efficient curve is made up of all points representing allocations for which:

$$mrs^A = \frac{u^A_{h^B}}{u^A_{h^A}} \quad = \quad \frac{u^B_{h^B}}{u^B_{h^A}} = mrs^B \tag{5.31}$$

Equation 5.31 shows the condition that the marginal rates of substitution must be equal at Pareto-efficient outcomes, meaning that their indifference curves are tangent.

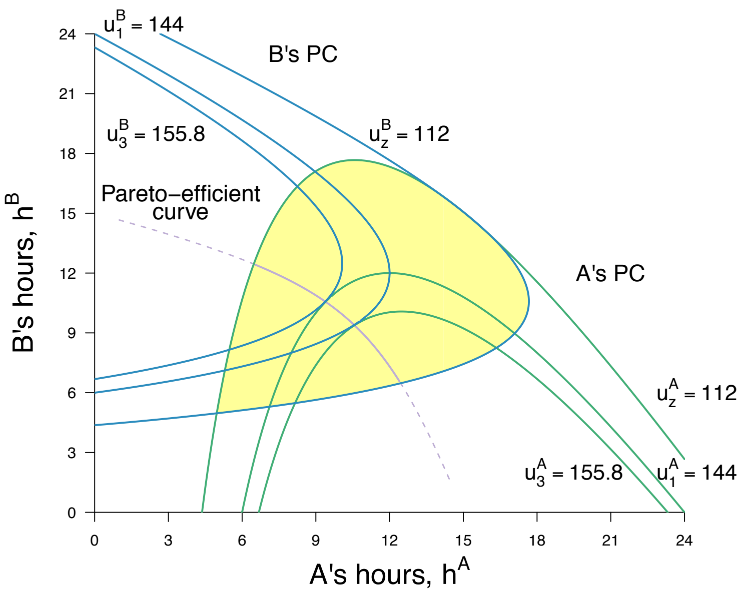


Figure 5.11: **Abdul's indifference curves and Bridget's indifference curves showing their fallback levels of utility (their participation constraints)**, u_z^A and u_z^B. At their fallback positions, they are not fishing at all and receiving a payment (in fish) of u_z^A and u_z^B from the government .

The figure clarifies the difference between *Pareto improvements* and *Pareto efficiency*:

- The points on the purple Pareto-efficient curve that are indicated by a dashed line outside the yellow Pareto improving lens are *Pareto-efficient but not Pareto improvements* over the fallback *no fishing* option.
- The points in the yellow Pareto-improving lens that are *not* on the purple Pareto-efficient curve are *Pareto improvements but not Pareto efficient*.

Checkpoint 5.6: Understanding the parameters

a. What would be the effect on the Pareto-improving lens if α increased?
b. What would be the effect on the Pareto-improving lens if β increased?
c. Why does β affect the person fishing even when no one else fishes? Why does it make economic sense?

5.9 A Pareto inefficient Nash equilibrium

We return now to the symmetrical interaction between Bridget and Abdul, in which the Nash equilibrium is the allocation at the intersection of their best-response functions. And we ask: is that allocation Pareto-efficient?

To answer, we combine two figures we have already introduced: Figure 5.11 showing the two fishermen's indifference curves and Figure 5.8 showing their best-response functions. The combination of these figures results in Figure 5.12.

REMINDER To understand Figure 5.12 it will help to remember that for Abdul *down* is better (his indifference curves have higher utility the lower they are) because the less Bridget fishes the better it is for him. Similarly, Bridget is better off on the indifference curves further to the left.

Figure 5.12 shows that the Nash Equilibrium is not Pareto-efficient: at the Nash allocation (point **n**) the indifference curves of the two *intersect* rather than being *tangent*. So allocation **n** *cannot* be Pareto-efficient.

How do we know that their indifference curves cannot be tangent at that point that is, how do we know that

$$mrs^A = \frac{u^A_{h^B}}{u^A_{h^A}} \quad \neq \quad \frac{u^B_{h^B}}{u^B_{h^A}} = mrs^B \tag{5.32}$$

The answer is that the Nash equilibrium is a point on both best-response functions, defined by $u^B_{h^B} = 0$ for Bridget's best response and $u^A_{h^A} = 0$ for Abdul's best response. At the best response each fisherman adjusts their own fishing time to maximize utility so that these two terms will be zero. If we substitute the zeroes for the marginal utilities in Equation 5.31, we find the following:

- The first expression is now zero divided by $u^B_{h^A}$ so the slope of Abdul's indifference curve is zero; it is flat (as in Figure 5.12), and
- The second expression is now $u^A_{h^B}$ divided by zero, so the slope of Abdul's indifference curve is infinite; it is vertical (as in Figure 5.12 too)

A flat line cannot be tangent to a vertical line, so the condition for Pareto efficiency is violated and the Nash equilibrium is not Pareto efficient.

A view from a Pareto inefficient status quo Nash equilibrium.

We now imagine Abdul and Bridget, fishing 12 hours each as indicated by the Pareto-efficient Nash equilibrium. They realize they could both do better. And they consider the options. They each might propose some different allocation. To agree on an alternative level of fishing, the proposal would have to implement a Pareto improvement. The Pareto improvement would need to be over the Nash equilibrium, *not* over their no-fishing fallback option. Remember that the Nash equilibrium is already better than their the fallback positions.

With allocation **n** the new fallback for the agreement, we now have a new yellow shaded Pareto-improving lens. There are two things to notice about Pareto-improvements over the Nash allocation:

- both fishermen spend *less* time fishing and both are *better off* (have higher utility than at the Nash equilibrium) and
- the new Pareto-improving lens is much smaller than the lens of Pareto improvements over the no-fishing fallback option.

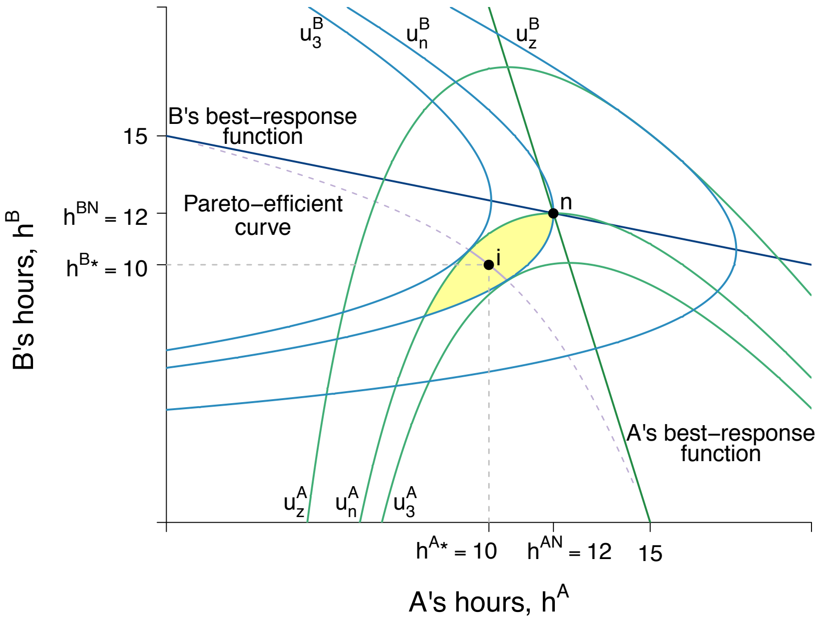


Figure 5.12: **The Nash equilibrium and the Pareto-improving lens.** The Pareto-improving fishing times (in which both fish less) are in the pale yellow lens. Notice that Abdul's indifference curve at the Nash equilibrium is flat, and Bridget's at the same point is vertical (their marginal rates of substitution are not equal). This being the case there must be a Pareto-improving lens and the Nash equilibrium cannot be Pareto-efficient.

Checkpoint 5.7: Checking the marginal rate of substitution

Use the values of $\alpha = 30$ and $\beta = \frac{1}{2}$ and substitute them into the marginal rate of substitution for Abdul and Bridget.

a. Confirm that when Abdul sets his $mb = mc$, the numerator is zero.

b. Also confirm that if you find the equivalent for Bridget, the $mrs^B = \infty$ (or it is undefined) when Bridget sets her marginal benefits equal to her marginal costs.

The reason why there exist allocations that are Pareto improvements over the Nash is as follows.

- *Reason 1:* Each of them would benefit a lot if the other were to fish less and
- *Reason 2:* at the Nash equilibrium each of them would experience very little lost utility by themselves fishing a little less.

Reason 1 concerns each fishermen's marginal utility with respect to the other's hours of fishing, and we can see that $u^A_{h^B} < 0$ and $u^B_{h^A} < 0$ because each fisherman's fishing time reduces the other's productivity.

Concerning Reason 2, suppose that, at the Nash equilibrium level of the fishing times, Bridget decided she would try to *bribe* Abdul to fish less. How much would she have to give him to fish a tiny bit less? The answer is "almost nothing" because at the Nash equilibrium, changes in his fishing time have no effect on his utility because the marginal benefits of fishing a little more *or less* equal the marginal costs of fishing a little more *or less* (that is how he chose that level of fishing to do).

So Abdul's fishing a little less would not matter much to Abdul but it would definitely benefit Bridget. A similar results is true for Bridget: Abdul could bribe her to fish a little less for a tiny portion of his fish. This being the case if they both could agree to fish less (and just forget about the bribes) they would both be better off.

The conclusion is that Bridget and Abdul need not lament their sorry condition at the Nash equilibrium. If a deal can be enforced – an agreement to limit fishing, maybe along with a bribe – there's a deal to be made that benefits them both.

We turn now to considering changes in the rules of the game that might reduce fishing times, keeping in mind that we are thinking about not just two people, but an entire community of people – perhaps the entire world's population if we are considering coordination problems such as climate change or the spread of epidemic diseases.

M-Note 5.8: The Nash equilibrium cannot be Pareto-efficient

To show that the Nash equilibrium is not Pareto-efficient we ask if they could agree each to fish an arbitrarily small amount less would they both be better off. If the answer is "yes," then the NE cannot be Pareto-efficient. We know that $u^A_{h^B} < 0$ and $u^B_{h^A} < 0$ each would be better off if the other fished less. We also know that $u^A_{h^A} = 0$ and $u^B_{h^B} = 0$ because these equalities define Bridget's and Abdul's best-response functions, and the Nash Equilibrium they are trying to improve on is a pair of strategies each of which is a best response to the other.

So for any change dh^A and dh^B, representing an agreement to change their fishing time, we can evaluate the change in each utility associated with change in the fishing times of each.

$$du^A = u^A_{h^A} dh^A + u^A_{h^B} dh^B$$
$$du^B = u^B_{h^A} dh^A + u^B_{h^B} dh^B$$

Eliminating the terms equal to zero in the expressions above, namely those involving $u^A_{h^A}$ and $u^B_{h^B}$ we have:

$$du^A = u^A_{h^B} dh^B < 0$$
$$du^B = u^B_{h^A} dh^A < 0$$

or, rearranging

$$\frac{du^A}{dh^B} = u^A_{h^B} < 0$$
$$\frac{du^B}{dh^A} = u^B_{h^A} < 0$$

Both expressions are negative: the utility of each would be enhanced by an agreement to fish a little less. The Nash equilibrium allocation of fishing times is not Pareto-efficient.

Checkpoint 5.8: Pareto-improvements

Use the numerical values of $\alpha = 30$ and $\beta = \frac{1}{2}$

a. How much utility would the players have if they both simultaneously reduced their hours of work from the Nash equilibrium values of 12 hours to 11 hours? Would they be better or worse off?
b. If one reduced their hours to 11 hours, what would the other's best response be? Would they actually reduce their hours to 11 or not?

5.10 A benchmark socially-optimal allocation

To provide a benchmark or standard against which we might evaluate the various rules of the game that might improve on the Nash equilibrium of the symmetric interaction above, we will reintroduce the Impartial Spectator, who we relied on for the same purpose in Chapter 4. The Impartial Spectator wishes to determine fishing time and distribute fish so as to maximize a social welfare function, which, because she values the utilities of the two equally, is just the sum of the utilities of the two:

M-CHECK In Chapter 4, we gave the Impartial Spectator Cobb-Douglas preferences over the two players' utilities. Here we have the Impartial Spectator sum the utilities of the two fishermen, as the Impartial Spectator did in Chapter 3.

$$\begin{aligned} \text{Total social welfare} &= \text{Abdul's utility} + \text{Bridget's utility} \\ W &= u^A(h^A, h^B) + u^B(h^A, h^B) \end{aligned} \tag{5.33}$$

She knows that the solution to this problem must be Pareto-efficient, because if it were not, then one of the two could be made better off without worsening the condition of the other, so this could not be the optimum for the Impartial Observer, who values the well being of both. This means that the socially optimal allocation must be somewhere along the Pareto-efficient curve in Figure 5.12. But where?

A socially optimal allocation

To answer the question, we transform the view of the problem in Figure 5.12, where the space in the figure is defined for hours of fishing, into a new graph, Figure 5.13, in which presents the same information in terms of the utilities of the two. The Pareto-efficient curve in Figure 5.12 appears in Figure 5.13 as the dark green curve that is frontier of the feasible set of utilities.

Its slope is the marginal rate of transformation of Bridget's utility into Abdul's utility. This provides the answer to the question: along the feasible frontier, how much does Bridget's utility have to fall in order for Abdul's to increase by one unit?

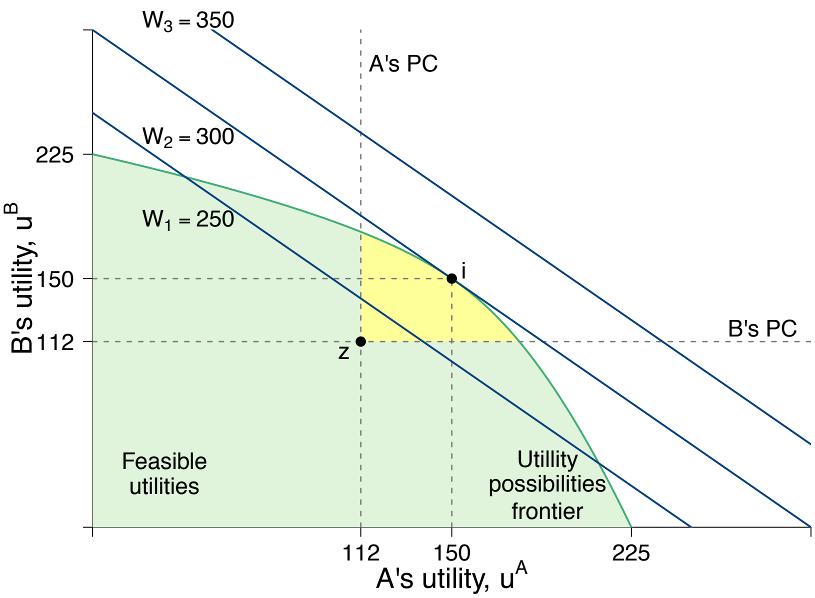


Figure 5.13: **Feasible utilities, the utility possibility frontier, and the Impartial Spectator's iso-social welfare indifference curves.** Here we show the utility possibilities frontier and feasible utilities for the Impartial Spectator. All points on the frontier are Pareto-efficient. The points above and to the right of the fishermen's participation constraints constitute the bargaining set, that is the outcomes that are Pareto-superior to their fallback options, $u_z^A = u_z^B = 112$. The Impartial Spectator's iso-social welfare indifference curves show her equal valuation of the utility of the two and the negative of the slope of her iso-social welfare curves is her marginal rate of substitution. The slope of her iso-social welfare curves is -1 indicating that she values the two utilities equally. The negative of the slope of the utility possibilities frontier is the marginal rate of transformation of Bridget utility into Abdul's. That is, it is the opportunity cost of Abdul having more utility in terms of the utility that Bridget forgoes as a result.

The impartial spectator will therefore choose point **i** where $mrs = mrt$ to maximize social welfare given the constraint of the utility possibilities frontier.

You can see that when Bridget has almost all of the feasible utility then it does not 'cost' Bridget much for Abdul to have a little more (the frontier is not very steep); but the marginal rate of transformation rises (the curve steepens) as Abdul gains more utility. The reason is that when Bridget has most of the utility, she is working long hours (almost 15) and incurring a *substantial disutility* of working time as a result. Fishing a little less would *not reduce* her utility much, but for Abdul fishing a little *more* would substantially *increase* his utility. So when Bridget is doing most of the fishing (and gaining most of the utility) the opportunity cost of increasing Abdul's utility (in terms of Bridget's forgone utility) is small.

The Impartial Spectator's values are expressed by her indifference curves (the blue lines), their slopes, her marginal rate of substitution, are a constant, namely -1, because she values the utility of the two equally.

The point **z** represents the fallback utilities of the two (namely 112), and the yellow shaded area is the set of feasible Pareto-improvements over this fallback position. You can see that the optimum point, **i**, is found where the highest feasible Impartial Spectators indifference curve is tangent to the utility possibilities frontier (the frontier of the feasible set). So this is another case of the familiar $mrs = mrt$ rule, but now for the Impartial Spectator, rather than Abdul or Bridget.

$$\text{marginal rate of substitution} = \text{marginal rate of transformation} \quad (5.34)$$

Rules that implement the social optimum

We know that acting on the basis of their best-response functions Abdul and Bridget *over-exploit* the resource. They could both do better if they adopted a different rule for deciding how much to fish. Before turning to institutions that might implement such a new rule for their decisions, lets think about a rule that would exactly implement point *i* in the figure.

To find the optimum, point **i** in Figure 5.13 the Impartial Spectator proposes the following rules that, if followed, will maximize her social welfare function Equation 5.33. We show in the M-Note 5.9 how these are derived.

$$\begin{aligned} h^A &= \alpha - 2\beta h^B - 2\beta h^A \\ h^B &= \alpha - 2\beta h^A - 2\beta h^B \end{aligned} \tag{5.35}$$

Focusing on the equation for Abdul, the socially optimal condition looks very similar to Abdul's best-response function (shown again below) when he was maximizing his own utility, except for one big difference.

$$\begin{array}{lccc} & \text{Marginal private costs} & = & \text{Marginal private benefits} \\ \text{Abdul's own optimality condition} & h^A & = & \alpha - \beta 2h^A - \beta h^B \end{array} \tag{5.36}$$

Comparing Equations 5.35 and 5.36 with the latter rearranged, or:

$$\begin{array}{lccc} & \text{Marginal social costs} & = & \text{Marginal private benefits} \\ \text{Abdul's social optimality condition} & h^A + \beta h^A & = & \alpha - 2\beta h^B - \beta h^A \end{array} \tag{5.37}$$

we see that the difference is that there is an extra $-\beta h^B$ in the socially optimal condition (Equation 5.35.) This is the negative external effect of Abdul's fishing on Bridget's utility.

PRIVATE AND SOCIAL COST The private cost (marginal or average) is the cost that the decision-maker bears as a result of some action that he or she takes. The social cost is the private cost plus any costs imposed on others as negative external effects.

In Equation 5.37 we have moved this term to the left-hand side of the equation, adding it to the marginal private cost of fishing more (namely the disutility of hours of fishing, h^A). So Equation 5.37, the condition for Abdul's fishing time to implement a social optimum says the following.

$$\text{marginal private cost} + \text{marginal external cost} = \text{marginal private benefit} \tag{5.38}$$

The left-hand side is called the *marginal social cost*. The **private cost** (marginal or average) is the cost that the decision-maker bears as a result of some action that he or she takes. The **social cost** is the private cost plus any costs imposed on others as negative external effects.

The Impartial Spectator reasons that if we are to implement an allocation that values the utility of both equally, then Abdul should act is if he is taking account of this cost – treating the costs he imposes on Bridget no differently than his own disutility of labor – when deciding on how much to fish. As you know from the previous chapter, this is called internalizing the negative external effect of his actions. We can think of these socially optimal responses in the following way:

$$\text{Optimal response} = \text{Own best response} + \text{Internalized cost to other} \quad (5.39)$$

Imposing the same condition on Bridget, the Impartial Spectator provides a rule where each fisherman *internalizes the negative external effect* of their hours of fishing on the other. As a result, we arrive at the levels of socially optimal fishing time for Abdul and Bridget, denoted by as h_i^A and h_i^B:

$$\begin{aligned} \text{Abdul's socially optimal fishing time :} \quad h_i^A &= \frac{\alpha}{1+4\beta} \quad (5.40) \\ \text{Bridget's socially optimal fishing time :} \quad h_i^B &= \frac{\alpha}{1+4\beta} \quad (5.41) \end{aligned}$$

Because $\beta > 0$, we see that each of the players' Nash equilibrium levels of fishing time are higher than the socially opimal levels:

$$h^{AN} = h^{BN} = \frac{\alpha}{1+3\beta} > \frac{\alpha}{1+4\beta} = h_i^A = h_i^B$$

These socially optimal levels of fishing time correspond to point **i** (for *i*mpartial) in Figure 5.12 where Abdul and Bridget have the same fishing time (10 hours) and the same level of utility.

The job description of the Impartial Spectator is not to figure out how this optimal allocation might be implemented ("way above my pay scale," she says). We leave that to a second (also imaginary) person who we will introduce in the next section.

M-Note 5.9: The Impartial Spectator's Choice

We know that the Impartial Spectator has the following social welfare function, and we can substitute the fishermen's utility functions into W as follows:

$$\begin{aligned} \text{Total Social Welfare} &= \text{Abdul's Utility} + \text{Bridget's Utility} \\ W &= u^A + u^B \\ &= h^A(\alpha - \beta(h^A + h^B)) + h^B(\alpha - \beta(h^A + h^B)) - \frac{1}{2}(h^B)^2 - \frac{1}{2}(h^A)^2 \\ &= \alpha h^A + \alpha h^B - 2\beta h^A h^B - \frac{1}{2}(h^A)^2 - \frac{1}{2}(h^B)^2 \quad (5.42) \end{aligned}$$

Next, the social planner needs to find the social welfare maximum, or the optimal social welfare. To do this, we partially differentiate W with respect to the hours of fishing of each fisherman, h^A and h^B, as follows to find the first order conditions for the social welfare optimum:

$$W_{h^A} = \frac{\partial W}{\partial h^A} = \alpha - 2\beta h^B - 2\beta h^A - h^A = 0$$

$$h^A = \frac{\alpha - \beta h^B - \beta h^B}{1+2\beta} \tag{5.43}$$

$$W_{h^B} = \frac{\partial W}{\partial h^B} = \alpha - 2\beta h^A - 2\beta h^B - h^B = 0$$

$$h^B = \frac{\alpha - \beta h^A - \beta h^A}{1+2\beta} \tag{5.44}$$

Notice that Equations 5.50 and 5.44 look similar to the best-response functions we found previously, but that each incorporates an additional $-\beta h^B$ for Abdul in the numerator of Equation 5.50 and $-\beta h^A$ for Bridget as shown in the numerator of Equation 5.44. These terms correspond to the cost of the external effect that each fisherman imposes on the other.

M-Note 5.10: Numerical Choice of the Impartial Spectator

The Impartial Spectator has a social welfare function, W which is the sum of the two fishermen's utilities. We substitute each fisherman's utility function (Equations 5.11 and 5.12) into the social welfare function and assume the parameter values of $\alpha = 30$ and $\beta = \frac{1}{2}$.

$$\begin{aligned} W &= u^A + u^B \\ &= h^A(30 - \frac{1}{2}(h^B + h^A)) - \frac{1}{2}(h^A)^2 + h^b(30 - \frac{1}{2}(h^A + h^B)) - \frac{1}{2}(h^b)^2 \\ &= 30h^A + 30h^B - h^A h^B - 2(h^A)^2 - 2(h^B)^2 \end{aligned} \tag{5.45}$$

The Impartial Spectator then needs to differentiate the social welfare function defined by Equation 5.45 with respect to the two hours of work to determine how much each person should work:

$$W_{h^A} = \frac{\partial W}{\partial h^A} = 30 - h^B - 2h^A$$

$$\therefore h^A = \frac{30 - h^B}{2} = 15 - \frac{1}{2}h^B \tag{5.46}$$

$$W_{h^B} = \frac{\partial W}{\partial h^B} = 30 - h^A - 2h^B$$

$$\therefore h^B = \frac{30 - h^A}{2} = 15 - \frac{1}{2}h^B \tag{5.47}$$

We can solve for the Impartial Spectator's choice of work hours for Abdul and Bridget by substituting Equation 5.47 for h^B into Equation 5.46. Following the process of substitution as usual, we find:

$$\begin{aligned} h_i^A &= 15 - \frac{1}{2}(15 - \frac{1}{2}h_i^A) \\ h_i^A &= 15 - 7.5 + \frac{1}{4}h_i^A \\ \text{Collect terms} \quad \frac{3}{4}i^A &= 7.5 \\ \text{Multiply by } \frac{4}{3}h_i^A &= \frac{4}{3}(7.5) = \frac{30}{10} \\ h_i^A &= 10 = h_i^B \end{aligned}$$

Each of them will work 10 hours at the socially optimal and Pareto-efficient allocation of hours that the Impartial Spectator made. At the allocation (h_i^A, h_i^B), each of the fishermen has a utility of 150, which is higher than they had at the Nash equilibrium (see Marshal Memo 5.6) and the total social welfare is $W_i = u_i^A + u_i^B = 150 + 150 = 300$.

Checkpoint 5.9: Changing things for the impartial spectator

Assume that $\alpha = 24$ and $\beta = \frac{1}{3}$ instead of the values you have used so far.

1. What would the number of hours be at the Nash equilibrium? How much utility would each player have and what would the total utility be?
2. How many hours would the Impartial Spectator choose? How much utility would each player have and what would the total utility be?
3. Sketch the best-response functions you found for (a) and sketch indifference curves for the Nash equilibrium and the Impartial Spectator's choice through the relevant points.

Remedies: Preferences, power, and policy

Stories about two fictional people such as Abdul and Bridget and their textbook lake are light years away from real fishermen in Rhode Island or Australia. John Sorlien, the Rhode Island lobsterman who we quoted at the start of the chapter, is in competition not with a single other fisherman, but with hundreds. Unlike Abdul and Bridget (so far) real-world fishermen and lobstermen do sometimes cooperate to pursue common objectives (Sorlien headed their association).

A friendly conversation and a handshake might be enough for Bridget and Abdul. But how might such an agreement be arrived at, and how might it be enforced in Rhode Island or Australia? An even greater challenge is how to design and enforce similar agreements – for example to burn less carbon in order to mitigate climate change – that span not thousands of actors but billions living under the jurisdiction of hundreds of independent governments.

But the parable of Abdul and Bridget has provided an important insight. Illuminating the basic source of coordination failures: the negative effect of their own fishing on the other person ($u^A_{h^B}$ and $u^B_{h^A}$, that is) is *not* part of the utility-maximizing process by which each choose how much to fish.

Addressing these external effects is where institutions come in, meaning changes in the rules of the game. There are three basic approaches whether the common property resource that is being over exploited be a fish stocks in a lake or the limited carbon emissions carrying capacity of earth's atmosphere.

HISTORY In 1968, Garrett Hardin wrote that "freedom in the commons means ruin to all" and as a result he advocated – "mutual coercion mutually agreed upon." But Hardin's pessimism overlooked the many non-coercive ways that local communities have prevented the tragedy.[5]

- *Regulation* of the exploitation of the resource by a government.
- *Private ownership* of the resource so that private incentives will deter over-exploitation.
- *Management* of the resource through local interactions among the the

resource users.

These three approaches are sometimes referred to as *states* (meaning governments), *markets*, and *communities*, or similar terms.[6]

5.11 Government policies: Regulation and taxation

To underline the fact that we are not modeling actual governments but instead an ideal of what a well informed and set of public officials seeking to implement better allocations might do, we will introduce a second hypothetical character, the Mechanism Designer. He is tasked with finding policies to implement outcomes that would be recommended by the Impersonal Spectator according to her values of efficiency and fairness. He is an economist, with an engineering mentality, she is a philosopher and her job is to identify good, better or best outcomes.

The mechanism designer's job then is to implement the best that can be done. The Mechanism Designer is the main character in the Chapter 16 which is about public policy. You than think about him as an economist advising a government about how to design and implement its policies.

A government has many options to address coordination failures such as common property resource over-exploitation, including educating the public about the costs and promoting basic research to find ways of making the resource more sustainable. But we focus on just two:

- *Fiat*: Governments can sometimes order the implementation of an allocation – for example, reduced fishing – that they or the voters who elected them prefer. A fiat is an order.
- *Taxation:* Governments regularly implement taxes as incentives: taxes make activities more costly and can therefore, discourage these activities, while still allowing each person or firm to choose how much of the activity to engage in given the increased costs.

FACT CHECK In response to the rapid depletion of fish stocks in 1992 the Canadian government simply banned fishing for cod in the Grand Banks region of the North Atlantic. Stocks have been recovering since then.[7]

Fiat power

With respect to fiat, the government, if it knew all the relevant information, could select $h^A = h_i^A$ and $h^B = h_i^B$ to maximize total utility. The government might implement this outcome by direct regulation, simply issuing a fishing permit allowing each fisherman a certain number of hours. Any deviation from the permitted hours would result in revocation of the permit and the fishermen would have to revert to their no-fishing fallback positions.

Point **i** in Figure 5.12 is the Impartial Spectator's efficient fishing time allocation that corresponds to what the government would choose. Assuming the government had no reason to favor one fisherman over the other from the standpoint of fairness, a Pareto-efficient and equal distribution of fishing times would be the fiat allocation.

Optimal taxes: Internalizing external effects

TAXES A tax is a charge the government enforces on the production or purchase of a good. A subsidy is a payment the government makes to the producer or purchaser, similar to a negative tax.

Rather than implementing the efficient fishing time plan by fiat, however, the government might want to let the fishermen each decide how much to fish, but to change their incentives in order to address the coordination failure. The government would levy what is called a **Pigouvian tax** on fishing designed to eliminate the discrepancy between the social and private marginal costs and benefits of fishing.

The problem for the government is to select a tax rate on fishing time that as an *intended* byproduct will motivate the fishers to implement an allocation that maximizes total utility while at the same time maximizing their own utility. This means bringing the fishermen's private incentives (the utility function that each maximizes) into alignment with the conditions laid out by the Impartial Spectator.

HISTORY Imposing taxes on particular behaviors which the government wants discourage because they impose negative external effects on others – over fishing fishing, smoking – the government is taking an approach pioneered by the early 20th century economists Alfred Marshall (1842–1924) and A.C. Pigou (pee-GOO) (1877–1959). In recognition of his contribution to the field of what is called welfare economics, these are sometimes called Pigouvian taxes.

The problem can be posed this way: find the tax rate that would transform the utility functions of the two fishermen so that their individual best-response functions are identical to those implied by the problem solved by the Impartial Spectator: maximizing total utility and internalizing the costs of the negative external effects.

To internalize the cost means to require each of them to pay (in taxes) for the reduction in the catch of the other that their additional fishing time imposes. We know that each additional hour that Bridget fishes means that Abdul catches βh^A pounds less of fish. So to force Bridget to take account of this negative external effect, she must be taxed at a rate of βh^A for every hour she fishes.

EXAMPLE The "golden rule" is a common ethical principle that people should treat each other as they themselves would like to be treated. A Pigouvian tax is designed to accomplish the same result by imposing on each decision maker the costs that their decisions impose on others.

Bridget's socially optimal tax rate depends on Abdul's fishing time, because the external effect of Bridget's fishing on Abdul's well-being depends on how much Abdul fishes. If Abdul is not fishing at all, for example, there is no need to tax Bridget's fishing, because it has no external effect.

The tax that induces the fishermen to choose the socially optimal levels of fishing time is just equal to the negative *external effect* they impose on others at the Pareto-efficient levels of fishing time. A Pigouvian tax is a change in the rules of the game that has the effect of *internalizing the external effect* that is the cause of the coordination problem. The tax is an indirect form of coordination: the fishermen as citizens elect a government which they delegate to impose on them a set of incentives to overcome the over-fishing problem.

M-Note 5.11: The best-response function of fishing with taxes

Equation 5.48 shows Bridget's utility function when her fishing time is taxed at the rate of τ per hour fished:

$$u^B(h^A, h^B, \tau) \quad = \quad h^B(\alpha - \beta(h^A + h^B)) - \tau h^B - \frac{(h^B)^2}{2} \tag{5.48}$$

To obtain the best-response function Bridget's fishing conditional on the tax rate and Abdul's fishing time, we differentiate the equation above and set the result equal to zero:

$$u^B_{h^B} \frac{u^B}{h^B} \quad = \quad \alpha - \beta h^A - 2\beta h^B - \tau - h^B = 0$$

We can rearrange this first order condition to say that (on the left-hand side of the equation below) the marginal benefits of fishing more (in fish caught) must be equal to (on the right hand side) the marginal *costs* including the the disutility of additional fishing time plus the taxes incurred by fishing more:

$$\alpha - \beta h^A - 2\beta h^B = \tau + h^B$$

Re-arranging this to isolate h^B so as to have a best-response function we have:

$$(1 + 2\beta)h^B \quad = \quad \alpha - \tau - \beta h^A$$

$$h^B(h^A, \tau) \quad = \quad \frac{\alpha - \tau - \beta h^A}{1 + 2\beta} \tag{5.49}$$

M-Note 5.12: A implementing the Impartial Spectator's choice

We know from M-Note 5.9 that the first order condition for Bridget's fishing time in the social optimum allocation proposed by the Impartial Spectator is:

$$h^B \quad = \quad \frac{\alpha - \beta h^A - \beta h^A}{1 + 2\beta} \tag{5.50}$$

The mechanism designer's job is to find the tax rate per hour of Bridget's fishing time that will induce her to act as if that were her private (self-regarding) first order condition too. We know from M-Note 5.11 that Bridget's true best-response function including taking account of the tax is the following:

$$h^B(h^A, \tau) \quad = \quad \frac{\alpha - \tau - \beta h^A}{1 + 2\beta} \tag{5.51}$$

The question that the mechanism designer must now solve is: what is the level of the tax rate τ, that will make Equation 5.51 look like Equation 5.50 so that Bridget's private incentives will lead her to implement the Impartial Spectator's social optimum.

Comparing the two equations you can see that setting the tax rate that Bridget pays $\tau^B \quad = \quad \beta h^A$ will make the two equations identical. So that is the optimal tax rate. The tax rate for Abdul would, by the same reasoning be $\tau^A = \beta h^B$.

Then we can calculate the tax rate that Bridget pays at the Nash equilibrium. Because we know that the optimal tax implements the social optimum recommended by the Impartial Spectator, we substitute the value for h_i^A into the expression for the tax rate so we have , $\tau = \beta h_i^A$ or

$$\text{Bridget's tax rate} \quad \tau \quad = \quad \beta h_i^A$$

$$\text{Abdul's hours worked (Nash)} \quad h_i^A \quad = \quad \frac{\alpha}{1 + 4\beta}$$

$$\text{Bridget's tax rate (Nash)} \quad \tau h^B \quad = \quad \frac{\alpha\beta}{1 + 4\beta}$$

For the parameters we have introduced to illustrate the model with $\beta = 0.5$ and $\alpha = 30$, the tax rate at the Nash equilibrium of the game with optimal taxes is 5 pounds of fish per hour of fishing time, meaning that every hour that Bridget fishes she must pay from her catch a total of five pounds of fish. Given that she is fishing 10 hours at the equilibrium of the game she pays 50 pounds of fish in taxes. Abdul pays the same amount.

We do not subtract this amount from the their utilities at the Nash equilibrium because we assume that the total tax revenues are redistributed to the population equally without regard for their fishing times.

Checkpoint 5.10: A partial tax

Assume $\alpha = 30$ and $\beta = \frac{1}{2}$. Consider that a government implements a tax where each person is charged $\tau = \frac{1}{4}$ for an hour of work.

a. What is Abdul's utility function incorporating this tax? What is Bridget's utility function with the tax?

b. What are their best responses to the tax policy?

c. What is the Nash equilibrium level of hours with $\tau = \frac{1}{4}$? Is the outcome Pareto-efficient? If it is not Pareto-efficient, is it a Pareto-improvement over the Nash equilibrium without the tax? Explain.

d. If $\tau = \frac{1}{4}$ is insufficient to produce a Pareto-efficient outcome, what level of τ would produce the Pareto-efficient outcome? Why? Explain.

5.12 *Private ownership: Permits and employment*

But government policies are not the only change in the rules of the game that might address the over-fishing coordination problem. Suppose that the property rights over the lake are changed such that the lake is no longer a common pool resource, but is *privately owned*. As a result, lake is no longer non-excludable, as a resource that is both excludable and rival, the it is a *private good.* The person who owns the lake, say Bridget, could exclude Abdul entirely (remember that is what private property means). But as an owner she now has bargaining power over Abdul, and may be able to do better by letting him fish under conditions favorable to her.

REMINDER While a single owner will take account of the costs of over-fishing and restrict fishing accordingly, the owner may also be a monopolist in selling the fish to others, and will restrict fishing even more than is socially optimal so as to sustain a high price of fish. We do not include the consumers of fish other than the owner and those fishing on the lake. But we will include them when we return to these effects of a monopolist in Chapter 9.

How do these new rules of the game change the Nash equilibrium?

- *Permits*: Bridget might sell Abdul a fishing permit allowing him to catch not more than a given amount of fish, setting the highest possible *fee* for the permit consistent with Abdul being willing to fish under those terms.

- *Employment*: Bridget might offer Abdul an employment contract under which Abdul would fish a given amount of time; the fish caught by Abdul would be Bridget's. Abdul's compensation would be a *wage* (paid in the fish caught by the two of them) which would be sufficient to offset the disutility of Abdul's fishing time and the opportunity cost of his fishing (and therefore to satisfy Abdul's participation constraint).

Selling permits to fish

To understand what Bridget will do as the owner of the lake, let us return to her utility function:

$$\text{Bridget's utility} \qquad u^B(h^A, h^B) \quad = \quad y^B - \frac{1}{2}(h^B)^2 \tag{5.52}$$

In Equation 5.52, when Bridget was one of the two fisherman and could *not* charge anyone to access the lake, her *production* of fish equalled her *consumption* of fish (y^B). Now, though, as she will be charging Abdul to access the lake, we separate her *production* of fish, $x(h^A, h^B)$, from her consumption of fish, y^B. When she is the owner, her consumption of fish equals her *own* production plus the fee she charges, F. Her utility therefore becomes:

$$\text{Owner's utility} \qquad u^B(h^A, h^B, F) = \underbrace{\overbrace{x(h^A, h^B)}^{\text{Bridget's production}} + F}_{\text{Bridget's consumption}} \overbrace{-\frac{1}{2}(h^B)^2}^{\text{Bridget's disutility}} \tag{5.53}$$

Equation 5.53 tells us that Bridget, as the owner, now has three variables to determine not just one (her own fishing time) when she interacted with Abdul in the symmetric game.

- her own fishing time, h^B
- Abdul's fishing time, h^A
- the cost F to Abdul of the permit allowing him to fish h^A hours.

Here, the fee for the permit will be an amount of fish that Abdul would transfer to Bridget, but the idea easily extends to thinking about monetary payments rather than payments in kind like this one, since fish in this scenario are effectively money.

Bridget, being self-regarding, will want to know: "what is the largest fee that I can charge Abdul?" Remember Abdul has the option of not fishing at all and and receiving a transfer of an amount y_z of fish; this is his fallback position with associated utility $u_z^A = y_z$. Agreeing to fish in Bridget's lake means foregoing the fallback option, so $u_z^A = y_z$ is the *opportunity cost* of fishing.

This is Abdul's participation constraint, limiting how much Bridget can charge for the permit: the fee plus the opportunity cost of fishing cannot be larger than Abdul's utility from fishing, $u^A(h^A, h^B)$:

$$\begin{array}{lrcl} & \text{Abdul's utility from fishing} & \geq & \text{Permit fee + Foregone fallback} \\ \text{Abdul's participation constraint} & u^A(h^A, h^B) & \geq & F + y_z \end{array} \tag{5.54}$$

Because Bridget would never consider charging Abdul less than she could, we can assume that equation 5.54 will be satisfied as an equality and so (re-arranging the equation) we have:

$$\text{Abdul's participation constraint (PC)} \qquad F = u^A(h^A,h^B) - y_z \tag{5.55}$$

Bridget's constrained optimization problem is now to vary h^A and h^B to maximize her utility from fishing plus the fee she charges Abdul, or:

$$\text{Maximize her total utility:} \qquad u^B(h^A,h^B,F) = x(h^A,h^B) + F - \frac{1}{2}(h^B)^2 \tag{5.56}$$

$$\text{subject to Abdul's PC:} \qquad F = u^A(h^A,h^B) - y_z \tag{5.57}$$

We can use the expression for F that is equation 5.57 to replace the F in equation 5.56, so that now Bridget's objective is to chose h^A and h^B to:

$$\text{Maximize:} \qquad u^B(h^A,h^B) = x(h^A,h^B) - \frac{1}{2}(h^B)^2 + u^A(h^A,h^B) - y_z \tag{5.58}$$

Once we have found the h_i^A and h_i^B that maximize 5.58, we can insert those values of h^A and h^B into equation 5.57 to determine F, the cost of the permit to charge Abdul.

Comparing equation 5.58 with equation 5.33 we can see that Bridget maximizes the same quantity that the Impartial Spectator maximized, namely the sum of the utilities of the two, except here y_z is subtracted. But because y_z is a constant (112 pounds of fish in our numerical examples) the solution of these two optimizing problems – the values of h^A and h^B chosen – must be the same.

This means when Bridget is the owner the hours worked, h_i^A and h_i^B will be *equal*, ten hours each in our numerical example – but the levels of utility realized will be maximally *unequal*. Abdul will get exactly 112, his fallback position, and Bridget will get 188 as shown by point **b** in Figure 5.14.

Because the allocation was determined by Bridget maximizing her utility subject to a constraint on Abdul's level of utility (that is the participation constraint) it has to be Pareto efficient, by definition. If Bridget as the owner implemented a plan in which she reduced Abdul's work hours and increased her own, she would obtain an allocation on the utility possibilities frontier at point b'. However, if she implements the optimal number of hours ($h_i^A = h_i^B = 10$) at point **i** and has Abdul pay her for the right to fish with a fee, then the total rents available are 300. She will charge a fee such that Abdul will receive just a bit more than his fallback, $u_z^A = 112$ and she will get $300 - 112 = 188$. This corresponds to a movement *along* the blue line with slope $= -1$ from point **i** to point **b**. The blue line therefore indicates movements from point **i** to alternative feasible trades when the fishermen are able to trade fish between them as payments.

The economic reason why the result is Pareto-efficient follows directly from the fact that Bridget knew in advance that she would capture all of the feasible

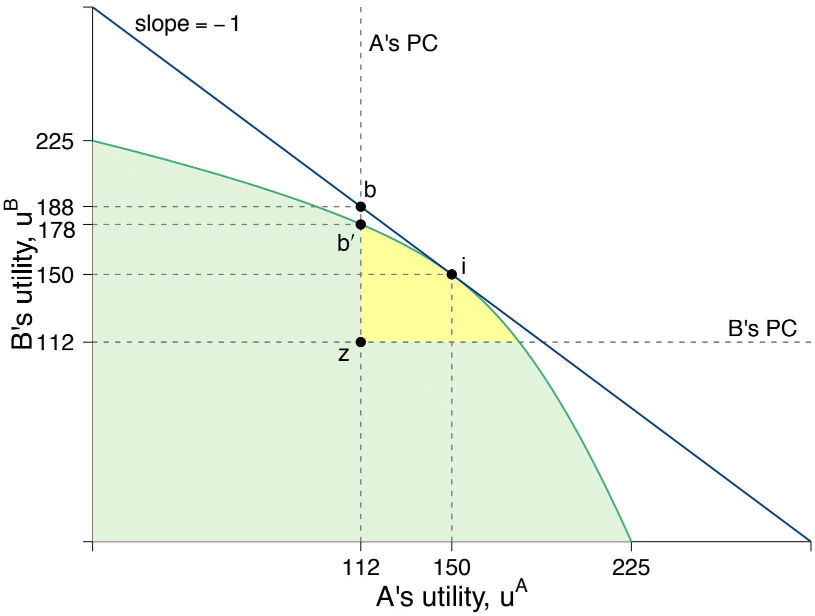


Figure 5.14: **Payments in fish takes the fishermen to allocations outside of the feasible set.** The blue line with slope -1 shows the allocations of utility that are possible if the two fish at the socially optimal times indicted by point **i** follow a transfer of fish from on to the other. The slope is -1 because the opportunity cost of, say Bridget having a kg more fish, is that Abdul has one kg less. If Bridget as the owner implemented a plan in which she reduced Abdul's work hours and increased her own, she would obtain an allocation on the utility possibilities frontier at point b'. However, if she implements the optimal number of hours ($h_i^A = h_i^B = 10$) at point **i** and has Abdul pay her for the right to fish with a fee, then the total rents available are 300. She will charge a fee such that Abdul will receive just a bit more than his fallback, $u_z^A = 112$ and she will get $300 - 112 = 188$. This corresponds to a movement *along* blue trade line with slope $= -1$ from point **i** to point **b**.

rents. Given that fact she had every interest in making the total rents as large as possible. Had she chosen a Pareto-inefficient outcome she would have foregone the opportunity to make herself better of without making Abdul worse off (than his participation constraint).

But why wouldn't Bridget select $h^A = 0$, and have exclusive access to the lake? The reason is that the marginal cost of compensating Abdul's fishing time is *very small* when he is not fishing much, or at all. So it is to Bridget's *advantage* to let Abdul fish in the lake and pay her for the privilege, rather than doing all the fishing herself.

Employing others to fish

Instead of issuing a permit, Bridget might hire Abdul to work for her. Employment differs from the permit system in that when Bridget employs Abdul, she owns all of the fish caught by Abdul, but must devote some of this to paying a wage w to Abdul sufficient to satisfy his participation constraint.

From our reasoning in the permit case we know that the participation constraint will be satisfied as an equality. This allows us to use the fact that the total wage paid (w) must offset Abdul's disutility of fishing time and the opportunity cost of fishing (namely his fallback option, y_z that he gives up if he fishes) or:

$$\text{Abdul's PC as an employee} \qquad w = \frac{(h^A)^2}{2} + y_z \qquad (5.59)$$

Bridget then must choose h^A and h^B to maximize her utility:

$$\begin{aligned} u^B(h^A,h^B,w) &= x^A(h^A,h^B)+x^B(h^A,h^B)-\frac{1}{2}(h^B)^2-w \\ &= \text{A's catch } + \text{B's catch} - \text{B's disutility} - \text{A's wage} \end{aligned} \tag{5.60}$$

Then using equation5.59 for Abdul's wage, what Bridget maximizes when she employs Abdul is:

$$u^B(h^A,h^B) = h^B(\alpha-\beta(h^A+h^B)) - \frac{(h^B)^2}{2} + h^A(\alpha-\beta(h^A+h^B)) - \frac{(h^A)^2}{2} - y_z \tag{5.61}$$

Equation 5.61 can be understood as follows:

- it is identical to what Bridget maximized in the permit case, namely equation 5.58, and
- identical to what the Impartial Spectator maximized, namely, equation 5.33 minus the constant y_z.

In both the permit and the employment cases, the outcome is Pareto efficient, but Abdul gains only an amount equal to his disutility of fishing time plus the opportunity cost of his fishing at all. The allocation proposed by the Impartial Spectator and that implemented by Bridget as owner of the lake does not differ in the fishing times of each, or the degree of exploitation of the fishing stock. In this sense private ownership of the lake has addressed the Pareto-inefficiency of the over-exploitation of the lake as a common property resource.

The only difference is that in the private ownership case there is transfer of rents (amounting to 88 pounds of fish in both cases) from the non owner to the owner:

- In the permit case the transfer took the form of the fee for the permit to fish (88) that Abdul paid to Bridget.
- In the employment case the transfer occurred because Bridget owned all of the fish that Abdul caught (200), 88 pounds of which she retained for her own consumption after paying him the wage (112).

This is general feature of social coordination problems. When one actor is sufficiently powerful to maximize their utility subject to the participation constraint of other actors, a Pareto-efficient allocation will result, and the powerful actor will get all of the economic rents. We have already seen this pattern in the TIOLI power scenario in Chapter 4.

Checkpoint 5.11: Wages vs. Permits

Refer to M-Note **??**. Assume the same values for the parameters of α and β.

a. What wage would Bridget pay Abdul if Abdul's fallback was zero?
b. What wage would Bridget pay Abdul if Abdul's fallback was to have a job working as an administrator for a utility of $u_z^A = 100$ rather than working for Bridget?

5.13 Community: Repeated interactions and altruism

Here and in previous chapters we have used two-person games to represent economic interactions among a very large number of people. But some of our interactions are with small numbers of people, for example, in our neighborhoods, families and workplaces, and even, in some cases, in exploiting a local common property resource like a forest or fishery.

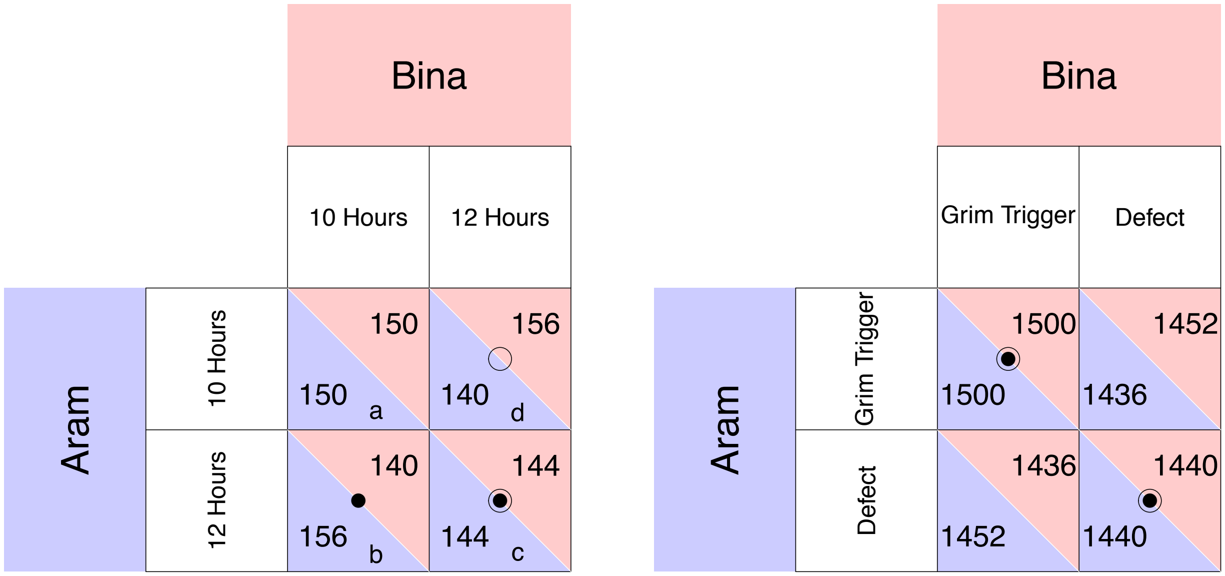
These small communities often address coordination problems in ways not possible when the number of people interacting is very large. This is possible because members of small communities:

- often have *information* about each other that is not available to governments or private owners who are not part of the community;
- *interact repeatedly* with each other repeatedly so that there are opportunities to retaliate against members who violate social norms or informal agreements; and
- often care about each other, and these *social preferences* can reduce conflicts of interest (as we saw in the previous chapter) and can provide the basis for addressing coordination problems.

These characteristics of small communities give them capabilities in solving coordination problems that are unavailable to purely government- or market-based approaches. As we have seen in Chapter 2, public goods experiments show that people are willing to punish fellow group members whose behaviors violate norms, even when inflicting the punishment is costly to the punisher. Lets see how a small community of fishing people – illustrated by Bridget and Abdul – might address the over-exploitation of the common property resource.

Repeated interactions

In the one-shot games we have introduced so far the strategies available to the players are limited: select some amount of fishing hours. One way to make the game more realistic is to let the interaction be repeated over possibly many periods with the same players. Then more complicated strategies are possible, even if in every period there are just two actions one can take, for example, fish ten hours or fish twelve hours. Importantly, strategies can now be conditional on what the other player has done in previous play.



(a) Stage game

(b) Repeated Interaction

Figure 5.15: **Repeated interactions can convert a one-shot Prisoners' Dilemma into an Assurance Game allowing or coordination on a socially optimal allocation**. Panel a is the payoff matrix for the one-shot (stage) game between Bridget and Abdul, that is played once only. Inspection of the payoffs shows that it is a Prisoners' Dilemma. You can confirm using the circle and dot method introduced in Chapter 1 that each player fishing 12 hours is a Nash equilibrium (the circles and dots show that this is also a dominant strategy equilibrium. Panel b gives expected payoffs for the game, if at the end of each period with probability $P = 0.9$ the game is played again (with the same payoffs per period as shown in panel a). In this case the circles and dots indicate that the repeated game has two Nash equilibria: the Nash equilibrium of the one-shot game with payoffs to each of 1440, and the socially optimal allocation with payoffs 1500.

One strategy might be play the strategy that would implement the social optimum (fish 10 hours) in the first round and on the next and successive rounds of the game, play whatever the other player played on the previous round. This strategy is called "nice tit for tat": nice because it begins with a strategy that could be mutually optimal, but "tit for tat" because it punishes the other player if she takes the over-fishing option.

Consider a repeated game between Abdul and Bridget with the following properties:

- *Actions*: In each period of the game, they may fish either 10 hours, the socially optimal amount or 12 hours, the over fishing level at the Nash equilibrium of the symmetric game in which they do not coordinate in any way.
- *Duration of the game*: After every period that the game is played it is continued with some probability $0 < P < 1$ We show in the Mathematics Appendix that this means that the expected duration of the game in number of periods is $\frac{1}{1-P}$.
- *Payoffs*: In each period the payoffs are given in panel a of Figure 5.15. The cell entries are from the analysis of their interaction we have carried out so far, with the parameters used in our numerical examples. Payoffs for the game are the sum of payoffs for each period the game is played.
- *Strategies*. Each may choose either to fish 12 hours in every period of the game (called "Defect") or fish 10 hours in the first period of the game and every subsequent period until the other fishes 12 hours, in which case fish

12 hours as long as the game lasts. This strategy is called "Grim trigger": the term *trigger* is used because on act of defection by the other sets of a a punishing defection by the actor, it is *grim* because the defections go on as long as the game lasts.

How will this repeated game be played? We will assume that both players are entirely self regarding. You can see that the single period payoff matrix – called the stage game – has the structure of a prisoners' dilemma. The repeated game will continue following each period with probability **P**, which for concreteness we set as $P = 0.9$. This means that the expected duration of the game is $\frac{1}{1-0.9} = 10$ periods. If the game were played among total strangers, it is unlikely that it would be repeated with such a high probability. But if the two players are neighbors or co-workers they are very likely to continue interacting.

M-CHECK We show in the appendix that the expected duration of an interaction is the inverse of the probability that at the end of a period the interaction will be terminated.

The payoffs in the repeated game are derived as follows.

Playing Defect against Defect The payoff will be the payoff (144) of the mutual defect option of the stage game (fishing 12 hours) for as long as the game continues (10 periods) so the payoff is 1440.

- *Playing Grim Trigger against Grim Trigger*: The payoff will be 150 (the stage game payoff to fishing 10 hours) for as long as the game lasts, or 1500.

- *Payoff to playing Grim Trigger against Defect*: In the first period, the Grim Trigger player fishes 10 hours, while the defector fishes 12 hours, and so receives 140 that period; and then he defects in the next period (should it occur) and until the game ends. The probability that the second period happens is P and if it does it can be expected to continue for 10 more periods so the payoffs from period two to the end of the game are $0.9 \times 10 \times 144$ or 1296 which, adding the first period's payoffs totals 1436.

- *Payoff to playing Defect against Grim Trigger*: This is calculated exactly is in the case immediately above. The defector gets 156 in the first period and then the mutual defect payoff as long as the game lasts, totalling 1452.

Looking at panel b of Figure 5.15 you can see that mutual Defect is still a Nash equilibrium in the repeated game: it is a best response to both Defect and Grim Trigger, as the circles and dots show.

But in this case the repeated game is not a prisoners' dilemma: the best response to Grim Trigger is not Defect, but Grim Trigger itself. And so if the two were to coordinate on fishing at the socially optimal level (10 hours) and had decided to play Grim Trigger they would continue doing so until the game ended.

This means the what was Prisoners' Dilemmas if played as a one shot can become an Assurance Game if played repeatedly if the game is repeated with a sufficiently high probability. What this means is that entirely self-regarding actors acting independently and without government regulation can escape the prisoners' dilemma.

You can confirm using the circle and dot method introduced in Chapter 1 that each player fishing 12 hours is a Nash equilibrium (the circles and dots show that this is also a dominant strategy equilibrium). Panel b gives expected payoffs for the game, if at the end of each period with probability P = 0.9 the game is played again (with the same payoffs per period as shown in panel a). In this case the circles and dots indicate that the repeated game has two Nash equilibria: the Nash equilibrium of the one-shot game with payoffs to each of 1440, and the socially optimal allocation with payoffs 1500.

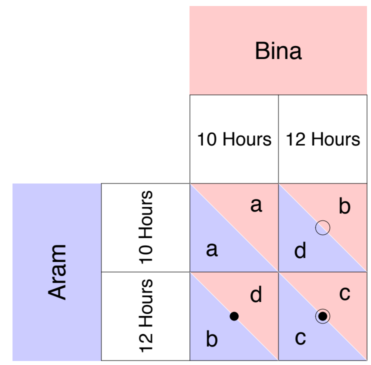


Figure 5.16: Payoffs for the one-shot Prisoners' Dilemma Game shown in Panel a of Figure 5.15 repeated here to accompany M-Note 5.13. For both players: $b > a > c > d$.

M-Note 5.13: Cooperation without agreements in a repeated game

The key to how game repetition converts a Prisoners' Dilemma stage game into an Assurance Game that can implement the socially optimal level of fishing is that Defect should not be a best response to Grim Trigger. This requires that the payoff to playing Defect against Grim Trigger should be less than the payoff to playing Grim Trigger against itself, or (using he letters in the payoff matrix of panel **a**):

$$b + c\left(\frac{P}{1-P}\right) < \frac{a}{1-P}$$

which, re-arranged to isolate the P gives us:

$$P > \frac{b-a}{b-c}$$

This means that(in the one shot game) the probability that the game will be continued after each round must be greater than the payoff advantage of defecting on a cooperator ($b - a$) relative to the payoff advantage to coordinating on 10 hours rather than 12 hours ($b - c$). This means that repeating the game is more likely to result in the Pareto superior symmetric outcome (both fishing less) if

- the incentive to exploit the cooperation of the other is less
- the joint benefit of mutual restricting fishing hours is more and if
- the interaction will be repeated with high probability

For the payoffs in 5.15 **a** this condition is satisfied for any $P > 0.5$

Social preferences: Altruism

The fact that the community of fishermen is *small* means two important facts about their context are likely to hold. First, people in small communities can more easily access information about one another and engage in repeated interactions as the basis for retaliation against those community members who defect. Second, small communities are also often the basis of the people having a concern about each others' well-being. such as altruism, fairness concerns, or reciprocity.[8]

To see how social preferences might help solve coordination failures, suppose that in choosing an action each participant puts some weight on the utility of the other as in Equation 5.62, so that Abdul and Bridget have social preferences like those we introduced in Chapters 3 and 4. Because we have used u^A and u^B to refer to the utility each gets from fishing, we now introduce v^A and v^B, which include their concern for the other's well-being. Because they are other-regarding, their evaluation of the outcomes they believe their actions will produce are based on these social preference utility functions, v^A and v^B.

$$\begin{aligned}
\text{Altruistic A:} \quad v^A(h^A, h^B) &= \text{Own utility} + \lambda \text{Other's utility} \\
&= u^A + \lambda u^B \\
\text{Altruistic B:} \quad v^B(h^A, h^B) &= u^B + \lambda u^A
\end{aligned}$$

We show in the M-Note that the best-responses maximizing these utility functions are:

$$\text{B's best response :} \qquad h^B(h^A) = \frac{(\alpha - (1+\lambda)\beta h^A)}{1+2\beta}$$

M-Note 5.14: The best-response function of altruistic people

Now, Abdul cares about Bridget. His utility function is:

$$\begin{aligned}
v^A(h^A, h^B) &= u^A + \lambda u^B \\
&= h^A(\alpha - \beta(h^A + h^B) - \frac{1}{2}(h^A)^2 + \lambda \left(h^B(\alpha - \beta(h^A + h^B)) - \frac{1}{2}(h^B)^2 \right)
\end{aligned}$$

To obtain his best response function, we differentiate Abdul's utility functions with respect to fishing time and set it equal to zero:

$$\begin{aligned}
\frac{\partial v^A}{\partial h^A} &= \alpha - 2\beta h^A - \beta h^B - h^A - \lambda \beta h^B = 0 \\
(1+2\beta)h^A &= \alpha - (1+\lambda)\beta h^B \\
h^A &= \frac{\alpha - (1+\lambda)\beta h^B}{1+2\beta}
\end{aligned}$$

Each takes account of a fraction, λ, of the external effect that they have on the other person. Concern for the well-being of other people thus might at least partially substitute for a tax or government regulation in alleviating the social coordination failure: when people care for each other they are willing to internalize the external effect of the cost they impose on other people.

What level of concern for the other would implement the Pareto-efficient outcome?

In order for altruism to implement the social welfare maximizing allocation proposed by the Impartial Spectator, these altruism based best response functions would have to to mimic those of the Impartial Spectator in Equations 5.40 and 5.41. This means that the $(1+\lambda)$ would have to take the value of

M-CHECK Notice here that $0 \leq \lambda \leq 1$ rather than $0 \leq \lambda \leq \frac{1}{2}$ in Chapters 3 and 4. This is because in those chapters the utility functions are Cobb-Douglas whereas in this chapter we are *adding up* the utility functions to display altruism (making these functions Cobb-Douglas would be very tough mathematically!).

2, meaning that we would have to have $\lambda = 1$. Each fisherman would have to be the perfect altruist that we defined in Chapter 2, caring as much for the other person as she does for her self (namely $\lambda = 1$). The difficulty of sustaining this level of altruism may suggest why most successful communities do not rely entirely on good will, but supplement it with mutual monitoring and punishment for transgression of norms.

EXAMPLE Even the most utopian, such as the contemporary Amish or Hutterite communities in the U.S., supplement altruistic action with social sanctions and monitoring of peers (which is easier to do in small communities).

Checkpoint 5.12

1. Refer to Equations 5.50 and 5.44. Why do we need $\lambda = 1$ to implement a fully Pareto-efficient outcome? Explain.
2. What would it mean if $\lambda < 1$? If people aren't only altruistic towards each other in a community, what else might they do to affect the utility others receive if they break social norms by over-fishing?
3. What would it mean if $\lambda > 1$. Would the outcome be Pareto-efficient? Why or why not?

5.14 Application: Is inequality a problem or a solution?

Recall that there two standards that we use to evaluate policies to address coordination failures:

- Does it result in a Pareto-improvement over the status quo, that is, does it improve efficiency by making at least one of the participants better off and none worse off?
- Is the resulting allocation more fair than the status quo, that is, are the rents (improvements over the status quo) that the players receive fair?

In some cases the two objectives can be jointly realized; in others they are in conflict.

The distribution of the economic rents resulting from coordination depend on the particular transformation of the game which makes coordination possible. Conflicts may arise about how best to address the coordination problems that people face: some participants may prefer an inefficient solution to the allocation problem because they get a larger share of the economic surplus at a Pareto-inefficient outcome.

Unequal solutions to local social coordination problems are generally based on the disproportionate wealth or power of one of the fishermen. It is easy to see that if one of the fishermen has a much larger net than the others and so can be assured of catching most of the fish, then his best response will approximate the allocation of a single owner of the lake. In this case, inequality in wealth among the fishermen would lessen the coordination failure.

Important inequalities may exist even among otherwise identical fishermen. To see this consider two possibilities:

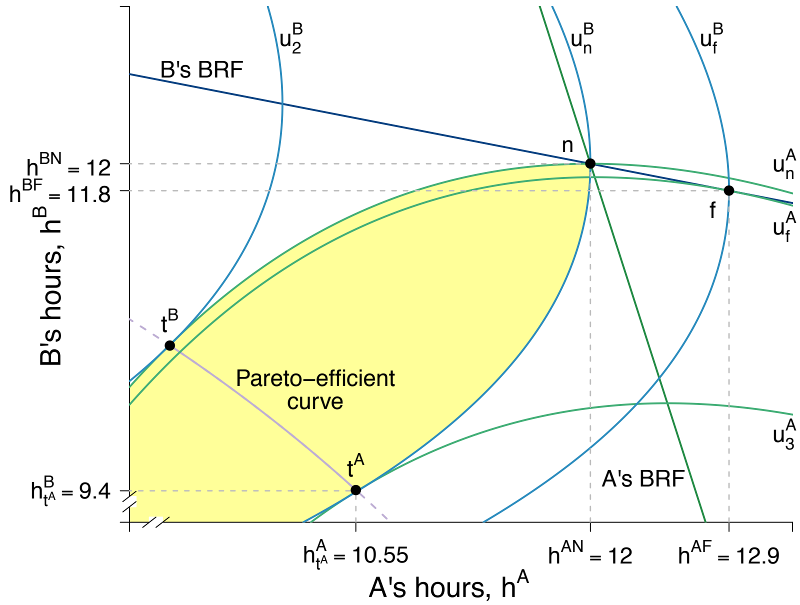


Figure 5.17: **First-mover advantage: Fishing time-setting power.** In the figure, Abdul is the leader with fishing time-setting power (first-mover power) and Bridget is the follower. Abdul doesn't have enough power to make a TIOLI offer, but he can make a credible commitment of his own fishing time such that Bridget will have to adapt in choosing her fishing time to Abdul's fishing time. Abdul, when choosing his fishing time, takes Bridget's best-response function as his *incentive compatibility constraint.* He maximizes his utility subject to satisfying her incentive compatibility constraint, finding the point at which his indifference curve is tangent to her best-response function as occurs at point **f** where Abdul exerts fishing time h_S^A and Bridget exerts fishing time h_S^B and the two fishermen obtain utilities $u_2^A = u_S^A$ and $u_1^B = u_S^B$. This Stackelberg or Leadership outcome is contrasted with the Nash equilibrium outcome of the simultaneous interaction where Bridget had higher utility $u_2^B = u_N^B$ and Abdul had lower utility $u_1^A = u_N^A$.

1. *Take-it-or-leave-it (TIOLI) power:* When one player's social position is such that they have substantially more power than the other making a TIOLI offer of both their own and the other's fishing time, then they may implement an allocation where they obtain all the economic rents and the other remains on their participation constraint. As you already know from the previous chapter, the outcome is Pareto-efficient.

2. *First-mover advantage*: When one player's social position is such that they can credibly commit to a fishing time such that the other must simply respond, they obtain more of the economic rent than the other player, but not as much as if they had TIOLI power. This is similar to the price-setting power in Chapter 4. The outcome is Pareto-inefficient.

We start with first-mover advantage: the power to commit to one's own fishing time.

First-mover advantage: fishing time-setting power

Suppose that Abdul can announce a level of fishing time and commit to it in such a way that Bridget understands that nothing she can do will alter Abdul's fishing activity. Bridget will then select her level of fishing to maximize her utility given what Abdul has committed to. In this situation, Abdul is the first-mover and has fishing time-setting power similar to the *price-setting power* (as in Chapter 4). Economists call Abdul in this situation the *Stackelberg leader*.

HISTORY Heinrich von Stackelberg (1905-1946) used this model to represent price-setting among duopolists (two firms in a market), which is why this type of model is named after him.

The big difference between Abdul having fishing time-setting power and our previous analysis is that the game is now *sequential* and the order of play matters: who gets to go first is important.

How would Abdul decide what level of fishing time to commit to as the fishing time-setter? As the first-mover, he will begin by determining what the second-mover will do in response to each of the his actions, and then select the action that maximizes his own utility given the best-response function. The second-mover's best-response function is the *incentive compatibility constraint*.

REMINDER In Chapter 4, when she had Price-setting power, A maximized her utility subject to B's price-offer curve. B's price-offer curve was A's *incentive compatibility constraint*, which is exactly what a best-response function is: a best-response function shows what action would be your *best response* to the action (e.g. price or fishing time level) the first mover commits to.

Abdul maximizing his utility subject to Bridget's incentive compatibility constraint is a simple but important change in the assumed behavior of the fishermen: Abdul now recognizes and takes advantage of the fact that by choosing various levels of fishing time he can affect the level of fishing time Bridget chooses. Abdul's behavior is *strategic* because it takes account of Bridget's reaction to his action.

In this first-mover case, Abdul is constrained not by a given level of Bridget's *utility*, but by Bridget's *maximizing behavior* as given by her best-response function. As a result, the first-mover outcome will not be Pareto-efficient because Bridget's indifference curve intersects Abdul's indifference curves at point **f** in Figure 5.17. Because the indifference curves intersect, the marginal rates of substitution are not equal and therefore the allocation is Pareto-*inefficient*. Similar to the analysis of price-setting power in Chapter 4, the fishing time-setting outcome is not Pareto-efficient, because when Abdul maximizes utility subject to Bridget's best-response function, he does not fully internalize the external effect. Abdul's first-mover advantage allows him to improve his position by comparison to the Nash equilibrium, in this case at the expense of Bridget whose outcome as second-mover is worse than the Nash equilibrium.

Take-it-or-leave-it power

Let us switch roles now and consider what would happen if Bridget had more power than Abdul. She has enough power to make Abdul a take-it-or-leave-it offer, specifying not only how much she would fish, but how much Abdul is to fish, too, along with the threat that should Abdul *not* accept the offer, then Bridget would simply fish at the level of the Nash equilibrium of the simultaneous move game.

REMINDER When either trader had TIOLI power in Chapter 4, the exercise of power led to a Pareto-efficient outcome.

Because Abdul will refuse her offer if it is worse for him than the Nash outcome, Bridget must make an offer to Abdul that satisfies Abdul's participation constraint. If she does so, the outcome will be Pareto-efficient.

Referring to Figure 5.12, we can see the outcome that Bridget would implement if she had take-it-or-leave-it power over Abdul. In the case where Abdul's fallback position is the Nash equilibrium, then his participation constraint is given by u_n^A. As a result, Bridget would find the allocation on Abdul's participation constraint at which she would maximize her utility. In Figure 5.12, Bridget's TIOLI offer is shown at point $\mathbf{t^B}$. At $\mathbf{t^B}$, her indifference curve u_2^B is tangent to Abdul's indifference curve u_n^B, so the allocation she chooses for her TIOLI offer is Pareto-efficient, unlike when Abdul was the leader and could set a fishing-time for Bridget. If Abdul had TIOLI power over Bridget, he would implement the allocation at point $\mathbf{t^A}$ and this allocation would also be Pareto-efficient.

Summing up, our model shows that the effect on unequal power will always benefit the more powerful and may, but need not, result in a Pareto-inefficient outcome. Abdul's fishing first-mover ability (his time-setting power), resulted in gains for Abdul, losses for Bridget, and increased over-exploitation of the lake. This resulted in greater inequality than the Nash equilibrium, but the outcome was not Pareto-comparable to the Nash equilibrium (one does better and the other worse at the Nash).

Positive effects on Pareto efficiency occurred when Abdul had take-it-or-leave-it power and the outcome was Pareto efficient, but probably regarded as unfair by Bridget or an Impersonal Spectator. The model is *hypothetical* but the problem is not.

Checkpoint 5.13: Numerical TIOLI power

Using the values of $\alpha = 30$ and $\beta = \frac{1}{2}$:

1. Find the take-it-or-leave-it offer that Abdul would make to Bridget at point $\mathbf{t^A}$. How many hours would Abdul fish? How many hours would Bridget fish?
2. What would Abdul's utility be at the TIOLI offer? What woudl Bridget's utility be at the TIOLI offer?

Evidence from field studies

A field experiment among forest commons users in rural Colombia underlines how inequality may be an impediment to achieving more satisfactory outcomes through coordination. Juan Camilo Cardenas implemented common pool resource behavioral experiments among villagers who rely for their living on the exploitation of a nearby forest.[9] So the subjects in the experiment were in real life playing the same game that the experimenter invited them to play.

In Cardenas's game, the subjects choose to withdraw a number of tokens from a common pool (these represented exploitation of the common property resource), and after all subjects had taken their turn the tokens remaining were multiplied by the experimenter and then distributed to the players, the tokens then being exchanged for money. (This is similar to the Public Goods Game experiment in Chapter 2 except that subjects decide how much to withdraw rather than how much to contribute to the pool). For an initial set of rounds of the game, no communication was allowed. But in the final rounds of the game, subjects were invited to converse for a few minutes before making their decisions.

Cardenas expected that communication would reduce the level of withdrawals from the common pool (as has been the case in similar experiments) despite the fact that it does not alter the material incentives of the game. Communication was indeed effective among groups of subjects with relatively similar wealth levels (measured by land, livestock and equipment ownership); their levels of cooperation increased dramatically in the communication rounds of the experiment. But this was not true of the groups in which there were substantial differences in wealth among the subjects.

In one group, one of the wealthiest subjects tried in vain to persuade his fellow participants (who in real life were his tenants and employees) to restrict their withdrawals to the socially efficient amount, in order to maximize their total payoff. But the wealthy subject's advice fell on deaf ears.

"I did not believe Don Pedro," one of the less well-off women in his group later explained, "I never look him in the face." She was right not to trust him: Don Pedro (not his real name) had withdrawn the maximal amount despite his contrary advice to the other players.

This is not an isolated example.

- A study of water management in 48 villages in the Indian state of Tamil Nadu found lower levels of cooperation in villages with high levels of inequality in landholding. Moreover, lower levels of compliance were observed where the rules governing water supply were perceived to be chosen by the village elite.

- A similar study of 54 farmer-maintained irrigation systems in the Mexican state of Guanajuato found that inequality in land holding was associated with lower levels of cooperative fishing time in the maintenance of the field canals.[11]

In other cases, inequalities based on traditional hierarchical have made a positive contribution.

- Another study of Mexican water management, for example, found that increased mobility of rural residents undermined the relationships that had been the foundation of a highly unequal but environmentally sustainable system of resource management.[12]

- And in the port of Kayar, on the Petite Côte of Senegal, a cooperative fishing time to limit the catch (to support higher prices, not to protect fishing stocks) owed its success in part to the leadership of the wealthy local

FACT CHECK In a recent study of participation in church, local service and political groups, as well as other community organizations providing local public goods by Alberto Alesina and Eliana La Ferrara found that participation in these groups was substantially higher where income is more equally distributed, even when a host of other possible influences are statistically "held constant."[10]

FACT CHECK Social differences among commons users affects outcomes in other ways. The fishing agreement in Kayar was threatened by conflicts between locals and outsiders using differing technologies, and other attempts to limit fishing failed due to the indebtedness of fishermen to fish sellers (who opposed the limits) and because the wives of many of the fishermen were fish sellers.

traditional elite of elders.[13]

5.15 Over-exploitation of a non-excludable resource

We stated at the beginning of the chapter that we would illustrate the problem of the common pool resource problem by the example of just two people. This was despite the fact that, as a non-excludable resource, there would be *no limit* on the number of people who could, if they wished, fish on the lake and compete with Abdul and Bridget for the available stock.

We have so far studied just one of the two aspects of the coordination failure resulting in over-fishing: the fact both Abdul and Bridget fished more hours than was Pareto-efficient. They both could have been *better off* had they been able to agree to fish *less*.

Now we introduce the second aspect of the problem: many more people could fish the lake. Because the lake is a common pool resource, its non-excludablity property of a means that there is open access.

How many people would use the lake? To answer the question we add more context to the initial problem. Abdul and Bridget are part of a large community of people who may make their livelihood fishing on the lake, or if not that, then doing some other kind of work yielding a utility of u_z (their fallback option). People will decide to make their living fishing as long as the utility they gain exceeds u_z.

In Figure 5.18 using the same values for the parameters as in the other numerical examples in this chapter, we calculate the maximum utility that could be attained by one person fishing alone, two (as in the case of Bridget and Abdul), three, and so on up to 11. However many there are fishing, they will receive the same utility in the Nash equilibrium because they are identical, and we have so far assumed that none has any advantage in bargaining with the others. The height of each bar is the utility attained.

When there are just two people fishing as in our previous examples involving Abdul and Bridget each receives a utility of 144, as we found in M-Note 5.6. The more people that fish in the lake the lower the utilities each of them receive will be. When there are ten they all have a utility of 21.3, barely greater than their fallback options.

Now think about some other member of the community who is *not* currently fishing but is thinking of doing so. Those fishing are doing better than the fallback options. But if the 11th person decided to fish they would *all* receive a utility of 18.4 (including the new fisherman). That is, they would all receive *less* than their fallback options. So the eleventh person would decide *not* to fish.

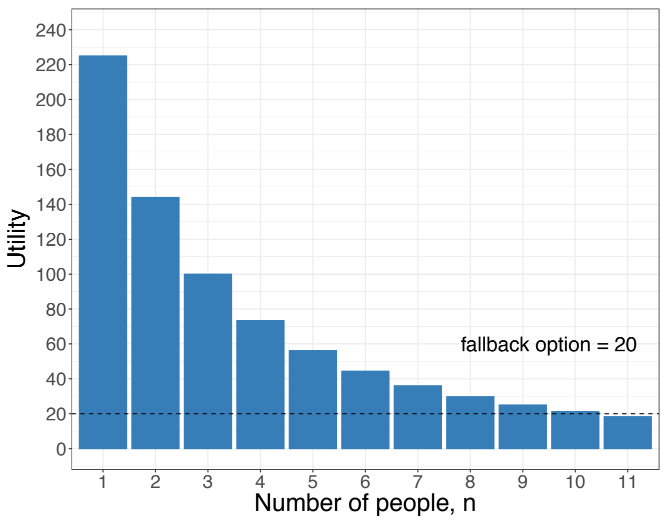


Figure 5.18: **The dynamics of over exploitation of a common property (non-excludable) resource.** All of the people who might fish on the lake have the same utility functions as Abdul and Bridget with the values of $\alpha = 30$ and $\beta = \frac{1}{2}$. The height of the bar for a given number on the x axis is the utility of each of the fishermen when there are the indicated number fishing on the lake. The fallback utility is $u_z = 20$. You can see from the figure that if the lake is a common property resource, so that no fisher can be excluded, the Nash equilibrium number fishing on the lake is 10 with each receiving a utility of 21.3. If the 11th person fished on the lake, she – and all of the rest of the fishermen there – would receive a utility of 18.37, that is, less than their fallback option). The mathematics on which this figure is based are shown in M-Note 5.15.

Generalizing from this example, the Nash equilibrium number of people fishing is the largest whole number of people fishing such that the utility that those fishing receive is greater than or equal to the utility they *would* receive at their fallback option.

MATH-CHECK The equilibrium number of people fishing must be a whole number because the entry of a "fractional fisherman" would not make much sense unless we allowed people to split their day between fishing and the fallback option, which we do not.

As a result, the Nash equilibrium of this game is that we have:

- $n^N = 10$ the number of people fishing and
- $h^N = 4.62$ the number of hours each of them works.

We use the N superscript for each of these quantities because *both* are Nash equilibria (but under different rules of the game):

- $n^N = 10$ is a mutual best response because none of those fishing could do better by *not* fishing, and none of those *not fishing* could do better by fishing, and
- given that ten people are fishing, then $h^N = 4.62$ is also a mutual best response because for each person fishing this is a utility-maximizing choice of hours, given the hours that everyone else is fishing.

The Nash equilibrium is Pareto-inefficient for two reasons: too many people are fishing too many hours each. Just as was the case with Abdul and Bridget, if each fished a little less they all would be better off.

And if fewer of them fished, all ten of them could be better off. Figure 5.18 shows that if 3 people fished they would each have a utility of 100. We call people who are already doing an activity, such as fishing or owning a firm, *incumbents*. We therefore call the existing fishermen, the *incumbents* or *incumbent fishermen*. Suppose this was the case, and that the incumbents could somehow agree to bribe the other 7 *not* to fish. Notice we have just changed the rules of the game to allow the incumbent three to *coordinate*.

The incumbent three would have to give the other 7 potential fishermen an amount of fish sufficient that each would be as well off as their fallback. This amount would be $u^N - u_z$ or $21.3 - 20$, or 1.3 each. The total payments by the three to the other seven would be $7 \times 1.3 = 9.1$, leaving each of the three better off (each receiving $300 - 9.1)/3 = 97$). If the incumbent three *increased* the 'bribe' just a little bit then all 10 would be better of than at the Nash equilibrium with open access.

M-Note 5.15: Nash equilibrium number of people fishing

Because access to the lake is open to all, the number fishing there will be the largest whole number (because we cannot have fractions of people fishing) such that the utility of those fishing is equal to than the fallback option (their utility if they are not fishing in the lake), which is $u_z = 20$. To determine this number, we first derive $h^N(n)$ the hours of fishing that each will do as a function of the numbers fishing, and use this result to determine $u^N(h^N(n))$ the utility of those fishing as a function of how many there are.

To determine $h^n(n)$ we study the utility maximization problem of person 1:

$$\begin{aligned}\max_{h_1} u_1 &= h_1(\alpha - \beta \sum_{i=1}^{n} h_i) - \frac{1}{2}h_1^2 \\ &= h_1(\alpha - \beta h_1 - \beta \sum_{i=2}^{n} h_i) - \frac{1}{2}h_1^2 \end{aligned} \tag{5.62}$$

To find the hours of fishing that maximizes the utility of person 1 we differentiate equation 5.62 with respect to h_1, and set the result equal to zero. This gives us the first order condition:

$$\begin{aligned}\alpha - 2\beta h_1 - \beta \sum_{i=2}^{n} h_i - h_1 &= 0 \\ \text{marginal benefit} = \alpha - 2\beta h_1 - \beta \sum_{i=2}^{n} h_i &= h_1 = \text{marginal cost}\end{aligned} \tag{5.63}$$

Rearranging Equation 5.63 we get person 1's first order condition giving the utility maximizing amount of fishing time:

$$\begin{aligned}(1+2\beta)h_1 &= \alpha - \beta \sum_{i=2}^{n} h_i \\ h_1 &= \frac{\alpha - \beta \sum_{i=2}^{n} h_i}{1+2\beta}\end{aligned} \tag{5.64}$$

All face the same first order condition so in the Nash equilibrium all fish the same amount of hours: $h_1 = h_2 = \ldots = h^N$. Equation 5.64 becomes:

$$\begin{aligned}h^N &= \frac{\alpha - \beta(n-1)h^N}{1+2\beta} \\ (1+2\beta)h^N &= \alpha - \beta(n-1)h^N \\ h^N + 2\beta h^N + \beta n h^N - \beta h^N &= \alpha \\ (1+\beta+\beta n)h^N &= \alpha \\ h^N &= \frac{\alpha}{1+\beta+\beta n}\end{aligned}$$

As in the rest of the chapter, we let $\alpha = 30$ and $\beta = \frac{1}{2}$. You can verify that, if $n = 10$, then

$1+\beta+\beta n = 6.5$ and so $h^N(10) = \frac{30}{6.5} = 4.61$. The utility of each fisher would be:

$$\begin{aligned} u^N(10) &= 4.61 \times (30 - \frac{1}{2} \times 10 \times 4.61) - \frac{1}{2} \times 4.61^2 \\ &= 21.30 \end{aligned} \tag{5.65}$$

If one more enters takes up fishing so that $n = 11$, then the hours of fishing would be $h^N(11) = \frac{30}{7} = 4.29$.The new utility of each fisher would be now:

$$\begin{aligned} u^N(11) &= 4.29 \times (30 - \frac{1}{2} \times 11 \times 4.29) - \frac{1}{2} \times 4.29^2 \\ &= 18.37 \end{aligned} \tag{5.66}$$

But this ($n = 11$) cannot be a Nash equilibrium, because everyone – including the new entrant – would then be worse off than with the fallback option, $u_z = 20$. So the 11th person would not enter (or if she did, others would leave). So the Nash equilibrium is $n^N = 10$. This is illustrated in Figure 5.18.

5.16 The rules of the game matter: Alternatives to over-exploitation

The new rules of the game allowing the incumbent three to bribe the others is just a thought experiment demonstrating that the Nash equilibrium with open access is *not* Pareto efficient. But commonly observed real-life rules of the game – like inequalities in bargaining power, cooperative management of the lake, or private ownership – could also address the over-fishing problem.

Checkpoint 5.14: Pareto efficiency and open access

a. Explain why open-access Nash equilibrium outcome with 10 fishermen is *not* Pareto efficient. What alternative, if any, is Pareto superior to it?
b. Given your reasoning for a., do you think there are alternative outcomes that are Pareto superior to, say, three fishermen bribing the other 10 not to fish? Explain what the dynamics for the situations you describe would be? How many fishermen? How many hours spent fishing? And so on.

TIOLI bargaining power

To see that the institutions governing the interactions among them matter, think about the case in which one of the ten people fishing on the lake has the power to make a take-it-or-leave-it offer to the rest. Here we have changed the rules of the game by giving one of the fishermen TIOLI power (which allows a kind of coordination). But the lake is still open access, so there are ten fishermen there.

The one with bargaining power – suppose it is Abdul – can now say to the others "each of you will fish x number of hours, and I will fish as many hours as I wish." This is the “take it" part of the offer. The "leave it"" part is: "and if you refuse, then I will return to fishing 4.61 hours." That is, return to the former Nash equilibrium hours that occurred when there was no coordination among the fishermen.

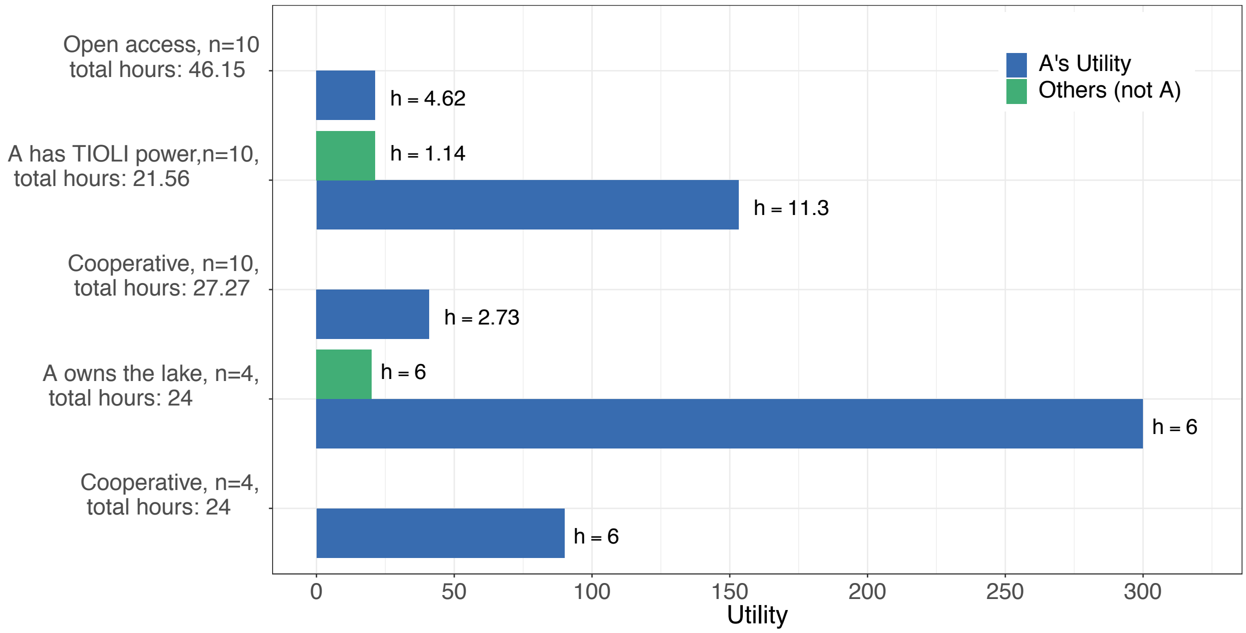


Figure 5.19: **The rules of the game: Non-cooperation, bargaining power, and ownership**. The bars show the utility of the fishermen (it is identical for all fishermen in the first and third row). The numbers at the end of the bars show the hours fished, where h_A is the fishing hours of the owner or person who has TIOLI power, h_{-A} is the fishing hours of the non-owner or people who do not have TIOLI power.

The other fishermen would know that without coordination the best they could do is to all fish 4.61 hours, gaining a utility of 21.3. This is the others' fallback option to the TIOLI offer. If accepting Abdul's offer made them worse off than their fallback they would refuse, and just fish 4.61 hours. This is their participation constraint; if it is violated – so that they would receive a utility of less than 21.3 – the others will not accept ("participate in") Abdul's offer.

Abdul would know, therefore, that he needs to find the hours of all 10 of them (his and the rest) that maximizes his utility subject to the participation constraint on the minimum utility the others can receive. Figure 5.19 shows the offer Abdul would make, and the utility that he and the others would experience. The first row of the figure shows, from the previous figure, the result for the unlimited access case *without* coordination.

When Abdul has TIOLI power the other fishermen work fewer hours (just 1.14 each, rather than 4.62 before), but get exactly what they had under the uncoordinated open access case. This is so because that level of utility – 21.3 – is the participation constraint on what Abdul can offer them.

Abdul himself works 11.30 hours and enjoys utility equal to 153.2. Notice that the total number of hours is reduced sharply compared to the uncoordinated Nash equilibrium: from 46.15 to 21.56 hours. This reduction in total hours is the reason why the others are able to fish less but still attain the same utility: they catch more fish in an hour due to the lesser total hours of fishing.

A democratic fishing cooperative

An entirely different set of rules of the game – a democratic cooperative of the fishermen – would implement a correspondingly different set of results. Suppose that *none* of the ten fishermen has any bargaining power advantage and that they jointly own the lake. They can decide jointly – democratically by unanimous consent – on the same number of hours that each of them will fish.

To figure this out they would think in the same way the Impartial Spectator did when she maximized total utility. They will maximize the *sum* of their utilities because this will also maximize the utility of each fisherman. The result is shown on line 3 of Figure 5.19. When there are 10 fishermen, they would each fish 2.73 hours and attain utilities of 40.9 each. Because their utility as coop members is now double their fallback option, others who are experiencing the fallback utility of $u_z = 20$ would wish to join the cooperative. But it might be difficult to persuade the members to admit others, as this would reduce the utility of the incumbent fishermen.

COOPERATIVE A cooperative is a business organization or other association whose members together own the assets of the organization; they share the income resulting from their activities and jointly determine how the organization will be run (possibly through the democratic election of a manager).

Their total fishing hours (27.3) is substantially greater than under the TIOLI power of Abdul, and so is their total utility (409.1 compared to 344.9).

We can conclude two things from this last fact:

- Suppose Abdul still had TIOLI power. If the other fishermen could co-ordinate their actions, they could 'bribe' Abdul to give up his bargaining power and join their cooperative; they could have offered him the 153.2 that he received under his TIOLI power and still be better off dividing up the rest of their utility (fish) amongst themselves. They would each receive $(409.1 - 153.2)/9 = 28.4$, far better than the 21.3 they had when Abdul had bargaining power. So shifting from Abdul holding TIOLI power to the democratic cooperative is a *Pareto improvement*.
- The reason why this is the case is that under Abdul's TIOLI power they were as a group *under*-exploiting the fishing stocks. Abdul forced them to do this because the less the other people fished the more fish he could catch, and that was the only way he could increase his utility.

The reason why the cooperative's decision results in a greater total utility than the TIOLI case is that the members of the cooperative were pre-committed to *sharing* the total utility equally. And so they each had an interest in making total utility be as large as possible.

Things would have been very different if Abdul had had the power to *take* some of the fish caught by the others (as in the employment and fee cases we dealt with earlier). In this case he would have done the same as the co-operative. He would have implemented the fishing times that maximized total utility. And then he would have taken fish from the others, leaving them just enough fish so that they did not decide to stop fishing. The TIOLI allocation was inefficient because Abdul's bargaining power was limited.

The TIOLI case was not inefficient because Abdul had some bargaining power. It was inefficient because he did not have enough power. As you will suspect from the 2 person case studied earlier, the allocation would have been Pareto efficient if Abdul had had all of the powers of a private owner of the lake.

Private ownership

Under these new rules – private ownership of the lake – the lake is no longer a common property resource because ownership means that Abdul can exclude anyone he wishes from fishing. Abdul would make three decisions:

- How many other fishermen should I *allow* to fish in the lake? That is, what is the total-utility-maximizing number of people who should fish the lake?
- How many *hours* should I allow them to fish?
- If I *employ* them, then what *wage* should I pay them? Or if I charge them a fee for fishing, how large should the fee be?

You know how to answer the second and third questions from the case earlier in the chapter when Abdul was the owner with just one other person Bridget on the lake.

The first question is similar to that asked in Figure 5.18 but the answer is very different. The number of people fishing the lake is no longer based on the fishermen's own decisions about where they can make a better livelihood. This is not their decision to make. The *owner* determines the number of others so as to maximize his utility.

How he would do this is shown on line 4 in Figure 5.20. Abdul will allow three other fishermen to access the lake (so that means $n = 4$ including himself). Going back to Figure 5.19 remember, there are now just four fishermen, *not* 10 as before. All four fishermen fish 6.0 hours with the owner receiving a utility of 300 and each of the others receiving the same utility as their fallback option, that is, 20.

The last line in Figure 5.19 shows what happens if Abdul is not the owner and instead if all four of the fishermen were members of a democratic cooperative. The members of the cooperative would implement exactly the same allocation of work time as occurred under private ownership: 6 hours of work time each. But the distribution of utility would be radically different, each of the four would receive 90.

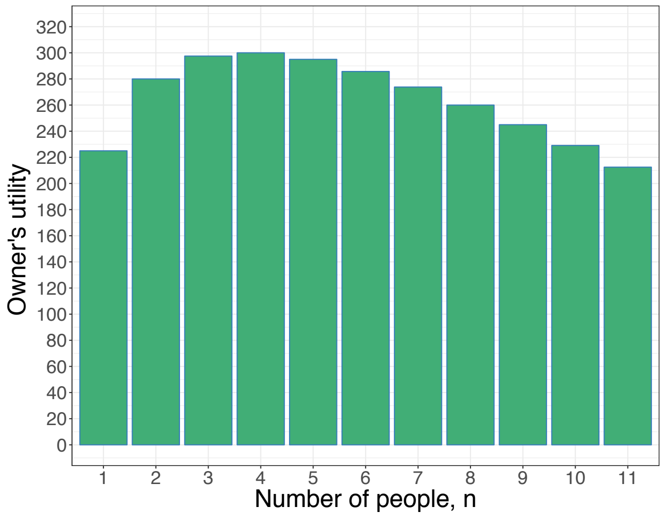


Figure 5.20: **Utility of the owner when the lake is privately owned.** On the horizontal axis are the total number of people fishing in the lake, including the owner. So, for example, where $n = 2$ we have Abdul as the owner and there is one other person, Bridget, the case we analysed earlier in this chapter. The height of each bar is the utility gained by the owner of the lake when he can both determine how any people fish there and dictate the terms under which they will work (as long as they receive utilities superior to their fallback position).

We know that each working 6 hours is the allocation that maximizes total utility. The reason why private ownership of the lake implements this outcome is that the owner is limited only by the participation constraints of the others, and this is a constant (their fallback position of 20). So he implements an allocation to maximize the total utility, from which he must subtract the amounts required to keep the three others "participating."

In sum, we can say the following:

- Open access leads to a Pareto inefficient over-fishing outcome in which all the fishermen receive the same utility.
- TIOLI power implements to a highly unequal and Pareto inefficient under-exploitation of the lake.
- Private ownership implements a Pareto-efficient and highly unequal outcome.
- A democratic cooperative implements a Pareto-efficient and equal outcome.

5.17 Conclusion

In practice, none of the approaches to addressing the common property resource coordination problem could be expected to work perfectly as:

- no government is likely to have the information about the people's preferences, production functions, and fishing times necessary to implement Pareto-efficient fishing levels by fiat, or to design the optimal taxes that would achieve the same result.
- private owners face some of the same problems due to lack of information

and moreover, while private ownership of a small lake or other common property resource is conceivable, for may important common property resources is infeasible (owning the oceans or the atmosphere for example) or undesirable (think of the unaccountable power that a private owner of such a vast common property resource could wield.)

- altruism towards close family and loved ones will lead us to take at least some account of the effects of our actions on their well-being, but we are less likely to know or care deeply care about how our actions affect total strangers or even yet unborn generations who will benefit or suffer the external effects of what we do.

The conclusion is *not* that the approaches to addressing coordination problems introduced here are ineffective. Each approach can contribute to making economic outcomes more efficient and fair.

The models we introduced simplify the rules of the game that regulate how we interact with each other in exploiting a common property resource, in contrast to the vast diversity and complexity of rules that we observe. What the models have done is *not* to represent the world as it is, but to identify key aspects of *how the world works* to provide a lens for understanding them better.

FACT CHECK Elinor Ostrom and her colleagues' field research in different parts of the world from Colombia to Switzerland uncovered twenty-seven different local rules for excluding others from access to common property resources. These were based on such things as residency, age, caste, clan, skill level, continued use of the resource, use of a particular technology, and so on.[14]

Making connections

Social interactions and external effects: The interactions that economists study include buying and selling in markets, but they also include non-market interactions, sometimes called 'social interactions' ("social" here means simply non-market). The social interactions studied here include an external effect: the fact that one person's fishing reduces the catch of another, and this effect is not taken into account when each of the fishermen decide how many hours to fish.

Public goods, common pool resources and club goods: All of these things have the property such that each person's actions have *external effects* on others and in the absence of social preferences or policies that internalize the external effects, these are not taken into account when people decide how to act, resulting in outcomes in which some potential mutual gains remain unexploited.

Policy: Government policies and institutions may be designed so that people take account of *external effects* when they act. An example, is a *tax* on fishing that imposes on each fisher the marginal costs that their fishing imposes on the other fisherman, inducing each to choose their hours of fishing as if they cared about these external effects as if it as the actor herself rather than the other who bore the costs.

Property: Converting a *common property resource* into a *privately owned* *resource* may result in a *Pareto-efficient outcome* in which the owner captures all of the potential *mutual gains (rents)*.

Power: When a single person has all of the bargaining power and so can make a binding *take-it-or-leave-it-offer*, he or she implements an outcome in which there are no unrealized mutual gains, and all of these gains (*rents*) go to the powerful person. Lesser forms of power – to commit to a particular fishing time, to which the other must respond, for example – advantage the powerful and can result in inefficient outcomes.

Mutual benefits from coordination and conflicts over their distribution Policies to address coordination failures differ in how the resulting rents are distributed; the resulting conflicts may make it difficult to agree on any policy.

Inequality: Differences in wealth, political connections and other sources of power can be both a source and a consequence of inefficient and unfair outcomes among people facing coordination problems. In some cases, these differences can also mitigate the inefficiencies arising from coordination failures.

Models and relevance: Models, we wrote in Chapter 3, are like maps – a simplified guide to the territory, not the territory itself. But the model of social interactions introduced here, though quite abstract can be directly applied to very concrete economic actions such as firms competing for customers and, suitably extended, can illuminate *global social interactions* and *coordination problems* such as climate change and the spread of epidemic diseases.

Important Ideas

utility	disutility	external effect
common property resource problem	private property	coordination failure
rivalness	excludability	interdependence
Impartial Spectator	altruism	reciprocity
symmetrical interactions	asymmetrical interactions	social identification
TIOLI power	time-setting power	stackelberg leader
fiat power	government policy	decentralized implementation
tax	employment	wage
permit	participation constraint	fallback
incentive compatibility constraint	best-response function	binding participation constraint
social preferences solution		

Mathematical Notation

Notation	Definition
h	fishing times
α	parameter regulating the productivity of fishing times
β	external effect of fishing time on the other's productivity
$u()$	utility function
$v()$	value function expressing an altruistic concern for the utility of another person
W	Impartial Spectator's social welfare function
w	wage in the employment solution
F	permit fee in the permit solution
τ	per unit tax in the government policy solution
a,b,c,d	payoffs in the repeated interactions game
λ	extent of altruism (valuation of the other's utility relative to one's own)

Note on super- and subscripts: A and B: people; N: Nash Equilibria; i: Pareto-efficient outcome; F: outcome with a first mover.

Discussion questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

See Reference List.

Part II

Markets for Goods and Services

> Enter the Royal [Stock] Exchange of London, that place more respectable than many a court; you will see there agents from all nations assembled for the utility of mankind. There the Jew, the Mohammedan, the Christian deal with one another as if they were of the same religion. There the Presbyterian confides in the Anabaptist, and the Churchman depends on the Quaker's word. ... They give the name infidel only to those who go bankrupt.

Voltaire, 1734, *Lettres philosophiques*, Melanges (Paris, 1961) pp 17-18

When you hear the word "market" what other word do you think of? "Competition" probably is what came to mind. And you would be right to associate the two words.

But you might have also come up with "cooperation". That is what impressed Voltaire about the London stock market: mutually advantageous interactions, even among total strangers "from all nations assembled for the utility of mankind." Markets allow us, each pursuing our private objectives, to work together producing and distributing goods and services in a way that, while far from perfect, is in many cases better than the alternatives. Markets accomplish an extraordinary result: unintended cooperation on a global scale, although often with a highly unequal distribution of the benefits.

HISTORY The Israeli historian Yuval Noah Harari explains why it is our capacity to cooperate in flexible ways with large numbers of other humans that makes us unique among all the animals. https://tinyurl.com/y3bpy4px

To better understand what markets do and how they work, begin with two workaday facts: We acquire skills as we produce things and,for this and other reasons, producing a lot of the same thing is often more effective in terms of time and other inputs per unit than producing just one or a few of many different things. This is called learning by doing.

Because of learning by doing and other advantages of large scale production people do not typically produce the full range of goods and services on which they live. Instead we specialize, some producing one good, others producing other goods, some working as welders others as mothers, teachers or farmers.

There are huge advantages to this pattern of specialization – called the division of labor. Those who are naturally better at some task, or have learned to be good at it by experience, or are in an environment in which it can be most productively done can devote themselves entirely to what they are relatively good at.

This is part of the explanation of why as a species we are so productive. The limited number of species that have adopted a highly developed division of labor – humans, ants and other social insects, for example – have out competed other species. The total biomass of humans and the livestock we have domesticated, for example, is estimated to be 23 times the weight of all the other mammals on earth. And throughout most of human history the biomass of ants – one of the most cooperative of species – has exceed that of humans by a considerable amount.

But the division of labor poses a problem for society: once they are produced by specialized labor, how are the goods and services to be distributed from the producer to the final user. In the course of history this has happened in a number of distinct ways from direct government requisitioning and distribution as was done in the U.S. and many economies during the Second World War, to gifts and voluntary sharing as we do in families today and was practiced among even unrelated members of a community by our hunting and gathering ancestors.

In a modern capitalist economy, the institutions that govern how the goods and services are distributed from producer to user include markets, firms, families and governments. In this second part of our book we study markets and the actors who make up markets: the owners (and managers) of firms, and other individuals (and the families of which they are a part).

To understand how markets facilitate specialization in chapter 6 we study the production process and how the division of labor and the exchange of products can be advantageous to all concerned. Then to understand the workings of markets we explain how individuals' valuation of goods and services is expressed in market demands (Chapter 7). Then, along with these market demands, we explain how firms' costs of production are expressed in their owners' and managers' decisions about how much to produce and supply to the market (Chapter 8).

We then study the process of competition among sellers and buyers, each seeking to enlarge their share of the mutual gains made possible from the division of labor and exchange. And we show how this so called rent-seeking process affects the movement of prices and the quantities produced (Chapter 9).

Taking these four chapters as a whole poses a tension that can be expressed by the following contradiction:

- The models and evidence on the advantages of large scale production provide a reason why we specialize.
- Competitive markets are essential to the process of specialization can be organized in ways that allow the mutual benefits of the division of labor to be widely shared, as Voltaire said "for the utility of mankind."
- But the advantages of large scale production can also promote the emergence of giant firms and a winner take all process that appears to be making markets less competitive.

Making market competition sustainable given the advantages of large scale production will have to be addressed by public policy.

6

Production: Technology and Specialization

The division of labor is limited by the extent of the market.

Adam Smith *The Wealth of Nations*

A technician glances quickly from one to the other of her three monitors and around the huge room many other technicians do the same. Occasionally a technician looks up at gigantic blue video screens on which news reports flash, international weather reports display, and flight conditions stream live. Twenty-four hours a day, translators stand ready to facilitate conversations in 28 languages.

What is this command center?

It's the Production Integration Center that coordinates the global production of the Boeing 787 Dreamliner four stories above the production floor at the company's plant in Everett, Washington, USA. There the super-jumbo airplanes are being assembled from components being flown in from around the world: parts of the wing from Japan, wing tips from Korea, the center fuselage and the horizontal stabilizer from Italy, passenger doors from France, cargo doors from Sweden, landing gear from the U.K., and the list goes on and on. Figure 6.2 shows where the components of the Dreamliner are produced. In 2015, Boeing contracted with over twenty-six thousand suppliers around the world.[1]

Boeing selected Rolls Royce, Mitsubishi, Saab, Fuji and other companies to design and build the components because they were – in Boeing's estimation – simply the best companies to do the job, anywhere in the world. The Japanese companies were global leaders in aircraft construction. The Italian partner Alena had critical intellectual property rights (patents) that Boeing would otherwise not have had access to.

Boeing modified four 747-400 aircraft – renaming them "Dreamlifters" – to deliver the wings, body, and other parts of the plane to Everett where American

DOING ECONOMICS

This chapter will enable you to do the following:

- Explain how learning-by-doing and economies of scale are reasons for the division of labor and specialization.
- Understand how markets allow specialization according to the principle of comparative advantage.
- See how in the presence of economies of scale and learning by doing an economy can benefit by specializing, but also may specialize in ways that perpetuate is low income as a result of a poverty trap similar to coordination failures studied earlier.
- Manipulate some commonly used production functions to study marginal and average products of the labor and capital goods and derive a production possibilities frontier.
- Describe the main dimensions on which technologies differ – the extent of substitution among inputs, productivity, factor intensity and economies of scale – and how these are represented in different production functions.
- Understand how a the owners of a firm can determine a set of inputs and a way of combining them to produce output (a technique of production) that will minimize the costs of a given level of output.
- See that owners of a firm will try to innovate to reduce the inputs required to produce a given output and therefore low costs and receive innovation rents (at least until the competition catches up).

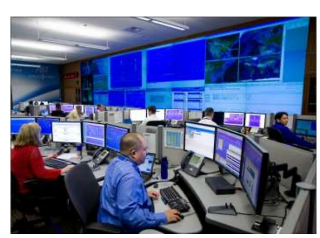


Figure 6.1: **Boeing's Production Integration Center in Everett, Washington, USA**.

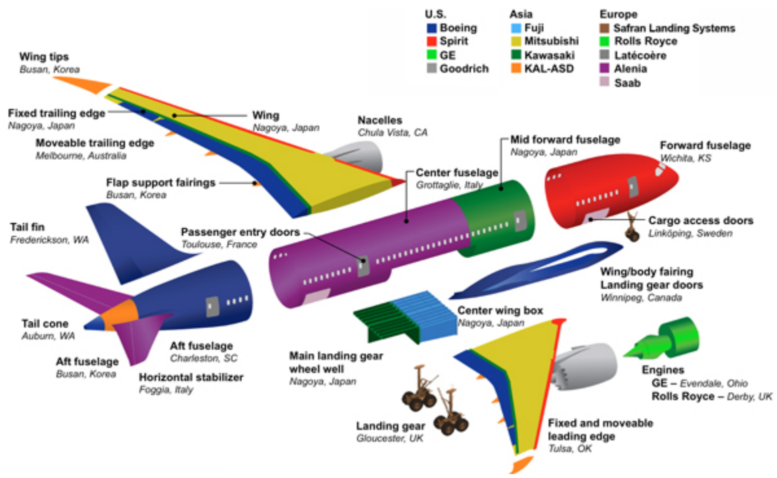


Figure 6.2: **The different parts of the Boeing 787 aircraft built by different specialized partners from all around the world.**

machinists and others assembled the planes. Four stories above them the engineers at the Production Integration Center kept minute-by-minute track of the movement of the components around the world.

6.1 The division of labor, specialization and the market

Boeing's globally integrated Dreamliner production process illustrates an important economic idea: the **division of labor**. The division of labor is an expression for the fact that people, organizations, or geographical regions specialize in particular tasks or the production of a limited range of goods or services. For Boeing, purchasing components of the Dreamliner from hundreds of other specialized firms was more cost effective than producing the entire plane in-house at their plant in Everettt, WA.

There are two consequences of specialization.

The first is *increased productivity*. The specialization allowed by the division of labor increases productivity for three reasons:

- *Comparative advantage:* specialization enables people, firms, and regions to focus on the tasks and products that they are comparatively good at (we will take up comparative advantage below).
- *Learning by doing*: People learn better ways of working both through the developing individual skills and discovering better ways to organize production among members of a team. Figure 8.2 b provides a dramatic example of learning by doing.
- *Economies of scale*: By allowing the production of a few things on a large scale rather than many things on a small scale, specialization raises the amount of output that is available for a given amount of inputs. Figure

8.6 presents some physical evidence for economies of scale based on engineering studies.

A second consequence of specialization is the *need for integration*. The advantages of the division of labor can only be realized if there are institutions to coordinate the many distinct production activities that take place when people specialize. The Boeing example illustrates the need for integration: somebody has to put the parts together to produce a Dreamliner.

This can be summarized: production is specialized, but the use of goods and services is generalized. Specializing in consumption is not biologically sustainable. As a result, the goods and services, somehow have to get from specialist producers to generalist users.

To grasp the scope of this problem imagine a 3D map of the world showing the stocks of goods produced annually in each location. There would be a hundred or so Dreamliners piled up in Everett Washington, billions of square meters of cloth stacked in Bangladesh and other textile producing countries, well over a billion barrels of oil piled over tiny Kuwait, mountains of computer components and other consumer electronics rising from coastal China were Dell is located, and so on.

Now imagine the same map, but showing where all of the goods are used. The second map would be different from the first in two ways:

- it would be much flatter, the goods would have been spread around to the entire population of the world and
- in any location there would be an assortment of a great many products, not towering stacks of a single product.

The coordination of specialized producers and generalist users is accomplished by a set of institutions that differ in importance both over time and across the economies of the world today. These include:

- *Market exchanges:* Selling the goods that specialized producers have made provides the budget for purchasing the general market basket of goods and services on which we live.
- *Government acquisition and provision:* Publicly provided services are based on the integration of the specialized producers goods and services to provide education, security and other government-provided services to generalist users of these services.
- *Families and other face-to-face communities:* Typically families exhibit a division of labor by age and gender: adult women, for example, biologically producing children and spending disproportionate time on raising them and caring for other family members (for example preparing meals). The goods

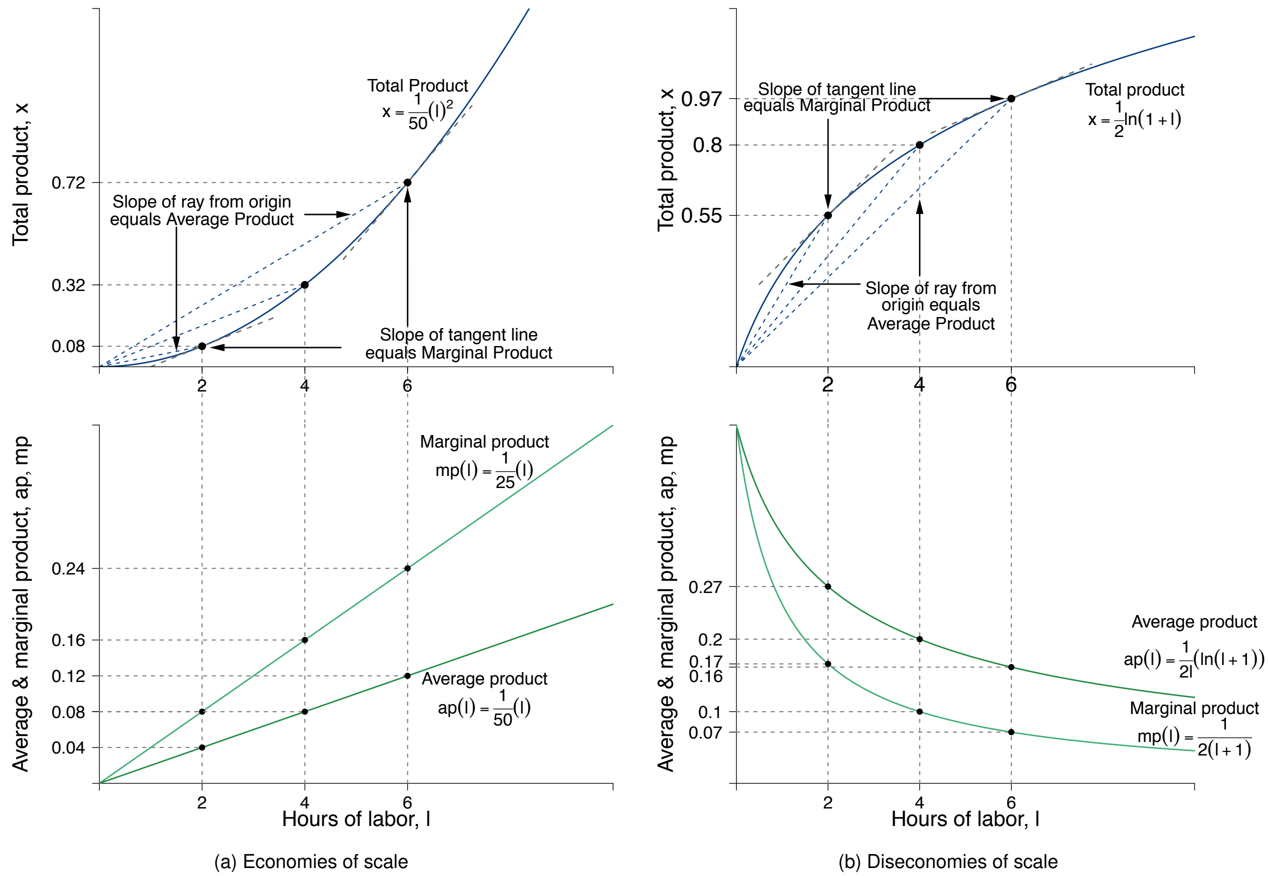


Figure 6.3: **Production functions with economies and diseconomies of scale.** With *economies of scale* (Figure a) doubling labor input more than doubles output, as can be seen by going from 2 hours of labor to 4 hours. Average product and marginal product increase with labor input as you can see from the slope of the production function (the marginal product), which *steepens* as the labor input increases and the slope of the ray from the origin to a point on the production function (the average product), which also steepens. The average and marginal products given by these slopes in the upper figure are shown in the lower figure of panel a. With *diseconomies of scale* (panel b), the opposite occurs: the average and marginal products of labor are both *decreasing*.

produced and tasks performed by adult men and women and by children are shared within the family or some other larger consumption unit.

These three ways of coordinating the division of labor have in common that they are like a two-sided platform that connects specialist producers with generalist users of goods and services. – like Air BnB that matches home owners to people looking for a place to stay, or Tinder, a dating app. This chapter's head quote by Adam Smith tells us that markets play a critical role in allowing the division of labor to expand to global proportions, leading to ever greater specialization. Here, in Part III of this book and also Chapter 14 we explain how markets work to coordinate the division of labor. We begin with the production process, and some aspects of it that favor specialization.

6.2 Production functions with a single input

In Chapter 3, we used information on the way that Aisha's study time translated into her learning to ask how much time she will choose to study. In Chapter 5 information about the relationship between fishing time and the amount of fish caught was a key idea in posing and then addressing the common property resource coordination problem. In both cases we were using, as you recall from Chapter 5, production functions.

To better understand the division of labor and specialization we now need to look more carefully at the properties of production functions. Think about another person, Alex, who has to choose how much of his time to spend producing one or more goods. Alex can devote more or less labor (l) to production and he can observe how his output (x) varies as he changes the number of hours he works. Alex may use a computer and a desk or a given plot of land and farming equipment or other inputs, but for now we assume that the only input is his labor time.

The relationship between the input of his labor and the output of the goods is described in a production function – $x = f(l)$ – taking the form:

REMINDER A production function is a mathematical description of the relationship between the quantity of inputs devoted to production on the one hand (the arguments of the function) and the maximum quantity of output.

$$x(l) = ql^{\alpha} \tag{6.1}$$

The exponent α measures the responsiveness of output to a change in the level of labor. The positive constant q measures the overall productivity of the production process, which will be greater the more skilled or hard-working Alex is. The top panel in Figure 6.3 a illustrates this production function (equation 6.1) with $q = \frac{1}{10}$ and $\alpha = 2$. The top panel in 6.3 b shows a different production function: $x = \frac{1}{2}ln(1+l)$.

ECONOMIES AND DISECONOMIES OF SCALE When production exhibits economies of scale, increasing inputs by a factor more than proportionally increases output; with diseconomies of scale, increasing inputs by a factor less than proportionally increases output.

In the top portion of both panels more hours of labor result in more output, but the panels differ in *how much* output is obtained for given inputs.

- **Economies of scale** (6.3 a): when Alex doubles all of the inputs – in this case that means just his labor input – the output *more than doubles*.
- **Diseconomies of scale** (6.3 b): when Alex doubles his labor input (assumed to be the only input), the output *less than doubles*.

The term **constant returns to scale**, not shown in the figure, refers to the case where when inputs double output doubles, so the production function is just a straight line as shown in the lower right panel of Figure 6.4.

In the lower figures of both panels we show two important statistics describing aspects of the two production functions in the above figures. The *ratio* of the amount of output to the amount of the input involved in producing it is the *average product* of that input (also called average productivity).

AVERAGE PRODUCT The average product of labor is the ratio of the output to the labor input.

Average product is measured by the *slope* of the line from the origin (called a ray) to a point on the production function. The bottom panel in figures **a** and **b** show the average product of labor associated with the production function shown at the top of those figures. When production exhibits economies of scale, average product *increases* as the scale of production increases through an increase in inputs.

MARGINAL PRODUCT The marginal product of labor is the ratio of the change in total output to a small change in input.

The ratio of the increase of output to an increase in labor input is the *marginal product* of labor (also called marginal productivity).

M-Note 6.1 summarizes the different cases when the output is x and the only input is labor, l.

Checkpoint 6.1: Production and labor inputs

Consider a production function: $x(l) = 10l^{0.5}$:

a. Sketch the production function.

b. Calculate $ap(l)$ and $mp(l)$. Sketch them.

c. Does the production function exhibit *economies* or *diseconomies* of scale?

M-Note 6.1: The average and marginal product

Here we summarize the concepts of total product, average product and marginal product. Because the marginal product is the slope of the production function, it is also the derivative of the production function with respect to a particular input, e.g. for a production function using only labor, $x_l = \frac{df(l)}{dl}$.

If there is just a single input, labor, and the marginal product of labor is greater than the average product of labor then the average product must be increasing as more labor is used. We will use equation 6.1 to illustrate why this must be the case.

We start calculating the average product ap and marginal product mp of the production function $x(l) = ql^{\alpha}$:

$$\text{Average product:} \quad ap = \frac{x(l)}{l} = \frac{ql^{\alpha}}{l} = ql^{\alpha-1} \tag{6.2}$$

$$\text{Marginal product:} \quad mp = \frac{dx(l)}{dl} = \alpha q l^{\alpha-1} \tag{6.3}$$

To analyze how the average product changes as more labor is put into production, we calculate the derivative of the ap function with respect to labor hours:

$$\frac{dap(l)}{dl} = (\alpha - 1)ql^{\alpha-2}$$

If $\alpha > 1$, the expression above is positive, which means that the ap increases with more labor. If $\alpha < 1$, it is negative: the ap is reduced if we add labor.

In summary:

- If $\alpha > 1$, then $mp > ap$, and so $\frac{dap(l)}{dl} > 0$ and
- If $\alpha < 1$, then $mp < ap$ and so $\frac{dap(l)}{dl} < 0$

6.3 Economies of scale and the feasible production set

Suppose that Alex can spend his time fishing and making shirts in some combination, including complete specialization (spending all of his time on one or the other). He prefers to have more of both shirts and fish: both are goods; and he needs at least some of each to survive. His labor time, as in Chapter 5 is a "bad" but we will set aside his choice of total hours of work by saying that he can work any amount up to ten hours a day, and that given how productive his labor is and how much he values the goods, he will choose to work the full 10 hours. As a result, the more time Alex devotes to producing one good, the more of that good he will have, but because Alex's time is limited, the less he can produce of the other. Therefore the opportunity cost of more fish is the amount of shirts he will have to forego if he shifts his work time from shirt-making to fishing. In order to pose the question – how much time will he spend on each? – we need two pieces of information:

PRODUCTION POSSIBILITIES FRONTIER (PPF) The production possibilities frontier for two goods shows the maximum *feasible* amount of one good that can be produced given the output of the other. The production possibilities frontier is the boundary of the producer's feasible set and is an alternative name for the feasible frontier when we study on production.

- the *feasible set* of combinations of fish and shirt amounts that are available to him, given his labor time and the production functions at his disposal; and
- the *indifference map* representing his valuation of each of the combinations of the two goods.

We derive the feasible set in this section and introduce the indifference map in the next.

We will assume that there are economies of scale in shirt making (as is common in manufacturing processes) but constant returns to scale in fishing. This means that the fish production function is linear: it is a straight line. As a result both the average and marginal products of labor are constant and equal to each other and also equal to the slope of the blue line through points **f** an **c** in Figure 6.4. In in Figure 6.4 we derive his feasible set of fish and shirts based on

M-CHECK Remember, the average product is the slope of a ray from the origin to a point on the production function, and the marginal product is the slope of the production function. So if the production function is just a straight line, both of these are equal and do not vary as more labor is devoted to production.

- the *total amount of time he will work* (shown in the lower left quadrant of the figure); and
- the *productions functions* for fish and shirts (shown in the lower right and upper left quadrants, respectively).

In the figure the horizontal axis to the left of the origin represents positive amounts of labor devoted to making shirts, and the vertical axis below the origin is positive amounts of labor devoted to fishing.

Alex needs to make a choice between three different ways to allocate his time in production to two types of output (fish and shirts):

a. Allocate 10 hours of work to producing only shirts

b. Allocate some hours of work to producing shirts and some to fish, and

c. Allocate 10 hours of work to producing only fish

Option a) is shown as point **a** in the figure where Alex dedicates all of his 10 hours of work time to producing shirts. We extend a line up to his production function for shirts and notice that 10 hours of labor results in Alex producing 50 shirts.

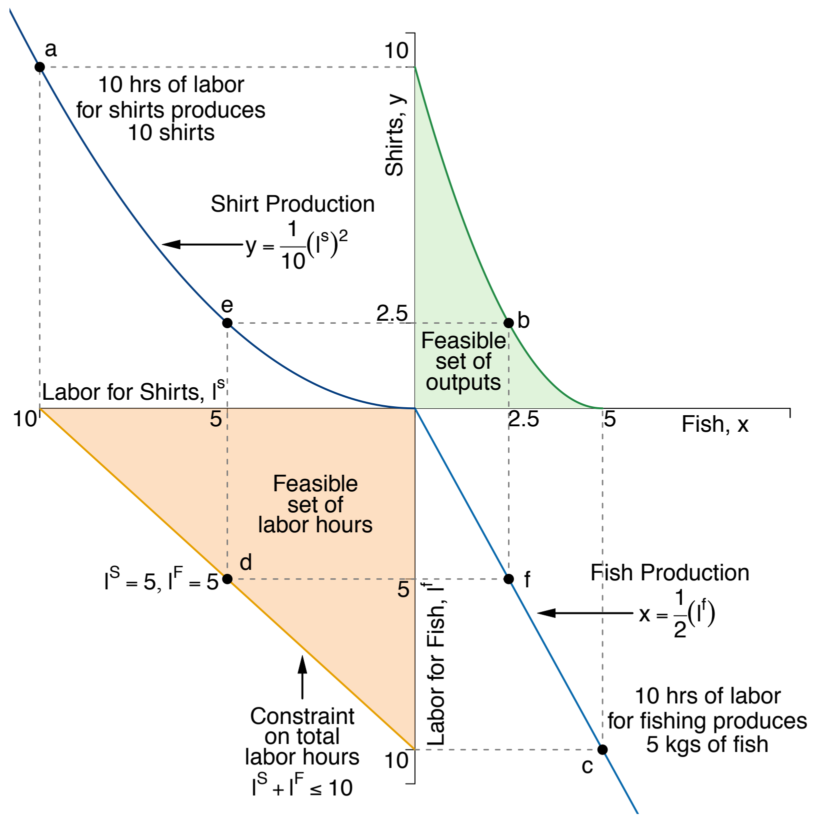


Figure 6.4: **Deriving the production possibilities frontier with economies of scale.** The lower left quadrant shows the constraint: a given amount of labor is available. The upper left and lower right show how the available labor can produce shirts and fish respectively. The upper left is an economies of scale production function similar to the panel **a** in Figure 6.3, but just rotated clockwise 280 degrees. The lower right production function has constant returns to scale. Points d,e,f, and b illustrate the production of shirts and fish that are possible if the labor time is divided equally among the two sectors. To check that you understand how the figure works, find the point on the feasible frontier associated with devoting 8 hours to fishing and 2 to shirt-making, (trace out the new points, **d'**, **e'**, **f'**, and **b'**).

We extend a dashed line to the y-axis to see what this would correspond to on the production possibilities frontier and see that Alex would produce 50 shirts and no fish as a result of dedicating all his labor to shirts. We could follow the same process for option c) corresponding to point **c** in the figure. He would produce 5 fish and no shirts by dedicating all his time to fishing.

Option b) (that is, point **b**) on the other hand, shows what Alex would produce by dedicating half his time to shirts and half to fishing. Because production in fishing is *linear*, if he dedicates 5 hours, he simply gets half of what he would have produced at 10 hours (2.5 fish). But, because there are economies of scale in shirt production, from dedicating half his time to shirts, he only gets a quarter of the output relative to 10 hours of labor for shirts (12.5 shirts vs. 50 shirts)

The top-right quadrant of Figure 6.4 illustrates economies of scale. The result is that Alex's production possibilities frontier is bowed *inward* toward the origin. This reflects the fact that with economies of scale in one or both production functions, dividing your work time between the production of both is not as good (it is closer to the origin) than devoting *all* your time to just one or the other.

MATH NOTE We describe the production possibilities frontier with economies of scale as *convex* toward the origin, that is, it is *bowed in* to the origin.

The (negative of the) slope of the production possibilities frontier shows the opportunity cost of acquiring more fish by shifting labor from shirt-making to fishing, in terms of the amount of shirts that must be foregone as a result. With economies of scale, as Alex shifts his labor input from producing clothing to producing fish, $mrt(x, y)$ declines, so he gives up smaller and smaller amounts of clothing to get larger and larger amounts of food.

This reflects the fact that with economies of scale the marginal product of his labor decreases the less labor he devotes to production of a good. This means that when he is doing little shirt-making, his marginal productivity in that activity is low, so doing a little less (so as to allow him to do more fishing) does not result in a large reduction in shirts produced.

REMINDER Remember that production functions with economies of scale are *convex* (output increases at an increasing rate with input) and production functions with diseconomies of scale are *concave* (output increases at a decreasing rate with input).

6.4 Economies of scale, specialization and exchange

In Figure 6.5 we combine the feasible set from the Figure 6.4 with an indifference map represented by the three numbered indifference curves. Recall that, as indicated by the numbering of the curves, farther away from the origin is better in Alex's evaluation of outcomes because both shirts and fish are goods.

Diversification in the absence of exchange

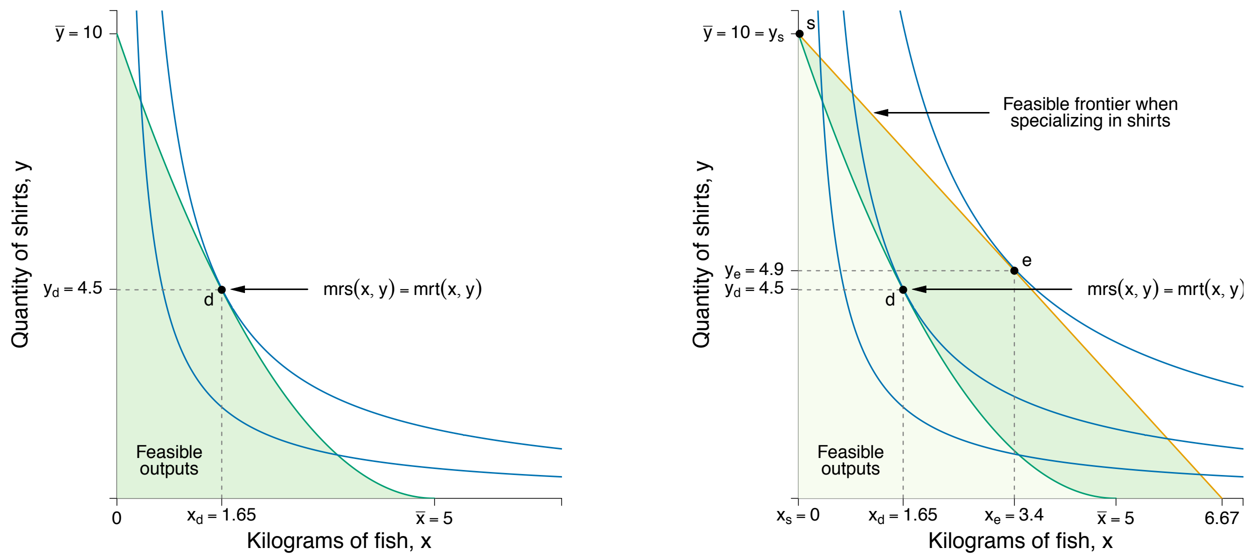
If Alex cannot exchange the goods with others, does the best he can by following the $mrs = mrt$ rule and finding the point on the production possibilities frontier that is tangent to the highest indifference curve, at point **d** (for **d**iversified production), and consuming x_d and y_d.

In this case Alex chooses to produce some of both goods because the marginal utility of each of them is diminishing the more he consumes, so having some of both is superior to having all of one kind. Remember this is why his indifference curves are bowed *inward* toward the origin.

"The division of labor is limited by the extent of the market"

But, if the producer can exchange the goods with others, then there is a second way that he can "transform shirts into fish." He does *not* do so by reallocating his time from shirt making to fishing. Instead he can spend *all* of his time making shirts and then exchange some shirts for some fish if he can find a willing buyer for his shirts.

Suppose such a trader is found, and she is willing to buy any amount of his shirts at a given price (p): in return for *p* shirts she is willing to provide 1 kg of fish. This is the second way of transforming shirts into fish, and the marginal rate of transformation is *p*: the quantity of shirts that one has to give us in exchange for a kg of fish.



(a) Economies of scale without trade

(b) Economies of scale and trade

Figure 6.5: **Production possibilities frontier (PPF) with economies of scale**. In Figure a, we present the producer's choice when they do not have the opportunity to trade. Using the $mrs = mrt$ rule, he producer chooses the point at which their indifference curve is tangent to their production possibilities frontier at point **d**. In Figure b, we show what happens if the producer can exchange shirts for fish at some constant price ratio. In this case, he can do better by specializing in shirt production (good y) and then acquiring the fish she desires through exchange, not by producing them. The green shaded area is the enlargement of his set of feasible levels of consumption of the two goods. He produces shirts only at point **s**, then exchanges the shirts on the market for fish, taking him to his higher indifference curve, u_3 at point **e**. Notice that in selecting point **s** the producer is not implementing the $mrs = mrt$ rule. The reason is that he does better at the corner solution, producing none of the x good at all. But in choosing point e by exchange, he is consuming both goods (not a corner solution) and so the $mrs = mrt$ rule implements his constrained utility maximum.

This opportunity for exchange alters the feasible set constraining what Alex can do, as shown in Figure 6.5. The orange line with the y-intercept at $\bar{y}$ (which is the maximum amount of shirts Alex can produce) represents his *new* feasible frontier with exchange. Its slope is *-p*, the (negative of the) opportunity cost of acquiring more fish in terms of the shirts foregone. This means that marginal rate of transformation of shirts into kg of fish is just p: giving up p shirts gets you 1 kg of fish. Here the process depicted by movements along the price line shows *exchange* in varying amounts, *not* shifting Alex's own labor from making shirt to fishing.

Because at the price p, 1 kg of fish is worth the same as p shirts, then the value of all of the combinations of quantities of fish and shirts along the orange price line in Figure 6.5 b. have the same value. This is because the value (expressed in number of shirts) of the fish purchased $p \cdot x$ must be equal to the value of the shirts sold (which is just the number of shirts sold, $\bar{y} - y$) or:

$$p \cdot x = \bar{y} - y \tag{6.4}$$

The constrained optimization problem that Alex faces comes in two steps:

- *Step 1:* Decide on whether to specialize and if so, in which good; to do this find the distribution of labor time between fishing and shirt making that *maximizes the value of one's output* then

- *Step 2:* Decide whether to exchange any of the goods produced, and if so how many; to do this *maximize utility subject the new feasible frontier* given by the goods produced and the relative price.

In Figure 6.5, Alex will decide to produce at point **s** (for **s**pecialized production) at the intercept of his production possibilities frontier with the shirt axis to maximize the value of his output at the relative price p. Then he will exchange the shirts he produces for fish at the price p to reach point **e** (for **e**xchange) on the highest indifference curve in his new feasible set, u_3^A. This is where:

$$\text{marginal rate of substitution} = p = \text{marginal rate of transformation by exchange}$$

Our example – Alex choosing what to produce – demonstrates two general truths:

- If one or more production function with economies of scale is available, it may make sense to specialize but
- this will be true only if there are others producing different goods and there are opportunities for exchange, integrating specialized producers with generalist users to coordinate the *division of labor*.

This is the basis of the interdependence of different producers within the division of labor.

EXAMPLE In modern economies a household may specialize in providing labor with some particular mix of skills, training and experience to an employer.

Checkpoint 6.2: The choice of what to specialize in

- Redraw Figure 6.5 with a higher relative price of fish (so p, the number of shirts that one must give up to get a kg of fish is now larger).
- Show that if p is sufficiently high, Alex will do better specializing in fish (indicate the amount he will produce, and the amount he will exchange).

Diseconomies of scale, diversification and exchange

In contrast with the production with economies of scale illustrated in Figure 6.4, Figure 6.6 illustrates the case of diseconomies of scale, in both fishing and shirt making. The result is that Alex's production possibilities frontier is *bowed outward* from, or concave to, the origin. With diseconomies of scale, as Alex shifts his labor input from producing shirts to producing fish, he gives up larger and larger amounts of shirts to get smaller and smaller amounts of fish. This reflects the fact that with diseconomies of scale the marginal product of his labor decreases the more labor he devotes to production of a good. With diseconomies of scale the $mrt(x, y)$ increases as labor is reallocated from shirts to fish reflecting the idea of *increasing opportunity costs*.

M-CHECK The production possibilities frontier with diseconomies of scale in both production functions is *concave* toward the origin, or bowed in.

If he did not have opportunities for trading goods, he would select point **d** in Figure 6.7 with utility u_2^A.

But he can do better if he decides what to produce knowing in advance that he be able to *exchange* the goods he produces. So he will use the two-step constrained optimization procedure outlined above: first decide what to produce

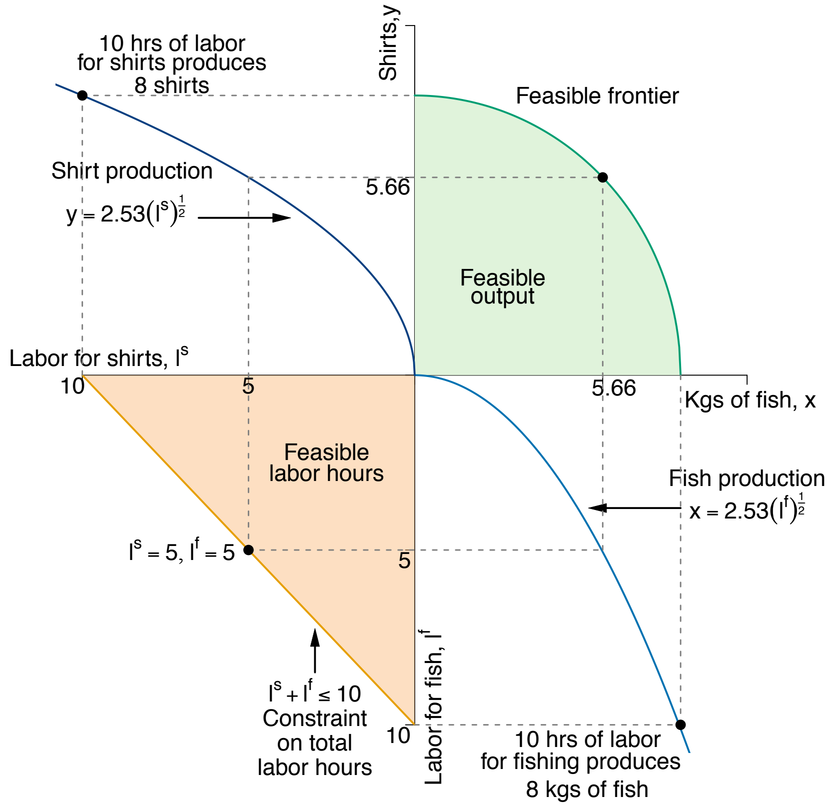


Figure 6.6: **Production possibilities frontier (PPF) with diseconomies of scale.** The four-quadrant graph shows a production possibilities frontier with diseconomies of scale in production in the top-right quadrant. The diseconomies of scale depicted in the production possibilities frontier arise from a relationship in the production technologies from the two different sectors: fishing (bottom right quadrant) and shirt-making (top-left quadrant). A worker is constrained by how much of their labor time they can dedicate to either producing fish or producing shirts. Their labor constraint is depicted in the bottom-left quadrant, shown as a limit of 10 hours of labor per day.

so as to maximize the value of his output, then exchange goods to maximize his utility.

But due to the diseconomies of scale the result is not complete specialization. He will decide to *diversify*. He will produce at point **s** somewhere in middle of his production possibilities frontier where his $mrt(x, y)$ is equal to the relative price p, putting labor into producing *both* goods, to maximize the value of his output at the relative price p.

But the possibility of exchange *expands* his feasible set: the orange line is the feasible frontier with production at point s and exchange possible at price p. He will then exchange one or the other of the goods he produces for the other – in this case *exchanging* fish for shirts – to reach the highest indifference curve that is in his feasible set, u_3^A.

6.5 Comparative and absolute advantage

The question "What should you specialize in?" seems to have an obvious answer: "Specialize in what you are best at." The same would seem to go for countries: they should specialize in what they are best at producing. But what,

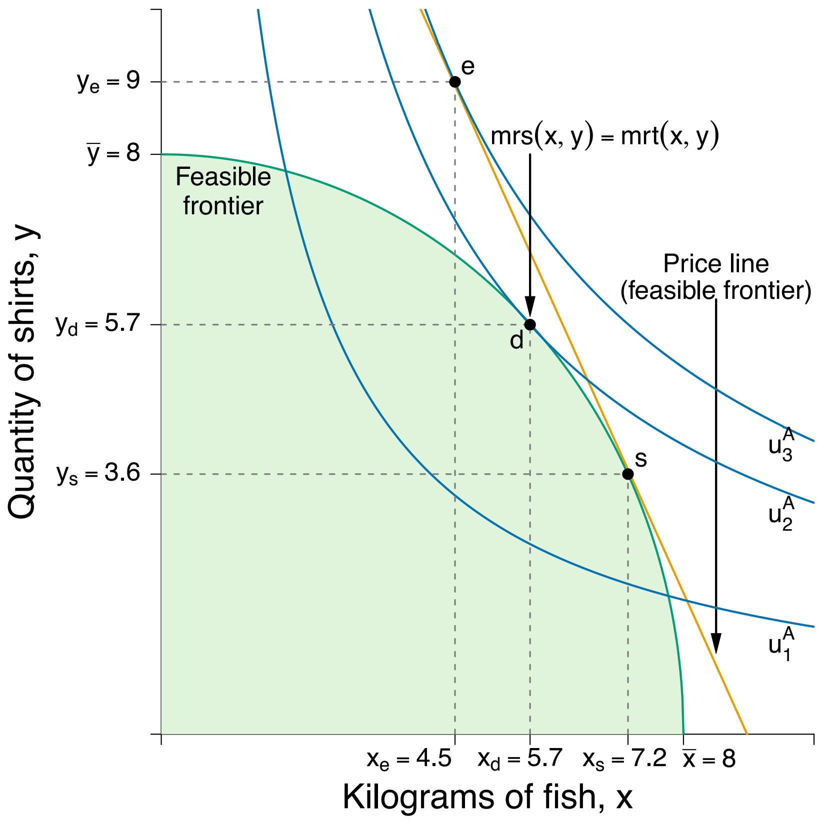


Figure 6.7: **The production possibilities frontier for fish and shirts when there are diseconomies of scale** in production. If the producer cannot exchange the goods with others, he does the best he can by finding the point on the production possibilities frontier that is tangent to the highest indifference curve, at point **d**, and consuming x_d and y_d. But, if the producer can exchange the goods with others, the producer chooses the production point with the highest value at that price, and then exchanges output to maximize utility at point **s** and then moves along the price line to point **e** on u_3.

exactly, does that mean? ”Better” than other people or other countries? What if you are not better than everyone else at anything? Should you not specialize in anything?

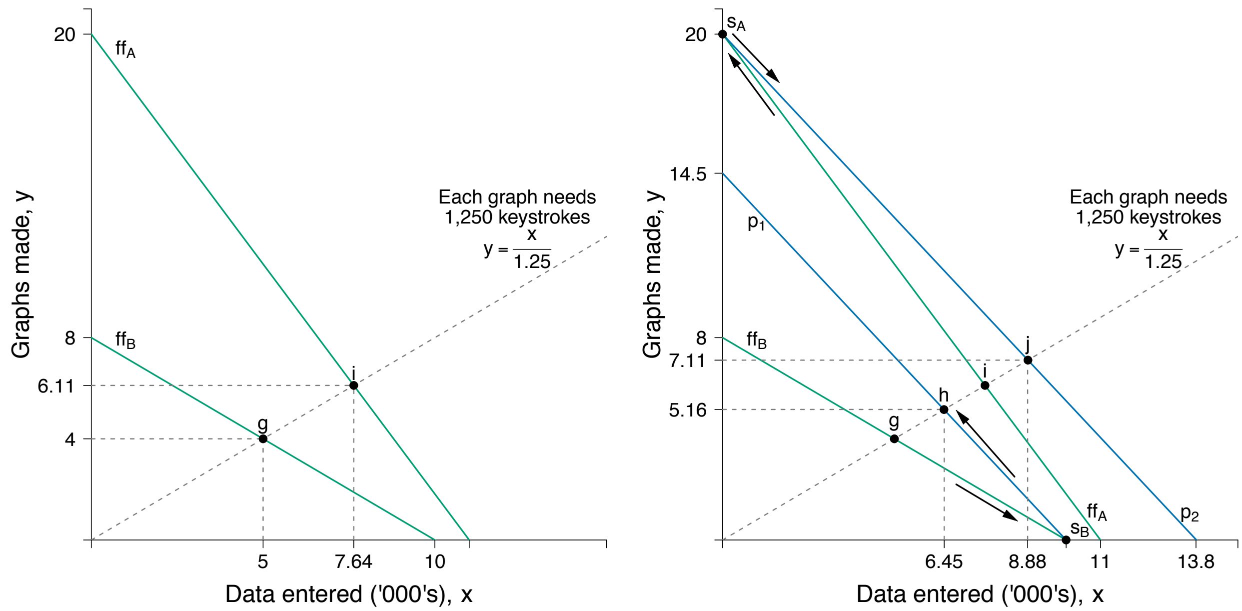
Comparative and absolute advantage

Or, does “better” mean ”better than you are at other things”? If that’s what it means – at least compared to how good others are in those same things – then we are talking about comparative advantage. This idea is more complicated than it at first seems, so let’s introduce a specific example:

Suppose that a recent graduate, Brett, has started a data science business. When Brett writes reports for his business, there are two tasks: entering data in some digital format and generating graphs to some detailed specifications using the digitized data. Each graph requires 1,250 keystrokes of data.

Brett has the option of doing both tasks, or doing either one of them himself and getting the other done for pay on Mechanical Turk (which calls itself ”the **online marketplace** for work”). There are many people like Brett, some of them offering their services on M-turk, as it is called, and with their pay

ONLINE MARKETPLACESMechanical Turk is one of many online marketplaces for work tasks for pay. Others would include Clickworker, Fiverr, UpWork and many others. People can be paid be for small, short tasks like data input (which is more typical for sites like M-Turk and Clickworker) or for more advanced jobs like Fiver and UpWork.



(a) Feasible frontiers

(b) Feasible frontiers and relative prices for exchange

Figure 6.8: **Feasible frontiers: absolute and comparative advantage.** April has an *absolute advantage* in the production of both goods because her feasible frontier is outside Brett's. Brett has *comparative advantage* in data entry because his feasible frontier is flatter than hers (lower opportunity cost of data entry). Without the possibility of exchange Bret completes 4 graphs (at point g). The two arrows show that instead, he could move to point s_B, specializing entirely in data entry, and then exchange some of his data entry with April in return for her making 5.16 graphs for him. The arrows at the top show how April could specialize and exchange.

purchasing other services from M-turk. But we will consider just one of these people named April (there are lots of people like her, lots like Brett).

Figure 6.8 a shows how good both April and Brett are at the two tasks. In panel a. we show the feasible frontiers for two potential workers: April and Brett. The area below and on the green line labeled ff_B is Brett's feasible set showing all of the combinations of keystrokes and graphs that Brett could make in an hour. In an hour he can produce 8 graphs and enter no data, or 10 (thousand) keystrokes, and no graphs, or 4 graphs and 5 (thousand) keystrokes, and so on. The green line is Brett's feasible frontier.

Figure 6.8 shows that April has an *absolute advantage* in producing both goods: data entry in thousands (x) and graphs (y). This means that, in one hour, if April devoted all of her time to data entry, she could enter more thousands of keystrokes of data than Brett (11 rather than 10) and likewise for making graphs (20 rather than 8). As her feasible set includes Brett's entire feasible set (her feasible frontier is farther from the origin), in any combination – complete specialization in one or the other or some ratio of data entry to figure making – she can produce more in an hour than Brett can.

COMPARATIVE ADVANTAGE A person has a comparative advantage in the production of a particular good if their opportunity cost of producing that good is lower than it is for some other person.

	Brett	*April*
Max possible data entry (thousands of keystrokes per hr)	10	11
Max possible graphs (graphs per hr)	8	20
Opportunity cost of 1 thousand keystrokes of data entry (in graphs)	0.8	1.82
Opportunity cost of making 1 graph (in thousands of keystrokes)	1.25	0.55

Table 6.1: **Absolute and Comparative Advantage: Number of bits of data and graphs created in one hours work**. The entries in red show that Bret has a comparative advantage in producing data entry (the opportunity cost of doing that for him is lower than for April) and similarly April has a comparative advantage in making graphs.

This raises the question: if April is better at both data entry and graph making, why would she want to trade with Brett at all? To answer this question, we need the concept of *comparative advantage*. A person has a comparative advantage in the production of a particular good if their opportunity cost of producing that good is lower than it is for some other person.

For Brett, the hour it would require to enter 10 thousand more keystrokes of data, would mean that he could not make 8 graphs. So 8 graphs is his opportunity cost of 10 thousand keystrokes of data entry. By contrast, for April, 10 thousand keystrokes of data entry requires just 55 minutes (she enters 11 thousand keystrokes per hour) and in that period of time she could have made 18.2 graphs. So for April the opportunity cost of 10 thousand. keystrokes is 18.2 graphs.

Brett, therefore, has a comparative advantage in data entry. This is not because he is so good at data entry; he is not. It is because he is so unproductive in producing graphs, so the opportunity cost of doing data entry (the graphs he otherwise could have made) is low. It can similarly be seen that April has a comparative advantage in producing graphs.

Here is a simple way to remember the difference between absolute and comparative advantage:

- If for a given axis (horizontal or vertical) the *intercept* of one person's feasible frontier is outside (farther from the origin than) the other's, then that person has an absolute advantage in the good on that axis.
- Comparative advantage is determined by the *slope* of the feasible frontier: The person with the flatter feasible frontier has a comparative advantage in the good on the horizontal axis. This is because the (absolute value of the) slope of the feasible frontier is the opportunity cost of the good on the x axis. The person with the steeper feasible frontier has a comparative advantage in the good on the vertical axis.

You can see from the second bullet that unless the two feasible frontiers have the same slope, the two persons must have a comparative advantage in different goods.

Table 6.1 summarizes Brett and April's absolute and comparative advantage in these tasks. Note that the maximum possible data entry/graphs made are not actually possible, as each graph requires 1250 keystrokes of data, but are tools to help us see who has the absolute advantage in each skill. Also note that, for example, Brett's two opportunity costs are simply the inverse of one another.

Remember: each graph requires 1250 keystrokes of data (also remember that the horizontal axis of Figure 6.8 is measured in thousands of keystrokes)

:

$$y = \frac{x}{1.25}$$

This means that they must be on the dashed line from the origin. The question is how far out they can get. This dashed line for $y = \frac{x}{1.25}$ lets us see how many complete graphs (data entry and graph preparation combined) each person could be by themselves in one hour.

6.6 Specialization according to comparative advantage

In Figure 6.8 we can see (in Panel a) that if Brett produced both data entry and graphs himself, he would get to point **g**– four completed graphs. Similarly, April herself could produce 6.11 graphs.

In Panel b of Figure 6.8 we show how specialization But (in Figure b) he specializes in data entry and then exchanges some data entry for graphs. Remember the (negative of the) slope of the exchange price line is the number of graphs that can be purchased with a thousand keystrokes of data entry, or 1.45 in our example. If we imagine Brett at point **g**, he could move in two steps to point **h** if he, first, specialized at point s_B and, then second, engaged in exchange to move to point **h**. April would move from point **i** if she produced by herself, to point **j** if she specialized at point s_A and engaged in exchange to take her to point **j**.

If they both specialize in what they have a comparative advantage in (Brett specializes in data entry and April specializes in graph-making), and if there is a price for exchange of your tasks with other work, they can both do better and obtain more graphs. The exchange opportunities are shown by the movement onto the exchange price lines p_1 and p_2 (which are parallel because *both* people face the same relative prices).

If Brett devoted the entire hour to just this task, for example, Brett could enter 10,000 keystrokes worth of data. But Brett could also divide his time. Point **g** in the figure indicates that in an hour Brett could make 4 graphs and record 5 thousand keystrokes worth of data. So, in an hour Brett could produce 4 completed graphs (the data entered in half an hour would be just sufficient for the graphs Brett would make).

How will Brett get his reports done?

	Brett	*April*
Data entry	10,000 (1 hour)	15,280 (1.39 hours)
Making graphs	8 (1 hour)	12.22 (0.61 hours)
Graphs submitted for the project	8	12.22

Table 6.2: Working without specialization: Number of graphs produced working alone for two hours.

To be a viable product, a data science report must have no fewer than 8 graphs and reports with more graphs will make more money. So, working alone Brett could do the project at least minimally in two hours, producing 8 graphs. April by herself (devoting quite a bit more than half of her time to data entry) could produce 12 graphs along with the required data entry in just under two hours.

But what if Brett specialized in data entry, the task in which he has a comparative advantage? This would only be possible if Brett could somehow get another person to do the task at which he is most bad, that is, making graphs.

This would be possible if there were a market – like M-turk – in tasks. If he traded on such a market, what Brett would be paid for his keystrokes would allow him to buy graphs using the data that he had entered.

Brett could spend his first hour inputting the data required for 8 graphs (that is 10,000 keystrokes), and then spend a second hour on mechanical turn inputting another 10,000 keystrokes of someone else's data. With the pay Brett would get for his second hour of data entry he would be able to pay for somebody on M-turk who has a comparative advantage in graph making to produce the graphs he requires for his business's reports.

Maybe even April might have figured out that she too can do better by specializing, even if she is as good or better than anyone in both tasks.

Let's think about how this would work out for Brett. In two hours, Brett would record a total 20 thousand keystrokes of data. Some would be the data for his project, and the rest would be for pay, with which Brett would purchase the graphs made with Brett's entered data. For example, Brett could enter a total of 10,250 keystrokes of his data, enough for 9 graphs. Brett would have been paid for the remaining 9,750 keystrokes entering someone else's data.

Now suppose that a thousand keystrokes sell for an amount sufficient to purchase 1.45 graphs from some other worker on Mechanical Turk (possibly even from April but it could be from any M-Turker supplying graphs). (We will not ask how this price is determined.)

Specialization and exchange at that price (1.45) allows Brett to submit a project with two additional graphs. The trade is shown by point **h** in Figure 6.8 b. on the exchange price line. Previously two hours of work earned him 8 graphs entering 10,000 strokes of data; now he gets 10.32 graphs entering all the data himself and buying graphs for the data he entered.

Let's think now about April. Like Brett, she would specialize in the task in which she has a comparative advantage. So, in two hours she could produce 40 graphs, some of them using her own data that she paid an M-turker (maybe Brett) to enter. The rest she would create for some other person (us-

	Brett	*April*
Data entry	20000 (2 hours)	0
Making graphs	0	40 (2 hours)
Work produced for own project	12,900 keystrokes	14.22 graphs
Work for pay to exchange with others	6,100 keystrokes	25.78 graphs
Others' work purchased	10.32	20,619 keystrokes
Project submitted	10.32 graphs	14.22 graphs

Table 6.3: **Specialization and exchange according to comparative advantage.** The price at which Brett and April exchanged tasks on M-turk is 1.45 graphs is equal to 690 keystrokes.

ing their entered data). If she could purchase data entry at the same price that Brett sold his for, she would get 690 keystrokes of data for each graph she made for someone else (a thousand keystrokes bought 1.45 graphs, so one graph buys $1000/1.45 = 690$ keystrokes). If she sold 1 graphs (we'll assume that a half-completed graph can be sold), the keystrokes she could purchase would be sufficient for over 14.22 graphs. So, she, too could add more graphs to her reports (she previously got 12.22 graphs).

Figure 6.8 a, the parallel blue lines show us the exchange opportunities of both Brett and April, when April specializes in graph making and Brett in data entry. The (negative of the) slope of the blue exchange price lines is the number of graphs that can be purchased with a thousand keystrokes of data entry, or 1.45 in our example. The two lines are *parallel* because, although they are buying and selling different tasks, April and Brett exchange at *the same* price. If Brett gets paid enough to buy 1.45 graphs when he does a thousand keystrokes of data entry, it means that what April gets for one graph is enough to purchase 690 keystrokes of data (0.69 is equal to 1/1.45).

To understand why specialization and exchange according to comparative advantage allows for mutual gains, remember that the green lines in Figure 6.8 are the feasible frontiers of the two prior to exchange and the blue lines are the exchange price lines that provide each of them an opportunity to do better through exchange. The possibility of exchange gives Brett a new feasible set, with the feasible frontier being the exchange price lines indicating his exchange opportunities when he can buy 1.45 graphs with the pay he gets for a thousand key strokes.

The example of the online task market Mechanical Turk shows that if they could have agreed on the price – 1.45 graphs for 1000 keystrokes – Brett and April could have simply traded directly with each other, without going online.

The reason why a mutually beneficial exchange would be possible is that the price at which the exchange took place (1.45 graphs per 1000 keystrokes) was greater than Brett's opportunity cost of keystrokes and less than April's opportunity cost of keystrokes. Or returning to Figure 6.8 the slope of the price line was greater than the slope of Brett's feasible frontier and less than the slope of April's.

What made specialization possible in this case is two things:

- *Differences*: Brett and April differed in their comparative advantage so there was some price – 1.45 per 1000 is just one example – at which they could both benefit form an exchange.
- *Opportunities for exchange*: There was a way to exchange one's completed tasks with others so as to obtain the right mix of data entry and graph making.

Checkpoint 6.3: The distribution of the gains from specialization and exchange

Using the right panel of the figure determine the price ratio (graphs per thousand keystrokes) such that Brett would not benefit at all from specializing and trading and also the price ratio such that April would not benefit.

6.7 History, specialization, and coordination failures

Brett and April simply decided to specialize in the tasks in which each had a comparative advantage. The existence of the online marketplace for tasks made this possible, and both people benefited by comparison to producing their reports without specializing. In the earlier example, Alex simply chose to produce shirts rather than fish, and he was able to feed himself because he could exchange shirts for fish. These personal examples have important lessons about specialization.

But comparative advantage is more often applied to what *countries* do, not to what people do. When it comes to countries, we cannot say that, for example, Germany "decided" to specialize in machine tools and Bangladesh in textiles. What countries specialize in is the result of decisions made by vast numbers of people independently choosing what kinds of skills they will learn, the jobs they will take, what kinds of products the firms they own will produce and similar decisions. Countries – unlike Brett and April – can sometimes end up specializing in such a way that they remain poor. Had they specialized in something else, they would have been rich.

To see how countries can specialize and stay poor, return to the feasible frontier in Figure 6.5. But now think about the figure as applying to an entire country, not just choices that Alex might make between fishing and making clothing.

In this case, the economies of scale in the production of shirts occur because in every firm labor is more productive in producing shirts the more shirts are being produced *in all of the clothing industry*. Industry-wide (rather than firm-level) economies of scale are called **economies of agglomeration** meaning that the productivity of labor is greater the larger is the total output of the many firms producing similar goods in one country or region.

ECONOMIES OF AGGLOMERATION refers to cases in which the roductivity of labor is greater, the larger is the total output of the many firms producing similar goods in one country or region.

Economies of agglomeration contribute to the geographical concentration of particular industries, for example:

- software engineering in Bangalore (Bengaluru), India
- finance in Hong Kong, London and New York City
- information technology and IT related production in Silicon Valley, California
- machine tools and motor vehicles in the Stuttgart-Munich region of Germany

Economies of agglomeratoin occur because when large numbers of people are employed in producing the same product, the skills and other knowledge particular to that industry are widely diffused in the population, resulting in higher levels of productivity across the board. Public policies favoring a locally dominant industry also reduce costs.

If the relevant economies of scale do not pertain to an individual firm, but instead are economies of agglomeration, then a single firm even if capable of operating at a large scale will have little reason to, for example, introduce a truck manufacturing plant in a finance agglomeration such as Hong Kong, or an IT region such as the Silicon Valley.

In Figure 6.9, a country could find itself at point **c** on indifference curve u_1^F where they have diversified production or at **b** where they have specialized production in fish, trade some of the fish along the price line at price p and arrive at bundle **b** on indifference curve u_2^F. They are better off specializing in fish than they would be if they tried to produce a diversified set of goods as shown by the specialized production resulting in higher utility at **b** than at **c**.

In contrast, if the very same country specialized in producing shirts and then *traded* some of their output to acquire fish, they would consume at bundle **a** on indifference curve u^S. Suppose that for as long as anyone can remember they have specialized in fishing. Why don't they just change that? In the case of April deciding to specialize in data entry, or Brett in making graphs, the two people would have quickly realized that they were ignoring their comparative advantage: they would quickly switch their specialization to what in which they have a comparative advantage. But in the case of an entire country how would they switch?

Economies of scale and poverty traps as an assurance game

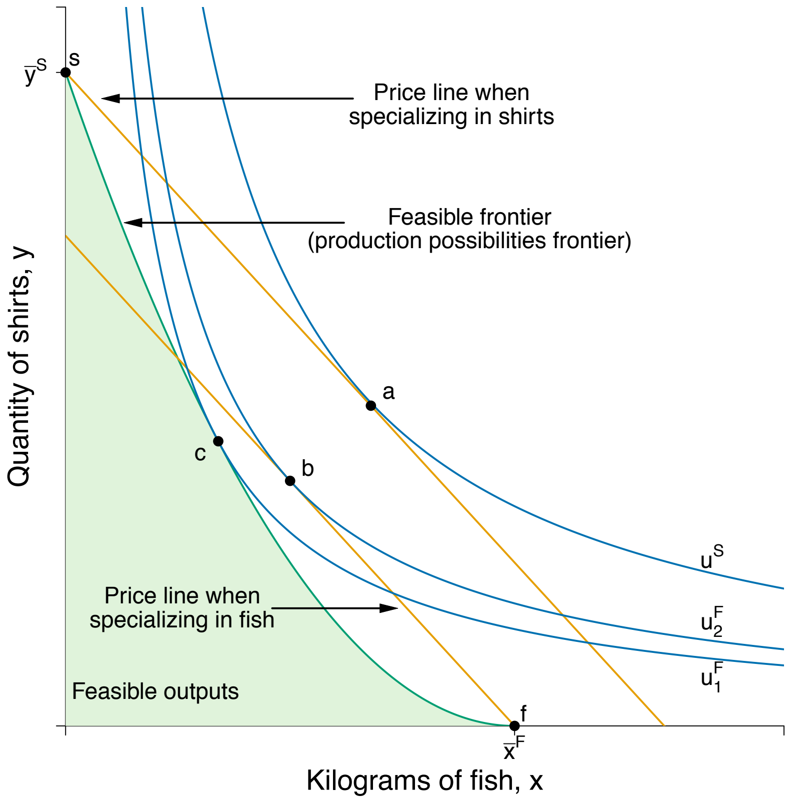


Figure 6.9: **Production possibilities frontier poverty trap.** In the figure, the country is specialized in fish and point **f** and producing $\bar{x}^F$ and obtaining welfare (utility) of u^F at point **b**, but would like to produce shirts (which would take them to a higher indifference curve). The country is better off specializing in fish than they would be if they produced a mix of goods, but they would prefer to be producing $\bar{y}^S$ shirts at point **s** where they would be on the higher price line and therefore obtaining higher utility at u^S at point **a** ($u^S > u_2^F > u_1^F$). The country cannot simply shift inputs to produce new outputs as that would require a dramatic re-purposing of production.

Suppose that in the fish-producing country a few people realized that everyone would be better off (be on a higher indifference curve) if they switched to specializing in shirts. What could they do? If people decided *individually* to produce shirts they would be much worse off than specializing in producing fish. You can see this because getting more shirts by producing them – that is moving along the feasible frontier away from point $\bar{x}^F$ – rather than producing fish and trading them for shirts is a losing proposition. With specialization, people would produce at $\bar{x}$ and trade to point **b** on indifference curve u_2^F But, if they clothed themselves by producing shirts rather than exchanging fish for shirts the best they could do would be point **c** (diversified production) on a *lower* indifference curve (u_1^F).

		Player B: Shirts	Player B: Fish
Player A	Shirts	3.33 (B), 3.33 (A)	2.5 (B), 1.67 (A)
Player A	Fish	1.67 (B), 2.5 (A)	2.5 (B), 2.5 (A)

Figure 6.10: Two people in a country have to choose whether they will play "Shirts" or "Fish". At given prices of $p = \$1$ for fish and $p = \$0.67$ for a shirt, they confront the following payoffs. The game is an assurance game with two Nash equilibria (Shirts, Shirts) and (Fish, Fish) where (Shirts, Shirts) is Pareto superior to (Fish, Fish) and Pareto efficient.

A country specializing in fish in this model is *locked in* to lower income. If they all decided to switch then they would all be better off. But as long as the decision about what each person will produce is taken independently, people would not specialize in shirt production. They are facing a coordination problem similar to those discussed in Chapters 1, 4, and 5.

To see this, imagine that the population of the country we have been modelling is composed of just two people Anjali and Budi. Budi's parents have urged him to take up fishing, and Anjali's parents, too, have urged her to continue with the family's traditional livelihood.

To determine if each will take up fishing or shirt making they will engage in the non-cooperative game shown in Figure 6.10. Assuming that each spend 5 hours a day working, we have calculated their output depending on their choice and the choice of the other, using the production functions in Figure 6.4. So:

- Fishing for 5 hours will produce 2.5 kg independently of what the other does, and the price of fish is 1, so the value of their output if either of them fish is just 2.5
- If both produce shirts, that is 10 total hours of shirt production resulting in 10 shirts, or 5 for each of them; at the price 0.67, they both receive a value of output of 3.33 (5×0.67)
- If one produces shirts and the other does not, the production function for shirts tells us that the output will be $\frac{1}{10}(l^s)^2$, with l^s = 5, this results in 2.5 shirts, with a value of 1.67 . 0.67)

You can use the circle and dot method (introduced in Chapter 1) to identify the Nash equilibria of the game. There are two: both fish or both produce shirts, and producing shirts Pareto dominates fishing.

How would the two play the game? That would depend on their beliefs. Budi might reason that for him taking up shirt making is risky because if Anjali does not make the same choice then there will be no economies of agglomeration and the payoff would 1.67. Fishing by contrast is a sure thing: 2.5. Anjali might well think the same way. Based on the traditions of their society they would probably believe that the other would take up fishing. And so they would both fish.

Of course if they could have agreed to both produce shirts, then they would have benefited from the economies of agglomeration and each produced twice as many shirts (5) as one of them working singly could do. But we are letting Anjali and Budi represent an entire population who are mostly strangers to one another, not two neighbors who could agree on a course of action.

So they have no way of coordinating their actions. Like the farmers of Palanpur – all planting late when they could all be better off by planting early – they will be less well off because of the poverty trap which they cannot escape because they lack institutions that would coordinate a joint decision.

This kind of self perpetuating specialization is part of the reason why so much of the world remained poor – the labor force of Africa, Asia and Latin America engaged in agriculture and other low productivity sectors – when Europe and its offshoots (North America, Australia, and New Zealand) became wealthier starting in the early 19th century, in some measure by producing shirts and other manufactured goods.

In the late 19th century other countries – starting with Japan but continuing in the late 20th century with South Korea, China, Singapore and Vietnam – have now shifted their specialization to sectors with higher labor productivity (analogous to shirts in our example, but including electronics, automobile production and shipbuilding). In all of these cases the change in specialization occurred as a deliberate government project, not as the result of countless people deciding to produce commodities like shirts rather than fish.

6.8 Application: The limits of specialization and comparative advantage

Economies of scale and opportunities to exchange are pervasive in modern capitalist societies, and, as a result, we live with an extensive (even global) division of labor in which many individual households and firms specialize in producing only one or a narrow range of products and meet their needs by exchanging these products through monetary transactions.

Figure 6.11: **A street-side dosa**. Courtesy Sachin Gupta. CC ShareAlike.

When we think of specialization, we often conjure images of Silicon Valley's engineering and technology hub or the City of London financial center. But, India is home to one of the most developed and specialized information technology industries in the world based in Bangalore. The Bangalore based IT firms InfoSys and Wipro exemplify the dynamics of an industry that grew from nothing in the early 1980s to become major global players by the early 2000s.[2] Specialization occurs, too, in older industries. Manufacturers in Bangladesh export a lot of shirts and hats, and very few bed sheets, whereas firms in Pakistan export a great number of bedsheets, but very few hats.[3] Neither is a particularly skill-intensive kind of production and there is no reason for us to expect that one of them *ought* to be better at bedsheets than hats. But they have specialized due to the advantages of learning by doing and economies of scale.

In contrast with this specialization, however, many households do still remain *diversified* rather than *specialized*. Many households cannot achieve the benefits of economies of scale due to insufficient wealth to sustain the training and investment required for specialization, and also because of the riskiness of starting businesses or engaging in just a single kind of work. As a result, many poor households diversify of rather than specialize.

EXAMPLE Abhijit Banerjee and Esther Dulfo won the Nobel Prize in Economics in 2019 (alongside Michael Kremer). They won the prize for their work on projects to alleviate poverty using the methods of randomized controlled trials, which they have advocated for worldwide to understand the impacts of policy.

For example, Abhijit Banerjee and Esther Duflo describe the economic lives of poor women in Guntur, a city in India.[4] The women spend time in the morning selling *dosas* (a rice and bean breakfast food), they make small amounts of money collecting trash, they gather firewood to sell, they sell fruit, vegetables and clothing (mostly saris), they make and sell pickles, or they work as short-term laborers. Similar patterns of diverse occupations occur in Cote d'Ivoire, Guatemala, Indonesia, Pakistan, Nicaragua, Panama, Timor Leste, and Mexico. An example from India shows one extreme: a survey by Nirmala Banerjee in West Bengal showed that the average family had three people who worked, sharing seven occupations among them.[5]

The economic analysis of these two different configurations of production – specialization or diversification – is based on the same fundamental concept – doing the best you can given a set of constraints. But as the examples above show, whether a person or family specializes or diversifies is not simply a matter of technology – economies or diseconomies of scale for example, or learning by doing or differential skills. For a family with limited or no wealth and exposed to uncertainty of their incomes in any single pursuit, risk mitigation becomes an important priority. As a result, diversification may be the best they can do. We show in Chapter 13 how this very common combination of limited wealth and exposure to uncertainty may contribute to the perpetuation of poverty.

6.9 Production technologies

PRODUCTION Production is the process by which we transform the resources of the natural world using already produced tools, facilities, and inputs to meet human needs.

In modern economies, production takes place in families, in governments, in privately owned firms and in other settings, each distinguished by a characteristic set of rules of the game determining who owns the goods produced, who directs the production process and so on.

Private owners of the buildings, machinery, intellectual property and other assets making up a firm aim to sell the output (which they also own) for more than their inputs cost, the difference between sales revenues and costs being the owners' profit. The owners (or managers) of the firm choose the methods of production, the amounts of inputs it hires (hours of labor, number of machines), and the level of output to maximize their profits, given the production methods available, the prices they pay for inputs, and the market prices for their output. For this reason owners of firms want to *minimize the costs* that they incur to produce any given level of output that they decide to produce. Here we explain how they choose cost-minimizing technologies to use in converting raw materials and some given level of output of products for sale. In Chapters 8 and 9 we turn to the owners' decision about how much to produce.

Inputs and outputs

Consider a firm producing an *output*, for example, cars, smartphones, or clothing.

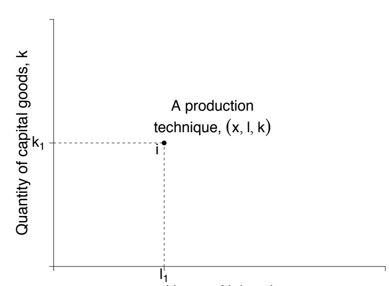


Figure 6.12: **A production technique, (x, l, k) producing some amount of output x using labor l_1 and capital goods k_1.**

FACTOR OF PRODUCTION Any input into a production process is called a factor of production. In the past economists often referred to land, labor and capital goods as

To produce its output, x, the firm needs to hire labor with the skills necessary for the production tasks and provide the workers with raw materials, tools and facilities. In a general model we could think of the inputs as a list describing the amounts of labor of each kind and of all the different raw materials (wood, steel, plastic, glass), tools (dies, drill presses, forges), and facilities (factories, vehicles) required to produce the output. These inputs to the production process are sometimes termed **factors of production**. We would measure all these inputs over the same time period as output: so many hours of each kind of employee per month, so much steel per month, so much factory space per month, and so on.

TECHNIQUE OF PRODUCTION A technique of production is a particular way of producing some given amount of output (x). In this case it is a combination of an output level, hours of labor input, and capital goods input, (x, l, k).

It's easy to see how the process works if we look at two dimensions on horizontal and vertical axes. The labor hired is l (on the horizontal axis), and k (on the vertical axis) is the quantity of capital goods that the firm uses – the machines, tools and facilities that the firm needs to hire or own to produce their output over the relevant time period. We can describe one way of producing a particular level of output by indicating in this space a level of the labor input l, and the capital goods input k that will produce the specified output. This combination (x, l, k) describes one of the possible firm's ***technique of production***. For a given level of x, we can describe the technique of production as a point in (l, k) space, as in Figure 6.12.

TECHNICAL EFFICIENCY A technique of production is technically efficient if there is no other technique with which the same output can be produced with less of one input and not more of any input.

Technology and feasible production

The firm is constrained by the *available technology*, which describes what techniques it can in fact carry out, given its state of knowledge, the skills of workers, and the conditions of work (health, safety and intensity) that the firm can legally and socially impose on its workers. Technology is therefore not just a question of engineering or scientific knowledge, but also involves relations between workers and management and among workers, and the legal and institutional framework within which the firm operates.

Figure 6.13 displays the feasible set for producing a hundred units of output. The green shaded region shows the set combinations of capital goods and labor, sufficient to produce a given output, $\underline{x} = 100$, of the good x. The dark green line is the border of the feasible set and is called an **an isoquant**. There are additional isoquants each associated with the differing level of output that the inputs produce. So there is a set of isoquants, called an isoquant map, each one of them derived from a production function and associated with a different level of output.

ISOQUANT An isoquant gives the combinations of two inputs that are just sufficient to produce a given level of output. 'Same quantity' is exactly what the two parts of the name isoquant mean: 'iso' for 'same' and 'quant' for quantity. The quantity that is the same on the production isoquant is the quantity of output.

Each of the four lettered points in Figure 6.13 is a particular combination of labor and capital goods that are sufficient to produce 100 units of good x. But the owners of a firm seeking to produce that amount would not be equally happy to use any of the four.

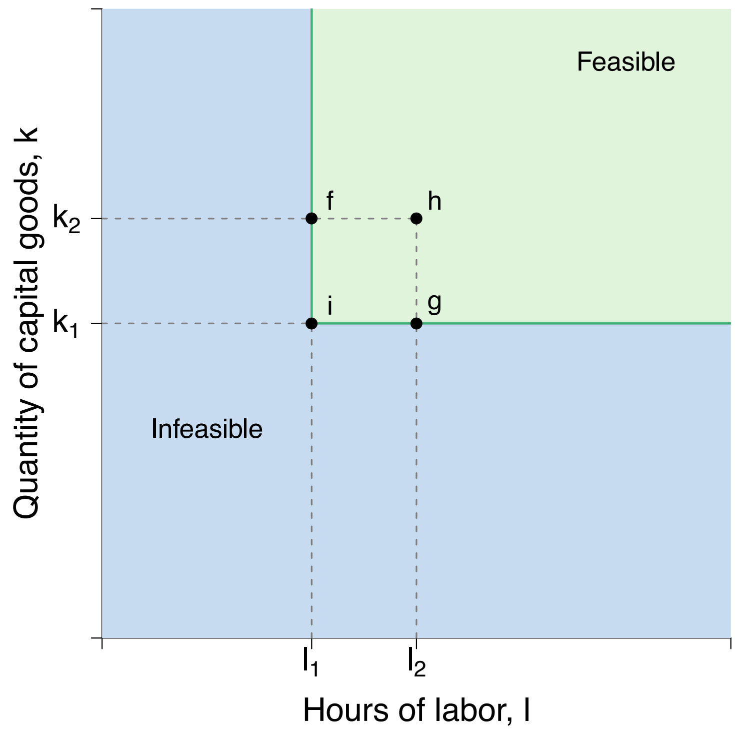


Figure 6.13: **Production techniques.** For a given level of output $x = 100$, we can describe any technique of production as a point showing the amount of labor, l, and the quantity of capital goods, k, required to produce output x. The area shaded in green shows the *feasible* combinations of labor and capital goods that can produce output x, which is equivalent to the feasible set introduced in chapter 3. The area in blue shows the infeasible combinations of capital goods and labor to obtain an output of $x = 100$.

Technique **i** *dominates* technique **h** because it uses less of both inputs to produce the same output. Similarly techniques **f** and **g** are dominated by point **i**: each use the same amount as does technique **i** of one input and more of the other to produce 100 units. Technique **i** is called technically efficient because (considering the alternatives, **f**, **g**, and **h**) there is no other technique that produces the required amount of output ($x = 100$) with less of one input and not more of any other.

REMINDER A Pareto-efficient allocation is one that is not dominated by any alternative, so there is no other allocation that is preferred by at least one person and not "dis-preferred" by any person. The definition of technical efficiency is similar but applying to techniques and inputs used rather than allocation s of goods, and people's utilities.

The isoquant map derived from a production function is analogous to the indifference curves based on the utility functions in Chapter 3. It is important to remember that an production isoquant is a *constraint* on the choice of inputs required to produce a particular level of output, rather than something to be maximized. A utility maximizer wants to get to the highest possible indifference curve given the set of feasible options. The cost-minimizing firm wants to get to the minimum cost point on the production isoquant for any given level of output.

Figure 6.14 shows a production isoquant with two techniques of production one of which uses more capital goods and less labor than the other. The production isoquant includes the points representing the two techniques and the line joining them, representing the possibility of doing some of the production with one technique and some with the other. The second technique of production provides some possibility of *substitution* of one input for the

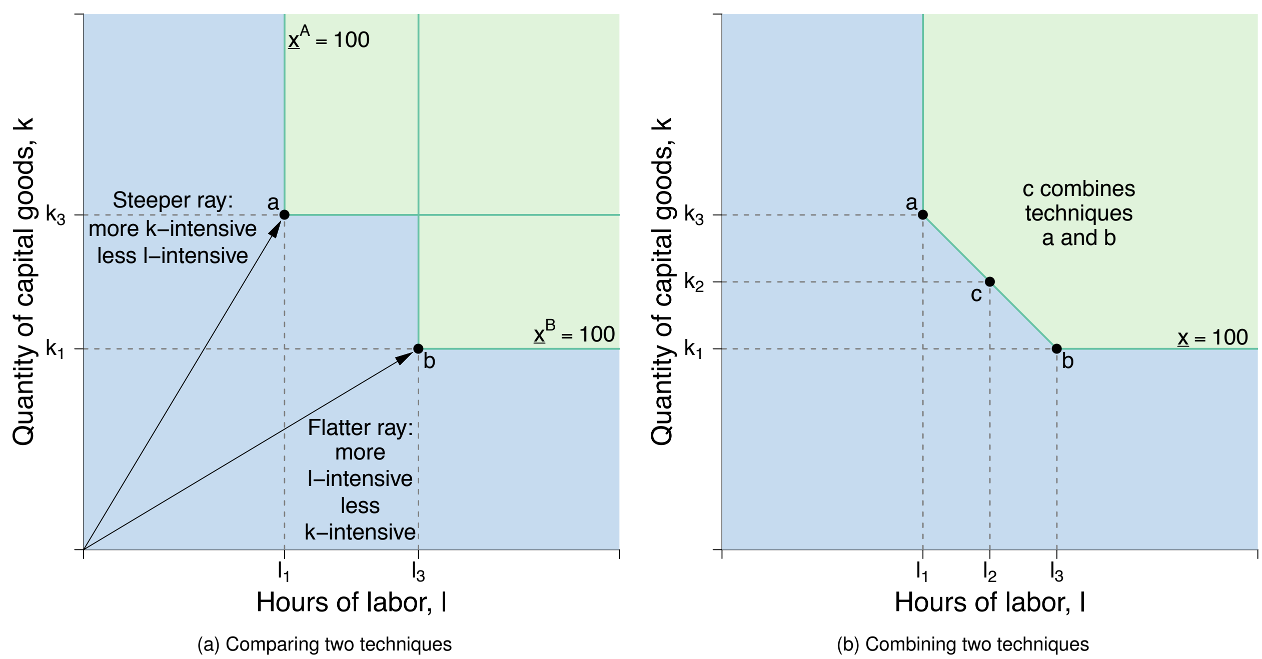


Figure 6.14: **Production isoquant combining two techniques** For a given level of output x, there may be more than one feasible technique of production. The production isoquant in this figure consists of the two techniques, and the straight line between them, representing production with a combination of the two techniques (as shown by point **c**). The availability of more than one technique implies that *substitution* of one input for the other is possible by shifting some production from a more capital-intensive technique to a more labor-intensive technique.

other by switching from one technique to the other, but this substitution is limited because there are only two techniques.

As shown in Figure 6.13 and 6.14, the production isoquant can be thought of as points corresponding to the various techniques of production, and the lines connecting those points (which correspond to mixing the techniques of production). The production isoquant is equivalent to the idea of the feasible frontier in earlier chapters as it defines what combinations of labor and capital goods can *feasibly* produce the level of output, x.

6.10 Production functions with more than one input

The techniques of production available are often described in a production function, which is a mathematical expression giving the least quantity of inputs – such as capital goods (k) and labor (l) – that are sufficient to produce any given level of output, x. The production function can also be thought of as specifying the *maximum* level of output attainable for each combination of inputs:

$$\text{Production Function } x = f(l,k) \tag{6.5}$$

You have already seen examples of the simplest production function in the *Leontief production function* in which there is but a single technique available for a given level of output x as in Figure 6.13 and Figure 6.15.

REMINDER We have already examined production functions, but so far they have only involved one input, such as labor as an input into studying in Chapter 3 or labor as an input into either fishing or shirt production in section 6.2.

PRODUCTION FUNCTION A production function $x = f(l,k)$ describes a firm's available set of techniques of production as a mathematical relationship. Here we present production functions with just two inputs – labor and capital goods – but production functions may describe the relationship between output and any number of inputs, labor with different skills, for example, or different kinds of capital goods (buildings, machines, and so on).

As in those figures what are called Leontief production isoquants are rectangular because the technology specifies a given ratio of capital goods to labor (at the point of the rectangular isoquant). If that particular ratio of inputs is in use, then adding more labor or more capital goods has no effect on production: their marginal products are zero. There are therefore no possibilities of substituting one factor of production for another.

To clarify what this "no substitution" assumption means with an extreme example, think about nuts and bolts: if you have n nuts and n bolts, then having $n+1$ bolts is no better than having n bolts. A bolt is useless without a nut, and a nut is useless without a bolt. You need to use the inputs in fixed proportion to each other to get "a nut and a bolt."

HISTORY Wassily Leontief (1906-1999) was a Russian-American Nobel Laureate in economics. He modeled the whole economy as what became known as an input-output system, with each industry being represented by a Leontief production functions. His work is valued by economists because it allowed a mathematical representation of the whole economy that could be estimated empirically (for example, engineers could determine how many tons of coal are needed to produce a ton of steel.)

M-Note 6.2: Leontief Production Function

The output of good x, is produced with l the amount of labor input used and k the amount of capital goods used. a_l and a_k are the minimum amounts of labor and capital goods required to produce a single unit of output.

Noting that $min(m,n)$ means *m* and/or *n*, whichever of m or n is least (or both of them if they are equal), the Leontief production function can be written:

$$x = f(l,k) = \min\left(\frac{l}{a_l}, \frac{k}{a_k}\right) \tag{6.6}$$

The equation can be read: "The number of units of x produced is the smaller ("min") of the ratio of the amount of the input used (the numerator in the two fractions) to the input required for a single unit of production (the denominator) Any capital goods input in excess of the minimum amounts required is of no use in production, and might as well be thrown away, and similarly for any labor input.

EXAMPLE Leontief's input-output models are today used, for example, to calculate the amount of CO2 emissions produced per unit of output of each industry, taking account of both the direct and the indirect inputs. That is counting for example not only the coal used to produce a ton of steel, but the coal used in producing the machinery and all of the other inputs required for a ton of steel.[6]

Checkpoint 6.4: Leontief Production

a. Using the Leontief production function in the M-Note, if $a_l = 2$ hours and $a_k = 1$ hour of machine use, what is the level of output in each of the following cases:
   i. $l = 10$ and $k = 10$
   ii. $l = 10$ and $k = 5$
   iii. $l = 16$ and $k = 5$
b. In each case above (i, ii, iii) how would output change if one more hour of labor or one more unit of machine time were devoted to production (this is the marginal product of labor and of machine time, respectively)?

Cobb-Douglas production function

Another representation of how inputs are combined to produce outputs is the *Cobb-Douglas production function*.

REMINDER The Cobb-Douglas production function has the same structure as the Cobb-Douglas utility functions we studied in Chapter 3.

$$\text{Cobb-Douglas Production Function} \qquad x(l,k) = ql^{\alpha}k^{\beta} \tag{6.7}$$

HISTORY Paul Douglas (1892-1976) developed the function with his colleague at Amherst College, Charles Cobb. Though a Quaker, Douglas was fiercely anti-fascist and during World War II volunteered for the U.S. Marine Corps as a private at the age of 50. He later won two purple hearts in recognition of the battle wounds he suffered in the Pacific theatre. He went on to be a prominent member of the Democratic Party and a U.S. Senator serving from 1949-1967.

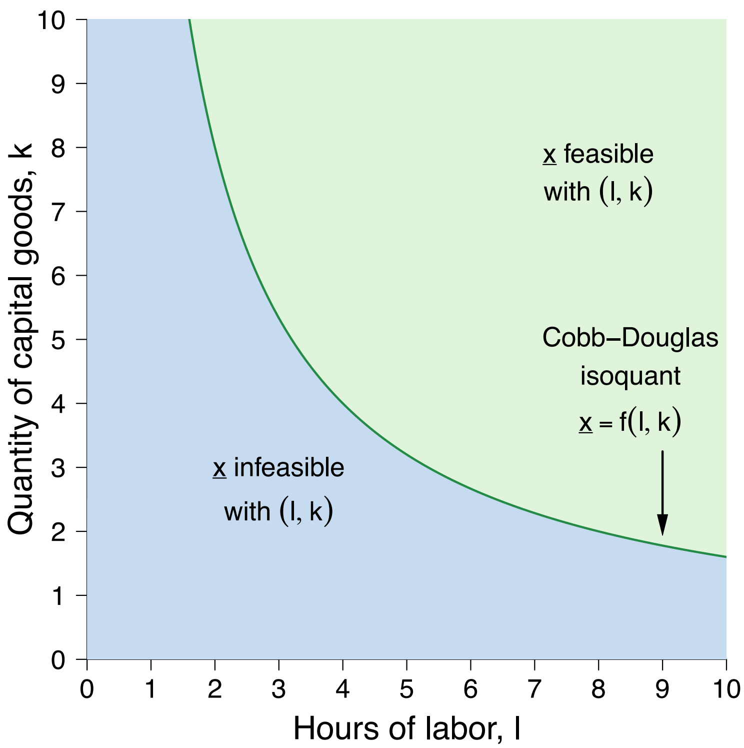


Figure 6.15: **Production isoquant given a Cobb-Douglas production function.** The feasible set of production for a given x is the set of techniques of production, combinations of labor input and capital goods, (l,k) that permit the firm to produce x. Given the Cobb-Douglas production function $\underline{x} = ql^{\alpha}k^{\beta}$ we can isolate k to find an equation for the production isoquant: $k = \left(\frac{x}{ql^{\alpha}}\right)^{\frac{1}{\beta}}$. The production isoquant has a negative slope.

The Cobb-Douglas production function requires that $l > 0, k > 0$ for production to take place: some of both inputs are essential, but their proportions used can vary.

- $0 < \alpha$ and $0 < \beta$ capture the contribution of labor and capital goods, respectively to producing output;
- The sum of α and β tells us how output responds to changes in proportional increases in both of the inputs indicating whether the firm experiences economies of scale, diseconomies of scale, or constant returns to scale.
- $q > 0$ is a positive constant that captures a level of productivity of the specific technology .

A Cobb-Douglas isoquant for $\underline{x} = 100$ is shown in Figure 6.16. The negative of the slope of an isoquant at any point is the ratio of the marginal products of labor and capital goods inputs, and is called the **marginal rate of technical substitution**, or $mrts(l,k)$.

The marginal rate of technical substitution

The negative of the slope of a production isoquant shows the ratio in which the two inputs can be substituted for each other while output remains constant, the *marginal rate of technical substitution* between the inputs. The

MARGINAL RATE OF TECHNICAL SUBSTITUTION The marginal rate of technical substitution is the rate at which labor and capital goods inputs can be substituted holding constant firm output. It is the negative of the slope of the production isoquant and equal to the ratio of the marginal products of the inputs.

marginal rate of technical substitution based on the isoquant is analogous to the marginal rate of substitution, the negative of the slope of an indifference curve, which we introduced in Chapter 3. But notice that in the case of the firm seeking to minimize the cost of producing a given level of output, the isoquant is the *co nstraint* not the firm's objective. It tells the owners of the firm what combinations of inputs will produce the given level of output (the constraint).

M-Note 6.3: The marginal rate of technical substitution and marginal products

The production isoquant is defined as the combination of inputs that can produce a given output, $f(l,k) = x$.

To find the slope of an isoquant we proceed as we did when finding the slope of an indifference curve. We use the property of the isoquant that the points on it made up of different amounts of l and k result in the same level of output x. So for small changes in l and k the following is true:

$$dx = f_l(l,k)dl + f_k(l,k)dk = 0 \tag{6.8}$$

Because along a production isoquant the difference in output is zero (just like along an indifference curve in earlier chapters the difference in *utility* is zero), Equation 6.8 can be understood as follows:

$$\underbrace{f_l(l,k)dl}_{\substack{\text{Change}\\\text{in x as}\\\text{l changes}}} + \underbrace{f_k(l,k)dk}_{\substack{\text{Change}\\\text{in x as \&}\\\text{k changes}}} = 0$$

which we can rearrange as $$mrts(l,k) = -\frac{dk}{dl} = \frac{f_l(l,k)}{f_k(l,k)} \tag{6.9}$$

Equation 6.9 can be stated as:

$$\text{Marginal rate of technical substitution} = \frac{\text{Marginal product of labor}}{\text{Marginal product of capital}}$$

This is the negative of the slope of the production isoquant.

Checkpoint 6.5: Isoquants and marginal rate of technical substitution

Using the same production function as Figure 6.16, calculate

a. the marginal product of labor and the marginal product of capital
b. determine the $mrts(l,k)$
c. choose 3 different points (not a, b and c) along the production isoquant at which to evaluate the $mrts(l,k)$ to confirm that $mrts(l,k)$ decreases as l increases.

M-Note 6.4: Cobb-Douglas economies of scale

We start with the following Cobb-Douglas production function:

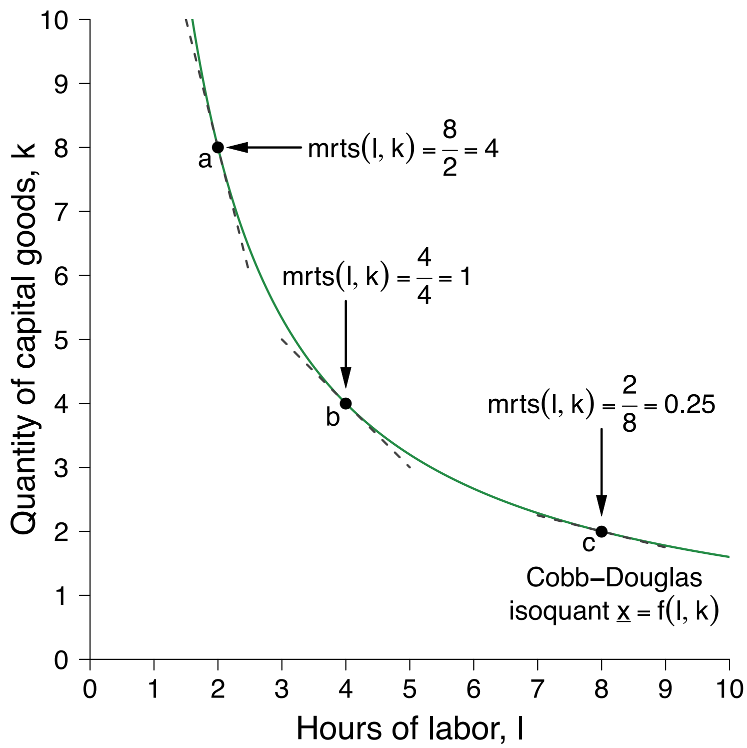


Figure 6.16: **A production isoquant for the Cobb-Douglas production function** $f(l,k) = l^{0.5}k^{0.5}$ for the output level $\underline{x} = 4$. The marginal rate of technical substitution is the ratio of the marginal products, $\frac{mp}{mp_k} = \frac{x_l}{x_k}$, which is the negative of the slope of the production isoquant, $\frac{dk}{dl}$. Three points along the isoquant curve are shown: **a**, **b**, and **c** illustrating how the marginal rate of technical substitution decreases moving rightwards down the production isoquant from $l = 4$, to $l = 1$ to $l = \frac{1}{4}$. Three gray dashed lines are tangent to the production isoquant, the slopes of which are the marginal rate of technical substitution, $mrts(l,k)$, at each point.

$$x(l,k) = ql^{\alpha}k^{\beta}$$

To confirm what a firm's economies of scale are we need to increase both inputs by some proportion, S. Therefore, increase l and k by the proportion S. That is, multiply each input by S before raising the input to the relevant power:

$$x(Sl,Sk) = q(Sl)^{\alpha}(Sk)^{\beta}$$

Now, take each S out of the parentheses:

$$\begin{aligned} x(Sl,Sk) &= qS^{\alpha}l^{\alpha}S^{\beta}k^{\beta} \\ &= S^{\alpha+\beta}ql^{\alpha}k^{\beta} \\ &= S^{\alpha+\beta}x(l,k) \end{aligned}$$

The final step occurs because we know that $ql^{\alpha}k^{\beta}$ is equal to our original production function, x. If $\alpha+\beta$ is greater than one, the output grows more than proportionally with an increase of l and k by the proportion S. Therefore, the production function has increasing returns to scale. If $\alpha+\beta$ is lower than one, the production function has decreasing returns to scale. If $\alpha+\beta$ is equal to one, it has constant returns to scale.

Diminishing marginal products of inputs

It is important not to confuse *economies and diseconomies of scale*, which describe what happens when *all* inputs are changed proportionally with *diminishing or increasing marginal productivity* of one input when the others are held constant (say, increasing labor, holding capital goods inputs constant).

A production function may have **diminishing marginal productivity** to any one input when the others are held constant, and still exhibit economies of scale when all the inputs are changed together.

M-CHECK For the Leontief production function we cannot compute the marginal rate of technical substitution from the slope of a production isoquant, because it's slope is undefined at the kink in the isoquant. But at the "kink" in the isoquant, adding more capital goods or more labor has no effect on output, so we could view the Leontief production isoquant as representing an extreme form of diminishing marginal products.

M-Note 6.5: Diminishing marginal productivity

To compute the marginal product of labor in the Cobb-Douglas production function we start with the production function:

$$x(l,k) = ql^{\alpha}k^{\beta}$$

To find the marginal product of labor, we calculate the first partial derivative of the production function with respect to labor, which gives us the effect on total output of a small change in the labor input, holding constant the level of capital goods input:

$$\begin{aligned} \frac{\partial x(l,k)}{\partial l} = x_l = MP_l &= \alpha q l^{\alpha-1}k^{\beta} \\ &= \frac{\alpha q l^{\alpha}k^{\beta}}{l} \\ &= \frac{\alpha x(l,k)}{l} \end{aligned}$$

For $\alpha > 0$ and $l > 0$, the marginal product of labor is positive: as you can see from the equation immediately above, it is equal to α itself times the average.

To work out whether the marginal product of labor is diminishing, we need to know whether the derivative of the marginal product of labor with respect to the labor input itself is positive, zero, or negative:

$$\begin{aligned} \frac{\partial^2 x(l,k)}{\partial l^2} = x_{ll} &= \alpha(\alpha-1)ql^{\alpha-2}k^{\beta} \\ &= \frac{\alpha(\alpha-1)ql^{\alpha}k^{\beta}}{l^2} \\ &= \frac{\alpha(\alpha-1)x(l,k)}{l^2} \end{aligned}$$

The sign of x_{ll} depends on the size of α.

Diminishing If $\alpha < 1$, $x_{ll} < 0$ because $\alpha - 1 < 0$, which implies *diminishing* marginal productivity of labor.

Constant If $\alpha = 1$, then $\alpha - 1 = 0$, and $x_{ll} = 0$, which implies *constant* marginal productivity of labor.

Increasing If $\alpha > 1$, $x_{ll} > 0$ because $\alpha - 1 > 0$, which imply *increasing* marginal productivity of labor

Checkpoint 6.6: Marginal products of factor inputs

Check your understanding by doing the following:

a. Repeat the steps in M-Note 6.4 to find the marginal product of capital goods with a Cobb-Douglas function.
b. With $\beta = 0.4$ is the marginal product of capital goods diminishing, constant or increasing?

c. Determine the values of β under which the marginal product of capital goods will be diminishing.

6.11 Cost-minimizing technologies

Having introduced a description of the production process – the production function – we now introduce the firm as a profit-maximizing entity. To determine the level of output that will yield the greatest profit for the owners of the firm, consider two pieces of information that the owners of the firm would need:

- *Cost minimization*: for every possible level of output, given the costs of using the inputs to the production function, find the technique of production that minimizes the costs of production;
- *Profit maximization*: using the resulting cost curve (describing the least cost at which each level of output can be produced) and the demand curve for the firm's product, determine the level of output to produce.

Here we describe cost minimization. We describe profit maximization step in Chapters 8 and 9. We call any particular combination of labor and capital goods used (l, k) as a bundle of inputs. Finding the minimum cost bundle for producing each level of output the firm's owners might want to produce requires three steps:

- *Step 1*: Calculate the cost of every input bundle that the firm might use
- *Step 2*: Identify bundles that cost the same, and use the resulting isocost line to distinguish between more costly and less costly bundles and
- *Step 3*: Use the isoquants based on the available production functions to determine, for each level of output, the least costly bundle.

Isocosts: Equally costly bundles of inputs

We assume that:

- the capital goods used by the firm are rented (for example, buildings and equipment) rather than owned; and
- the firm's own demand for labor and capital goods *does not* influence the price it pays for these inputs (as would be the case, if the firm is small in relation to the markets for its inputs, labor and various types of capital goods).

Then the cost using any particular combination of labor and capital goods depends on:

- Wages (w) paid per hour for the for the hours of labor hired (l)

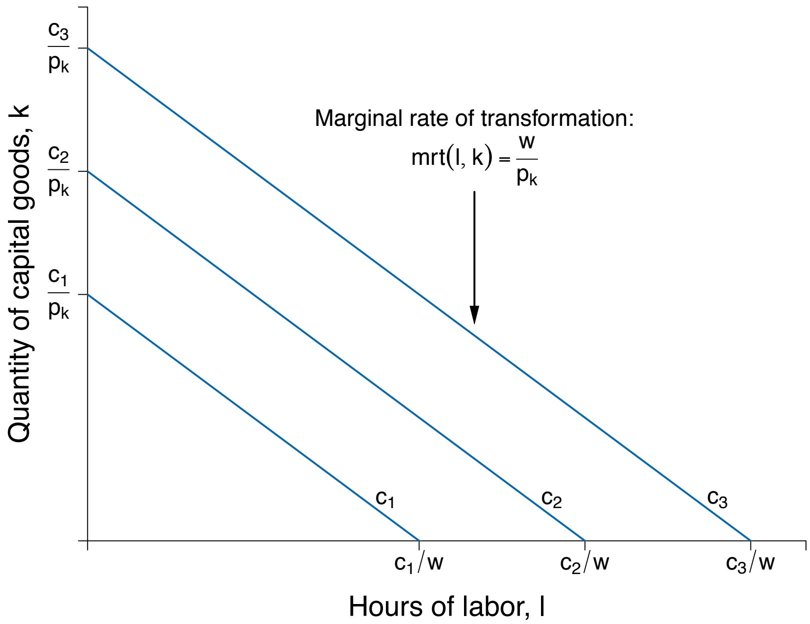


Figure 6.17: **Three isocost lines** are presented: c_1, c_2 and c_3. Isocost curves closer to the origin are made up of less costly input bundles. The equation for an isocost curve is given by $c = p_k k + wl$, where p_k is the cost per unit of renting capital goods, k is capital goods input, w is the wage, and l is the quantity of labor input. We can re-arrange this equation in terms of the capital goods input, $k = \frac{c}{p_k} - \left(\frac{w}{p_k}\right) l$. The slope of the isocost line is determined by the marginal rate of transformation of capital goods into labor, $mrt(l,k) = \frac{w}{p_k} = \frac{dk}{dl}$, which is the opportunity cost of using more labor in terms of the lesser quantity of capital goods that can be used, in order to hold constant the cost of the resulting bundle.

- The rental cost of the capital goods (p_k) times for the quantity of capital goods used (k).

Then the cost of a bundle of inputs is:

$$c(l,k) \quad = \quad wl + p_k k \tag{6.10}$$

Using Equation 8.2, we know the cost of every input bundle so we can construct an **isocost line**, a line showing all the possible combinations of amounts of labor and amounts of the capital good that result in a constant or equal ("iso") level of costs. Re-arranging equation 8.2, we can find the equation for an *isocost line*:

$$\text{Isocost line } k \quad = \quad \frac{c}{p_k} - \left(\frac{w}{p_k}\right) l \tag{6.11}$$

The isocost lines represent the *objectives* of the owners of the firm. The owners would like to find the way of producing their product that (for some given amount of output) will put them on the *lowest* isocost line, that is, the one close to the origin. The *constraint* limiting the owners decision is the available technology or technologies as described by the production isoquant for the given level of output.

ISOCOST LINE The line through an input bundle (l,k) when the wage is w and the price to hire capital goods is p_k is the line through the input point with slope equal to $-w/p_k$, and represents all the input combinations that have the same cost.

REMINDER In earlier chapters the indifference curves bowed in towards the origin (like the green isoquant in Figure 6.18) represented the objectives of the *person*, that is the thing that she wished to *maximize* based on their preferences, subject to some constraint, for example a limit on how much she could spend. Here the blue isocost lines represent the objective of the firm's owners, that is the thing they wish to *minimize*, while the curved isoquant is the *constraint* based on the feasible set of production techniques that produce at least the given amount of output.

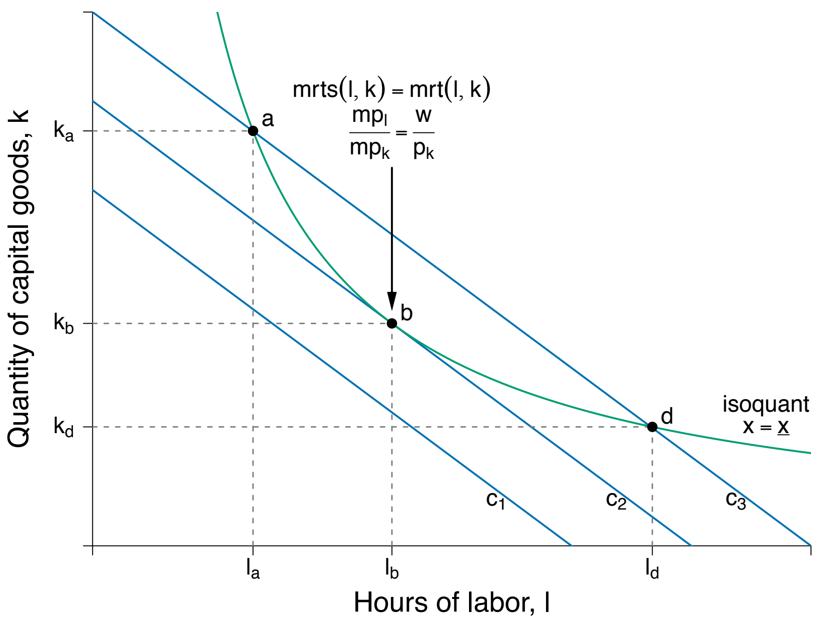


Figure 6.18: **The minimum cost of producing a given level of output.** To produce the output given by the isoquant $\underline{x}$, the least cost input bundle is indicated by point **i**.

The general principle of cost minimization

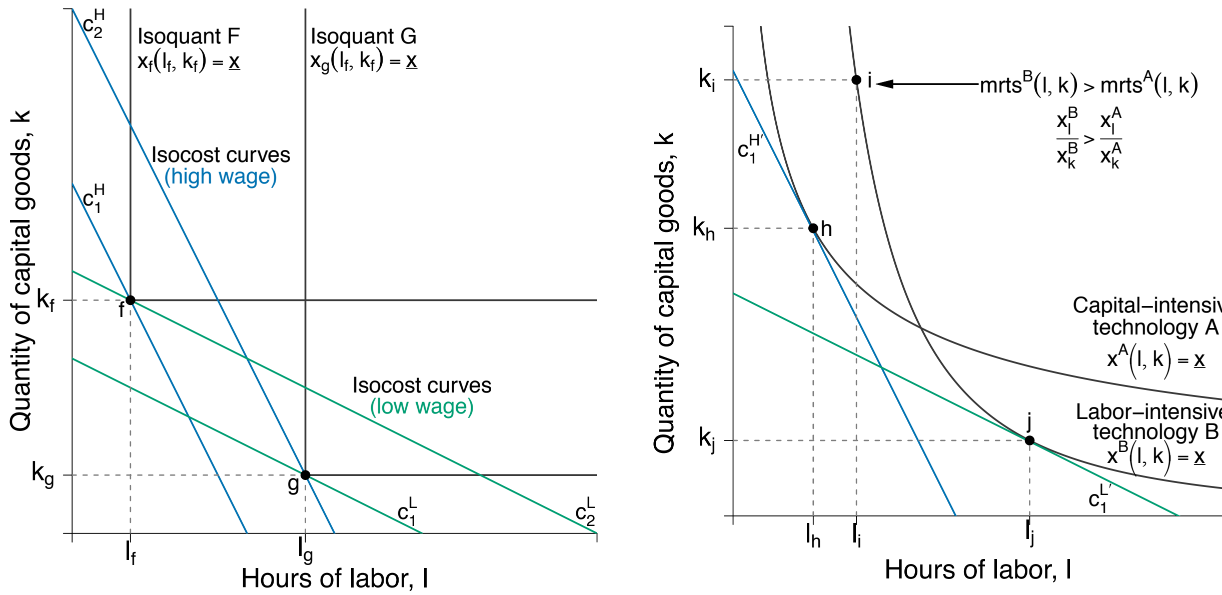
Contrast point **b** with points **a** and **d**. At **a**, the marginal rate of technical substitution is *high* (the isoquant is steep), meaning that it can reduce capital goods inputs substantially and still sustain the same level of output with only a modest addition of the labor input. At the going cost of renting capital goods and the wage rate and the firm would *decrease* its costs if it employed *fewer* capital goods and *more* labor. The effect of this is to lower the marginal product of labor and raise the marginal product of capital goods. It would continue to substitute labor for capital goods until the point where the ratio of marginal products equals the ratio of prices for the inputs at **b**.

PRINCIPLE OF COST MINIMIZATION A firm with a production isoquant consisting of a continuum of techniques of production defined by a production function $x = f(l,k)$ will *minimize* its costs at the point where its marginal rate of technical substitution of capital goods for labor, ($mrts(l,k) = \frac{f_l}{f_k}$), equals its marginal rate of transformation, or the price ratio of labor for capital goods, $mrt(l,k) = \frac{w}{p_k}$. The principle of cost minimization is satisfied where the production isoquant is tangent to the lowest isocost line.

The isocost line and production isoquant toolset allows us to understand *cost minimization*. The firm wants to choose the lowest possible isocost line to produce the output necessary to produce the target output, the ***minimum cost technique of production***.

Minimizing the cost of producing some hypothetical level of output involves the principle of constrained optimization from Chapter 3. The *constraint* is the isoquant for this particular level of output and the objective is to reach the *lowest* isocost line.

Parallel to the principle of demand in chapter 3 for people maximizing utility, we have a ***principle of cost-minimization*** for firms choosing techniques of production. The firm will produce where $mrts(l,k) = mrt(l,k)$ or where the ratio of marginal products equals the ratio of input prices, $\frac{f_l}{f_k} = \frac{w}{p_k}$.



(a) Two Leontief technologies

(b) Two Cobb-Douglas technologies

Figure 6.19: **Choosing a capital-intensive or labor-intensive technology to minimize costs.** The cost-minimizing choice of technology depends on the wage and the cost of capital goods. Higher wages (a steeper blue isocost lines) will lead the owner to implement the more capital goods intensive technology. In panel **b** the coefficients for the labor-intensive Cobb-Douglas technology B are $\alpha = 2/3, \beta = 1/3$, and for the more capital goods-intensive Cobb-Douglas technology, A $\alpha = 1/3, \beta = 2/3$.

The cost-minimization graph and the graph describing the utility-maximizing choice of a consumer facing a budget constraint are similar. But it is important to recognize that the meaning of the elements is different. The constraint in the production case is that the firm produce some given amount $x = \underline{x}$, *not* the isocost line. The owners of the firm are trying to move as close to the origin as possible within the constraint that it produces the specified amount, while the consumer is trying to move to as high an indifference curve as possible within the constraint of the budget available. The iso-cost lines are analogous to the indifference curves of the consumer (costs are being *minimized*, like utility was being *maximized*), and the production isoquant is analogous to the feasible frontier (it is the constraint).

Checkpoint 6.7

Make sure you understand Figure 6.18 by explaining why if the firm were producing at point **d** it could *reduce* costs of producing the given amount of output by using more capital goods and less labor.

Input prices and the choice of a labor-intensive or a capital-intensive technology

Now, think about a firm that sells some product and is considering which of two technologies to use producing it. One uses some powerful machinery (the capital good) and little labor while the other technology uses lots of labor and a smaller machine. For concreteness, think of the two technologies as similar to plowing a field using a powerful tractor or with a small garden type roto-tiller.

If these two alternative ways of producing the good were described by a Leontief technology, then we could say that the one using the roto-tiller is the more labor intensive, or what is the same thing (because there are just two inputs) the less capital goods intensive. In the Leontief technology, the ratio of inputs of labor to the inputs of the capital good required to produce a unit of output, that is the ratio of the amounts a_l / a_k, is a measure of the labor intensity of the technology. While the more accurate expression is to refer to capital *goods* intensive technologies, to save words we sometimes refer to technologies as "capital-intensive."

Figure 6.19 a. illustrates this case first with Leontief technologies and Figure 6.19 b illustrates Cobb-Douglas technologies. In Figure 6.19 b, the Cobb-Douglas technology indicated by point **g** is more labor-intensive than the technology at point **f**.

Where substitution between inputs is possible – as with the Cobb-Douglas technology the distinction between labor-intensive and capital-intensive technology is not so simple. The basic idea, however, is the same: the labor-intensive technology is the one that the owners of a firm would choose to would minimize costs *if wages were low* relative to the cost of capital goods. A capital-intensive technology, analogously, is one that would be used by a cost-minimizing firm *if wages were high* relative to the costs of capital goods.

Point **f** in Figure 6.19 a. shows the inputs required to produce a single unit of output using the capital-intensive technology. Point **g** shows the same information for the labor-intensive technology. Which technology the firm will adopt in order to produce its product at the lowest cost depends on the *relative* cost of labor and capital goods, as indicated by the isocost lines in green and blue.

If wages are low, then the isocost lines are flatter, as shown in the figure with the green isocost lines. If the firm uses the labor-intensive technology it will incur costs of c_1^L which is less than the cost it would incur if it used the capital-intensive technology when there are low wages (along c_2^L).

Higher wages (for the same rental cost of the capital good) are indicated by the steeper isocost lines in blue. Using the labor-intensive technology with higher wages (along c_2^K) would incur higher costs than using the capital-intensive technology (along c_1^K).

Figure 6.19 b. shows an analogous situation with greater substitutability between the two factors of production with Cobb-Douglas technologies. Once again, the relative costs are shown by two iso-cost lines, $c_1^{L\prime}$ and $c_1^{K\prime}$. The unit isoquants show the different combinations of capital goods and labor that would produce the same output, $\underline{x}$. The owners of the firm would choose point **h** if wages were high and point **j** if wages were low. Along the ray going through points **h** and **i**, the ratio of $\frac{k}{l}$ shows that Technology A is capital-intensive. At points **h** and **i**, the technical rate of substitution differs between the two isoquants (which we can see with the labor-intensive technology B having a much steeper isoquant at point **i** than the capital-intensive technology has at point **h**).

Checkpoint 6.8: Capital-intensive and labor-intensive technologies

- In Figure 6.19 Panel a, show that there is one ratio of wages to the cost of capital goods such that the least cost of producing x will be the same using the two technologies.
- Show that the input price ratio the of firm using technology F will use less labor and more capital goods than the firm using technology G.
- Show that if a firm had just two technologies to choose from, the Leontief technology F from Panel a and the Cobb-Douglas technology A from Panel b, it would choose the Cobb-Douglas technology if wages were either very high relative to the cost of capital or very low. But for some input price ratio in between, it would choose the Leontief technology.
- Explain why this means that it is not always possible to designate a technology as more labor-intensive or more capital-intensive.

6.12 Technical change and innovation rents

Firms and their owners compete not just by adjusting output levels but also by seeking to *innovate*, either by finding new lower-cost techniques of production, or by creating new products that open up new industries and new sources of demand.

Viable innovations

The theory of cost minimization suggests an important insight into the causes of innovation in production. A new technique of production will be of interest to owners only if it lowers costs of production *given current input prices*, the wage, w and the rental price of capital goods, p_k. The introduction of a new machine or a new organization of the production process may improve on some existing available technique of production but it will be irrelevant to the firm unless it results in lower costs than the existing *minimum cost* technique of production.

EXAMPLE Forbes magazine produces a list of the most innovative firms in the world (https://www.forbes.com/innovative-companies/list/. In 2018 Netflix, Tesla (electric vehicles), Facebook, and Amazon were in the top ten as was Hindustan Unilever (a consumer goods producer and marketer in India), and Naver (selling computer and web services based in South Korea).

How would we represent technological progress with production isoquant curves? With technological innovation the firm should be able to produce *the same* amount of output at lower total costs. We present production isoquants with technological innovation in Figure 6.20. The initial technology is shown

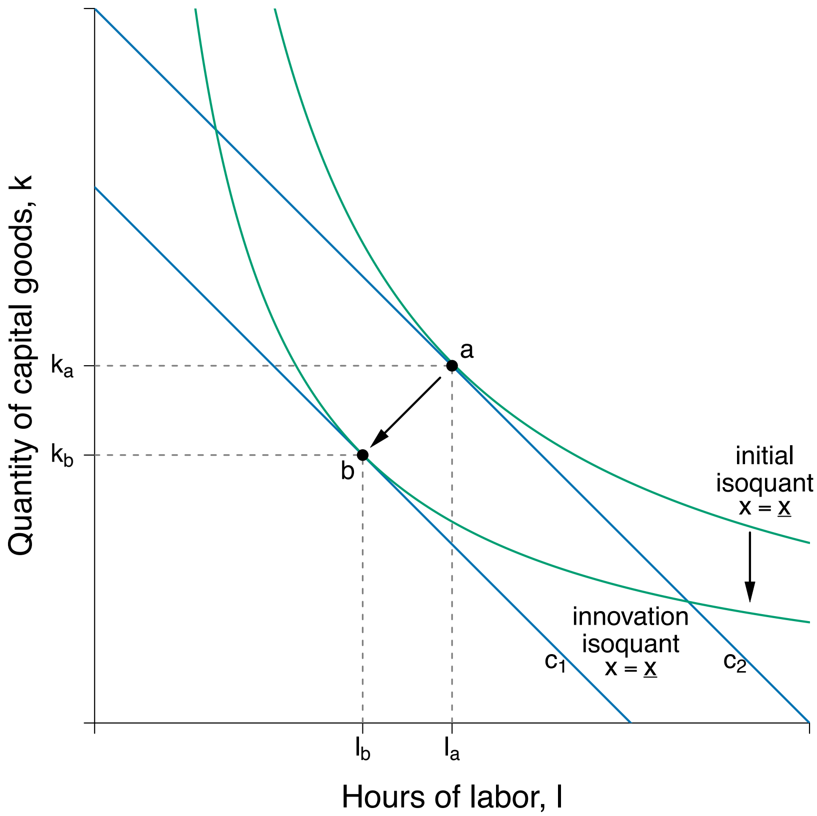


Figure 6.20: **Isocosts and technological progress.** A firm innovates to reduce its costs. The firm starts at **a** with its initial labor and capital goods combination (l_a, k_a) on isocost c_2. With innovation the firm's constraint is eased resulting in a new production isoquant with an expanded feasible set of production. As a result, the firm can, at going prices of labor and capital goods, w and p_k, employ a lower quantity of capital goods and less labor to produce output $\underline{x}$ at a lower cost total cost, moving to a lower isocost line. Following the principle of cost minimization, the firm chooses the point at which its new production isoquant is tangent to the lowest possible isocost at **b**, employing (l_b, k_b).

with the cost-minimizing point **a** where the firm employs the combination of labor and capital goods (l_a, k_a). With technological progress, two things occur:

- The firm is able to produce the *same* amount of output with *fewer* inputs. Its feasible set enlarges, resulting in a new production isoquant *closer* to the origin.
- As a result of the new production isoquant, the firm will minimize its costs *at existing input prices* and move to a lower isocost c_1, finding the point of tangency of the isocost and the new production isoquant at point **b**.

Notice that the new technology has made both labor and capital goods more productive.

Innovation rents

Other things equal – importantly the prices of the inputs and its output – if the firm produces the same quantity of x with lesser amounts of inputs per unit of x, it would necessarily increase its profit. The firm would therefore obtain an *innovation rent*. This is a rent because the firm's next best alternative –

its fallback position – would be to not innovate. The innovating firm can lower its prices and capture a larger share of the market. Other firms will either make losses and exit the industry or innovate as well. As other firms imitate the innovating firm, they would experience lower costs and would compete with the innovator resulting in a lower price for the good and therefore lower economic profits for the firm that initially innovated.

Innovation may result in new technologies that are more capital-intensive or more labor-intensive. If the production technology is Cobb-Douglas, then the innovation could be represented not only by an increase in q but also by new levels of α and β, which would indicate the change in the importance of the corresponding input.

A firm that successfully innovates by lowering its cost of production raises its maximum profit. The firm's increase in profit is an *economic rent*, or an *innovation rent*. It is an innovation rent because the firm's *fallback* is the profit it would have made at the initial technology, such as that resulting from point **a** in Figure 6.20. The innovating firm will continue to obtain higher profits until its competitors adopt the same or equivalent cost-reducing technical improvements. Once firms producing identical or similar products have matched the innovator's lower costs, if competition among firms is sufficient, some of them will reduce prices to gain a larger market share, forcing other firms to do the same. This will reduce innovation rents.

EXAMPLE Innovation rents play the key role in determining the profitability and survival of firms. Apple, for example, keeps ahead of its competitors by being the first to introduce important innovations like the iPad or the iPhone X (with facial recognition). Business history also provides dramatic examples, such as IBM in the 1980s where a firm that has managed to maintain innovation rents for many product cycles loses its position by misjudging the next turn of the technological revolution.[7]

6.13 Application: What does the model of innovation miss?

Our model of innovation captures essential parts of the process by which technical change revolutionizes an economy. But as this example shows, it misses important aspects too.

A cluster of small firms in Sialkot, Pakistan produce about forty percent of the worlds soccer balls - 30 million of them per year - including the match balls for the 2014 World Cup. The industry is highly competitive not only among the hundred or so firms in Sialkot, but also on a world scale, with Chinese firms recently challenging the Pakistani dominance in the field. Firm owners are constantly on the lookout for ways to slash costs. As the artificial leather that the balls are made from constitute almost half the cost of a soccer ball, they are particularly on the lookout for waste-saving methods of cutting the pentagons and hexagons that make up the balls.

Figure 6.21: [
0cm]**A soccer ball.** Using the new technology the white hexagons and black pentagons making up the ball could be cut with less leather wasted.

An Italian architect and her husband, an American economist, discovered a way to cut the pentagons and hexagons from the large sheets of leather that would allow a considerable saving of leather. (Unwittingly they had "discovered" what is called a "packing" principle already known by mathematicians.) They found a tool and die-maker in Sialkot to make some test dies (a cutting tool) using the new technique, expecting that they would quickly be taken up by the cost-conscious firms.

In May 2012 they gave 35 firms the new technology. They calculated that the new technology would increase profits of the companies adopting it by 10 percent. Fifteen months later only 5 of the firms had made any substantial use of the new cutting dies.[8] While the new design was easily copied and would have increased profits considerably if introduced, only one of the firms not given the new technology had copied it.

The reason, it seems, is that the employees who would have used the new dies (cutters and printers) were paid piece rates, that is, the employees were paid per panel they cut. The payment method mattered because the new technology did not speed up the process of cutting, which would have increased the pay of the cutters. Instead the cost reduction came from saving leather, which would enhance the profits of the owners, but would *not* have benefited the workers.

Because the cutters and printers did not stand to gain by saving leather, they had no interest in adopting the new technology. This was especially the case given the initial learning period in which the number of panels cut would actually be *lower* than before, meaning the workers would, for a short period, make less money. So they complained to their employers that the new dies did not work very well. Owners, lacking any independent way of verifying the competing claims of the Italo-American couple and their own cutters showed little interest in the new technology.


Figure 6.22: **A worker at the firm that adopted the new technology.**

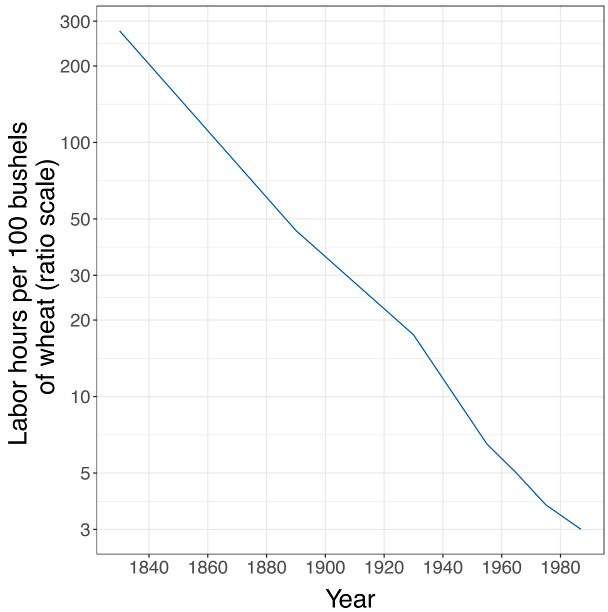
Except one. One of the larger firms had a different pay system – the cutters were paid a fixed monthly salary rather than per panel that they cut. This firm purchased (and used) 32 of the new dies, apparently without resistance from the cutters. As long as none of the other firms adopted the new technology, this firm would then have been making substantial innovation rents due to the reduced cost of materials.

If the competitive process worked in Sialkot that way economists think that it should, then this firm should have expanded its share of soccer ball production, eventually forcing other firms to either adopt the new technology, or to drop out. We do not know if that is what happened.

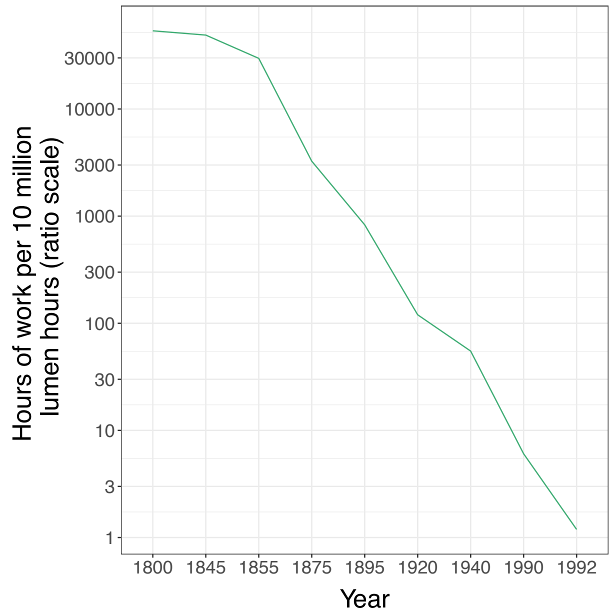
This case makes it clear that firms are made up of people, and the sometimes incomplete information and conflicting interests among them constitute impediments to improvements that in principle at least would allow for mutual gains to be shared among employees and owners.

6.14 Characterizing technologies and technical change

Production technologies shape how we live, and ongoing changes in technologies are revolutionizing the world. The Industrial Revolution and changes in



(a) Improvements in agricultural technology



(b) Improvements in light technology

Figure 6.23: **Improvements in farming and lighting technology over time.** In both panels, improvements in technology show the reduced number of hours of labor required to obtain the indicated output. The vertical axis measures what is called a ratio scale so that, for example, the distance between 20 and 100 is the same as the distance between 10 and 50 (the ratio of the first to the second number is the same in both cases). This is equivalent to a logarithmic scale, so the rate of change of the measure is the slope of the lines shown. Sources: Nordhaus (1996) and Spielmaker (2018).

technology since have transformed the economies of Europe and North America from largely agricultural production to manufacturing and later service-based livelihoods. Included were the shift of most work out of the home and into the factory or office, the enormous increase in the scale of production of typical firms, the widespread replacement of human labor by machines and vast increase in the quantity of goods and services available along with a decline in the amount of time in one's lifetime spent working.

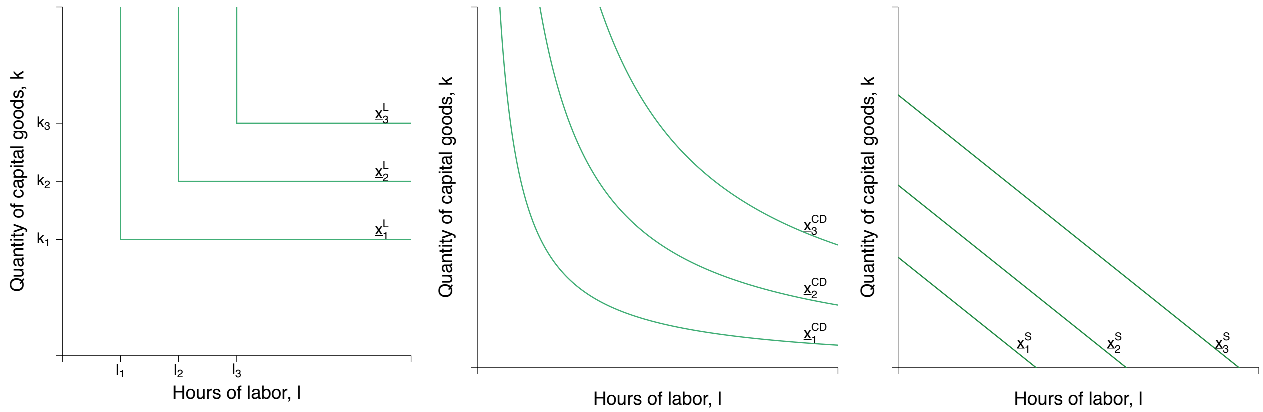
Figure 6.23 shows the scale of these productive improvements for two technologies: agricultural output and light. Panel a. shows the change in the the number of hours required to produce 100 bushels of wheat. In 1830, farmers needed 275 hours to produce 100 bushels of wheat or over thirty-four 8-hour workdays in contrast with merely 3 hours required to get the same amount of wheat in 1987 (and even less time today).

In panel a, we show how the number of hours of labor time have decreased to obtain 100 bushels of wheat. A bushel of wheat is approximately 60 pounds or about 27.5 kilograms of wheat. The graph shows the amount of time required to obtain 100 bushels, decreasing from over 275 hours in 1830 to merely 3 hours in 1987 – a 90-fold improvement in productivity.

REMINDER A *production technique* is one particular way of producing an output, x, l, k. A *production function* (for example, the Cobb-Douglas) describes a *technology*, that is, a set of techniques.

In panel b, we show the amount of labor to obtain 1000 lumen hours. A lumen is a standard measure of light intensity equivalent to the light of one candle.

The increasingly steep slope of the line in the panel b indicates an acceleration of the rate of decline in the amount of labor required to produce a given



(a) A Leontief Technology (b) A Cobb-Douglas Technology (c) A perfect substitute technology

Figure 6.24: **Substitutability between labor and capital goods with different unit isoquants.** Panel a. Leontief isoquants illustrating an **elasticity of substitution** of zero Panel b. Cobb-Douglas isoqaunts illustrating an elasticity of substitution of one. Panel c. three isoquants of a technology in which the elasticity of substitution is infinite.

mount of light. Panel b shows a vast increase in productivity of human labor demonstrated by the decline in how much labor time was required to produce 10,000,000 lumen-hours of light.

Two hundred years ago, to make even one candle required immense amounts of work to benefit from the output: a modest amount of artificial light.

In the contemporary world, though, light arrives at the flick of a switch and the labor required to produce the electricity and the advanced technology of super efficient light-emitting diodes (LEDs) is measured in minutes not days or weeks. Both of these outputs – wheat and light – show the ways in which productivity has increased over time as a consequence of human ingenuity. You can see from the figure that the pace of productivity improvements is accelerating in lighting (the decline in labor required is steepening over time) and slowing down in farming.

M-CHECK A production function **A** is more labor-intensive than production function **B** if for any given ratio of wages to the price of capital goods, the cost minimizing choice of inputs will be to hire more labor hours when using **A** than when using **B**.

Interpreting technological change with production isoquants

To understand how technologies impact how we work and live, and why new technologies continue to revolutionize our economy and society, think about five dimensions of a technology.

- *Economies of scale:* Does increasing all inputs by a factor of S increase output by more than a factor of S?
- *Overall productivity:* For a given set of inputs how much output is produced?
- *Input intensity:* Does the production process rely more on the input of labor (as in caring for children or the elderly, or doing scientific research), or capital goods (as in manufacturing) or natural resources, information, or some other input? Labor intensive and capital-goods intensive production is illustrated in Figure 6.19.

M-CHECK Figures 6.19 a. and b. illustrated the isoquants of Cobb-Douglas and Leontief technologies, one of which is more labor-intensive (less capital-intensive) than the other. You can check the input-intensity by looking at the slope of the isoquants (that is the, the marginal rate of technical substitution or the ratio of the two marginal products) for a given ratio of the inputs (that is, along a ray from the origin). Technology A (shown by x_1^A) is more capital goods

- *Complements and substitutes:* If the amount of one input used increases and this raises the marginal product of the other input then the inputs are complement. Example: computer driven welding robots increase the marginal productivity of the engineers who program them. If the increase in one input used reduces the marginal product of another input, the two inputs are substitutes. Example: computer driven welding robots reduce the marginal product of manual welders (possibly to zero).

- *Input substitutability:* Must inputs be used in some fixed proportions or can one input be substituted for another. An example of fixed proportions: a truck needs a driver, adding a second driver or a second truck does not add much to the transportation services delivered. An example of substitution: calculations done "by hand" with pencil and paper can also be done by a computer, using more capital goods and less labor.

The extent to which one input can be substituted for another in production is termed the *elasticity of substitution* defined as the percentage change in the minimum cost input proportions associated with a percent change in the ratio of the wage rate to the price of capital goods. The elasticity of substitution ranges from:

- *zero* in the Leontief production function (zero change in input proportions, no matter how much the input prices change (illustrated in panel a of Figure 6.24) to

- *one* in the Cobb-Douglas production function (if the ratio of the wage to the price of the capital good doubles, the ratio of amount of the capital good to the amount of labor used by a cost minimizing producer will double, illustrated in panel b of Figure 6.24), to

- *infinity* where the inputs are called perfect substitutes, and using one of the two inputs, but not both will generally be cost minimizing (illustrated in panel c of Figure 6.24).

ELASTICITY OF SUBSTITUTION is the percentage change in the minimum cost input proportions associated with a percent change in the ratio of the wage rate to the price of capital goods.

The elasticity of substitution tells us how linear the isoquants are, ranging from perfectly linear (perfect substitutes), to extremely kinked (the Leontief case).

Examples of how these dimensions of technologies altered ways of life in the past, and continue to do so today are given in Table 6.4. Key questions about technology today include:

- Are capital goods in the form of robots substitutes for workers doing routine tasks and at the same time complements of engineers who operate them? If so, the marginal products of the two kinds of workers will diverge, possibly generating greater inequality between engineers and routine task workers.

Characteristic	**Example: Cobb-Douglas** $x = ql^{\alpha}k^{\beta}$	**Example from Industrial Revolution**	**Examples from today and future**
Economies of scale	$\alpha + \beta > 1$ economies of scale $\alpha + \beta < 1$ diseconomies of scale	Industrial revolution increased economies of scale leading to larger firms	New technologies (e.g. 3-D printers) may reduce economies of scale but large first copy costs ("prototyping") imply economies of scale (e.g. R&D for producing a drug)
Overall productivity	q	Increases in productivity allowed for improved living standards including less work	Is a long term slowdown in productivity growth in our future?
Labor intensity	α	α fell and the ratio of capital goods to labor input rose	Labor with engineering and networking skills may be replacing both capital goods & other labor
Substitutes or complements	Inputs are complements	Capital goods were substitutes for some kinds of labor (manual, routine) and complements for others (engineering, design)	Artificial intelligence (AI) may become a substitute for even highly trained engineering and other labor
Elasticity of substitution	Elasticity of substitution = 1	For many early machining and production line processes, substitution was very limited.	If the elasticity of substitution is low, then the continuing increase in the quantity of capital goods per worker could allow wages to rise relative to profits.

Table 6.4: **The ways in which technologies differ and why it matters**.

- As the ratio of capital goods to labor input continues to increase, this may depress the marginal product of capital goods and raise the marginal product of labor as would be expected if capital goods and labor are complements. This could be the basis of greater income equality between owners of capital goods and the workers they employ. But by how much?

- Sectors of the economy with labor intensive production functions – such as education, security services, entertainment, child and elder care, and health services – are increasing their share of the economy at the expense of capital goods intensive sectors such as manufacturing. Will this result in greater scarcity of labor relative to capital goods and an increase in workers' bargaining power?

- Will devices and algorithms associated with artificial intelligence become a substitute or the work of engineers and other professionals, driving down their marginal products?

6.15 Conclusion

Economics (as you read at the beginning) is the study of how people interact with each other and with our natural surroundings in producing and acquiring our livelihoods. The technologies studied in this chapter – summarized mathematically by production functions– describe how we can produce our livelihood by transforming nature – crops mineral resources and energy – in order to provide the goods and services that make up our standard of living.

The available technologies and the ways owners of firms seek to maximize their profits by choosing techniques of production that minimize their costs of production have important effects on how we interact with each other in this process (the other part of the definition of economics). If technologies are highly efficient, then people will have the opportunity for high living standards for all. But if there are important economies of scale in production then it is likely that the economy will be dominated by a limited number of large firms whose owners may be able very disproportionate share of the high levels of production. If labor is highly productive relative to capital goods, then those who own the capital goods will not be able to earn such high incomes as those – virtually everyone – who provide labor to the system of production.

But technologies and the division of labor that results when people specialize according to comparative advantage is just one part of economic knowledge. Equally important are the wants and needs of people and how these are expressed in our willingness to pay for goods when they are supplied in markets. We turn to market demands next.

Making connections

Economies of scale and learning by doing are among the main reasons for the the division of labor and specialization, which makes important contributions to human well being; but we will see in Chapter 8 that economies of scale and learning by doing a may also limit the degree of competition in markets.

Markets as a means of coordination: The opportunity to exchange of goods expands the set of feasible outcomes available to people and nations by providing facilitating specialization and the division of labor.

External effects, coordination failures and poverty traps: The positive external effects associated with economies of agglomeration result in many possible Nash equilibrium patterns of specialization; countries may specialize in goods that keep them poorer than had they specialized in some other way.

Constrained optimization: the choice of technology Minimizing cost subject to a production function constraint and maximizing utility subject to a budget constraint have many features in common. They are both examples of maximization (or minimization) under constraints.

Innovation rents A firm that succeeds in finding a new technology that lowers costs of production at existing input prices can make substantial economic profits, called innovation rents, until others adopt the same or similar innovation.

Important ideas

specialization	production possibilities frontier	diversification
technique of production	production function	production isoquant
division of labor	average product	marginal product
economies of agglomeration	marginal rate of transformation	relative price
cost minimization	equalization of marginal products and input prices	marginal rate of technical substitution
economies of scale	constant returns to scale	diseconomies of scale
wages	isocost line	rental price of capital goods
diminishing marginal productivity	short run/long run	technical efficiency
opportunity cost of capital		

Mathematical Notation

Notation	**Definition**
x, y	goods produced using labor and capital (or just labor)
$\bar{x}, \bar{y}$	maximum feasible production of goods x and y, given the current technology
p	price
l	firm input: labor
k	firm input: capital
$f()$	production function
a_l	required amount of labor inputs in Leontief production function
a_k	required amount of capital goods in Leontief production function
α	intensity of labor in Cobb-Douglas production
β	intensity of capital in Cobb-Douglas production
q	parameter of productivity, Leontief and Cobb-Douglas production
$\underline{x}$	isoquant, a fixed amount of good x that can be produced by different combinations of labor and capital
S	scale factor for increases in all inputs
w	wage
p_k	cost of renting capital goods
c	cost function

Note on superscripts: L: related to labor; K: related to capital.

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

Works Cited

See reference list.

7

Demand: Willingness to pay and prices

> This division of labour... is the necessary... consequence of...the propensity to truck, barter, and exchange one thing for another. It is common to all men, and to be found in no other race of animals... Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog.
>
> Adam Smith, *The Wealth of Nations*, Book 1 ch 2

Ancona is a town on the Adriatic coast of Italy southeast of Venice. It hosts one of the many daily fish markets that sell to European restaurants and fish dealers.

Because fish (notoriously) spoil rapidly even with refrigeration, the price of fish on any one day depends largely on the amount of fish brought to the market that day (since none can be carried over from previous days). Economists view fish markets as a kind of ideal experiment for studying how *supply and demand* determine the prices at which goods are bought and sold .

Figure 7.2 shows that the average daily prices and the average daily quantities of fish sold in the Ancona market:

- if the price per kilogram of fish is high, the quantity of fish bought and sold is less, and
- if the price per kilogram of fish is low, more kilos of fish are transacted.

One explanation for the downward-sloping line in the figure summarizing this relationship is that typical buyers in the Ancona fish market will buy more fish if the price is lower. Another is that the greater quantity of fish brought to market on any given day, the lower will be the average price per kilogram of fish.

To understand how the price and quantity of fish purchased is determined, an essential concept is demand, as measured by the amount a person is willing purchase at any given price. For fish and other goods, knowing how

DOING ECONOMICS

This chapter will enable you to:

- Apply constrained utility optimization to the the problem of demand, relating a person's willingness to pay to their purchases of goods.
- Derive a person's demand curve from a utility function describing the person's preferences.
- Understand that consumption is often a social activity, so our preferences for particular goods (for example smoking tobacco products) often depend on what others are consuming.
- Explain how people change their purchases when prices or income change.
- Understand how these responses reflect both income and substitution effects and use these concepts to explain the effects of a proposed carbon tax and citizen dividend.
- Use the concept of consumer surplus and understand the conditions under which it makes sense to sum the consumer surplus of many people.
- Explain how market demand curves can be be derived from individual demand curves.
- Use the price elasticity of demand to explain the effects of price increases for example resulting from policies such as placing a tax on sugary drinks.

Figure 7.1: **Mosquito nets save lives; how widely used they are depends on the price. See Figure 7.3.**

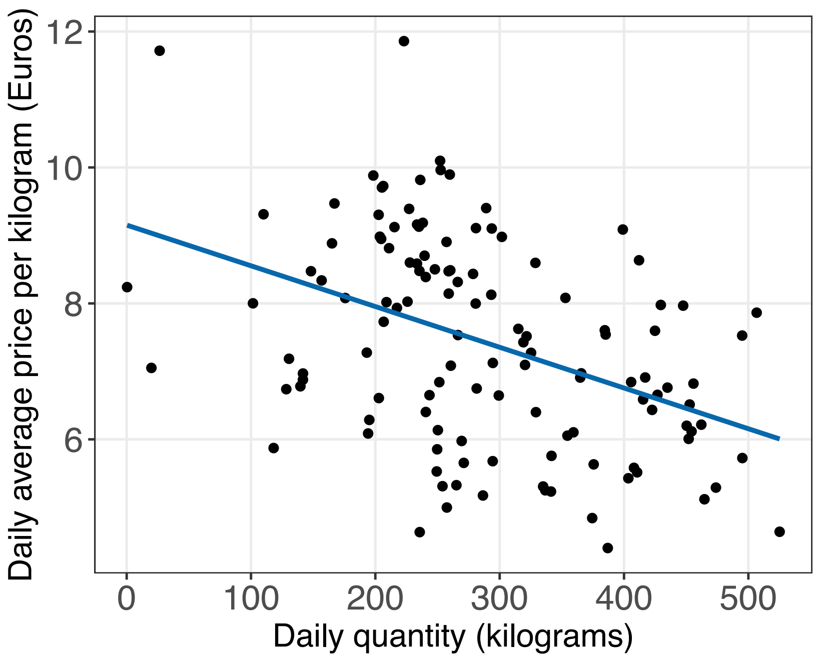


Figure 7.2: **Prices and quantities of fish bought and sold in the Ancona market.** The plot of daily average prices of fish in the Ancona market against the same day's quantity of fish sold can be summarized by a downward-sloping curve.

the amount purchased depends on the price is also an important piece of information in the design of economic policies.

Here is an example. There are now many low-cost life-saving preventative health products such as insecticide-treated mosquito nets, tablets to eradicate parasitic stomach worms, and water purification products. In many countries in Africa, Asia and Latin America, these products prevent illness and death of their users, and also limit the spread of infectious diseases to others, but they are used sparingly if at all. Some policy-makers think these products should be provided free of charge to low-income families to encourage the use of the products.

Other policy-makers disagree, suggesting that there should be a cost to acquiring these products to discourage wasteful use through better targeting of who gets the products. The question then arises: how will the take-up of the products depend on the price? Will charging even a small price significantly discourage use?

FACT CHECK Adam Smith's claim in the head quote that humans are unique among all animals in our division of labor and exchange of goods is probably right about dogs. But Smith is certainly wrong about the many other species such as ants and other social insects that practice a very advanced division of labor and specialization. Different species of fish exchange services in what are termed "biological markets."

Economists have conducted experiments in 8 countries to find answers to these questions. In the experiments, potential users are randomly selected to be offered the goods free or at one or more different prices. The average use of the products at each price (including zero) is then recorded. Some of their results are shown in Figure 7.3.

Figure 7.3 shows that the effect of charging higher (even if very low) prices

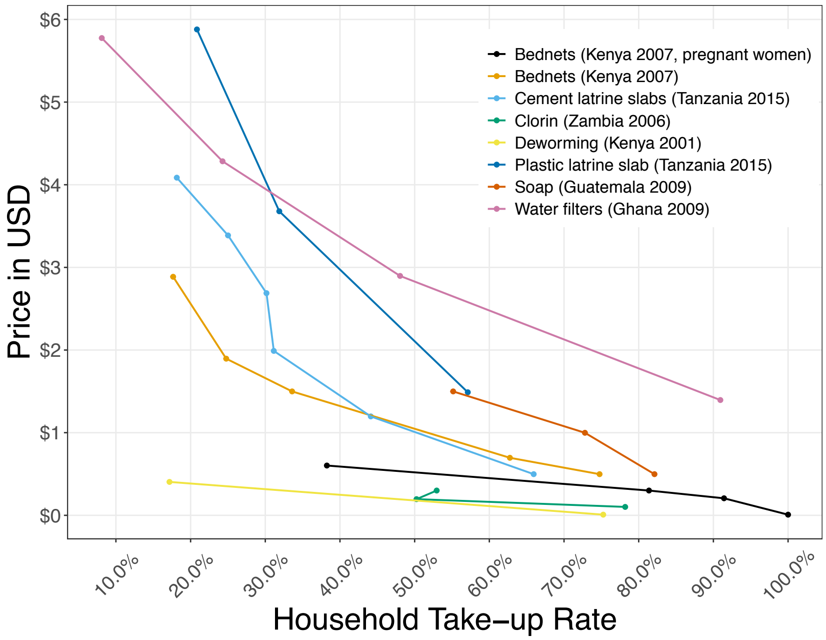


Figure 7.3: **The demand for preventative health products: Take-up rates at various prices and when available for free.** Our measure of demand is the take up rate, that is the fraction of the population that acquires the product (whether free or for a price). For most products the demand is substantially less when even a small price is charged compared with when the good is available for free; and higher prices are also associated with substantially less demand than lower prices. Source: Dupas and Miguel (2017).

was to reduce the amount of the product used, in some cases by a substantial amount.

- In Zambia, for example, increasing the price of a disinfectant tablet from nine cents to twenty-five cents reduced the fraction of the population using the product from 76 to 43 percent.
- Only 43 percent of a group of pregnant Kenyan women purchased insecticide treated mosquito nets when the price was 60 cents; virtually all used the nets when they were provided without charge.
- A program in Kenya that had initially given away de-worming tablets for children, but later introduced a charge of thirty cents per child found that usage of the tablets fell from 75 per cent of the affected population to just 18 percent.

REMINDER Economists have researched preferences based on observing people's behavior in real situations and in experiments. We reviewed some of the key findings from this work in Chapter 2 and we will review more experimental data in later chapters.

On the basis of this information, the Poverty Action Lab at MIT, led by economics Nobel Laureates Abhijit Banerjee and Ester Duflo, suggested that there are good reasons to make these products available without charge or highly subsidized to ensure very low prices.

We begin our analysis of how people spend their money – whether on fish or mosquito nets – with a basic fact: there are limits to how much a family or person can spend.

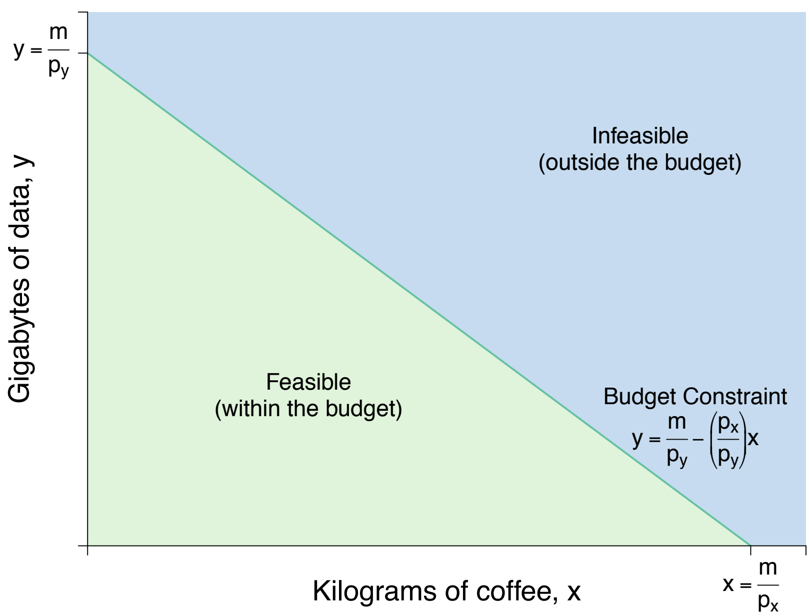


Figure 7.4: **Budget constraint for coffee and data.** The budget set is shaded in green and the budget constraint is the dark green line on the border of the budget set. Consumption bundles (x, y) in the budget set and on the budget constraint can feasibly be obtained with the current budget (m) at going market prices for x and y, p_x and p_y. Outside the budget constraint, in the shaded blue area, the bundles of x and y cannot feasibly be obtained at going market prices with the existing budget.

7.1 *The budget set, indifference curves and the rules of the game.*

To understand how prices influence the take up of one of the life saving health products in Figure 7.3 or the amount of some good that we will consume, think about someone who has a total amount of money to spend m that she has in cash, savings, available credit, and so on.

REMINDER We saw budget constraints in Chapter 3: the slope was a price ratio that represents the opportunity costs of spending money on one good in terms of how much less of the other you can afford depending on your budget.

The budget constraint

We shall consider a person, Harriet, and the decisions she needs to make. A person's budget set states what bundles (x, y) are *feasible* for her to consume given her budget and market price of the goods:

$$m \geq p_x x + p_y y \quad (7.1)$$
$$m \geq \text{Prices of Goods} \times \text{Quantities of Goods Purchased}$$

Expressing this inequality as an equality – assuming that Harriet would not consume less than her budget allowed – we have the budget constraint

$$m = p_x x + p_y y \quad (7.2)$$
$$\text{Budget} = \text{Prices of goods} \times \text{Quantities of goods purchased} \quad (7.3)$$

Equation 7.3 is a statement about prices and budgets. But it is also a statement about the rules of the game and preferences. They are as follows:

- *No gifts, thefts or consumption as a matter of citizen rights:* You can consume only what you pay for; so no gifts or goods provided by government, or acquired by theft.
- *No altruism or concerns about environmental sustainability:* You consume the most that your budget allows and you consume it yourself, rather than giving it to or sharing it with others.

We can rearrange the budget constraint, Equation 7.8, to obtain a line we can draw on the x and y-axes we use for indifference curves:

$$y = \frac{m}{p_y} - \frac{p_x}{p_y}x \tag{7.4}$$

BUDGET SET & BUDGET CONSTRAINT The *budget set* is the set of all feasible purchases of the bundles of x and y with current budget m, such that $m \geq p_x x + p_y y$. The budget constraint is the border of the budget set showing all combinations that exhaust the budget, i.e. for which the constraint holds with equality, $m = p_x x + p_y y$.

We plot the budget constraint in Figure 7.4. Examining the two terms on the right-hand side of Equation 7.4, we can see that if Harriet were to consume only good y and no good x, then she would consume $\frac{m}{p_y}$ units of good y which is the intercept of the budget constraint with the y-axis. As Harriet buys more of good x, she moves along the budget constraint with the slope $-\frac{p_x}{p_y}$ indicating the rate at which she can sacrifice good y for good x. If she were to buy only good x, she could afford $x = \frac{m}{p_x}$ units of good x.

RELATIVE PRICE A relative price shows the price of one or more goods relative to another good, as such it is indicated by a *ratio* of the one price relative to the other. Relative prices show the opportunity cost of having more of one good in terms of the lesser quantity of the other imposed by the budget constraint.

The negative of the slope of the budget constraint is a **relative price** and measures the *opportunity cost* of obtaining good x in terms of the amount of good y that Harriet must sacrifice because her funds are limited. The (negative of the) slope of the budget constraint is another marginal rate of transformation, it tells us the terms on which a reduced amount of good y can be "transformed into" additional amounts of good x while just satisfying the budget constraint.

M-Note 7.1: Budget for coffee and data

For particular values of m, p_x and p_y we can graph the budget constraint. Consider the following example:

- Harriet has a budget ($m = \$50$) to spend on kilograms of coffee, x, and gigabytes of data, y.
- The price of a kilogram of coffee, p_x, is \$10.
- The price of a gigabyte of data, p_y, is \$5

Putting these pieces of data together, therefore, the budget constraint is $50 = 10x + 5y$. We can re-arrange the budget constraint as we did in Equation 7.4:

$$\begin{aligned} y &= \frac{50}{5} - \frac{10}{5}x \\ y &= 10 - 2x \end{aligned} \tag{7.5}$$

Equation 7.5 is a line with an intercept at $\frac{m}{p_y} = 10$ on the y-axis, an intercept of 5 on the x-axis and a slope of $p = -2$. Such a curve would look like bc_1 in Figure 7.5.

M-CHECK From the budget constraint, we know: $y = \frac{m}{p_y} - \frac{xp_x}{p_y}$. We can take the first derivative to see that the slope of the budget constraint is:

$$\frac{dy}{dx} = \frac{-p_x}{p_y}$$

That is the opportunity cost of x in terms of y.

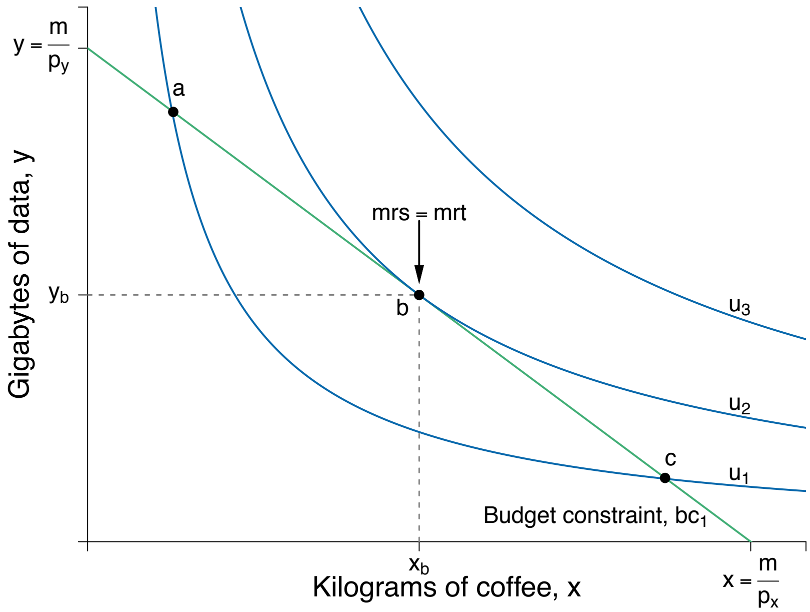


Figure 7.5: **Utility-maximizing consumption bundle.** Harriet maximizes her utility subject to her budget constraint bc_1. At point **a** she consumes too little of x and too much of y (her marginal utility of y is much lower than her marginal utility of x, or her $mrs(x,y)$ is too high, and she would be better off if she consumed less y and more x. Conversely, at **c**, she consumes too little of y and too much of x (her marginal utility of x is much lower than her marginal utility of y, or her $mrs(x,y)$ is too low, and she would be better off if she consumed less x and more y. She maximizes her utility at **b** where her marginal rate of substitution, $mrs(x,y) = \frac{u_x}{u_y}$, equals her marginal rate of transformation or the price ratio of x to y, $mrt(x,y) = \frac{p_x}{p_y}$.

Checkpoint 7.1: Sketching a budget constraint

a. Consider two goods: vegetables (x) which have a price of 4 euros per kilogram and meat (y), which has a price of 10 euros per kilogram. You have a budget of 50 euros a week for meat and vegetables for your family. Sketch your budget constraint.

b. The price of vegetables increases to 5 euros per kilogram. What happens to your budget constraint? Sketch and explain.

Budget constraints, indifference curves and the amount demanded

In Figure 7.5 we show three of Harriet's indifference curves. Remember at any given point on the indifference curve, the negative of its slope tells how much Harriet values the good on the x-axis compared to her valuation of the good on the y-axis, that is, her marginal rate of substitution $(mrs(x,y))$. Her marginal rate of substitution is her willingness to pay to get more of good x, namely how much of good y she would be willing to part with, in order to get one more unit of good x. So,

$$(\text{Negative of})\text{ the slope of an indifference curve} = mrs = \frac{u_x}{u_y}$$

REMINDER We reason here in the same way we did in Chapter 3 that the utility-maximizing choice is the point of tangency between the highest attainable indifference curve and the feasible frontier or budget constraint.

Harriet wants to get to the highest indifference curve that she can, given her budget. This is the point at which the budget constraint is *tangent* to her highest attainable indifference curve. For the two curves to be tangent, the slope of her indifference curve must equal the slope of the budget constraint. Or, the marginal rate of substitution must equal the marginal rate of transformation.

$$mrs(x,y) = \frac{u_x}{u_y} \quad = \quad \frac{p_x}{p_y} = mrt(x,y) \tag{7.6}$$

This is the principle of constrained utility maximization that you learned in Chapter 3 which when we are considering buying in markets we also call the **principle of demand** because it determines how much of a good people will demand or want to buy at given prices. When Harriet satisfies the principle of demand, her trade-offs of one good for another in terms of utility ($mrs(x,y)$) equal the opportunity costs of the two goods in terms of each other ($mrt(x,y)$), where the opportunity costs are given by their prices. Remember that money she spends on one good means money she cannot spend on another good: capturing the essential idea of an opportunity cost.

PRINCIPLE OF DEMAND The principle of demand states that if both goods are consumed, then the utility-maximizing bundle is a point on the budget constraint at which the marginal rate of transformation (the negative of the slope of the budget constraint) is equal to the marginal rate of substitution (the negative of the slope of an indifference curve).

Equation 7.6, the principle of demand, states that if Harriet consumes some of both goods and maximizes utility subject to a budget constraint then the relative valuation of x and y along her indifference curve must equal the opportunity cost of x in terms of y along her budget constraint. Her marginal rate of substitution of y for x equals her marginal rate of transformation of x in terms of y.

The point where Equation 7.6 holds, is Harriet's *utility-maximizing consumption bundle* or the total quantity of goods x and y that Harriet will buy. A useful interpretation of the marginal rate of substitution occurs when good y is not data, but instead the amount of money left over from the budget after purchasing good x. In this case, the marginal rate of substitution can also be thought of as Harriet's **willingness to pay** for the good x, the amount of money she will pay for an additional unit of x when she has already bought the quantity x.

WILLINGNESS TO PAY A person's *willingness to pay* for a good x in terms of y (for example budget left over to buy other goods) is their marginal rate of substitution between y and x when they are already purchasing the bundle (x,y).

We need to separate two ideas that we can derive from the principle of demand:

- *Quantity demanded*: The point (x,y) (the consumption bundle) at which Harriet's $mrs(x,y) = \frac{p_x}{p_y}$ are her *quantities demanded*, these are the amounts of each good that she consumes at the given prices and given income.
- *Demand function*: Using the principle of demand, we can derive a demand function for each good, where the quantity demanded depends on income (m) and prices (p_x and p_y).

M-Note 7.2: The $mrs = mrt$ rule applied to demand for goods

The problem the person faces is to maximize their utility subject to their budget constraint:

$$\text{Vary } x \text{ and } y \text{ to maximize} \quad u(x,y) \tag{7.7}$$

$$\text{subject to the constraint} \quad y = \frac{m}{p_y} - x\frac{p_x}{p_y} \tag{7.8}$$

Substituting Equation 7.8 into Equation 7.7, the problem becomes:

$$\text{Vary } x \text{ to maximize} \quad u(x, \frac{m}{p_y} - x\frac{p_x}{p_y})$$

Taking the partial derivative of the utility function with respect to x, the first order condition for maximum utility is:

$$\begin{aligned}
\frac{\partial u}{\partial x} &= u_x - \frac{p_x}{p_y}u_y = 0 \\
\therefore \frac{u_x}{u_y} &= \frac{p_x}{p_y} \\
mrs(x,y) &= mrt(x,y)
\end{aligned}$$

In the constrained utility maximum is a set of purchases such that the marginal rate of substitution is equal to the relative prices or the marginal rate of transformation.

7.2 *Income, prices and offer curves*

To study how prices and incomes (budgets) affect the demand for goods we ask a hypothetical "what-if" question: how much of good x would someone purchase if her budget were m and the price of good x were p_x and the price of good y were p_y.

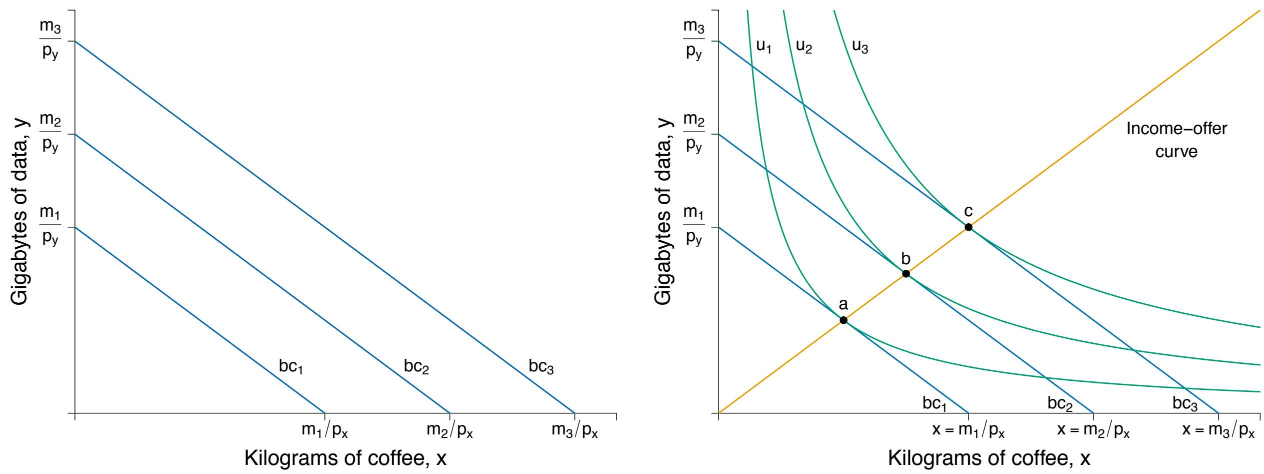
A **demand function** shows the quantity purchased of x that results for the various values of the prices of both goods and the budget, p_x, p_y and m. So $x(m, p_x, p_y)$ is the demand for x as income (m) or the price of x (p_x) and the price of $y(p_y)$ change. We use the term **demand curve** when we refer to the simpler 2-dimensional graphical relationship $x(p_x)$ where we see how the amount purchased varies its price (p_x) varies holding constant all of the other influences on the demand for x.

DEMAND FUNCTION, DEMAND CURVE
A demand function provides an answer a hypothetical "what-if" question: how much of good x would a person purchase if her budget were m, price of goods x and y were p_x and p_y? A demand curve is a 2-dimensional graphical representation of a demand function showing the purchases of x that result for the various values of p_x (with the other influences on demand held constant).

We sometimes use a demand curve in which, instead of quantity sold depending on the price $x = x(p)$, *price* depends on the *quantity* sold, $p = f(x)$. This is called an **inverse demand curve** based on the **inverse demand function**, because it is the mathematical inverse of the conventional demand function.

INVERSE DEMAND FUNCTION, INVERSE DEMAND CURVEThe inverse function answers the hypothetical question: what is the highest price that a person be willing to pay in order to purchase a given amount of some good, given her budget and the prices of other goods? The inverse demand curve is the simplified 2-dimensional graphical representation of this function as $p_x = f(x)$.

The inverse demand function contains exactly the same information and the demand function and the inverse demand curve looks identical to the conventional demand curve (it is downward-sloping). What differs is the hypothetical question for which the inverse demand function provides an answer. Instead of asking how much of a good will be purchased at a given set of prices and a budget, the inverse demand function answers the question: if the budget and the price of the other good are m and p_y what is the maximum price p_x that the buyer would be willing to pay to purchase an amount x of the



good?

Figure 7.6: **Harriet's budget constraint with shifts in income & her income-offer curve.** In panel a. Harriet's budget constraints with three levels of income are shown (m_1, m_2 and m_3) with the corresponding budget constraints bc_1, bc_2 and bc_3 shifting outwards as income increases. In panel b. Harriet's budget constraints are shown tangent to three indifference curves, u_1, u_2, and u_3. The points where they are tangent are where $mrs(x,y) = mrt(x,y)$. The curve joining all the points at which Harriet maximizes her utility as her income changes illustrate her income-offer curve.

A change in income: The income-offer curve

To understand these changes, therefore, we examine Figures 7.6 a. and 7.6 b. As Figure 7.6 a. shows, as Harriet's income changes, her budget constraint shifts. That is, the intercept with both axes, $\frac{m}{p_y}$ and $\frac{m}{p_x}$, shifts up as income (m) goes up and shifts down as her income goes down. Consider the three budget constraints in Figure 7.6 a where only income changes, but the prices of the two goods do not change.

- *Status quo*: She starts with an income of m_2 with intercepts $\frac{m_2}{p_y}$ and $\frac{m_2}{p_x}$.
- *Income decrease*: If her income *decreases* to m_1, then the intercepts of her budget constraint shift downwards and to the left to $\frac{m_1}{p_y}$ and $\frac{m_1}{p_x}$, so she can buy *less* of both goods.
- *Income increase*: If her income *increases* to m_3, then the intercepts of her budget constraint shift upwards and to the right to $\frac{m_3}{p_y}$ and $\frac{m_3}{p_x}$, so she can by *more* of both goods.

INCOME-OFFER CURVE The income-offer curve is the path traced out by the points (x,y) that maximize the decision-maker's utility as money-income, m, increases, holding the price of good x, p_x, constant.

Considering different levels in Harriet's income we can superimpose Harriet's indifference curves to find the consumption bundle for each income level that would maximize Harriet's utility and satisfy the principle of demand. The path traced out by the points (x,y) as m increases is called her *income-offer curve*. Her income-offer curve is also called her *expansion path* because it shows the effect of expanding her feasible set (by increasing her budget). In Figure 7.6 b, her income-offer curve is upward-sloping, showing the effect of an increase on her income on her consumption of the goods, x and y. As she gets more income, she would consume more of both goods.

The income-offer curve allows us to understand the *types* of goods people consume.

- *Normal goods*: normal goods are goods like coffee and data where people buy *more* as their income *increases*, or *less* of them as their income *decreases*.
- *Inferior goods*: inferior goods are goods like cheap staples, like white sandwich bread, basic rice, or instant noodles: people tend to consume *less* of inferior goods as their income *increases* and *more* of them as their income *decreases*.

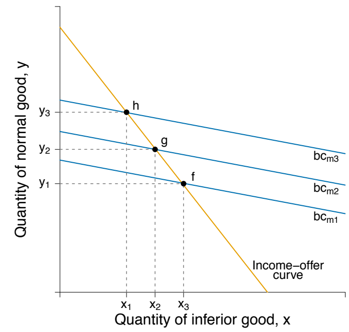


Figure 7.7: **Inferior Goods and An Increase in Income.** The downward-sloping income-offer curve for an inferior good (x) and a normal good (y). As the person's budget increases, they consume less of good y, showing that y is an inferior good. Consumption of good x increases as income increases, showing that x is a normal good.

Figure 7.7 shows a situation in which Harriet's income increases, but her consumption responses for the two goods differs. For good y, on the vertical axis, Harriet consumes more of it as her income *increases* from m_1 to m_2: she *increases* her consumption from y_1 to y_2. For good x, on the horizontal axis, on the contrary, Harriet consumes less as her income *increases* from m_1 to m_2: she *decreases* her consumption from x_2 to x_1 as her income increases. As a result, she has a *downward-sloping* income-offer curve.

Checkpoint 7.2: Inferior indifference curves

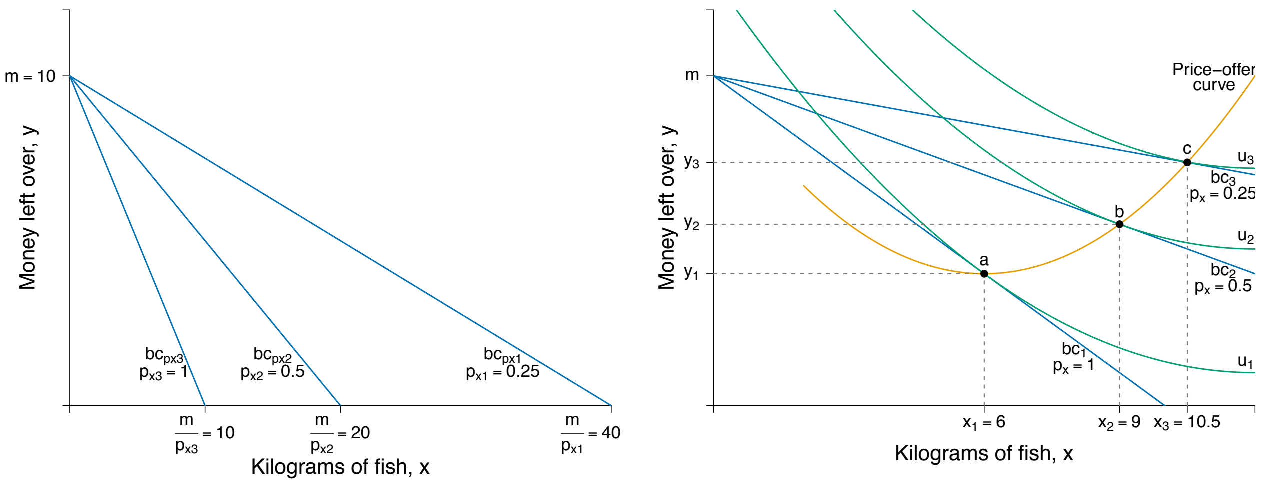
1. On your own set of axes, re-draw Figure 7.7. What condition must be true at each of points **f**, **g** and **h**?
2. Add the relevant indifference curves to your figure. What do you think they look like? Explain.

A change in prices: the price-offer curve

As prices increase, Harriet has to reduce her purchases of at least one good, and possibly both goods because she'll have less money to spend on the goods if they become more costly. As the price of a good changes, holding the price of the other good constant, the slope of the budget constraint will change. For example, if the price of good x changes while the price of good y remains the same, the budget constraint will pivot around the y-axis intercept depending on whether the price of x increases or decreases. In Figure 7.8 a. on the left-hand side we show budget constraints based on three prices for good x:

- An initial price along budget constraint bc_2.
- A price *increase*, which steepens the slope of the budget constraint to bc_1.
- A price *decrease*, which flattens the slope of the budget constraint to bc_3.

At the point of tangency of each budget constraint with a corresponding indifference curve in panel b., we can see how Harriet would respond to a change in the price of goods x and y, other things equal. The path traced out by the



(a) Budget Constraint & changes in the price of x

(b) The price-offer curve

Figure 7.8: **Harriet's budget constraint with changes in price & her price-offer curve.** In panel a. Harriet's budget constraints for three prices for fish, good x, p_{x1}, p_{x2} and p_{x3} are shown with the corresponding budget constraints pivoting inwards for a price increase (such as for bc_{px2} to bc_{px3}) or pivoting outwards for a price decrease (such as for bc_{px2} to bc_{px1}). The other good, y, is money for other goods she can purchase. The intercept for good y is $\frac{m}{p_y}$. Because the price of money for other goods is $p_y = 1$, this simplifies to m as in Figure b. Note the x-axis range differs in panels a and b. In panel b. Harriet's budget constraints are shown tangent to three indifference curves, u_1, u_2, and u_3. The points where the budget constraint and indifference curves are tangent are where the marginal rate of substitution $(mrs(x,y))$ equals the marginal rate of transformation $(mrt(x,y))$. The curve joining all the points at which Harriet maximizes her utility as the price of good x changes is her **price-offer curve**.

number $(x(m, p_x), p_y)$ as p_x changes (holding the budget, m and p_y constant) is called her price-offer curve. We illustrate the price-offer curve in Figure 7.8 b.

Points **a**, **b** and **c**, all map her utility-maximizing choices and the curve that connects these points is the price-offer curve as we introduced in Chapter 3.

7.3 *Cobb-Douglas utility and demand*

You already encountered Cobb-Douglas utility in Chapter 3. We build on that base and explore a person's choice of her utility-maximizing consumption bundle using indifference curves based on a Cobb-Douglas utility function and budget constraints. The Cobb-Douglas utility function has the following general form:

$$u(x,y) = x^{\alpha}y^{1-\alpha} \tag{7.9}$$

With α and $1-\alpha$ indicating the relative strength of preference for the two goods and the intensities sum to 1.

Using Cobb-Douglas utility, we can illustrate indifference curves for each good as the price of the good changes and we can derive a demand curve for each good. In Figure 7.9, we show the indifference curves for two goods: kilograms of coffee (x) and gigabytes of data (y). The marginal rate of substitution is:

$$mrs(x,y) = \frac{u_x}{u_y} \quad = \quad \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{y}{x}\right) \tag{7.10}$$

REMINDER The marginal rate of substitutions means that for a given ratio of the quantities of the two goods x and y making up a bundle, how much the person values a small increment in the amount of x compared to how much they value a small increment of y is given by the ratio of the exponents. For example, if $\alpha = 0.75$, and $x = y$, then the marginal utility of x is three times the marginal utility of y (because $\alpha/(1-\alpha) = 3$).

The principle of demand, then requires that,

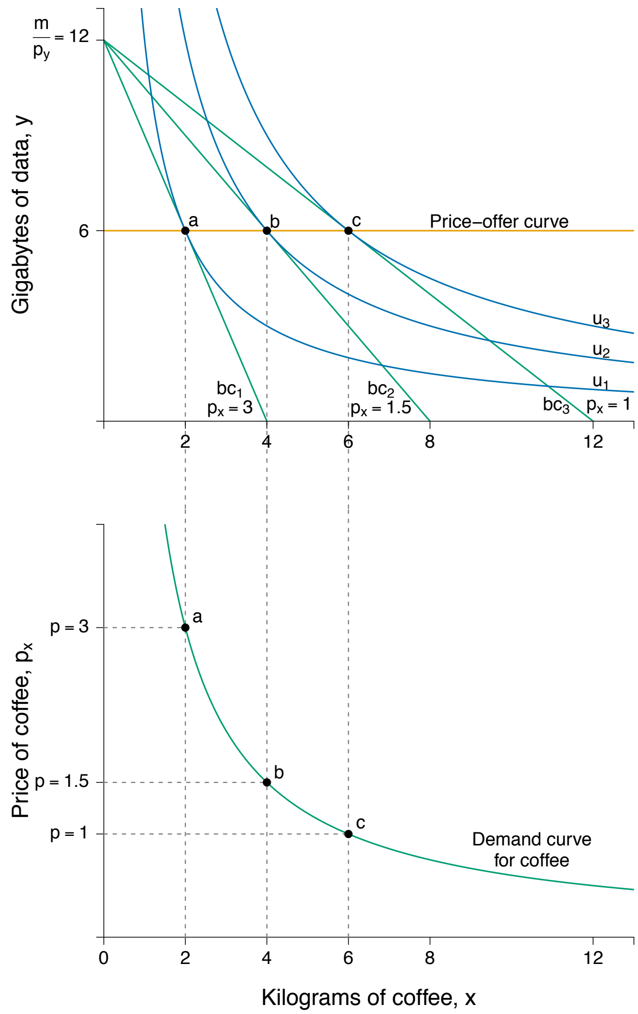


Figure 7.9: **Harriet's price-offer curve & demand curve for coffee.** In the upper panel we show three of Harriet's budget constraints corresponding to three different prices: $p = 3$ for bc_1, $p = 1.5$ for bc_2 and $p = 1$ for bc_3. At points **a**, **b**, and **c** tangent to three indifference curves, u_1, u_2, and u_3 The curve joining all the points at which Harriet maximizes her utility for different prices of good x is her **price-offer curve**.

In the lower panel, we see how the principle of demand in the top figure is the basis of the demand curve. The prices shown on the vertical axis are the (negative of the) slopes of the budget constraints in the top figure. We have assumed that $\alpha = 0.5$, and that her budget for coffee and data is \$100.

$$mrs(x,y) = \left(\frac{\alpha}{(1-\alpha)}\right)\left(\frac{y}{x}\right) \quad = \quad \frac{p_x}{p_y} = mrt(x,y) \tag{7.11}$$

From this relationship and the budget constraint as we show in M-Note 7.4, we can derive a demand curve for each good. The demand function for kilograms of coffee, good x is:

$$\text{Demand} \qquad x(m, p_x) = \frac{\alpha m}{p_x} \tag{7.12}$$

Equation 7.12 shows a relationship between quantity demanded (x) and price (p_x) such that the quantity demanded *decreases* as the price increases, or the quantity demanded *increases* as the price *decreases*.

Rearranging equation 7.12 as $\frac{x \cdot x}{m} = \alpha$ with a Cobb-Douglas utility function, a person conforming to the principle of demand will spend a *fraction* of their total budget on x, that is:

- equal to the exponent of x in the utility function namely, α a constant, and therefore is,
- independent of the price of x and the price of y.

The fraction of the budget spent on the *other* good is also independent of changes in the price of x. So, because the budget m has not changed, the amount spent on y will also remain the same. Because the price of the other good has not changed, the amount of good y purchased is also unchanged.

Thus, With Cobb-Douglas utility for a given price of y, p_y and income (m), the only thing that differs with different prices of x, is the quantity demanded of good x. This is why the price-offer curve in Figure 7.9 is horizontal.

Similarly the amount spent on x does not depend on the price of y, which you can confirm from the fact that p_y does not appear in the demand function for x. These are not general features of demand functions, they are specific to the Cobb-Douglas utility function.

M-Note 7.3: Marginal rate of substitution, Cobb-Douglas utility function

Consider a Cobb-Douglas utility function:

$$u(x,y) = x^{\alpha}y^{1-\alpha}$$

The marginal utility with respect to of each good is:

$$\begin{aligned} u_x = \frac{\partial u}{\partial x} &= \alpha x^{\alpha-1}y^{1-\alpha} \\ u_y = \frac{\partial u}{\partial y} &= (1-\alpha)x^{\alpha}y^{-\alpha} \end{aligned}$$

Therefore, the marginal rate of substitution of x with respect to y is:

$$mrs(x,y) = \frac{u_x}{u_y} = \frac{\alpha x^{\alpha-1}y^{1-\alpha}}{(1-\alpha)x^{\alpha}y^{-\alpha}} \tag{7.13}$$

Note that:

$$\begin{aligned} \frac{x^{\alpha-1}}{x^{\alpha}} &= \frac{1}{x} \\ \text{and} \quad \frac{y^{1-\alpha}}{y^{-\alpha}} &= y \end{aligned}$$

Thus, the marginal rate of substitution (Equation 7.13) becomes:

$$mrs(x,y) = \frac{u_x}{u_y} = \left(\frac{\alpha}{(1-\alpha)}\right)\frac{x}{y}$$

The inverse demand function

We can re-arrange the function and find the **inverse demand curve**. The inverse demand curve is:

$$\text{Inverse demand} = p_x(x,m) = \frac{\alpha m}{x} = \frac{\text{amount spent on x}}{\text{amount of x purchased}} = \text{price}$$

Here, we have a downward-sloping demand curve where price *decreases* as the quantity demanded *increases*. We show this demand curve in the lower panel of Figure 7.9. As previously, the demand curve is derived from a combination of points shown by the price-offer curve from each point satisfying the principle of demand from constrained utility maximization providing the price-offer curve in Figure 7.9.

Checkpoint 7.3: Changes to prices or income with Cobb-Douglas utility

Harriet not only buys fish, but she also buys coffee and cookies to fuel herself while running her fish-buying business. Her utility function for cookies (x) and cups of coffee (y) is given by the following utility function:

$$u(x,y) = x^{0.6}y^{0.4} \tag{7.14}$$

We assume that Harriet has a weekly budget of \$10 to spend on coffee and cookies, where the price of a cup of coffee is \$3 and the price of a cookie is \$0.50.

a. Sketch Harriet's indifference curves, her budget constraint, and find her utility-maximizing consumption bundles of cookies and coffee.
b. Assume that the price of cookies now increases to \$2 per cookie. What would happen to the quantity of cookies that Harriet demands?
c. Harriet now has more money to spend on cookies and coffee on her way to work. Assume that her budget for the goods has increased to \$40 and the price of a cup of coffee is \$3 and the price of a cookie is \$0.50. How much of each would she consume with her greater budget?

M-Note 7.4: Cobb-Douglas Demand Functions

The Cobb-Douglas utility function is:

$$u(x,y) = x^{\alpha}y^{1-\alpha}$$

where $0 < \alpha < 1$. The individual maximizes this function subject to a budget constraint:

$$m = p_x x + p_y y \tag{7.15}$$

The principle of demand tells us that at the utility-maximizing bundle, the slope of the indifference curve – the marginal rate of substitution – equals the marginal rate of transformation – the slope of the budget constraint which is also the ratio of the prices of the two goods. We found in M-Note 7.3 the following:

$$mrs^{CD}(x,y) = \left(\frac{\alpha}{1-\alpha}\right)\frac{y}{x} \tag{7.16}$$

So the utility-maximizing bundle must satisfy the following equation:

$$mrs^{CD}(x,y) = \left(\frac{\alpha}{1-\alpha}\right)\frac{y}{x} = \frac{p_x}{p_y}$$

$$\therefore p_y y = p_x x\left(\frac{1-\alpha}{\alpha}\right) \tag{7.17}$$

To find the demand function, substitute Equation 7.17 into the budget constraint, Equation 7.15, to isolate a value for x, which we then use to find y:

$$\begin{aligned} m &= p_x x + p_x x\left(\frac{1-\alpha}{\alpha}\right) \\ m &= \frac{(\alpha+1-\alpha)p_x x}{\alpha} \\ \therefore x(m,p_x,p_y) &= \frac{\alpha m}{p_x} \end{aligned}$$

Substitute $p_x x$ into Equation 7.17 to find $p_y y$ and y:

$$p_y y = (1-\alpha)m$$

y(m,p$_x, p_y) \;=\; \frac{(1-\alpha)m}{p_y}$ We have therefore found the demands $(x(m,p_x,p_y), y(m,p_x,p_y))$ as functions of the budget, m, and the prices of the goods, p_x and p_y, given the intensity of preferences for the goods, α. Notice that the demand for each good is independent of the price of the other good. The demand function for each good in terms of its own price is a hyperbola.

Checkpoint 7.4: Income-offer and price-offer curves

1. Sketch the income-offer curve for Cobb-Douglas utility. Draw three Cobb-Douglas indifference curves with three income levels and sketch the corresponding income-offer curve.
2. The price-offer curve for Cobb-Douglas utility is a horizontal line in (x,y) space. Make sure that you can sketch a similar figure to the one we used in Figure 7.9. Notice, though, that in that figure $\alpha = 1-\alpha = 0.5$. In Chapter 3 we used $\alpha = 0.6$ and $1-\alpha = 0.4$ to consider different preferences for Living (x) and Learning (y). How would these different exponents change the slopes of the indifference curves? Sketch *approximately* the corresponding price-offer curve for two goods, coffee (x) and cookies (y) for three different prices of coffee.

7.4 Application. Doing the best you can dividing your time

We can apply the Cobb-Douglas utility function to a problem we all face: how to divide up the limited number of hours in our day between all of the things we would like to do, or must do to make a living. We simplify the problem by limiting our objectives to only two things: free time and consumption (similar to the problem involving Living and Learning in Chapter 3). Because we pay for our consumption with the wages we receive for working, and working means not having free time, we face a trade-off: more free time means less consumption (and of course vice versa).

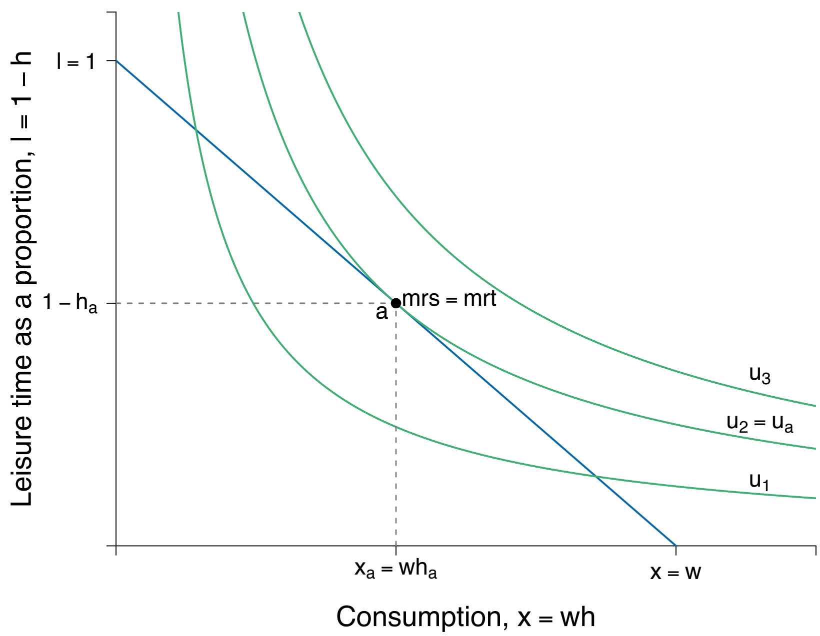


Figure 7.10: **Feasible set and indifference curves for working time and consumption** The choice of working time (and hence free time) and consumption. The worker chooses a level of consumption wh where their consumption is equal to their working hours (h) multiplied by their wage (w). The worker balances their consumption (x) against their leisure time as a proportion of their day.

A trade-off between free time and consumption

Consider a worker, Scott, deciding about how much leisure and consumption he would like. We define h as the fraction of the day that Scott spends working for wages, with $l = 1 - h$ the fraction of that day that is free time (l). Scott consumes his entire income, so his daily consumption, x, is the total income he would receive if he worked 24 hours, w, times the faction of the day that he works.

$$\begin{aligned} \text{Consumption} &= \text{Wages} \times \text{Proportion of day spent working} \\ x &= wh \end{aligned} \tag{7.18}$$

Scott's utility is given by the following Cobb-Douglas function that expresses how he values consumption (x) and free time ($l = 1 - h$):

$$u(x,h) = x^{\alpha}(1-h)^{1-\alpha} \tag{7.19}$$

Where, as before, $0 < \alpha < 1$ and the size of α indicates the relative intensity of preferences for the two goods. Because $x = wh$ we can rewrite this as

$$u = (wh)^{\alpha}(1-h)^{1-\alpha} \tag{7.20}$$

M-CHECK We can rewrite the budget constraint as

$$1 - h = 1 - \frac{x}{w}$$

From which we see that

$$\frac{d(1-h)}{dx} = \frac{-1}{w}$$

Which is the effect of greater consumption on feasible free time.

In Figure 7.12 we show Scott's feasible set of choices between consumption and free time and three indifference curves representing Scott's preferences for the two goods. The feasible frontier is your budget constraint. The maximum that Scott could spend on consumption is to have *no* free time and to set working time at 1 allowing a total expenditure of w on consumption. So w is analogous to m in the previous budget constraints and this limits expenditure on consumption x and the time not working, valued at the wages Scott would have received had Scott worked the entire day:

$$\text{Budget constraint:} \qquad w \geq x + (1-h)w \tag{7.21}$$

We can re-arrange Equation 7.21 as follows:

$$\text{Re-arranged constraint} \qquad x \leq wh \tag{7.22}$$

Equation 7.22 therefore requires that the value of Scott's consumption x not be greater than wages for time worked wh.

The negative of the slope of the budget constraint is the marginal rate of transformation of reduced consumption into free time. This is how much additional free time Scott is able to have by giving up one unit of consumption. The equation for the budget constraint (Equation 7.21) and Figure 7.10 show that the marginal rate of transformation is as follows:

$$-\frac{d(1-h)}{dx} = mrt(x, 1-h) \quad = \quad \frac{1}{w} \tag{7.23}$$

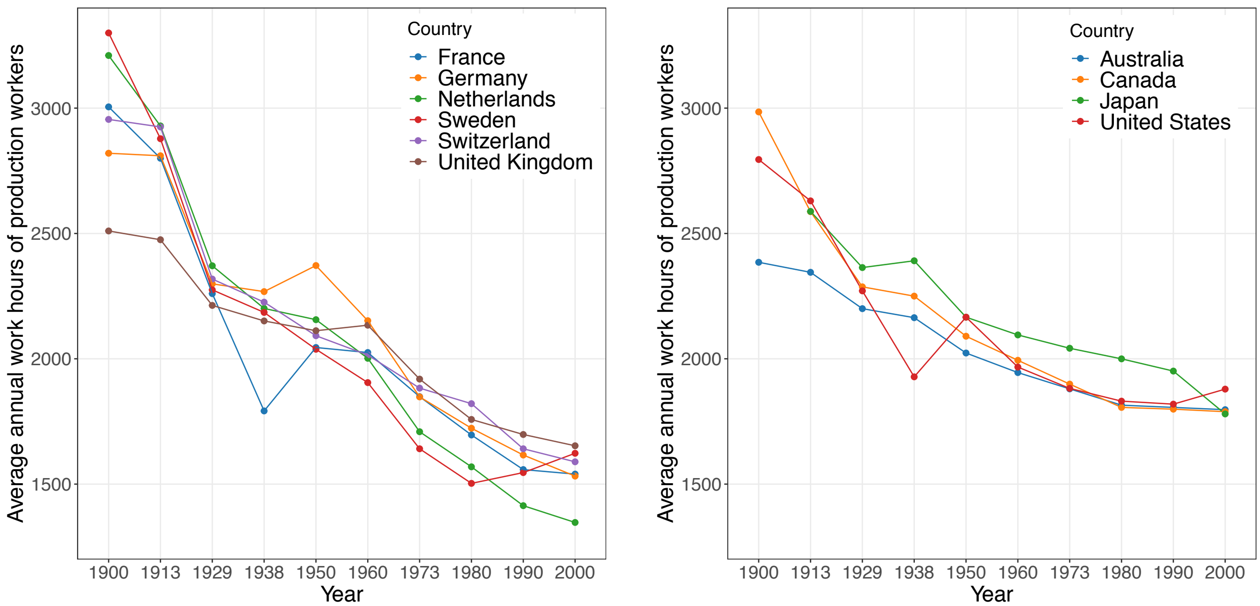
The slope of Scott's indifference curves is the marginal rate of substitution between consumption and free time. As shown in Equation 7.16 the marginal rate of substitution derived from a Cobb-Douglas utility function is the ratio of the x exponent to the y exponent times the ratio of the value of the y variable to the x variable or:

$$mrs(x, 1-h) \quad = \quad \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{y}{x}\right)$$

The y variable here is the amount of free time $1-h$ and x is the level of consumption which is $x = wh$, so

$$mrs(x, 1-h) \quad = \quad \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{x}\right)$$

The best Scott can do in this constrained optimization problem, maximizing his utility subject to his budget constraint, is to select the bundle $(x, 1-h)$ such that the marginal rate of substitution is equated to the marginal rate of transformation. At his utility-maximizing bundle the fraction of the day Scott will work, h, is equal to α (see M-Note 7.5). As a result, the fraction of Scott's day that is free time will be $1-\alpha$, because $1-\alpha$, the exponent of "free time" in your utility function is a measure of how important free time is to you.



(a) Work hours over time in European Countries

(b) Work hours over time in countries outside of Europe

Figure 7.11: **Work hours over time for a variety of countries**. The data refer to annual average work hours for full-time production workers (meaning, excluding supervisory personnel) Source: Huberman (2004)

7.5 Application: Social comparisons, work hours and consumption as a social activity

In Chapter 3 we looked at data on how men and women spend their time, and the increase in the amount of time women spend working for pay (the female labor force participation rate). Here we use the constrained optimization model to help understand another dramatic change in time use over the last century.

Figure **??** shows that in every country on which we have data people have been working less. But there are important differences among the countries:

- In the Netherlands work hours fell from the equivalent of 62 hours 52 weeks of the year to less than 27 hours per week. In Sweden, where work hours also declined dramatically
- Work hours declined much less in the U.S. than in most other countries – a decline of 32.77 percent compared to a decline of 58.04 percent in the Netherlands.
- In the US, as in Sweden there was a slight increase in work hours at the end of the last century.

How can our model help explain what explains these differences? We modify the model of choice of work hours to help us understand the differences among the countries and the changes over time shown in Figure 7.13.

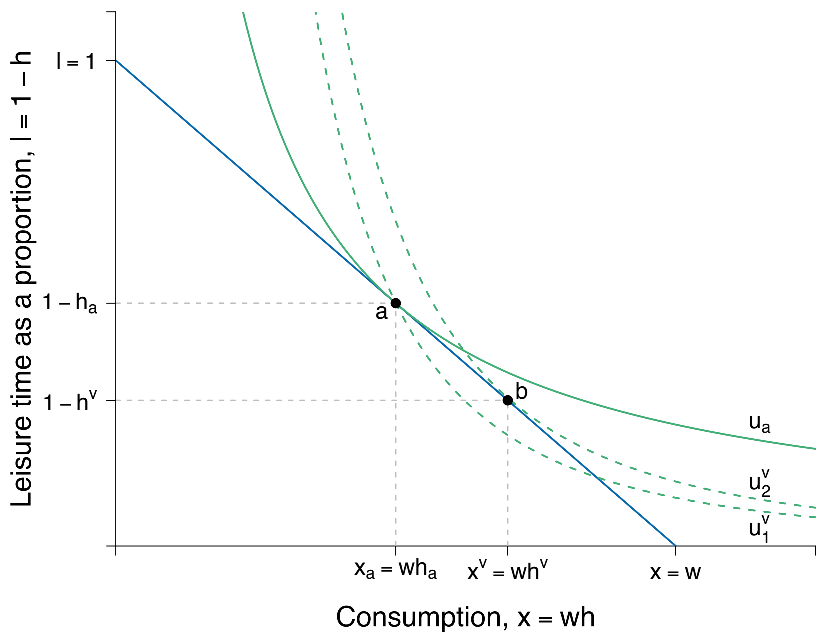


Figure 7.12: **Feasible set and indifference curves for working time and consumption: Veblen effect.** The worker now experiences a "Veblen effect" which changes their preferences over work time and leisure, altering the slope of their indifference curves and therefore resulting in a new utility-maximizing point. At the new utility maximum, they consume more goods, work more and consume less leisure.

The new idea is that what people consume - the quality, quantity and expense of what someone wears, or drives, or eats - is a signal to others and to themselves about where they stand in society *relative* to other people. That is, people judge their own level of consumption relative to other people's consumption, not based on the level of consumption alone.

HISTORY The term "conspicuous consumption" comes from the American economist and sociologist Thorsten Veblen (1857-1929). Well over a century ago he described exactly the trade off we are modeling here. "The means of communication and the mobility of the population now expose the person to the observation of many persons who have no other means of judging his reputability than the display of goods... the present trend of the development is in the direction of heightening the utility of conspicuous consumption as compared with leisure."[1]

Veblen effects, conspicuous consumption and working time

A particular variant of this view of consumption as social signaling terms the things we purchase in order to impress ourselves or others *conspicuous consumption* and it is the high income people who set the standards for everyone else. One way to model this is to say that we compare our consumption to that of the very rich, and the closer our consumption is to the consumption of the rich, the better we feel.

To do this we now define "effective consumption" as how our consumption feels to us given what others are consuming. To capture this idea, we define effective consumption as follows:

$$\begin{aligned} \text{Effective consumption} &= \text{Consumption} - \text{Veblen Effect} \times \text{Consumption of the Rich} \\ x^v &= x - v\underline{x} \end{aligned} \tag{7.24}$$

Where, x is Scott the worker's consumption, $\underline{x}$ is the consumption of the rich, and v is a positive constant representing the Veblen effect.

The negative effect of the consumption of the rich on our utility is captured in the term v (named after Thorstein Veblen). Effective consumption defined in this way expresses the idea that the consumption of the rich has the effect of diminishing the adequacy that we feel for any particular level of consumption. This raised the marginal utility of greater consumption to compensate. Scott's utility now includes this idea:

$$u = (x - v\underline{x})^{\alpha}(1-h)^{1-\alpha} = (wh - v\underline{x})^{\alpha}(1-h)^{1-\alpha} \quad (7.25)$$

Using equation 7.16 for the $mrs(x, 1-h)$ with a Cobb-Douglas utility function, Scott's marginal rate of substitution is now:

$$\begin{aligned} mrs(x, 1-h) &= \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{x^v}\right) \\ &= \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{1-h}{x - v\underline{x}}\right) \end{aligned} \quad (7.26)$$

HISTORY The consumption standard setters, according to Veblen, are the rich. He wrote: "all canons and reputability and decency and all standards of consumption are traced back... to the usages and thoughts of the highest social and pecuniary class, the wealthy leisure class."[2]

Because increased consumption by the rich ($(\underline{x})$) has diminished your level of effective consumption, it has raised the $mrs(x, 1-h)$ (you can see this in the second term in the expression above). So how much Scott values consumption relative to how much he values free time is now greater than before. This means that in Figure 7.12 it has made the indifference curves steeper.

Figure 7.10 shows an initial choice that the worker faces *without* the Veblen effect, that is, when he does *not* worry about what others consume. Figure 7.12, as a contrast, shows how the "Veblen effect" of the consumption of the rich affects the worker's choice. The Veblen effect does not alter the feasible frontier, but it changes the marginal rate of substitution between consumption and free time. The indifference curves are steeper because at any level of *actual* (*not* effective) consumption and free time, the marginal utility of effective consumption has risen. Why? Because when Scott compares himself to the rich it makes him feel as though he has *less* effective consumption than he actually has.

Figure 7.12 shows that when Scott maximizes utility by setting the $mrt(x, 1-h)$ to the $mrs(x, 1-h)$ Scott works longer hours, enjoys less free time, and increases the amount of consumption he purchases (see the M-Note 7.5). The time Scott works now is greater than before, as shown in Figure 7.10 (see the M-Note 7.5):

$$h^v = \alpha + \frac{(1-\alpha)v\underline{x}}{w} \quad (7.27)$$

Equation 7.27 shows the working time *without* the Veblen effect (namely, α) *plus* more time that Scott is now motivated to work due to the Veblen effect (namely, the second term on the right hand side of Equation 7.27).

FACT CHECK In 2001 the tax authorities in Norway began posting income tax records online, so that anyone could find out the income of their neighbors, friends and co workers. Huge numbers accessed the site. Ricardo Perez-Truglia studied the statistical relationship between Norwegian's income and measures of their "subjective well-being" – happiness and life satisfaction. Higher income people were happier and had greater life satisfaction. But after incomes became

Did Veblen effects and falling inequality explain declining work hours in the 20th century?

How does this model help us understand how working hours have changed over time and how work hours differ across countries? The model predicts that the more that rich people consume, the longer other people will work. So we would expect people to work longer hours in countries in which the rich are especially rich, and people to work less where the rich are only modestly richer than the rest.

Figure 7.13, presents the average annual working hours and a measure of the fraction of all income received by the richest 1 percent of households in ten countries over the 20st century. The fraction of income received by the very rich is a measure of the key variable in the model, the consumption of the very rich divided by the wages of typical workers or $\frac{x}{w}$.

The figure shows that this prediction of the model with Veblen effects is borne out by the data: longer working time is associated with a larger share of income going to the very rich.

But it shows more: the decline in the relative incomes of the very rich is closely associated with the decline in work hours. Notice that Sweden is both the most unequal and "longest working" nation (in the early years of the data set) and also the most equal and country with the most free time (in the more recent years). The countries that became more equal over this period also saw the greatest drop in work hours. The increase in work hours in both Sweden and the U.S. at the end of the last century was associated with an increase in inequality in both countries.

In this model, conspicuous consumption by the very rich is a kind of "public bad." It is experienced equally by all members of the society or at least can be by anyone with access to TV and social media (it is *non-excludable*). It is a "bad" and not a good because it reduces the utility of those it affects (it is *non-rival* in the disutility it creates). And the people affected then respond in ways that generate further negative external effects because the Veblen effect induces them to work and consume more, increasing the use of our limited environmental resources.

The Veblen effects model suggests some of the reasons for differing working hours among countries. But by itself it misses some important parts of the story. The most important missing element for the decline in work hours in the 20th century is that voting rights were extended to include most adults early in the century. When overworked employees got the right to vote, in virtually all countries their trade unions and political parties demanded reductions in working hours.

The model with Veblen effects and the data provide an illustration of a more

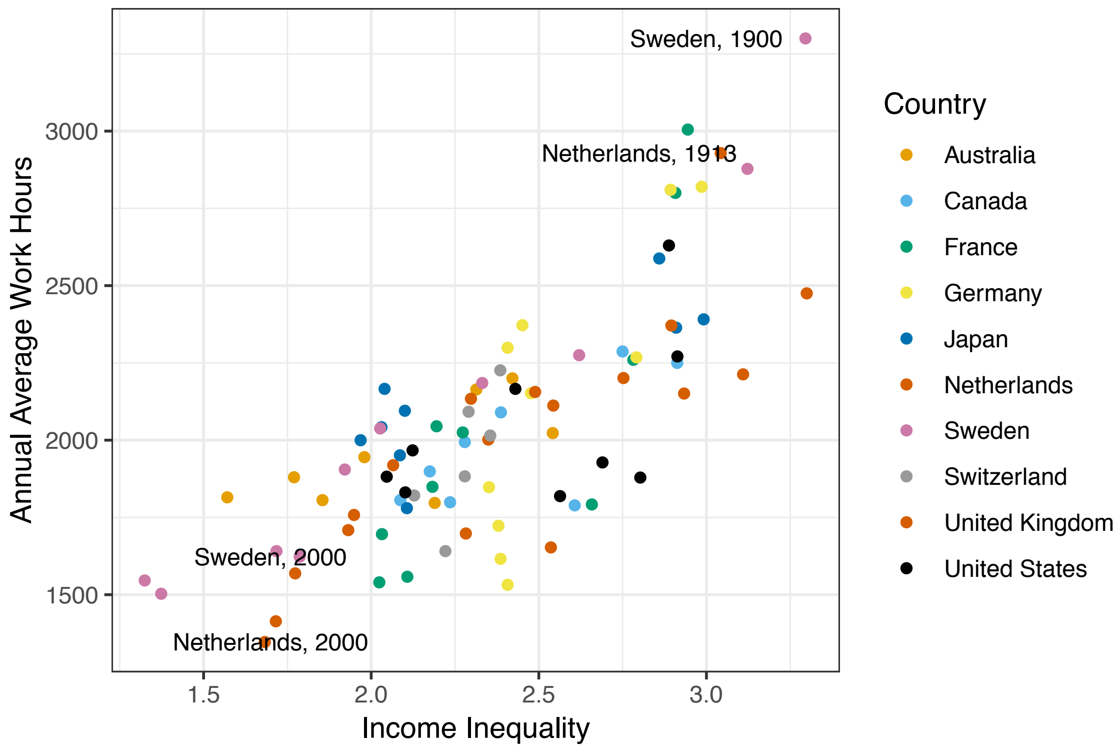


Figure 7.13: **Inequality and work hours 1900 to 2000.** The hours data are annual average work hours for full-time production workers. The income data are based on the share of total income received by the top 1 percent of households. Source: Oh, Park, and Bowles (2012).

general point: consumption is not just a biological activity. Eating is not just nutrition. Clothing is not just keeping warm. Your home is more than four walls to keep out the weather. Consumption is a social activity. As Veblen said our consumption is a *signal* to others and to ourselves about who we are. It is also a social activity in which we engage, for example, for the pleasure of the company of our friends.

M-Note 7.5: The choice of work hours with & without a Veblen effect

Using equation 7.25, Scott's utility function when deciding between leisure and consumption may be represented as:

$$u = (x - v\underline{x})^{\alpha}(1-h)^{1-\alpha}$$

With $v \geq 0$. When $v = 0$, there is no Veblen effect, and we have same utility function as in Equation 7.19 that we used in the previous section. When $v > 0$, there is a Veblen effect.

To calculate the time worked, we need to equate the $mrs(x, 1 - h)$ to the $mrt(x, 1 - h)$.

Following 7.26 and 7.23:

$$\begin{aligned}
\text{mrs} \quad \frac{\alpha}{1-\alpha}\frac{(1-h)}{(x-v\underline{x})} &= \frac{1}{w} \quad \text{mrt} \\
\text{Using } x = wh \quad \frac{\alpha}{1-\alpha}\frac{(1-h)}{(wh-v\underline{x})} &= \frac{1}{w} \\
\alpha w(1-h) &= (1-\alpha)(wh-v\underline{x}) \\
\alpha w - \alpha wh &= wh - \alpha wh - (1-\alpha)v\underline{x} \\
wh &= \alpha w + (1-\alpha)v\underline{x} \\
h^v &= \alpha + \frac{(1-\alpha)v\underline{x}}{w}
\end{aligned}$$

In the absence of Veblen effect ($v = 0$), the hours worked h depends only on the importance of consumption relative to and free time in the utility function, α. A positive Veblen effect ($v > 0$) reduces effective consumption. Because there are diminishing returns to effective consumption (marginal utility of effective consumption is lower the more of it you have) the effect of there being less effective consumption is to increase the marginal utility of it. As a consequence, Scott increases his working hours (and consumption) and reduces his leisure.

You can also see that the higher the consumption of the rich, the lower is *effective* consumption, and therefore the higher the hours worked (See Figure 7.12).

Checkpoint 7.5: Veblen effects

- What do you think explains the magnitude of v, the parameter governing the size of the Veblen effect?
- Juliet Schor, an economist, found that people who watch TV more save less. Saving is not included in our model, but how do you think this result might be explained by some kind of Veblen effect?
- How could our model explain the evidence in the Fact Check about the effects of making incomes public on the relationship between income and subjective well being? How would that affect v the Veblen effect coefficient?

7.6 Quasi-linear utility and demand

QUASI-LINEARITY When a function is quasi-linear it depends *linearly* on one variable, e.g. y, and *non-linearly* on another variable, e.g. x, and has the form $u(x,y) = y + g(x)$. Hence it is *quasi* or "partly" linear.

We don't always want to know how people trade off the benefits of two commodities – like coffee and data – in their utility functions. Sometimes, we find it useful to think about how a person will trade off money left over for other purchases (y) and a commodity (x), like we saw earlier in the example of kilograms of fish (x) and money for other goods (y). We explore this idea using a Quasi-linear (QL) utility functions with the form:

$$u^{QL}(x,y) = y + g(x) \tag{7.28}$$

Utility in the quasi-linear function depends linearly on y. The more y Harriet has, the higher her utility. Her marginal utility for y is always 1 and it does not decline as she gets more y. These properties make y more like wealth measured in terms of money than like a particular good such as data or coffee, so

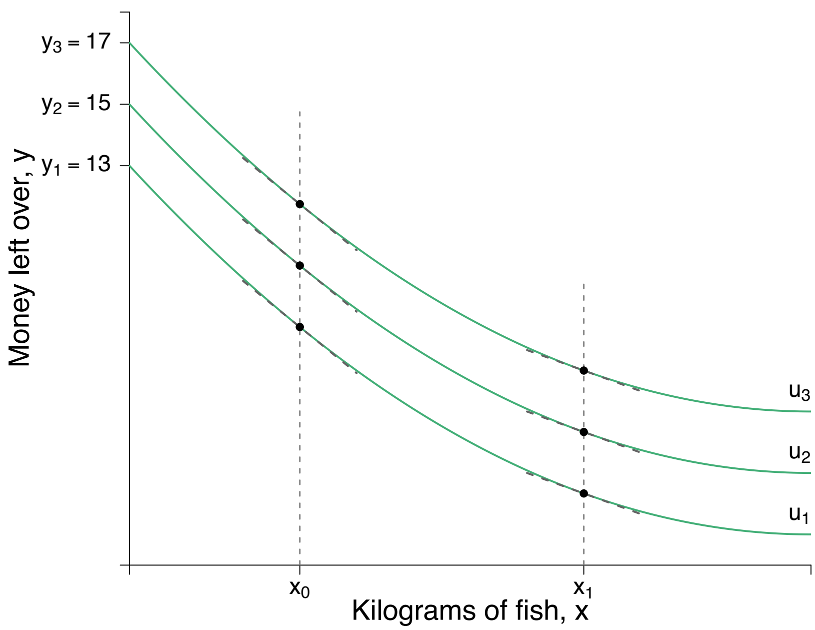


Figure 7.14: **Harriet's indifference curves: quasi-linear utility.** With quasi-linear utility, marginal rates of substitution depend only on the amount of the good x, and not at all on the amount of money left over to buy other goods, y. As a result indifference curves with different levels of utility are vertical displacements of a single curve. This means that the slopes of the indifference curves when consuming x_0 amount of fish are the same independently of the amount of money she has for other purchases.

we will often refer to y as *money*, understanding that it is really generalized purchasing power that can be spent on many other things possibly in many periods.

The analysis of demand is greatly simplified if the marginal rate of substitution depends only on the amount of the good someone purchases, and not on the amount of money she has left over. Here are the simplifications:

QUADRATIC, QUASI-LINEAR UTILITY In the case of *quadratic*, quasi-linear utility, the non-linear part of the utility function, $g(x)$ is a quadratic function of x such as $g(x) = \bar{p}x - \frac{1}{2}(\bar{p}/\bar{x})x^2$, where $\bar{x}$ is the maximum amount of x someone is willing to consume and $\bar{p}$ is their maximum willingness to pay for x when they have not yet consumed any x. The quadratic quasi-linear utility is *quadratic* in x (it has a "squared" term in x) and *linear* in y.

- *Prices*: When y is money left over for other purchases, then $p_y = 1$ (the price of a dollar is a dollar), so we can simplify $mrt(x, y) = \frac{p_x}{p_y} = p_x = p$. That is, when the other good is money, we shall simply refer to the price of good x as p, which is the opportunity cost of x.

- *Marginal utility*: when utility is quasi-linear in y, then $u_y = 1$, that is, the marginal utility of money is constant and equal to 1. Therefore the marginal rate of substitution $mrs(x, y) = \frac{u_x}{u_y} = u_x$ is just the marginal utility of x, which is the trade-off of money for x.

- *Willingness to pay*: Because with the quasi linear utility function the mrs is simply $u_x(x, y)$, that is, the marginal utility of x when consuming the bundle (x, y), the mrs is also the maximum amount (in money units) that the person would be willing to pay to have a small increase in the amount of x.

If we use a quasi-linear utility function in which money left over is the linear good, then utility is measured in money. To see this suppose the person

Figure 7.15: **Harriet's marginal rate of substitution (demand): quadratic quasi-linear preferences.** The figure shows Harriet's $mrs(x, y)$ for a good, x, and money for other goods, y. That is, the downward-sloping line is Harriet's *willingness to pay* in money for an additional unit of good x for different levels of the quantity she has of good x changes and is therefore also her *demand curve* because it shows the relationship between the price of the good and how much of it she will buy at different prices. The vertical intercept, $\bar{p}$, is the person's maximum willingness to pay when they currently consume zero units of good x. The horizontal intercept, $\bar{x}$, is the person's satiation point or bliss point beyond which Harriet's marginal rate of substitution is negative so that x changes from being a good to a bad.

spent nothing on the other good; then Equation 7.28 tells us that $u = y$. Utility is the amount of money the person has to spend. It is also the case that if the amount of money the person has to spend increases by \$10, then utility increases by 10. So the units in which utility is measured is money.

This does not mean that the *only* thing the person cares about is money. Whatever x is may matter a lot. It is just that how much it matters will be measured in money equivalents.

Quadratic, quasi-linear utility

Many of the examples in this book use the particular class of *quadratic quasi-linear utilities*, where the function $g(x) = \bar{p}x - \frac{1}{2}\left(\frac{\bar{p}}{\bar{x}}\right)x^2$ is a quadratic function of the good x, and as a result the utility function is:

$$\text{Quadratic, quasi-linear utility} \qquad u(x,y) = y + \bar{p}x - \frac{1}{2}\left(\frac{\bar{p}}{\bar{x}}\right)x^2 \tag{7.29}$$

In Equation 7.29:

- **Satiation:** The parameter $\bar{x}$ represents the level of x at which the buyer is *satiated* with x and would consume no more even if the price were zero.
- **Maximum willingness to pay:** The parameter $\bar{p} > 0$ represents the buyer's maximum willingness to pay for the first unit of x when they do not have any x, i.e. when $x = 0$.

$\bar{p}$, $\bar{x}$, SATIATION & BLISS

$\bar{p}$ is the person's *maximum willingness to pay* for good x. She won't pay more than $\bar{p}$ to get a unit of x.

$\bar{x}$ is the person's *satiation point* for x, beyond which her marginal utility of x is negative. She would prefer *not* to consume $x > \bar{x}$.

The point at which you are *sated* (verb) is where you reach *satiation* (noun) from consuming a good, like x. The intuition is easily seen with food: you reach satiation at that point where you do not want to eat another mouthful (the marginal utility hits zero) or, if you do, you know you'll regret it (the marginal utility will be negative). Or, it is the point at which you have reached **bliss**, which is perfect happiness or great joy, and at which, if you consumed or did any more, it would detract from that bliss, joy or wonder.

The marginal utility of x with the QQL utility is:

$$\frac{\Delta u}{\Delta x} = u_x(x,y) = \bar{p} - \frac{\bar{p}}{\bar{x}}x \tag{7.30}$$

Equation 7.30 tells us the following:

- When $x < \bar{x}$, the buyer's marginal utility of x is positive, and she regards x as a *good*.
- When $x > \bar{x}$, the buyer's marginal utility of x is negative, and she regards x as a *bad*.

If y is budget left over to buy other goods, then the marginal utility of y is always 1, regardless of the levels of x and y. As a result, the marginal rate of substitution is equal to the marginal utility of x:

$$mrs(x,y) = \frac{u_x}{u_y} = u_x = \bar{p} - \frac{\bar{p}}{\bar{x}}x \tag{7.31}$$

We can think about Equation 7.31 in the following way:

- Equation 7.31 is the equation for a line.
- Equation 7.31 has vertical intercept $\bar{p}$, which is the buyer's maximum willingness to pay.
- Equation 7.31 has a horizontal intercept $\bar{x}$, which is the point beyond which the buyer does not want to pay for good x (at $x = \bar{x}$, the buyer's willingness to pay is zero). $\bar{x}$ is the buyer's bliss point. To get the buyer to consume more than $\bar{x}$ of the good, you would have to *pay her*, rather than expecting her to pay you.
- Equation 7.31 shows that the marginal rate of substitution has a slope of $-\frac{\bar{p}}{\bar{x}}$, which is the negative of the ratio of the maximum willingness to pay to the satiation point.

Equation 7.29 shows that Harriet's utility for the good x depends on how much she has relative to a target level $\bar{x}$. When $0 \leq x \leq \bar{x}$, Harriet's utility increases if she has more x. But when $x \geq \bar{x}$, Harriet's utility decreases as she gets more x. We can plot quadratic, quasi-linear marginal rate of substitution or willingness-to-pay as a function of x as in Figure 7.15.

The principle of demand, equation 7.6, is particularly simple mathematically for QQL preferences, since it results in an *inverse demand function* , saying the highest price Harriet will pay for any given total amount of the good:

$$\text{Inverse demand function:} \qquad p(x,m) = \bar{p} - \frac{\bar{p}}{\bar{x}}x \tag{7.32}$$

We will often simplify the inverse demand function such that the slope of the function, $\frac{\bar{p}}{\bar{x}}$ is represented by a simple slope coefficient, β. We will therefore represent the demand curve as follows:

$$\text{Demand curve:} \qquad p(x,m) = \bar{p} - \beta x \tag{7.33}$$

Suppose Harriet has quadratic quasi-linear preferences between a good x and money left over to buy other things, with the willingness-to-pay $p(x) = \text{mrs}(x,y) = \bar{p} - \frac{\bar{p}}{\bar{x}}x$. She starts out with budget m and has the opportunity to buy any amount of the good x at the price p. If $p(0) = \bar{p} > p$, Harriet will buy at least some of the good. If she buys x_0 units of the good, Harriet's willingness-to-pay will have fallen to $p(x_0)$. If $p(x_0) > p$, then Harriet will buy more of the good, and so on, until $p(x) = p$.

At this point Harriet doesn't buy any more of the good, and spends the rest of her budget on other things. So from Harriet's willingness-to-pay (or marginal rate of substitution), we can derive the *quantity demanded* at a market price p, which tells us how much of the good she will buy when the market price for the good is p.

$$x^{QQL}(p) = \bar{x} - \frac{\bar{x}}{\bar{p}}p \tag{7.34}$$

M-CHECK For example, when the market price is: $p = \$10$, $\bar{p} = \$20$, and $\bar{x} = 10$, then Harriet would like to buy 5 apples since her willingness-to-pay is the following: $mrs(5,y) = \$10$ for any y.

We can also re-write Equation 7.32 as a demand function

$$\text{Demand function} \qquad x(m,p) \quad = \quad \bar{x} - \frac{\bar{x}}{\bar{p}}p \tag{7.35}$$

Equation 7.35 says that for $p > 0$ a person will consume an amount less than their point of satiation ($\bar{x}$) given the price of the good p and their maximum willingness to pay $\bar{p}$. As the price of x increases, then the quantity demanded of x will decrease.

Figure 7.16 demonstrates this relationship by showing Harriet's utility maximizing choices between x and y with her indifference curves and budget constraints for three prices of x in the top panel, while also showing her marginal rate of substitution of money for the good in the lower panel. The lower panel also shows how Harriet's marginal rate of substitution corresponds to a demand curve, by showing three different price levels and how the given price determines the quantity demanded at that price.

M-Note 7.6: The demand for x and y with with quadratic quasi-linear preferences

We begin with the principle of demand, and then rearrange it to give us the demand for good x

$$\begin{aligned} mrs^{QQL}(x,y) &= \bar{p} - \frac{\bar{p}}{\bar{x}}x = p = mrt(x,y) \\ \frac{\bar{p}}{\bar{x}}x &= \bar{p} - p \\ x(m,p) &= \bar{x} - \frac{\bar{x}}{\bar{p}}p \end{aligned} \tag{7.36}$$

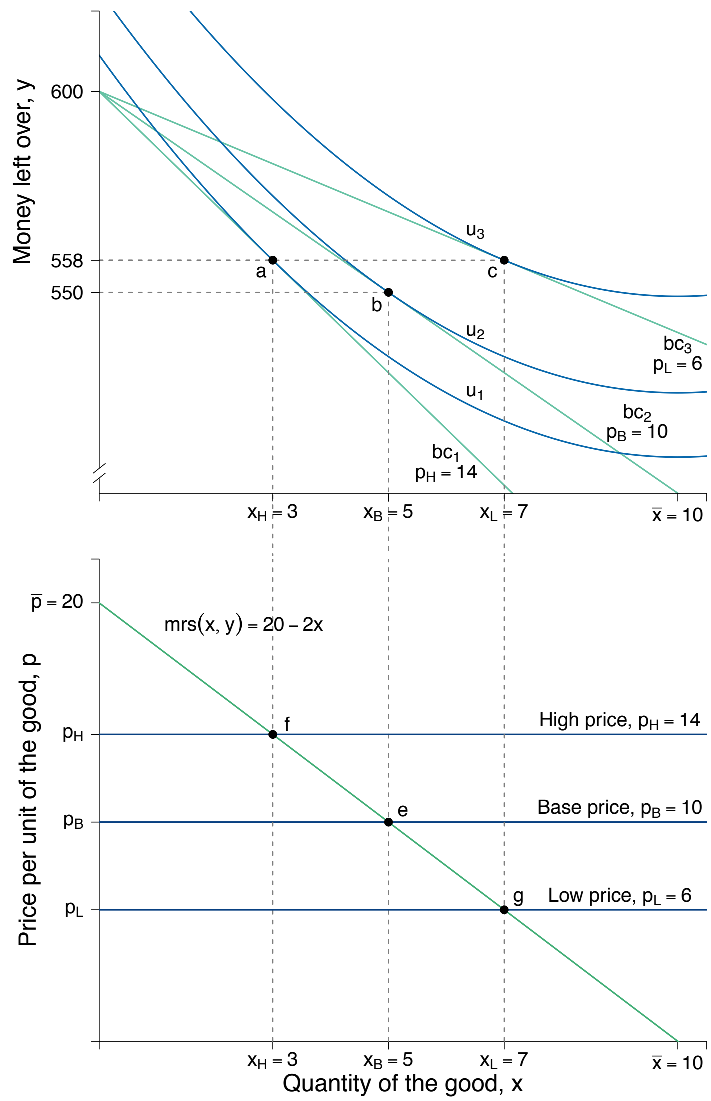


Figure 7.16: **Harriet's utility-maximizing choice and marginal rate of substitution (demand): quasi-linear preferences.** The top figure shows Harriet's indifference curves for kilograms of fish (x) and money left over for other goods (y). The figure shows her utility-maximizing choices at three levels of prices for a kilogram of fish. Harriet's utility function is $u(x,y) = y + 20x - \frac{1}{2}\left(\frac{20}{10}\right)x^2$. Her budget constraint is $y = 600 - p_x x$. The lower panel shows Harriet's $mrs(x,y)$ for a good, x, and money for other goods, y. Harriet's marginal rate of substitution is therefore $mrs(x,y) = 20 - 2x$. Her marginal rate of substitution, as her willingness to pay in money (y) for goods (x) is her *demand function* for x. She has a y-intercept of $y = 20 = \bar{p}$ (her maximum willingness to pay) and her x-intercept is $x = \bar{x} = 10$ (the amount of fish that satiates her appetite for fish, which is also the maximum quantity of fish she would consume were the price of fish zero). The slope of her marginal rate of substitution suggests she will exchange money (y) for fish (x) until her $mrs(x,y) = p_x$, i.e. when $20 - 2x = 10$, which implies $x = 5$ when $p_x = 10$.

The person will then use whatever remains of their budget (m) as money to spend on other goods, y, given what they spent on x at its price, p:

$$\begin{aligned} y &= m - px \\ y &= m - p\left(\bar{x} - \frac{\bar{x}}{\bar{p}}p\right) \\ y(m,p) &= m - p\bar{x} + p^2\frac{\bar{x}}{\bar{p}} \end{aligned} \tag{7.37}$$

Equation 7.37 shows that once we determined the demand for x, we can derive the demand for the other goods as well.

7.7 Price changes: income and substitution effects

When the price of a good changes, the consumption of the goods changes as we saw when deriving the demand curve from the offer curve in Figure 3.13. The total amount of the change is made up of two components:

- The **income effect**
- The **substitution effect**

The income effect occurs because a person has more or less *purchasing power* as a consequence of the change in prices. A person's real budget is a measure of their purchasing power with a given budget – the maximum amount they can purchase of a good at given prices. If the price of the good increases the consumer can buy less of the good than before the price increase. They have less *purchasing power* at their fixed budget. Generalizing this to two goods, if the price of one good decreases while the price of the other good remains the same, the person can buy more of both goods when the price of one good decreases.

INCOME EFFECT When the price of a good changes, a person's real income changes and they will switch to a new indifference curve with their new income as their budget constraint pivots with their new real income. A person's real income is their income given the prices of goods, for example, if the price of a good goes down, you can buy more of the good, which means your money goes further than before the price decrease. Conversely, as prices increase, your budget can't purchase a much, so your real income is lower.

The *substitution effect* occurs because as the prices of goods change, people will substitute *away* from goods that are relatively expensive and they will substitute towards goods that are relatively cheap. So, when choosing between two goods that both provide positive utility, when the price of one sharply increases relative to the other good, the person will substitute away from that good toward the other good (assuming some degree of substitutability between the goods). When the price of the good decreases, the person will substitute toward that good and away from the relatively more expensive good.

SUBSTITUTION EFFECT When the price of a good changes, that good's price relative to other goods changes, so people will move *along* an indifference curve in response to the opportunity costs at the new price ratio.

Income and substitution effects for normal goods

Consider a change in the price of good x, for example an increase in p_x. Harriet will change the amount of good x that she buys. We can *decompose* or separate this change into a *substitution effect*, which represents how Harriet would change her choice if she had to move along the original indifference curve, and an *income effect*, which represents the change in her behavior due to moving to another (in this case, lower) indifference curve. Harriet starts at point **a** with bundle (x_a, b_a) on her initial budget constraint bc_a before the price increases. When the price increases, the budget constraint pivots inwards to bc_b and creates a new equilibrium at **b** with bundle (x_b, y_b). The income effect can change both the amount of good x and good y, as Figure 7.17 shows.

COMPENSATED BUDGET CONSTRAINT The compensated budget constraint takes the new prices of goods as given (it is parallel to the budget constraint after the price change), but it gives the person sufficient income to return to their original indifference curve, therefore creating a new point of tangency with the original indifference curve. In the case of a price *decrease* a compensated budget constraint would take money away from a person, for example via a lump sum tax.

To decompose this change into an income and a substitution effect, we need to construct a counter-factual. We take Harriet back to her original indifference curve, but we retain the slope of the *new* budget constraint to see how

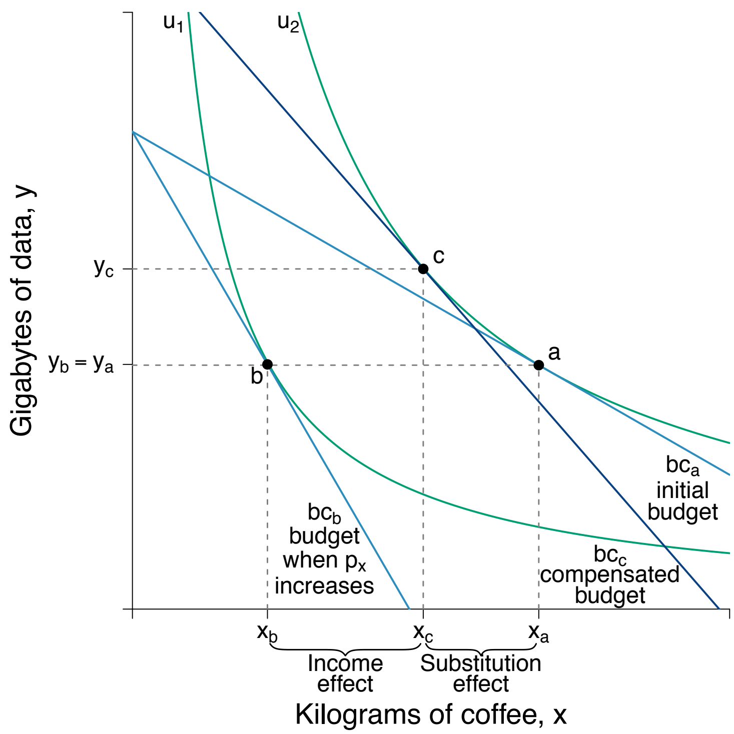


Figure 7.17: **Income and substitution effects: Cobb-Douglass utility.** Income and substitution effects with Cobb-Douglas utility with both a substitution effect and an income effect. The substitution effect is shown by the movement along u_2 from **a** to **c**. The income effect is shown by the movement to another indifference curve, shown by the distance between points **b** and **c**.

the new prices *would have* affected her. To break down the effects, we use the idea of a **compensated budget constraint**. The compensated budget constraint could hypothetically be implemented by a policy-maker who wants to ensure that a person who has lost real value of their available budget because of price increases can be compensated in some way, for example, by a government transfer.

The compensated budget constraint bc_c allows us to consider how Harriet would respond to the new price for good x if she were limited by the compensated budget constraint that provides Harriet with just enough money to be on her original indifference curve, u_2. The new tangency is at point **c** at bundle (x_c, y_c).

EFFECTS OF A PRICE CHANGE The *total effect* of a price change is the change in quantity demanded. The decomposed effects shows how the total effect is broken up (decomposed) into the two parts of the *substitution effect* (a movement along the indifference curve) and the income effect (a movement to a new indifference curve).

The difference between x_a and x_b is the *total effect* of the price change or the substitution effect plus the income effect. By construction the substitution effect causes a movement along the original indifference curve to point **c**. The difference between x_c and x_a is the *substitution effect*. The difference between x_b and x_c is the *income effect*, which is what drives Harriet's choice to point **b** with bundle (x_b, y_b).

Insert links to online appendices with *no income effect* and *no substitution effect*.

Utility Function	*Income Effect*	*Substitution Effect*
Cobb-Douglas	Yes	Yes
Quadratic, quasi-linear	No	Yes
Perfect Complements	Yes	No

Table 7.1: Utility Functions and their income and substitutions effects. Remember that the substitution effect is captured by a movement along an indifference curve as prices or real incomes change, whereas the income effect is captured by a movement to a new indifference curve.

Complements and substitutes in consumption

The size of the income and substitution effect will depend on whether the goods "go together" or have an "either/or" quality. If two goods are more enjoyable consumed *together*, then they are *complements* (coffee and cookies). "Either/or" goods are called *substitutes*: they are consumed *instead* of each other (tea and coffee). At the extreme, *perfect* substitutes have a constant marginal rate of substitution (linear indifference curve), whereas perfect complements (right shoes and left shoes) are only valuable when consumed together. Indifference curves for perfect complements are L-shaped.

REMINDER In Chapter 6 we introduced the idea that in inputs to a production process may be either substitutes (computer driven machine tools and skilled machine operators) or complements (computer driven machine tools and engineer-programmers).

COMPLEMENTS AND SUBSTITUTES IN CONSUMPTION Goods are complements in consumption if an increase in the quantity consumed of one raises the marginal utility of the other. Goods are substitutes in consumption of an increase in the quantity consumed of one reduces the marginal utility of the other. The property of being complements or substitutes is symmetrical: If good x is a complement of good y, then y is also a complement of good x. The same is true for substitutes.

M-Note 7.7: Complements and substitutes in consumption

Complements: Cookies and coffee. In the Cobb-Douglas utility function $u = x^{\alpha}y^{1-\alpha}$, x (coffee) and y (tea) are complements because

$$\frac{\partial u}{\partial x} = \alpha x^{\alpha-1}y^{1-\alpha}$$

which, because $0 < \alpha < 1$, means that

$$\frac{\partial u/\partial x}{\partial y} = (1-\alpha)\alpha x^{\alpha-1}y^{-\alpha} = (1-\alpha)\alpha(\frac{x}{y})^{\alpha} > 0$$

so the greater is the consumption of y the higher is the marginal utility of x. By the same reasoning, the greater is the consumption x, the higher is the marginal utility of y.

Substitutes: Coffee and tea. Here is a utility function in which x (coffee) and y (cookies) are substitutes:

$$u = (x + \varepsilon y)^{\alpha}$$

where ε is a positive constant measuring how much the person prefers tea to coffee and $0 < \alpha < 1$. Therefore, finding the marginal utility of x by taking partial derivatives:

$$\frac{\partial u}{\partial x} = \alpha(x + \varepsilon y)^{\alpha-1} > 0$$

The marginal utility of x is positive. But how does the marginal utility of x change as consumption of y changes? We can work that out by taking the partial derivative of the marginal utility of x with respect to y:

$$\frac{\partial u/\partial x}{\partial y} = \varepsilon(\alpha-1)\alpha(x + \varepsilon y)^{\alpha-2} < 0$$

It is negative because $\alpha < 1$. This shows that the marginal utility of coffee is less the more tea the person consumes. The same reasoning shows that the marginal utility of tea is less, the more coffee the person consumes. For this particular utility function, tea and coffee are what is called perfect substitutes. In this case,

$$\frac{\partial u}{\partial y} = \varepsilon\alpha(x + \varepsilon y)^{\alpha-1}$$

so the individual's marginal rate of substitution, the ratio of the two marginal utilities, is as follows:

$$\frac{\partial u/\partial x}{\partial u/\partial y} = \frac{\alpha(x+\varepsilon y)^{\alpha-1}}{\varepsilon\alpha(x+\varepsilon y)^{\alpha-1}} = \frac{1}{\varepsilon} = \text{mrs}$$

Because the $mrs(x,y)$ is the negative of the slope of an indifference curve, the fact that it is constant means that the indifference curves are linear. This is what the fact that x and y are perfect substitutes means.

7.8 *Application: Income and substitution effects of a carbon tax and citizen dividend*

The decomposition of the results of price changes into income and substitution effects can be illustrated by the proposed carbon tax to reduce emissions of carbon dioxide and other greenhouse gases that contribute to climate change.

The New York Times

Opinion

OP-ED CONTRIBUTORS

A Conservative Case for Climate Action

Figure 7.18: **A conservative case for climate action.** In their 2017 op-ed (opinion editorial) in *The New York Times* three economists made "A conservative case for climate action," proposing the carbon tax and citizen dividend that we analyse here. They included Martin Feldstein and Gregory Mankiw, chief economic advisors to U.S. Presidents Ronald Reagan and George W. Bush, respectively.

REMINDER In Chapter 5 we explained that the social cost equals the private cost plus the (negative) external cost imposed by a person's action. In that case, a person's marginal private cost of fishing included their disutility of fishing, but the social cost included not only the private costs but also the negative external effect the fishermen imposed on others.

The prices of petroleum, coal, natural gas and other fossil fuels do not include the costs of the environmental and climate-change external effects of their use . This means that people pay a *private cost* of using fossil fuels that is *lower* than the social (private plus external) cost of using them. The result – as in the case of over-fishing in Chapters 1 and 5 – is overuse of fossil fuels.

We now consider tax imposed on the sale of fossil fuels, in conjunction with a transfer of the tax resulting revenues in equal amounts back to the members of the population, called a citizen's dividend. We ask: how would this so-called carbon tax and citizen dividend policy reduce consumption of fossil fuels and affect citizens' consumption of other goods.

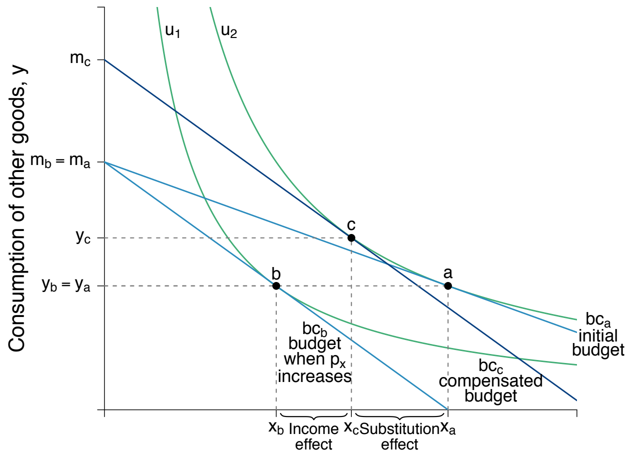
We consider two steps in our policy process:

- The substitution and income effects of the increased price of fossil fuels.
- The income effect of the citizen dividend.

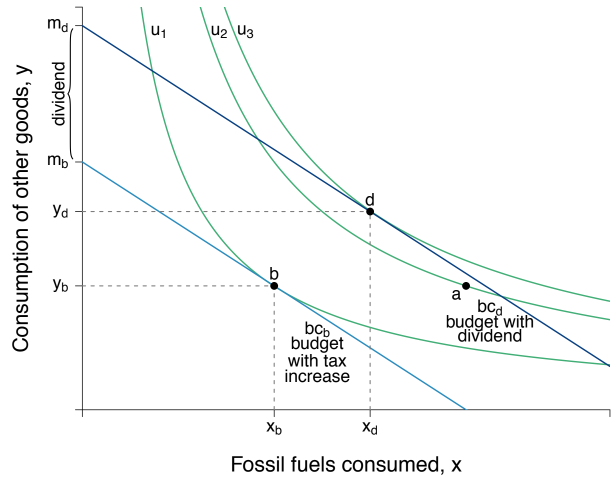
Reducing carbon emissions by imposing the tax

In Figure 7.19, the fossil fuel consumption (x) of a citizen is plotted on the horizontal axis, and the consumption of others goods measured in some currency is plotted on the vertical axis (y). Before the tax, a citizen is at point **a** in Figure 7.19 where the marginal rate of substitution equals the marginal rate of transformation, which is the existing price of fossil fuels, p_a. At **a**, the citizen has a utility of u_2 on the corresponding indifference curve.

The government then imposes a tax on fossil fuels. The tax increases the price of fossil fuels. With the increase in the price from p_a to p_b due to the the tax, the citizen's budget constraint becomes steeper. It pivots inward round its y-axis intercept because the amount of the budget itself is unaffected, but



(a) Substitution and income effects of the price increase



(b) Second income effect due to citizen dividend

Figure 7.19: **Carbon tax with dividend.** A citizen decides on consumption of fossil fuels (x) other goods (y). Prior to the introduction of the carbon tax the citizen's budget is m_a and the budget constraint is the line labeled bc_a, with slope $-p_a$. The utility-maximizing allocation is on the budget constraint at point **a** where the citizen's indifference curve u_2 is tangent to the budget constraint. A carbon tax is proposed that increases the price from p_a to p_b steepening the budget constraint. The result, shown in panel a, is the budget constraint pivots inwards. The effect of the price change – a reduction in fossil fuel consumption from x_a to x_b – is the sum of the income and substitution effects. In panel b the citizens' dividend increases the household's budget so the budget constraint shifts outwards and has a vertical intercept (the budget itself) $m_d > m_b$. The indifference curves shown here are based on a Cobb-Douglas utility function with $\alpha = 0.5$.

the price has *increased*. So bc_b is the new budget constraint. As before, citizen now maximizes her utility now consuming at point **b** where her marginal rate of substitution equals the new marginal rate of transformation, p_b. At **b** the constrained utility maximum, the citizen has decreased her consumption of fossil fuels from x_a to x_b, consistent with the policy goals of the tax to decrease consumption of fossil fuels.

Setting aside the value that the citizen places on the overall mitigation of the the climate change crisis that the population wide effects of the policy accomplished, the policy has lowered her utility because she is on a lower indifference curve, $u_1 < u_2$.

Is it fair?

This will be true of all citizens, but the effect will differ across by levels of income. In the U.S., the reduction in real income imposed by the carbon tax on will be a larger among poorer households. This is because, as Figure 7.20 shows:

- while (panel a) higher income people spend much more than lower income households on carbon costs (think about air travel, heating and air conditioning large houses)
- ...expenditure on carbon costs *as a fraction of their total expenditure* is greater for lower income households (panel b).

As a result high income people will pay more of the tax than low income people, but the tax will lower the real income of poor people by a larger percentage that will be the case for higher income people. In the U.S. the carbon tax is regressive, meaning that the amount paid as a fraction of a household's income is greater for lower income households.

This is where the citizen's dividend comes in.

Increasing income and ensuring fairness through a citizen's dividend

FACT CHECK Using detailed data on household expenditures economists Anders Fremstad and Mark Paul calculated that a $ 50 per ton of CO_2 would leave 56 percent of all U.S households with a higher real income (the same would be true of 84 percent of households with incomes less than the mean.)[4]

To see how the citizens' dividend alters the result, turn to Figure 7.19 b. As in panel a, the citizen is at point **b** with lower utility than before the tax. But now suppose the total carbon tax revenues collected are divided equally and distributed equally to each household as in the amount m_b to m_d. This is the citizen's dividend. As you can see the effect is to shift upwards the budget constraint by the same amount.

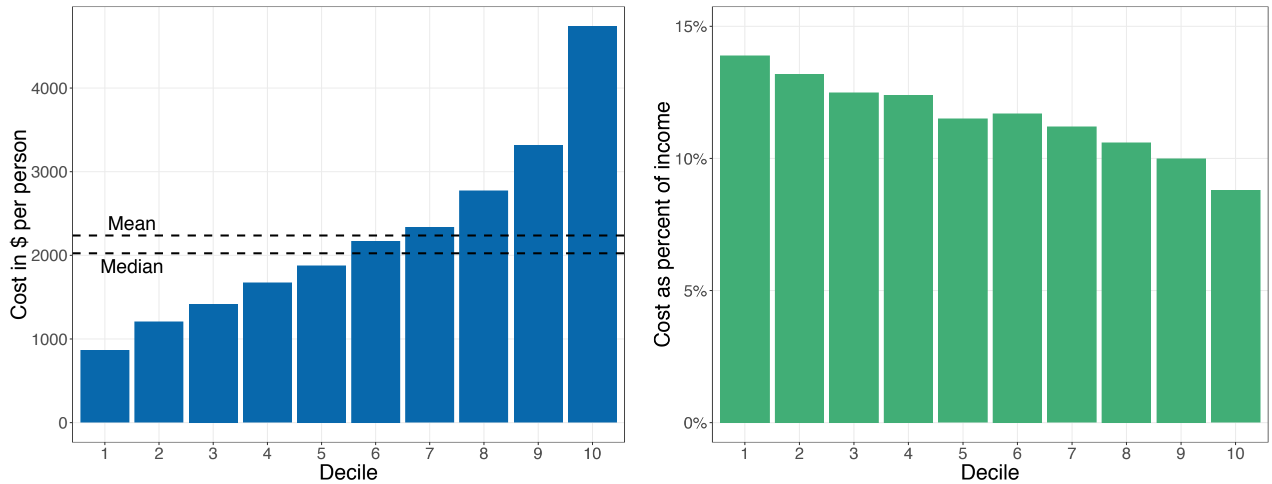
With the higher income, the citizen maximizes her utility at point **d**. For the citizen we have modeled, the dividend m_b to m_d is large relative to her previous budget. The result is an increase in consumption of other goods, so that her level of utility is higher than it was before the tax, that is, comparing points **d** and **a**, $u_3 > u_2$. At **d**, they have higher consumption of other goods ($y_d > y_b$) and they consume the a lower level of fossil fuels than they did previously, but greater than before the dividend ($x_a > x_d > x_b$). With the greater consumption of other goods, the citizen has higher utility and they obtain utility of $u_3 > u_1$.

All citizens will receive the same dividend. This means that more than half of the households – those poorer than the mean income – will experience an increase in their disposable income (income after paying taxes and receiving the dividend) as a result of the carbon tax and citizens' dividend policy. They will pay less in taxes (which are proportional to the cost of the carbon they consume) than they receive in the citizen's dividend which is proportional to the mean carbon costs consumed in the population and hence equal for all citizens.

While the carbon tax alone is regressive,the carbon tax and the citizen's dividend taken together is progressive. The case we modeled in Figure 7.19 illustrates the increase in disposable income and utility of a poorer than average citizen.

International differences

In some countries, however, the picture is reversed, with poor people spending a small fraction of their budget carbon related consumption and wealthy families spending a larger share on fuel as a relative proportion of their income. The data in Figure 7.21 for Mexico show exactly this, at least for motor fuels. As with carbon costs as a whole, the U.S. data (in the left panels) show that the fraction of the household budget spent on electricity and motor fuels respectively falls dramatically as income rises. But this is not the case for Mexico. The fraction of the household budget spent on electricity is only modestly lower for the upper income households. And for motor fuels (car and



(a) Average carbon costs per person by income decile

(b) Average proportion of expenditure on carbon by income decile

Figure 7.20: **Dividend distribution and carbon costs.** The left-hand figure shows the absolute amount spent on carbon for each of the ten household income deciles (1 is poorest, 10 is richest). The right-hand figure shows the proportion of household expenditure on carbon by for the ten income deciles. Source: Fremstad and Paul (2019) using data from the Energy Information Agency, the Bureau of Economic Analysis, and the Bureau of Labor Statistics.

truck use) the fraction of the budget spent is much greater for high income households than for those with lower incomes (among whom car and truck ownership is limited).

7.9 Application: Giffen Goods and The Law of Demand

The demand curves you have seen all slope downward: a lower price is associated with more purchases. This is called the **Law of Demand**, and the movement of prices and quantities purchased in opposite directions that it predicts is widely observed. But there is a special kind of good – called a **Giffen good** – for which the law of demand is violated. For Giffen goods, a higher price is associated with a greater amount of purchases.

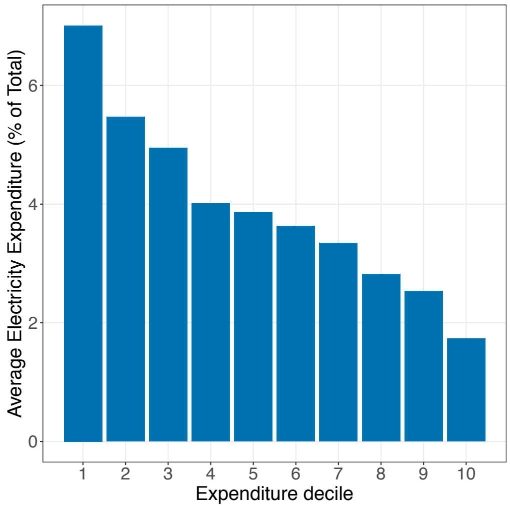
You already know that for an inferior good the amount purchased will decline as income rises. This is not really surprising, some of the low cost foods that people eat when they have very limited budgets will not be purchased at all when they have more income to spend. A Giffen good really is surprising because less is purchased when *its own price* decreases.

How could this be?

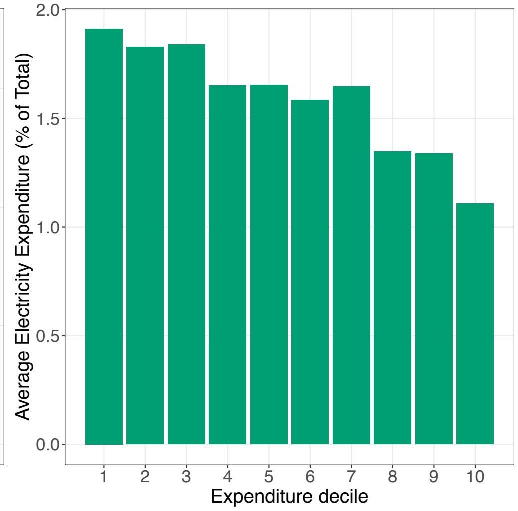
Think about a poor family consuming a large amount of some inferior good. When the price decreases, as you know, there is both a price effect motivating the family to purchase more and an income effect resulting from the decrease in price. Because the good is inferior, the higher real income of the family motivates them to purchase less of the good. If the negative income effect is greater than the positive substitution effect, purchases will decline in response to a decline in the price. Here is an example. In China, for very poor households, rice is the main staple food and if they have enough money, they add other foods that make rice taste better, such as shrimp or beef. However,

LAW OF DEMAND The law of demand holds that a decrease in the price of a good will be result in an increase in the quantity of the good purchased. Giffen goods are an exception to the law.

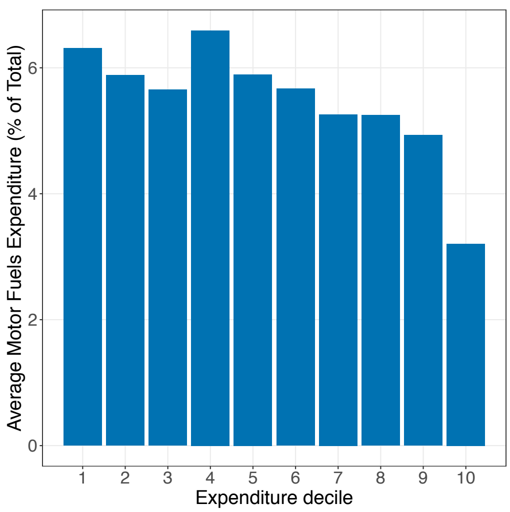
REMINDER The principle of demand states that if both goods are consumed, then the utility maximizing bundle is a point on the budget constraint at which the marginal rate of transformation (the negative of the slope of the budget constraint) is equal to the marginal rate of substitution (the negative of the slope of an indifference curve).



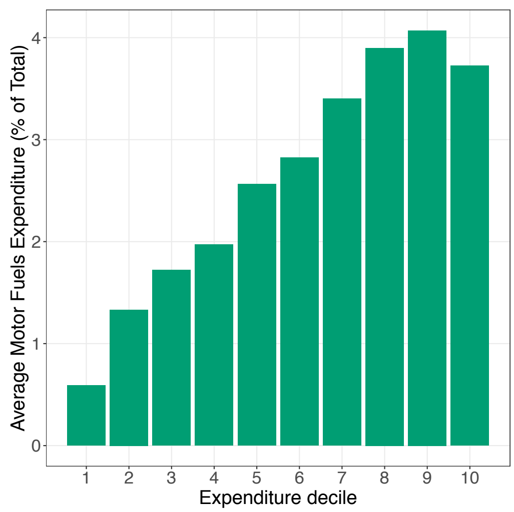
(a) U.S. Consumption of Electricity



(b) Mexican Consumption of Electricity



(c) U.S. Consumption of Motor Fuel



(d) Mexican Consumption of Motor Fuel

Figure 7.21: **U.S. and Mexican consumption of motor fuel.** Each county's income distribution is divided into deciles from poorest (1) to richest (10). The average consumption of motor fuel as a proportion of total family income for each decile is shown by the size of the bar for that decile. In the U.S., the consumption as a share of income is higher for lower deciles than for higher deciles. In Mexico, the consumption as a share of income is *lower* for lower deciles than for higher deciles. Source: Pizer and Sexton (2019) using data from the U.S. Consumer Expenditure Survey (Bureau of Labor Statistics) and Mexico's National Survey of Income and Expenditure (Instituto Nacional de Estadística y Geografía).

REMINDER: An inferior good is one for which purchases fall as income rises. A Giffen good must be inferior, but not all inferior goods are Giffen goods.

when the price of rice increases, this means the households have little money left over to buy beef or shrimp. Consequently, they will consume more rice even though it's price has increased. As a result, over some range of prices the demand curve for rice for these families is upward-sloping. Of course if the price of rice rose so high that the household purchased only rice, then further price increases would have to reduce the amount purchased, so the demand curve would then be downward sloping as the Law of Demand requires. Just such a demand curve is illustrated in Figure 7.22.

GIFFEN GOOD Over some range of prices, purchases of a Giffen good increase if the price rises, and fall if the price falls. Giffen goods are an exception to the law of demand.

This is exactly what economists who studied subsidies of rice observed in Hunan, a region of China. They ran an experiment by subsidizing the price of rice, lowering the price the families actually paid. When they provided the subsidy, very poor households *reduced* their consumption of rice. That is, when rice was cheap, households consumed less of it. When they removed the rice subsidy so that prices rose, the households consumed *more* rice. For these households rice was a Giffen good.[5]

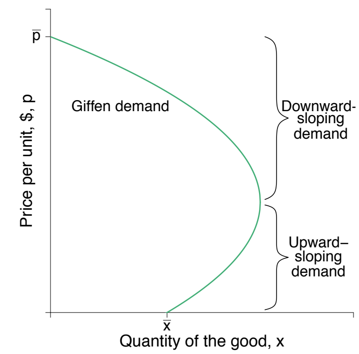


Figure 7.22: **Demand for a Giffen good.** The demand for a Giffen good is upward-sloping when the price is low. Over this region, a higher price is associated with more purchases. At a sufficiently high price, though, demand becomes downward-sloping.

7.10 *Market demand and price elasticity*

The **market demand** for a good at any given price is the sum of the demands at that price of all the people making up the *demand side* of the market. We can compute the *market demand* by adding up the individual demand curves. If we plot all the demand curves with the quantity demanded of x on the horizontal axis and the price p on the vertical axis, this requires the horizontal summation of the individual demand curves. We use an upper-case X for market demand and a lower-case x for an individual demand.

MARKET DEMAND CURVE The market demand curve is the horizontal summation of individual buyers' demand curves. That is, for each price (on the vertical axis) we add together each person's quantity demanded at that price

Figure 7.23 shows how (on the left) an individual market demand curve is (on the right) summed over ten people to produce the market demand curve, that is $X(p) = x_1(p) + x_2(p) \ldots + x_{10}(p)$.

A Linear Market Demand Curve (Quadratic quasi-linear utility)

If there are n identical buyers, each of whom has the same quadratic, quasi-linear utility for the good, with the same parameters $\bar{x}, \bar{p}$, each individual has the demand (from Equation 7.31):

$$\text{Individual demand} \qquad x^{QQL}(m,p) = \bar{x} - \frac{\bar{x}}{\bar{p}}p \tag{7.38}$$

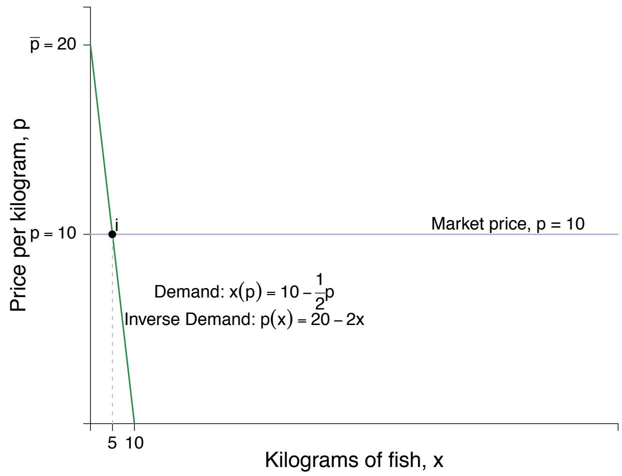
The market demand is then the sum of all the individual demands. But, since they are all equal for the identical people, this is the same as the the number of people (n) multiplied by the individual demand curves. In the quadratic quasi-linear case, therefore, the market demand curve is a given by the following:

$$\begin{aligned} \text{Market demand} \qquad X^{QQL}(p,m,n) &= \text{Number of people} \times \text{Individual demands} \\ &= n\left(\bar{x} - \frac{\bar{x}}{\bar{p}}p\right) \\ &= \bar{X} - \frac{\bar{X}}{\bar{p}}p \end{aligned} \tag{7.39}$$

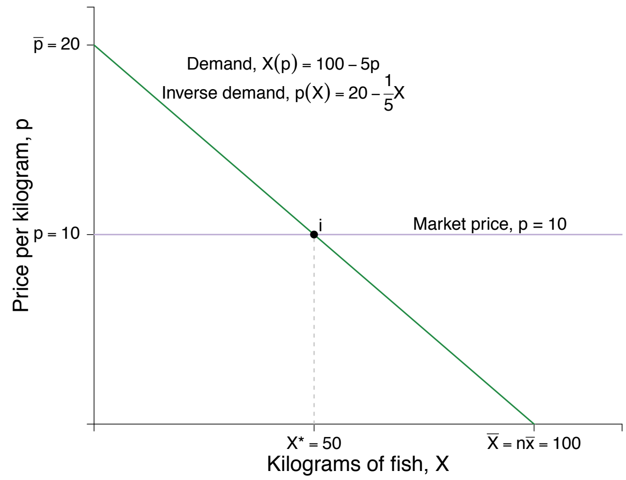
Because the market demand is the summation of individual demands, the market demand function is also downward-sloping: quantity demanded falls as the market price increases.

Re-arranging Equation 7.39, we can find the inverse market demand function:

$$\text{Inverse market demand} \qquad p(X) = \bar{p} - \frac{\bar{p}}{\bar{X}}X \tag{7.40}$$



(a) Individual demand



(b) Market demand

Figure 7.23: **Individual and market demand.** In figure a., we present Harriet's demand at different prices per kilogram of fish. On the right, is the *market demand* for fish, which is the sum of ten identical fish-buyers' demands for fish (including Harriet). Notice that Harriet's individual demand curve is much steeper than the market demand curve. The change occurs because, for example, for every \$2 decrease in the price Harriet will buy one more unit of fish, for the market as a whole, each of 10 people would buy one more unit of the good.

The inverse market demand curve is linear with a vertical intercept of $\bar{p}$ (the maximum willingness to pay of buyers like Harriet), a horizontal intercept of $\bar{X}$, and a slope of $\frac{\Delta p}{\Delta x} = -\frac{\bar{p}}{\bar{X}}$.

M-Note 7.8: Market demand with 10 buyers

Let us assume that the fish market is made up of Harriet and 9 other buyers who are identical to her (a total of ten buyers). Harriet's quadratic, quasi-linear demand function was:

$$\text{Harriet's Demand:} \quad x(p) = 10 - \frac{1}{2}p \tag{7.41}$$

If all the fish buyers are identical to Harriet, then we can sum their demand functions (quantity as a function of price), $x_i(p)$ to get the market demand, $X(p)$. This is the same as multiplying the demand function by the number of people, $n = 10$, to get the market demand function:

$$\begin{aligned} X(p) &= n(x_i(p)) \\ &= n(10 - \frac{1}{2}p) \\ &= 10 \times (10 - \frac{1}{2}p) \\ &= 100 - 5p \end{aligned} \tag{7.42}$$

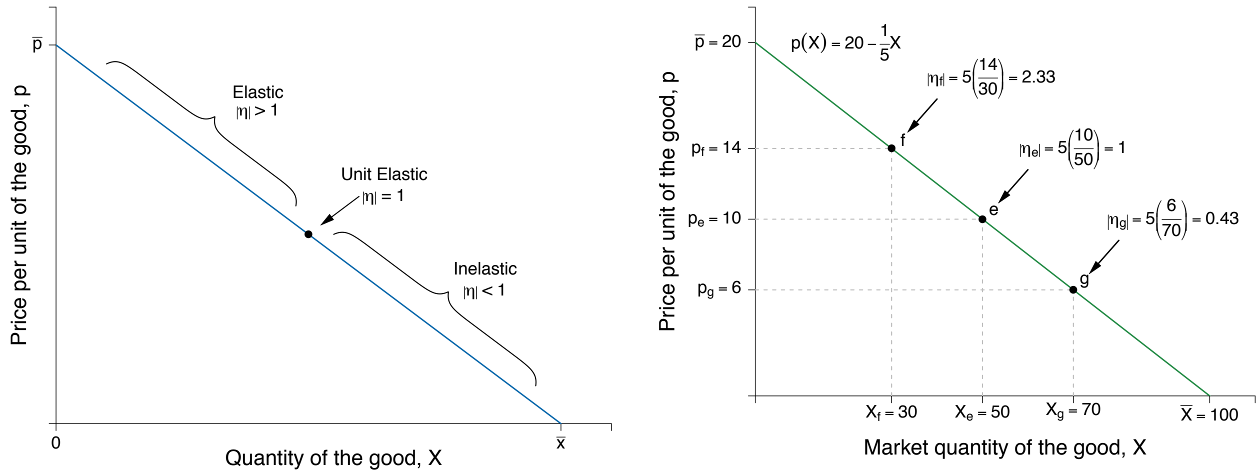
Recall, though, that we typically graph price as a function of quantity, or the *inverse* demand function. We use the market demand curve to find the *inverse* market demand curve with price as a function of quantity by re-arranging Equation 7.42 and similarly for Harriet with re-arranging Equation 7.41:

$$\text{Harriet's inverse demand:} \quad p(x) = 20 - 2x \tag{7.43}$$

$$\text{Market inverse demand:} \quad p(X) = 20 - \frac{1}{5}X \tag{7.44}$$

Contrasting Equations 7.43 and 7.44 we can see that they have identical vertical intercepts equal to $\bar{p}$, but the slopes of the two functions differ.

To see why, notice that $x = \frac{X}{n}$, and substituting this expression for x into Harriet's in inverse demand (Equation 7.43) we get the equation for Market inverse demand (Equa-



(a) General values of price elasticity

(b) Price elasticity for 10 People like Harriet in a market

Figure 7.24: **Price elasticity of demand: general and specific cases.** In figure a. on the left, we present the general relationship between the demand curve and the value of price elasticity of demand, η. The figure shows how price elasticity varies from a high value to a low value as you move left to right along the demand curve. In figure b. on the right, we present the Market demand curve for 10 buyers like Harriet whose preferences are the horizontal sum of Harriet's resulting in a market demand curve of $p(X) = 20 - \frac{1}{5}X$. Consequently, we can calculate three values for price elasticity of demand using the formula of $\eta = \frac{\delta X}{\delta p}\frac{p}{X}$. The slope of the curve, $\frac{\Delta p}{\delta X} = -\frac{1}{5}$, therefore inverting that value we see that $\frac{\Delta X}{\delta P} = -5$. We can substitute in the values for p and X at each of the price quantity combination to find the value of price elasticity at each of the points **e**, **f**, and **g** as shown in the figure.

tion7.44. This is why the market inverse demand curve has a slope equal to the slope of Harriet's inverse demand namely -2, divided by the number of total buyers (10), for a slope of $-\frac{1}{5}$ for the market inverse demand curve. The market demand curve is therefore flatter than Harriet's relatively steep demand curve.

Price elasticity and the slope of the demand curve

For many questions of both firm strategy and public policy an important question is: how much change there is in the quantity demanded when there is a change in price or $\frac{\Delta X}{\Delta p}$. But this is expressed in two different units: the units of the good (kilos of fish) and the monetary unit (dollars). But we often need to compare responsiveness across commodities – is the demand for restaurant meals more or less responsive than the demand for motor vehicle fuel?

To allow for comparisons across commodities, we need a measure of responsiveness to price that does not depend on the units in which it is measured, whether the quantity demanded is in kilos of fish or liters of Coca Cola, whether the price is in Yen or Euros. We therefore describe the response of market demand to a change in price as the ratio of the *percentage* change in quantity demanded to the *percentage* change in price, $\frac{\Delta X/X}{\Delta p/p}$. This ratio is called the *price elasticity of demand* and often represented by the Greek letter η (pronounced "ai-ta").

PRICE ELASTICITY OF MARKET DEMAND
The *price elasticity* of market demand with respect to price at a point (X, p) is the ratio of the percentage change in quantity demanded to the percentage change in price,$\eta_{Xp} = \frac{\Delta X/X}{\Delta p/p}$.

$$
\begin{aligned}
\text{Price elasticity of demand} \qquad \eta_{Xp} &= \frac{\text{\% Change in price}}{\text{\% Change in quantity}} \\
&= \frac{\Delta X/X}{\Delta p/p} \\
\eta_{Xp} &= \frac{\Delta X}{\Delta p}\frac{p}{X} \qquad (7.45)
\end{aligned}
$$

The **price elasticity of demand** at any point on the demand curve is equal to the slope of the demand curve multiplied by the ratio of price to quantity at that point.

M-CHECK The slope of the demand curve is negative, so the elasticity is also a negative number. We usually refer to the absolute value of the elasticity $|\eta|$; so a "larger" elasticity means at any given price and quantity a flatter demand curve.

The price elasticity of demand falls into three categories:

$|\eta| > 1$ Demand is *price-elastic*, which means that the quantity demanded responds more than proportionally to a change in price.

$|\eta| = 1$ Demand is *unit price-elastic*, which means that the quantity demanded responds exactly proportionally to a change in price.

$|\eta| < 1$ Demand is *price-inelastic*, which means that the quantity demanded responds less than proportionally to a change in price.

Slope and price elasticity with a linear demand curve (the QQL case)

As you already know the slope of the quadratic, quasi-linear demand curve is constant. But its price elasticity changes as price and quantity change along the demand curve:

$$\eta^{QQL} = -\frac{\bar{X}}{\bar{p}}\frac{p}{X} \tag{7.46}$$

The term $\frac{\bar{X}}{\bar{p}}$ is constant, but the term $\frac{p}{X}$ is very large when p is close to $\bar{p}$ and X is close to zero, and goes to zero when p is close to 0 and X is close to $n\bar{x}$ or $\bar{X}$. It is tempting to think of the price elasticity of demand for a particular good as a single constant number, but in general the price elasticity of demand changes with price and quantity demanded.

M-Note 7.9: Price elasticity along a Linear Demand Curve

When there are 10 people with demand functions as in M-Note 7.8, then we can evaluate the elasticity of demand as follows. Remember the following parameters:

- $\bar{p} = 20$
- $\bar{x} = 10$
- $n = 10$, therefore
- $\bar{X} = n\bar{x} = 100$

We now evaluate price elasticity of demand at two different (X, p) points. Using Equation 7.46 when price is 14 and quantity demanded is 30 units (refer to Figure 7.24):

$$\begin{aligned} \eta^{QQL} &= \frac{\Delta X}{\Delta p}\frac{p}{X} \\ &= -\frac{\bar{X}}{\bar{p}}\frac{p}{X} \\ &= -5\left(\frac{14}{30}\right) \\ &= -2.33 \end{aligned}$$

Therefore, we would say that for 10 people in the fish market, the price-elasticity of demand is *elastic* because $|\eta^{QQL}| > 1$.

Now consider an alternative point with a lower market price, $p = 6$, and corresponding higher quantity demanded equals $X(p) = 100 - 5 \times 6 = 70$. We then calculate elasticity as follows:

$$\begin{aligned}\eta^{QQL} &= \frac{\Delta X}{\Delta p}\frac{p}{X} \\ &= -\frac{\bar{X}}{\bar{p}}\frac{p}{X} \\ &= -5\left(\frac{6}{70}\right) \\ &= -0.43\end{aligned}$$

At the lower price, $p = 6$ (which is on the lower portion of the demand curve), the price-elasticity of demand is *inelastic* because $|\eta^{QQL}| < 1$, which means that quantity demanded responds less than proportionately to a change in prices.

Checkpoint 7.6: The price elasticity of demand for preventative health products

- Return to Figure 7.3 and identify the products (and their price ranges) that are most price *elastic* and most price *inelastic*.
- Identify three points on the demand curve in Figure 7.24 for which an increase in the price will raise, lower or leave unchanged total revenue.

7.11 Application. Empirical estimates of the effect of price on demand.

Why are some goods more price elastic than others?

Because we can observe price changes and how the quantity purchased changes as a response, we can estimate the price elasticity of demand for various goods. Some estimates are illustrated in Figure 7.25.

The demand for a good will be highly inelastic if it is "something that you cannot do without" and it also does not constitute a large fraction of your budget.

Generalizing from this intuition we expect goods to be price inelastic if:

- there are few substitutes for the good in question (e.g. brand loyalty, addiction, or prescription medication)
- it is considered to be a necessity not a luxury (e.g. rice, not expensive liquor or Coca-cola)
- it is not a large fraction of your total expenditures (e.g. fish)
- the person making the decision to buy is not the person paying for the purchase (e.g. when a doctor prescribes a drug that the patient will pay for)

TEXT HERE ABOUT UBER.

FACT CHECK The demand for another sugary drink - Mountain Dew – is even more price elastic that Coca-cola, namely, $|\eta| = 4.39$. Coke and Mountain Dew are close substitutes. The price elasticity of demand for sugary drinks *as a whole* is estimated to be much lower, i.e. $|\eta| = 1.4$. This is because other drinks, like milk or tea are not really close substitutes for sugary drinks. As a result a price increase for Coke might get you to switch to Mountain Dew, but not to tea or milk.

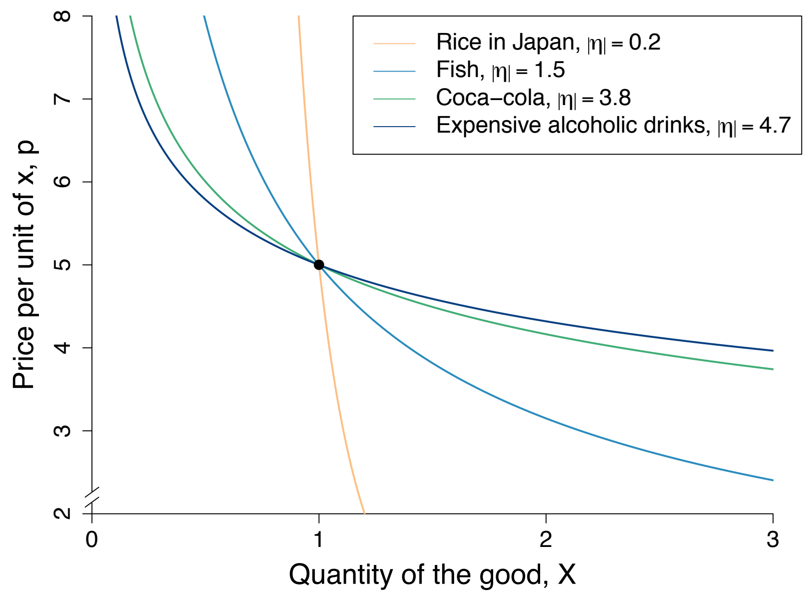


Figure 7.25: **Comparison of elasticities for different goods.** The demand functions shown are called iso-elastic, meaning that unlike the case for linear demand functions, the price elasticity of demand is the same at every point on the curve. (Recall that iso means equal). The functions have the form $x(p) = kp^{\eta}$. What the figure shows is that, for example, if the quantity of fish demanded when the price per kilo is 5 is 1 kilo, then if the price fell to 3 per kilo the demand would approximately double (increase from 1 to 2).

Checkpoint 7.7: Cobb-Douglas Price-elasticity

1. Using the Cobb-Douglas demand curve we found previously find the price elasticity of the good, x, in general (find $\frac{\Delta X}{\Delta P}$ to substitute into the formula).
2. Assume that $\alpha = 0.5$ and $m = 100$ as in the figures drawn in the examples and then consider the three prices used in Figure 7.9 to calculate the price elasticity of demand at each bundle.

The price elasticity of demand for sugary drinks and the effect of a tax

Obesity and its associated illnesses inflict extraordinary suffering and mounting health care costs around the world. Among high income countries, the U.S. and the UK are especially hard hit while obesity is rare in Japan and Singapore. Among middle income countries Mexico has one of the highest obesity rates.

Among the contributors to the epidemic rise in obesity rates in recent years, economists have proposed, are two facts:

- As economies shift from farming and manufacturing to services the amount of calories we use in a day's work has declined and
- the cost of calories, relative to other things we might spend our money on, has fallen.

Governments around the world have addressed the second economic pro-

FACT CHECK The fact that obesity rates (BMI > 30) in the U.S. are ten times the rate in Japan suggests that these economic factors – which apply with about equal force in the two countries – cannot be the entire story. Differences in culture and public policies also matter.[6]

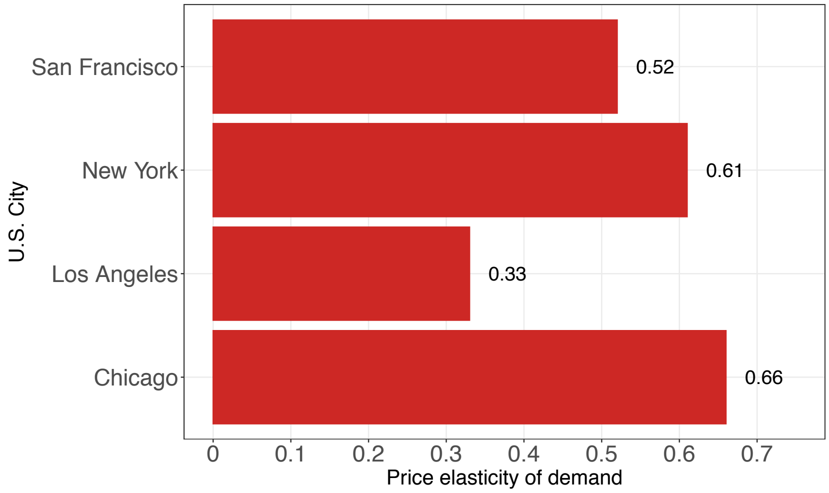


Figure 7.26: The price elasticity of demand for Uber rides in four US cities. Source: Cohen et al. (2016).

posed cause of obesity– reduced cost of calories – by instituting so-called "fat taxes" either to tax the consumption of saturated fats or to tax the consumption of sugar. As of 2019, 7 U.S. cities and 34 countries had implemented such policies.

These taxes do not aim to increase government revenue. Instead, the government wishes to discourage citizens from consuming the goods because of concerns over the citizenry's health. Similar reasoning applies to "sin taxes", which are taxes on cigarettes and liquor to discourage excessive consumption of those goods.

Sugary drinks that are commonly taxed in many countries include:

- fruit drinks (which includes sports drinks and energy drinks),
- pre-made coffee and tea (for example, bottled iced coffee and iced tea),
- carbonated soft drinks, and non-carbonated soft drinks (which includes cocktail mixes, breakfast drinks, ice pops, and powdered soft drinks)

The average American household consumed an average of 156 liters of these drinks per year during the years 2007-2016.

The demand curve and the price elasticity of demand derived from it provide essential pieces of information to assess the likely effects of the tax on sugary drinks. Because retailers frequently change prices of their drinks it is possible to estimate the price elasticity of demand. To do this, a team of economists recorded sugary drink sales, prices, and a long list of other possible influences on the individual's purchases (including health information, how much they "liked" sweet drinks, and more).

On this basis they estimated that the price elasticity of demand for sugary

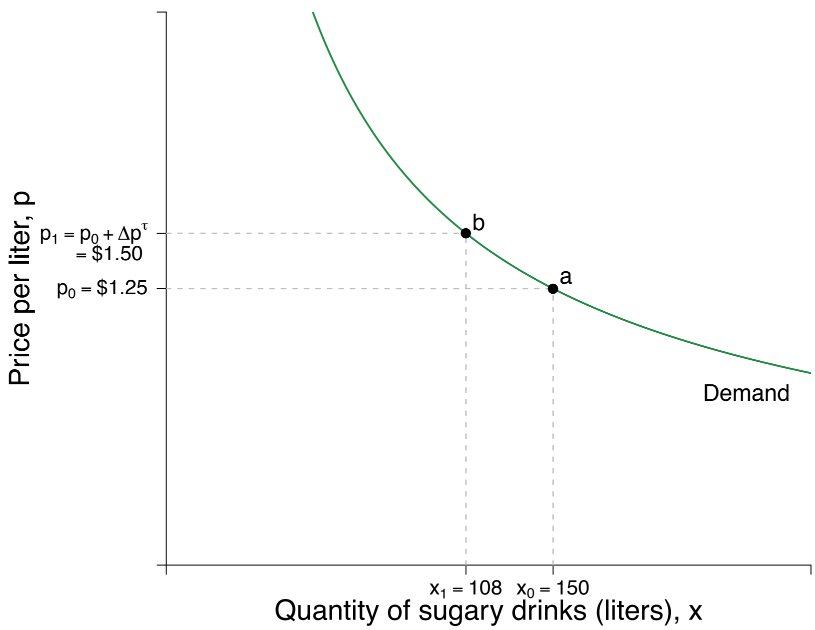


Figure 7.27: **The effect of a price increase on the demand for sugary drinks when the price elasticity of demand is $|\eta| = 1.4$.** The demand curve shown is iso-elastic with a price elasticity of demand of $|\eta| = 1.4$. There are two prices, the pre-tax price (p_0) and the post-tax price ($p_1 = p_0 + \Delta p$) where $\Delta p = 0.20xp_0$. At the higher price, the consumption of sugary drinks is less, comparing x_0 at point **a** to x_1 at point **b**.

drinks is about - 1.4, meaning that a ten percent increase in the price of the drinks would result in a 14 percent decrease in demand. This estimate – for sugary drinks as a whole – is much less than for one particular drink (Coca Cola) because there are many *other* sugary drinks that are close substitutes for Coca Cola.

Figure 7.27 illustrates what an elasticity of this magnitude implies. In this example we do not ask what determines the price per liter p_0. Instead we ask the hypothetical question: how many liters would be demanded at various prices? We can study the effect of a sugary drinks tax on the amount consumed by comparing the price per liter without the tax to amount consumed at the the (higher) price when the tax is imposed.

Suppose the price of sugary drinks was initially $1.25 dollars per liter. At this price we can see that the typical person would have demanded about 150 liters. Figure 7.27 depicts this interaction. The starting price and quantity combination are shown by (x_0, p_0). If the effect of the tax was to raise the price of sugary drinks by 20%, the the price after the imposition of the tax would be 1.50 per liter.

Recall that with a price elasticity of $\eta = -1.4$, this means a 10% increase in the price will result in a 14% decrease in quantity demanded. So, the effect of a 20% increase in the price is that the quantity demanded will decrease by 28% , down to 108 liters (x_1 in the figure) after the tax.

Checkpoint 7.8: Why is demand for sugary drinks price elastic?

Explain why the demand for sugary drinks as a whole in the U.S. is less price elastic than the demand for Coca Cola or Mountain Dew.

7.12 Consumer surplus and interpersonal comparisons of utility

Figure 7.28: **Harriet's willingness to pay and consumer surplus**. The height of the steps in the step function is the maximum Harriet would pay to have the first, second, third, and so on unit of the good. For each unit she buys, she pays $10. Her consumer surplus is the vertical distance between the price line at $p = 10$ and her willingness to pay for each additional unit, summed over the number of units she purchases. Her utility-maximizing consumption bundle is $x = 5$, when her willingness to pay equals the price: $mrs(x) = p = 10$. Summing over the bars, she receives consumer surplus, $CS = 10 + 8 + 6 + 4 + 2 = 30$.

When a person, call her Harriet as we did earlier, buys a good she does so because she expects to derive a benefit that exceeds the price of the good. The difference between the most she would be willing to pay for the good and what she actually pays for it is called the *consumer surplus* that she received as a result of that purchase.

CONSUMER SURPLUS is the difference between a person's willingness-to-pay and what they actually pay for each unit of the good that they consume. Because *not* purchasing the good is the buyer's fallback option, this definition shows tht consumer surplus is a rent. If we assume that the marginal utility of a unit of expenditure – the value to the buyer of having one more Euro to spend, for example – is the same across all individuals, the this quantity can be summed over all purchasers of the good. This total consumer surplus summed over all buyers of the good is used as a measure of the increase in well being of the consumers that is made possible by the availability of the good.

Because it is measured as the difference between the maximum willingness to pay in money and the money that is paid, consumer surplus gives a measure in *monetary* terms of the benefits (or " *consumer welfare*") that a person derives from a purchase of a good. If we consider not buying the good as the person's fallback position, then we see that consumer surplus is an economic rent, namely a measure of what they get above and beyond her next best alternative.

Figure 7.28 illustrates the consumer surplus available to Harriet by her purchases of 5 units of good x. The maximum she would pay for the first unit – think: access to a workout at the gym during a week, a film on Hulu – is $20. But each successive unit is worth less to her, and if she already has

4 – workouts, films – the most she would pay for the 5th is \$12. The sixth would not be worth more than she'd pay for it. So she will purchase 5 units. Adding up her willingness to pay for each and subtracting what she actually paid – \$10 in each case she has a consumer surplus of \$30 (that's $10+8+6+4+2=30$).

The graph of her willingness to pay – called a step function – is her demand curve. When we think of people's purchase over a longer period of time, or the purchases of many people, we smooth out the 'steps' and make a smooth curve (not necessarily a straight line).

We are often interested in a measure of how much people as a group benefit from the opportunity they have to purchase some good. A natural way to do this is to add up the consumer surplus enjoyed by each buyer. For this to make sense it must be that a dollar's worth of consumer surplus is as valuable to one person as to another. Unless we assumed this, we could not add the consumer surplus of one person – some dollar amount – to some other person's consumer surplus.

There are really two parts of this key assumption:

EXAMPLE To measure the utility gained by making a purchase and to sum this across individuals we need a measure of utility that is similar to money. If one person has a thousand U.S. dollars and someone else a hundred U.S. dollars we can say that the first person has ten times as much money as the second, irrespective of whether we measure their wealth in dollars, or in pennies.

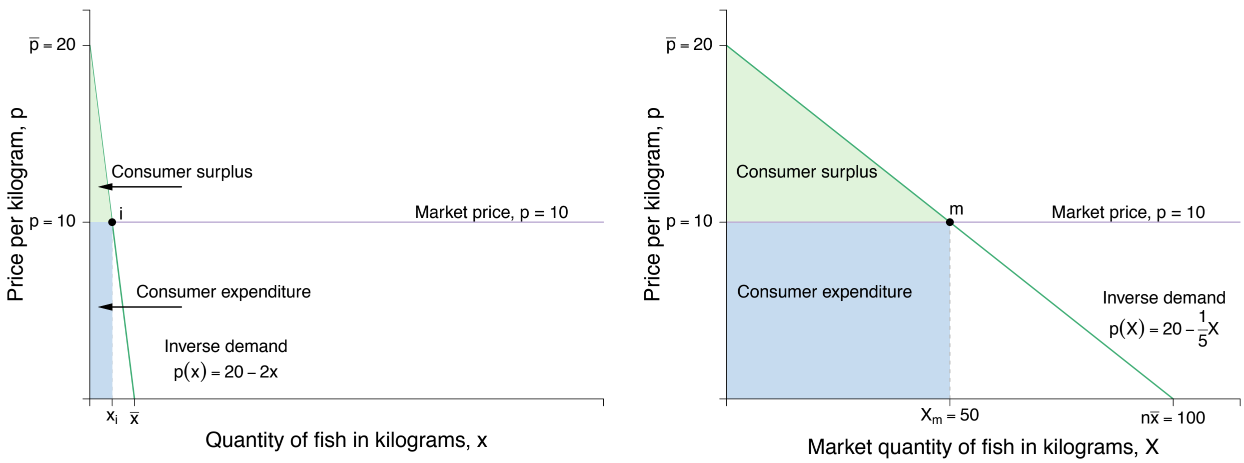
- *We can make inter-personal comparisons of utility*: we can compare one person's well being (or utility) with another. Recall that this means that we consider utility to be a cardinal measure that can be compared across individuals (like for height, *how much taller* is Simon than Harriet) rather than an ordinal measure (Simon is taller than Harriet).

- *The marginal utility of money left over for other purchases is the same to all people*: an additional dollar makes the same contribution to the well-being of one person as to another. This means that what Harriet would purchase with a dollar of money left over after spending on workouts contributes as much to her well being as a dollar's worth of additional expenditure by a less fortunate person.

This would almost certainly not be true if one of them were very poor, so that a dollar would be worth a lot (it would be used to purchase food, or other essentials), and the other was very rich (the additional dollar would be spent on a luxury good).

In the left panel of Figure 7.29, the individual consumer surplus (Equation 7.47) is the area of the light-yellow triangle above the price and below the demand (marginal rate of substitution) function for Harriet. We can see that consuming $x=5$ units of fish provides Harriet with consumer surplus of \$25. Notice that this is exactly the same as if we had substituted Harriet's values for $\bar{x}$, $\bar{p}$, p and x into equation 7.47.

The consumer surplus for all buyers is the area shaded in yellow in the right panel of the figure and because we have assumed all buyers are identical this



(a) Individual consumer surplus

(b) Market consumer surplus

Figure 7.29: **Individual and market consumer surplus.** On the left, we present a version of figure **??** showing the utility-maximizing kilograms of fish Harriet buys and the consumer surplus she derives as a consequence. She consumes output, x, and therefore her $cs(x) = \frac{1}{2}(\bar{p} - p)x = \$25$. Her expenditure is the price she paid multiplied by the number of units she bought: $ce(x) = p(x) \times x_i$. On the right, is the *market demand* for fish with the *market consumer surplus* which is the net benefits that all people obtain from paying prices beneath their maximum willingness to pay, thus $cs(X) = \frac{1}{2}(\bar{p} - p)X_m = \$250$.

is exactly $n = 10$ multiplied by the individual consumer surplus for Harriet. Notice that the scale of the x-axis is ten times larger, which is consistent with Equation 7.48.

M-Note 7.10: Consumer surplus with quadratic, quasi-linear utility

The demand curve based on quadratic, quasi-linear preferences is linear. Therefore, we can use basic algebra and geometry to calculate the value of the consumer surplus. We use the following three data points:

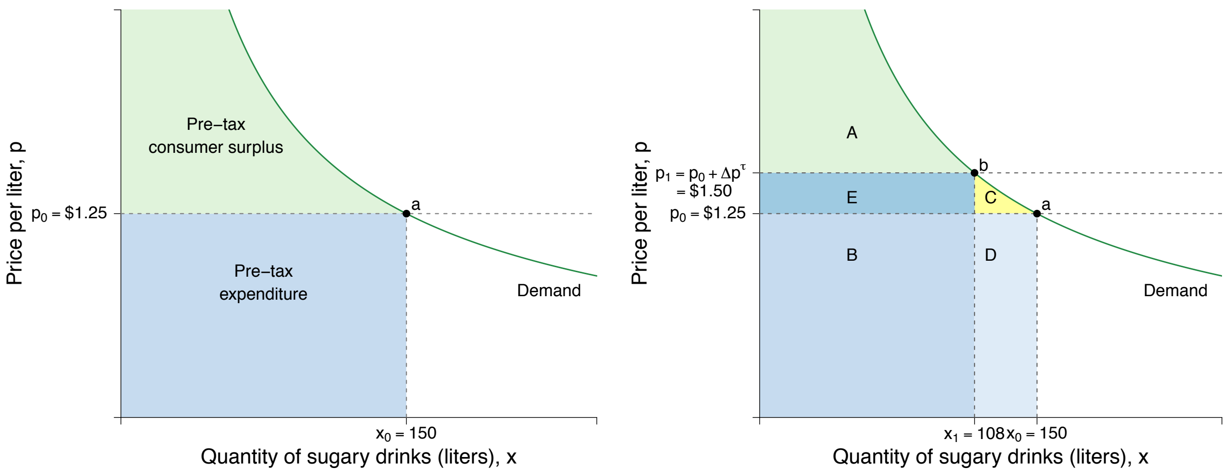
- The person's maximum willingness to pay is $\bar{p}$.
- The person actually pays the price $p < \bar{p}$.
- Therefore the difference between what they pay and what they are *willing to pay* is $\bar{p} - p$ for the first unit consumed.
- For units after the first unit consumed, the willingness to pay decreases, and the consumer surplus equals the difference between their willingness to pay at that unit and the price they actually pay, i.e. $CS(x) = mrs(x) - p$.
- They consume x_i at the price, p.
- Therefore the consumer surplus of what they consume is the area of the triangle between $\bar{p}$, p and x_i, which has an area of $\frac{1}{2}(\bar{p} - p)x_i = cs_{x_i}$

Therefore, consumer surplus for the individual and the market is given by:

$$\text{Individual} \quad cs(x) = \frac{1}{2}(\bar{p} - p)x \tag{7.47}$$

$$\text{Market} \quad CS(X) = \left(\frac{1}{2}(\bar{p} - p)x\right) \times n$$

$$= \frac{1}{2}(\bar{p} - p)X \tag{7.48}$$



(a) Initial consumer surplus and expenditure

(b) Consequences of the tax

Figure 7.30: **Effects of a tax on consumer surplus and expenditure**. There are two prices, the pre-tax price (p_0) and the post-tax price (p_1). As the price increases, the consumption of sugary drinks decreases from x_0 at point **a** to x_1 at point **b**. In Figure a. the area shaded in green below the demand curve and above the price, p_0, is the pre-tax consumer surplus. The area shaded in blue is the pre-tax expenditure, equal to the price multiplied by the quantity people consumed, i.e. $p_0 x_0$. Area **A** is the consumer surplus that remains after the imposition of the tax. Area **B** is the expenditure that was common before and after the imposition of the tax. Area **D** is the decrease in expenditure by people who are unwilling to purchase sugary drinks at the higher price, p_1. Area **E** was consumer surplus before the tax, but is now part of expenditure. Area **C** plus area **E** is the decrease in consumer surplus as a consequence of the tax.

7.13 Application: The effect of a sugar tax on consumer surplus

Some taxes like the so-called sin taxes on cigarettes and liquor – do not aim primarily to raise government revenues – the usual motive for taxation. Instead sin taxes aim to alter people's behavior: to discourage smoking, drinking alcoholic beverages, and consuming foods that contribute to obesity.

We showed earlier in Figure 7.27 when discussing price elasticity of demand that an increase in prices through a tax will decrease quantity demanded by people. We now analyze the consequences of that tax for people's utility.

Figure 7.30 shows the consequences of the tax for consumer welfare (measured in terms of prices and quantities consumed). Figure 7.30 a. shows how, at the initial price, consumer surplus is given by the combined area in green, shown by the total area A + C + E in Figure 7.30 b.

After the tax at the new higher price, consumers will be left with consumer surplus equal to area A. Consumers will lose consumer surplus indicated by the area E + C. Area B is the portion of consumer expenditure that is unchanged by the tax. Consumer expenditure will decrease by the area D, but increase by the area E. That is, before the tax, consumers spent areas B + D. After the tax, consumers spend areas B + E.

In Chapter **??** we will return to the sugary drinks tax, looking at its impact on on others, including firms' owners who will lose economic profits as a result.

Is it fair? Sugary drink taxes are regressive

In 2017 voters in Santa Fe, the capitol of the U.S. state of New Mexico, voted overwhelmingly to reject a proposed tax on sugary drinks. The measure had been put forward by a popular mayor and would have directed the resulting revenue towards expanding pre-school educational opportunities for the less well off. It was opposed by the American Beverage Association.

Opponents of the measure held that unfairly placed a burden on the less well off. To address the potential unfairness of the tax the Santa FE advocates of the sugar tax had linked the measure to the provision of a particular public service that was very much in demand among lower income Santa Feans. But the very real substantial negative income effect apparently outweighed the promise of better educational opportunities. Source: Reporting in The Albuquerque Journal,, the Santa Fe New Mexican, and The Sante Fe Reporter (e.g. "Sugar Tax Fails" 2 May, 2017.

Figure 7.31 provides evidence about the consumption of sugary drinks in households of differing incomes based on matched data on purchases of sugary drinks and household income. It is clear that households with lower incomes consume larger quantities of sugary drinks than households with higher incomes do. As a result, a per-unit tax on sugary drinks is regressive poorer households will pay more as a share of their household income. At the same time, however, decreasing the quantity of sugary drinks consumed by members of those households could be quite beneficial for health and for medical costs that those household incur.

We will return to the analysis of a sugary drinks tax in the next chapter, taking account not only of the consumer surpluses lost but also the profits lost by owners of sugary drinks producing firms and the benefits made possible by the government revenues raised.

Checkpoint 7.9: Policies to mitigate the income losses of less well off people imposed by the regressive sugary drink tax

The citizens' dividend – returning the tax revenues collected to citizens as an equal lump sum payment to each family – is proposed as a way to counteract the regressive nature of the carbon tax. Explain why something similar would not accomplish this purpose in the case of the sugary drinks tax.

Experiences around the world of sugar and fat taxes

Denmark instituted a per-kilogram tax on saturated fats in 2011. Hungary introduced both sugar and fat taxes in 2011, where the percentage of the tax is proportional to the amount of sugar or fat in the good. In 2012 France introduced a tax on both added-sugar and artificially sweetened drinks of €0.075 per liter (in 2015). Chile adopted a tax in 2015. In the United States, several

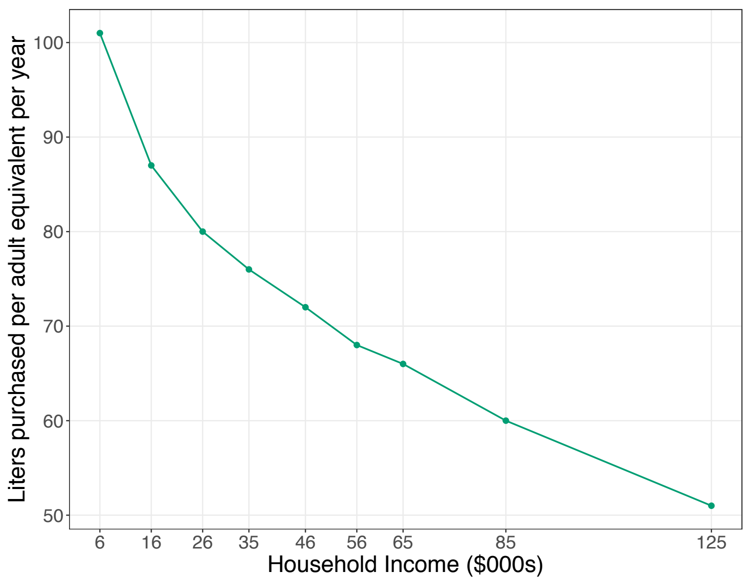


Figure 7.31: **Consumption of sugary drinks in households of differing incomes**. The figure shows the amount of sugary drinks purchased per 'adult equivalent' per year by income (measured in thousands of dollars). The term 'adult equivalent' means that children in the households have been counted as some fraction of an adult. The data indicate that the quantity of sugary drinks consumed in the poorest households is about double that consumed in the richest. Source: Allcott, Lockwood, and Taubinsky (2019).

states and cities have implemented soda taxes, such as the tax implemented in San Francisco, CA in 2014. The aim of the tax is not primarily to obtain tax revenues but to *reduce* consumption of the offending foods, so as to improve individuals' health and to reduce the cost burden of healthcare provision, including by the government.

What has happened as a consequence of these taxes? Did the taxes achieve the governments' aims?

Preliminary results from a study in Mexico showed that a 6 Peso-per-liter tax on beverages with added sugar reduced the quantity demanded by between 6% and 12% over the year of the study (2014). Consumption decreased more among low income families with the proportional decrease being between 9% and 17% over the year. The evidence also suggests that consumers switched to close – un-taxed – substitutes that did not contain added sugar such as diet sodas, 100% fruit juices, and sparkling and plain water (with between 7% and 13% increases in these categories).[7]

In Hungary, the tax has had several effects. The tax has reduced consumption, it has also caused firms to change the recipes of their food items, a sensible response because the tax is proportional to the amount of the sugar or saturated fat the food item contains.[8]

Denmark's case is more complicated. The "fat tax" definitely reduced consumption of butter, margarine and similar products, by 10 to 15%. People also changed their buying habits in terms of *where* they bought their butter and margarine: they switched to buying at discount stores. But, because these stores were aware of these buyer responses and the resulting positive shift in the demand curves they faced, they increased their prices on butter and fatty products more than high-end supermarkets did.[9]

The tax was unpopular in Denmark and was eventually repealed. Why? People had been crossing the nearby Swedish and German borders to do their shopping: one study showed almost half of Danish shoppers had gone across a border to avoid the tax.

These results illustrate the complexity of tax policy when the goal is to reduce consumption of a good. But, in places like Mexico, Denmark and Hungary, we've seen significant and important decreases in the consumption of sugary drinks and fatty foods. In Mexico, particularly, this is important for many poor people who are disproportionately affected by health problems caused by high sugar consumption, especially when they cannot afford proper treatment of cardio-vascular diseases or obesity.

The experience with taxes designed to motivate healthier consumption shows that these are one tool, among many, in the economist's toolboox to help curb unhealthy consumption while providing additional funds for public education about diet.

Checkpoint 7.10: Salt taxes & Sin Taxes: Putting the elasticity of substitution to work

Centuries ago in China, France and the British colony of India the salt tax was one of the major sources of government revenue. What is it about salt that made this tax a good way of raising revenue? Explain why sin taxes levied on goods with price elastic demand (alcoholic spirits, for example) will be effective in changing peoples behavior, but not in raising revenue, while the opposite is true for goods with inelastic demand (for example cigarettes).

7.14 Application. Willingness to pay (for an integrated neighborhood)

In Chapter 1 we illustrated the idea of a Nash equilibrium and the process by which a group of people might arrive at such an outcome by the buying and selling of homes among "Blues" and "Greens." We showed that:

- The equilibrium composition of the neighborhood – one in which none of the residents wished to switch their location and were able to do so – could be complete "segregation" of the Blues and Greens, even though everyone preferred an integrated outcome.
- Which of the multiple equilibria that would be realized was *path dependent*, like whether the farmers in Palanpur planted early or late: which equilibrium occurred depended on the recent history of the neighborhood.

These two characteristics – Pareto inefficiency and path dependence – will be results in the model we now introduce. But here we explicitly introduce a market in homes and people's willingness to pay.

In Milwaukee, Los Angeles, and Cincinnati towards the end of the last century over half of white residents, when asked, said they would prefer to live in a neighborhood in which 20 percent or more of their co-residents were African-American (one in five preferring equal numbers of each).[10] But few residents of these cities lived in integrated neighborhoods. Their preferences were elicited as part of court records in litigation concerning housing segregation in these and other cities. Most African Americans preferred fifty-fifty neighborhoods.

There are many reasons why members of a society might *not* want their residential communities to be highly segregated. Segregated living leads to racially segregated schools, friendships and other social networks. Because group members would then be unlikely to have friends in the other group, segregated living could encourage group stereotypes and intolerance leading to conflicts between groups.

The respondents in the above surveys may have misrepresented their preferences, of course, but those sincerely seeking integrated neighborhoods would have been disappointed. The housing market in these cities produced few mixed white-African American neighborhoods even though these were apparently in substantial demand.

In Los Angeles, for example, virtually all whites (more than 90 percent) lived in neighborhoods with fewer than ten percent African American residents, while seventy percent of Blacks lived in neighborhoods with fewer than 20 percent whites. Why was the result at the neighborhood level so seemingly at odds with the distribution of preferences of the individuals making up the neighborhoods? Imagine your surprise had we reported that one in five wanted a back-yard swimming pool and were prepared to pay the price for a pool, yet almost none had pools.

Why does willingness to pay get you a pool if you want one, but not an integrated neighborhood? To answer these questions we need an explanation of how highly segregated neighborhoods result, even when preferences are such that members of all groups would be better off with greater integration. In other words we need to understand why the housing market produces a Pareto inefficient level of segregation.

Residential segregation is the result of many aspects of how credit and housing markets work, and these differ across countries and even within the U.S among cities and states. But there is another less obvious and perfectly legal way that segregated neighborhoods are sustained, even when most people would prefer a more integrated community.

Preferences for integration or segregation

We will explain why this is true by modeling a single neighborhood (one of many in a large city) in which, when considered in isolation, all houses are equally desirable to all members of the population (they're identical). Peoples' preferences for living in this neighborhood depend solely on the racial composition of the neighborhood.

As before "greens" and "blues" are two population groups that are equally numerous in the city. Greens prefer to live in a mixed neighborhood with slightly more greens than blues and blues correspondingly do not prefer segregation, but prefer a neighborhood with somewhat more blues than greens.

We have *normalized* the size of the neighborhood, setting it equal to 1, so we can refer either to the fraction of greens or the number of greens by f (for fraction), which can vary from 0 to 1. with for example $f = 0.3$ meaning that the neighborhood is 30 percent green and 70 percent blue.

HISTORY This way of thinking about segregation in residential neighborhoods was developed by Thomas Schelling, a Nobel Laureate in economics. You can run a computer simulation of how a population may segregate itself, even with very modest preferences for segregation, here ncase.me/polygons/.

We will express the preferences of the greens and the blues by the maximum prices p^G greens and p^B blues would be *willing to pay* for a house in the neighborhood, each depending on the fraction of homes in the neighborhood occupied by greens f. The following willingness to pay (WtP) equations are a way to express the preferences described above:

$$\text{Blues' WtP} \quad p^B(f) = \frac{1}{2}(f+\delta) - \frac{1}{2}(f+\delta)^2 + p \quad (7.49)$$

$$\text{Greens' WtP} \quad p^G(f) = \frac{1}{2}(f-\delta) - \frac{1}{2}(f-\delta)^2 + p \quad (7.50)$$

where p is a positive constant reflecting the intrinsic value of the identical homes. Figure 7.32 shows the willingness to pay equation for the greens, with a low willingness to pay for a house in an all blue neighborhood ($p^G(f = 0)$), a greater willingness to pay for a house in an all green neighborhood ($p^G(f = 1)$), but the greatest willingness to pay in an integrated but green-majority neighborhood (with sixty percent greens). The term δ is a measure of the preference for segregation. We assume that greens and blues have similar preferences to live with their own group members, so δ is the same for the two groups.

To see how δ measures the degree of preferences for segregation, think about what would be the ideal neighborhood for a green and for a blue. If the ideal neighborhoods of members of the two groups are very different, then segregationist preferences are strong. Because the willingness to pay for a home – and therefore its value – depends on the composition of the neighborhood, the ideal neighborhood has a group composition that maximizes the value of owning a house in the neighborhood (or what is the same thing, that maximizes willingness to pay for a home there).

What would each type of person's ideal neighborhood look like? (M-Note 7.11

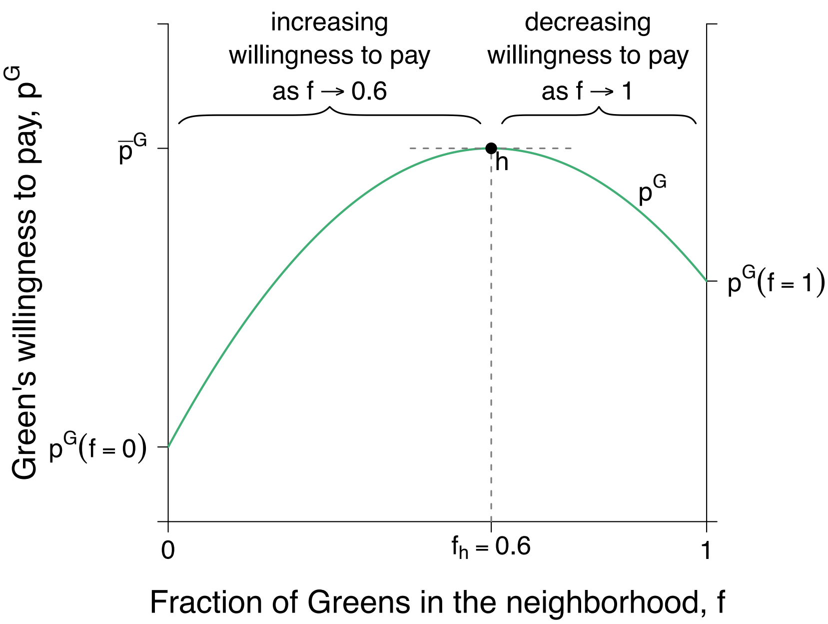


Figure 7.32: **An illustration of the Willingness to Pay of Greens, p^G.** Their willingness to pay reaches a maximum at point **h** where the proportion of greens $f_h = 0.6$. Between $f = 0$ and $f = 0.6$, Greens' willingness to pay is *increasing* as they move from being a minority to become a slight majority at 60% of the population. Between $f = 0.6$ and $f = 1$, Greens' willingness to pay is *decreasing* as they move from being a slight majority at 60% of the neighborhood to being 100% of the neighborhood. For this and the next figure we used $\delta = 0.1$.

explains how these are derived).

- **Greens:** The ideal neighborhood for greens (that which maximizes p^G) is composed of $f = \frac{1}{2} + \delta$ per cent greens
- **Blues**: Blues prefer an ideal neighborhood with $f = \frac{1}{2} - \delta$.

As the difference between the ideal neighborhoods (that for which they would pay the highest price of a home) of the greens and the blues is 2δ we will refer to δ as the preference for segregation of the two types (δ could differ between the two groups, or one group might not care about the racial composition at all, of course).

The willingness to pay curves and the degree of preferred segregation they express provide the essential building blocks for understanding how the housing market will work. But to put that information to work that we need to turn to how the market will change or not depending on its composition.

This means we need to identify the Nash equilibria of the market (where there would be no forces changing the situation) and the points that are not Nash equilibria, in which people could do better by buying or selling a house in a way that changes the composition of the neighborhood. This is called an analysis of market *dynamics* , that is how markets *change*.

REMINDER We discussed *dynamics* in Chapter 5 when exploring how the fishermen reached Nash equilibrium by comparing their marginal benefits and marginal costs.

M-Note 7.11: Finding the preferred proportions

We would like to find the proportion of greens in the neighborhood that would maximize each type's willingness to pay. We already can see that in the figure this is 60 percent for the greens. To see how we got this number we differentiate Equations 7.49 and 7.50, with respect to f and set the result equal to zero. This gives the value of f that maximizes the greens and blues respectively willingness to pay for a house in the neighborhood.

$$\text{Greens:} \quad p_f^G \equiv \frac{dp^G}{df} = \frac{1}{2} - (f - \delta) \tag{7.51}$$

$$\text{Blues:} \quad p_f^B \equiv \frac{dp^B}{df} = \frac{1}{2} - (f + \delta) \tag{7.52}$$

Now, to find the f that maximizes the house value for the two groups, we set each of Equations 7.51 and 7.52 equal to zero and isolate f:

$$\text{Greens:} \quad f_{max}^G = \frac{1}{2} + \delta \tag{7.53}$$

$$\text{Blues:} \quad f_{max}^B = \frac{1}{2} - \delta \tag{7.54}$$

As can be seen, on either side of $f = 0.5$ lie the two types' preferred proportions of green. They are separated by 2δ, that is, twice the degree of preferences for segregation.

Checkpoint 7.11: Color-blind and other preferences about segregation

With the willingness to pay curve of the blues as shown in the figure, draw a new willingness to pay curve of the greens based on an alternative assumption: greens do not care at all about the composition of the neighborhood and that the fixed value they place on homes there is greater than the value that blues place on an all-blue neighborhood, but less than the value that blues place on their ideal neighborhood. If this were the situation, what color-compositions (meaning values of f) would you expect to see?

7.15 Application: Market dynamics and segregation

Remember, an equilibrium is defined by the absence of change. So to determine what level of integration or segregation we would expect to observe (the equilibrium) we need to better understand the process by which the neighborhood composition will change as a result of the dynamics of the market.

Home sales: A pathway to segregation

DYNAMICS refers to how some market or other economic entity changes.

To do this, we now consider the conditions under which a house inhabited by a green might be sold to a "blue family", or vice versa. Imagine that prospective buyers from outside the neighborhood visit the neighborhood and just knock on the door of a randomly selected house. A sale takes place as long as the house is worth more to the visitor than it is to its current owner. If the current owner values it s much or more highly, no sale takes place. So houses never change hands among the same types (because they value the houses identicallly.)

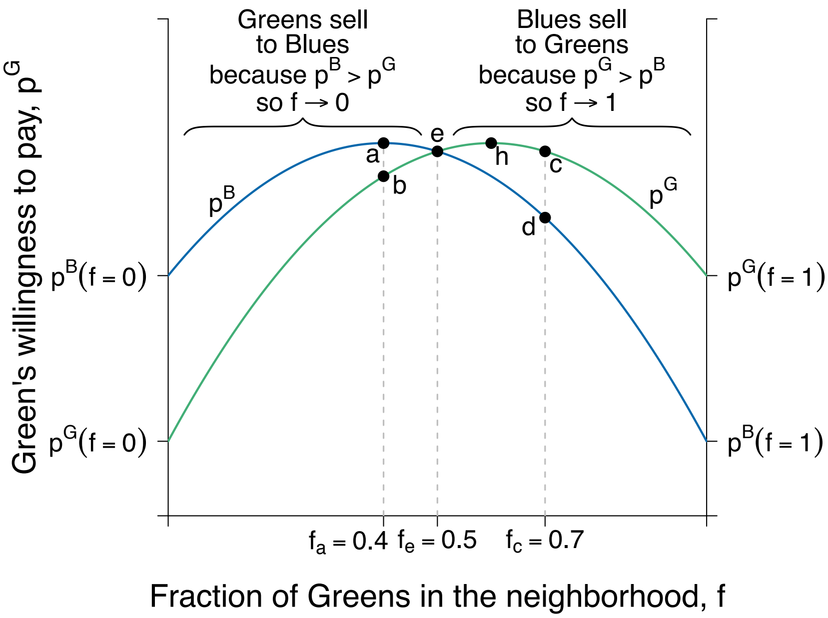


Figure 7.33: **Dynamics of residential segregation**. At **e**, the residents are at an unstable equilibrium. If the history of the neighborhood is such that it starts off at $f = 0.4$, corresponding to points **a** and **b** on the green and blue price curves, then, as a result of $p^B > p^G$, greens will sell to blues and the greens will leave the neighborhood. This will continue until there are no greens left. The neighborhood will then remain at $f = 0$, that is, no greens in the neighborhood. With $f = 0$ a house is worth more to a blue than to a green, but there are no "green houses" left for a blue to buy. A similar process, but in the opposite direction works if in some period $f = 0.7$, then, as a result of $p^G > p^B$, blues will sell to greens and blues will leave the neighborhood until $f = 1$, that is, no blues in the neighborhood.

But if a green visits the house of a blue, a sale will take place if $p^G > p^B$ and not if $p^G \leq p^B$. Remember that the homes are identical. While the residents in the neighborhood care about the composition of the neighborhood, they are "color blind" when it comes to buying or selling houses: they sell if they are offered a price above what their home is worth to them, irrespective of the color of the buyer.

REMINDER This is exactly how we model buying and selling in our bargaining model of general equilibrium in Chapter 14.

Figure 7.33 illustrates this. To check that you understand how it works, imagine that the fraction of greens in the neighborhood at the moment is a bit greater than f_h (such as at $f_c = 0.7$). What would you anticipate happening? Check the values that greens and blues would then place on their homes: greens would value a home in the neighborhood *more* than a blue would value living in that home.

So a visiting green would be willing to pay more for a blue's home than the home was worth to the blue. A transaction would take place – a home once occupied by a blue would now be a "green home." What would happen when the next green showed up? The same thing. When would it stop? When there were no "blue houses" left. The neighborhood would be entirely green.

Of course had the initial value of f been a bit below f_h the process would have run in reverse, and we would have ended up with an entirely blue neighborhood. Complete segregation would be the result in either case. The all green segregated neighborhood is a Nash equilibrium because there is no incentive for a blue to buy in the all green neighborhood given how much homes there are valued by the greens living there and therefore the minimum price at which they would be willing to sell.

Think about it this way: individual neighborhoods in a city containing green and blue people will have neighborhoods of greens and neighborhoods of blues that are totally segregated. We could say that each neighborhood is *locally homogeneous*, but the city is composed of neighborhoods of the different types and is therefore *globally heterogeneous*.

Which composition any particular neighborhood will exhibit will depend on history: if, in the recent past, f was less than $f = \frac{1}{2}$, we would expect to find $f = 0$. If f had been greater than $\frac{1}{2}$, we would expect to find $f = 1$. In other words we would expect to see a completely segregated neighborhood, one way or the other.

M-CHECK The outcome $f = \frac{1}{2}$ need not even be possible, as would be the case if there were an odd number of homes!

There is one composition of the neighborhood that is both integrated and a Nash equilibrium, namely $f = \frac{1}{2}$. At that composition the values of the homes to Greens and Blues do not differ. But we would not expect to see such an outcome. The reason is that if for any reason the composition moved a bit higher or lower, then the dynamic of buying and selling that propelled the neighborhood to complete segregation above would begin, with the same outcome of a completely segregated neighborhood. The composition $f = \frac{1}{2}$ is therefore an unstable Nash equilibrium: a small movement above or below it will set in motion further moves away from the equilibrium.

STABLE EQUILIBRIUMAn equilibrium is stable if a sufficiently small displacement away from the equilibrium is self correcting leading movement back towards the equilibrium.An equilibrium is unstable if the reverse is true.

Pareto-inefficient segregation: A coordination failure

A fully segregated neighborhood (either $f = 1$ or $f = 0$) is not the preferred outcome. You can see from point **e** in Figure 7.33 that both Greens and Blues prefer a neighborhood with an equal number of each group to a neighborhood in which they are the only type living there. Greens and Blues alike would *prefer* to live in the *other* group's ideal neighborhood than to live in their own segregated neighborhood. All of the points **a** through **e** in the figure are Pareto superior to either of the complete segregation outcomes.

To understand why this Pareto-inefficient Nash equilibrium occurs, set aside the Greens and Blues for a moment and think about what kind of economic entity a neighborhood is. Take the age composition of a neighborhood, whether mostly young people live there, or mostly middle aged or elderly, or mixed and so on.

The age composition of the neighborhood has some of the aspects of a *public good*. It is *non-rival* because a person visiting the neighborhood and experiencing its composition does not subtract the experience of its composition from anyone else. It is *non-excludable* because the change applies to everyone in the neighborhood. The sale of a home – say from a middle-aged family to a young family – also has external effects on those not involved in the sale because it changes the age composition of the neighborhood for everyone.

At the beginning of this section we asked: Why does willingness to pay get you a swimming pool if you want one, but not an integrated neighborhood? The answer, referring back to the distinctions we made in Chapter 5, is that:

- the pool is a private good, for which your willingness to pay provides sufficient motivation for some person or business to make the pool available
- the composition of the neighborhood is a public good, resulting from the Nash equilibrium of a large number of peoples actions; without coordination among the actors no person's willingness to pay for a more integrated neighborhood will result in the desired public good being realized.

The coordination failure represented by the completely segregated Nash equilibria does not occur because people do not want to live in integrated neighborhoods. To see this, imagine that the δ that we used in making the above figures been 0.01 rather than 0.10. This means that the ideal composition of the neighborhood would have been virtually identical for the Blues and the Greens. But the result would have been the same: total segregation.

This occurs because buying or selling a home has external effects on others, that buyers and sellers to not take into account. For example, consider the neighborhood with composition f_c Figure 7.33. With 70 percent Greens both Greens and Blues in the neighborhood would prefer a "Bluer" composition. But, as you can see by comparing points **c** and **d**, homes in that neighborhood are worth more to Greens than to Blues. So Blues will sell to Greens, and the neighborhood will not become more Blue (as both Green and Blue residents prefer). It will move in the other direction, becoming Greener. As a consequence of exchanges, all of which are voluntarily entered into, the value of the housing in the entire neighborhood falls: everyone is worse off except for the two who make the sale and purchase.

Like the coordination problem of over-fishing, or planting late in Palanpur this undesired outcome occurs because the residents are engaged in a non-cooperative game. They have no way of jointly agreeing on a neighborhood composition that they would all prefer to the segregated outcome.

Achieving a desirable and enduring neighborhood composition is a challenge that is not readily addressed by the measures we considered, for example, in the case of fishing.

In Chapter 5 we

Checkpoint 7.12: Self-correcting segregation?

a. Be sure you can explain why complete segregation is stable, that is, self-correcting.
b. Suppose all of the houses in the neighborhood were owned by a single individual, who could rent them for a monthly fee equal to some fixed fraction of the value of the house. What value of f would the owner implement, assuming that it was legal and otherwise acceptable for the owner to consider the type of the renter (Blue or Green) in offering a rental.

7.16 Conclusion

The concepts you have learned – especially the price elasticity of demand – are essential tools for understanding the effects of taxing sugary drinks, carbon emission, and other policies. We have learned, for example, that raising the price of a good with a substantial price elasticity of demand – such as sugary drinks – can have a substantial effect on consumers' purchases. Correspondingly raising the price of a good with a small price elasticity of demand – rice, medical care – will have more modest effects on how much is purchased.

But if we are interested in public policy – such as sin taxes – then we need to know how the tax will affect the price. And this depends on the costs of producing the good, and how these vary with the amount of the good produced. In Chapter 6 we explained how the owners of a firm will select a mix of inputs of labor capital goods or other inputs to minimize the cost of producing any given amount of their product. In the next chapter we turn to how costs of producing a good depend on how much of the good is produced, as described by what economists call cost curves.

Making connections

Feasible sets and indifference curves: These are the basis of our using constrained optimization to analyze the consumer's willingness to pay and resulting demand curves.

Social preferences: Our analysis of consumption as a social activity and Veblen effects takes account of the fact that people care not only about their own consumption but also about the consumption of others to whom they are compared in assessing their social status.

Economics as an empirical social science: Models provide a lens for refining how we look at some aspect of a complex empirical reality whether it be the differing patterns of work hours over the 20th century or the effect of a price change on the demand for fish.

Models: The map is not the territory. Our data on fish markets and declining work hours reveal a much more complex empirical reality than our models capture, which necessarily leave out aspects of the problem that might explain, for example, why Asian buyers paid less for their fish purchases, or why work hours fell so much more in Sweden than in the U.S.

Rents: As with other voluntary exchanges, people purchase goods because they expect to be better off by making the purchase compared to their fallback position (not making the purchase). Consumer surplus is a form of rent.

Ordinal and cardinal preferences: Our analysis of individual demand did not require utility to be measured cardinally; an ordinal ranking was sufficient. But when we aggregate the consumer surplus of the people making up a market, we are assuming that the marginal utility of money left over for other purchases is a) the same for everyone involved and b) is measured in the same units as money itself, a cardinal measure in which we can say, for example, that Harriet has twice as much utility as her brother.

Economics and public policy: The analytical tools for the analysis of demand – willingness to pay, principle of demand, demand curve, income and substitution effects – provide the basis for better understanding public policies, for example sugar and fat taxes to address the rise in obesity and carbon taxes to mitigate global climate change.

Fairness: These policies may place greater burdens on the less well-off and thus raise questions of distributive justice. The dubious assumption on which consumer surplus is based – that an additional dollar is of the same value to a poor individual as to a rich individual – makes it difficult to apply this concept to cases in which individuals of differing wealth levels are involved.

Important Ideas

marginal rate of substitution	principle of demand	demand curve
quasi-linear utility	quadratic, quasi-linear utility	linear demand curve
change in demand with change in price	substitution effect	income effect
market demand	horizontal summation of individual demands	slope of market demand curve
slope of demand curve	price elasticity of demand curve	linear demand curve
theory and data	noise in data	interpreting evidence

Mathematical notation

Notation	**Definition**
x, y	individual quantities of a good
p_x, p_y	market prices of two goods x and y
m	budget available for buying goods
$u()$	utility function
α	Cobb-Douglas intensity of preference for good x
$(1-\alpha)$	Cobb-Douglas intensity of preference for good y
$g()$	the non-linear part of a quasi-linear utility function
h	fraction of the day spent working for wages
l	fraction of the day that is free time
w	wage
$\underline{x}$	consumption of the very rich
v	Veblen effect parameter
β	slope of the demand curve
ε	intensity of preference for good y in utility function of substitutes
n	number of individuals in a market
X	market demand
η	elasticity of market demand with respect to price
f	fraction of Greens in a neighborhood
δ	degree of preferences for segregation

Note on superscripts: *: Pareto-optimal outcome; v: outcome with Veblen effect; QL: quasi-linear utility function; QQL: quadratic quasi-linear utility function; sub: substitution effect.

Discussion questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

See reference list.

8

Supply: Firms' costs, output and profit

> We might as reasonably dispute whether it is the upper or the under blade of a pair of scissors that cuts a piece of paper, as whether [prices are] governed by utility or cost of production . . . We have next to inquire what causes govern supply prices, that is prices which dealers are willing to accept for different amounts.
>
> Alfred Marshall, Principles of Economics, 1890 (Book V Chapter 3)

In the summer of 1943 Allied troops under General George Patton landed on the island of Sicily and began the battle to retake Italy from Italian Fascist forces. Hitler soon ordered German units to the south of the peninsula to reinforce the retreating Italians. Palermo, the capital of Sicily, fell to American units on July 23.

At the same time, in support of the war effort, the Ford Motor Company had just launched the production of a new armoured truck the M20-GBK in Chicago. A total of 3791 M20s would be built over the next two years. General Patton accompanied by his pet bull terrier Willie would inspect one of the first to arrive on the European front.

Because quickly getting the M20s built and transported to the fronts was a top priority, Ford kept careful account of the person days that were devoted to the production of the trucks, tabulating "man days of labor" (a great many of Ford's workers were women during the war) and total units produced each month. Although costs other than labor were of course involved – machinery and materials important among them – and we lack data on these inputs. Ford's records nonetheless provide a glimpse into how the costs of producing goods vary with the amounts produced.

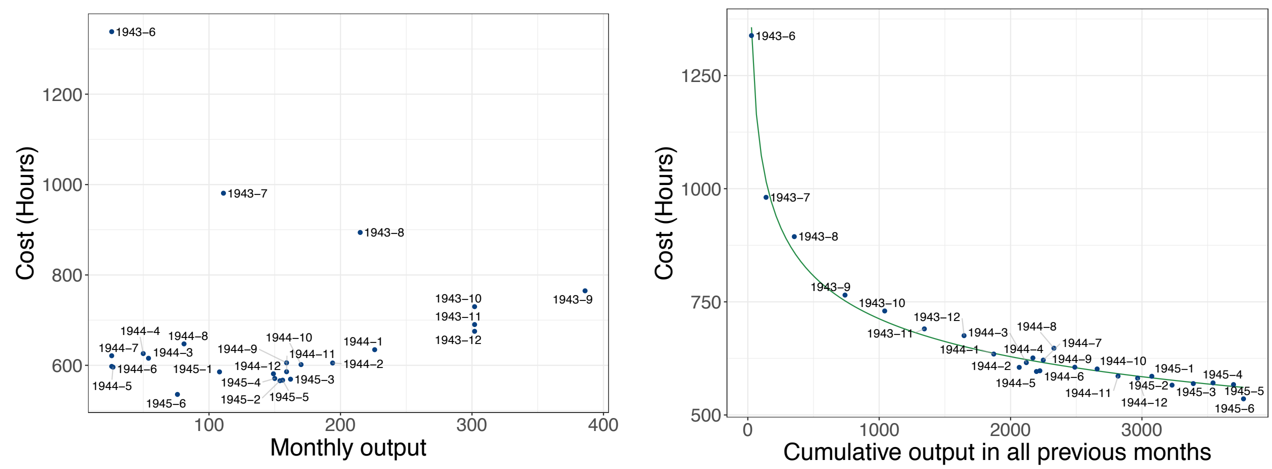
Figure 8.2 shows the evidence (photographed from records in Fords historical archives). In panel a., there is surprisingly little evidence that costs were lower for larger production runs. There is no relationship at all between the amount produced and the cost per unit. For example, we see that while the most

DOING ECONOMICS

This chapter will enable you to :

- Understand the difference between accounting profits and costs and economic profits and costs.
- Summarize the facts about how the cost per unit produced varies with the level of output as a result of economies of scale and other influences.
- Explain why, if competition in labor and product markets is limited, a firm's owners profit by restricting the firm's hiring and sales.
- Show why the price markup over costs is greater if there is less competition among firms.
- Explain how in some cases we can derive a firm-level and market supply curve based on firms' willingness to sell.
- See how production and sales of a good typically allow rents for both consumers and owners in the form of consumer surplus and economic profits.
- Understand why the Nash equilibrium price and quantity transacted in the model of competition among price taking buyers and sellers is Pareto efficient.
- Use the model of competitive supply and demand to show the effect of a tax on consumer surplus, economic profits, tax revenues and deadweight loss.
- Contrast two benchmark models: a price taking firm constrained by rising costs and a price making firm constrained by demand.

Figure 8.1: **The M20 GBK armored truck built by Ford Motor Company 1943-1945.**



(a) Costs were unrelated to number of units produced that month

(b) Costs declined as the total number of units ever produced increased

Figure 8.2: **Labor time to produce one M20 GBK Armored Truck.** The horizontal axis in Panel a is the usual dimension used for a cost curve: number of units produced in a given time period (in this case a month). The vertical axis is average labor hours devoted to the production of a unit in the given month. The horizontal axis in Panel b is the cumulative totl number of M20 GBK's produced since June 1943, when production began. Source:

costly month was June 1943, the beginning of the production run, when only 26 units were produced, the fourth most costly period was just four months later when the greatest number of units were produced.

Panel b. tells an entirely different story. Here we show the cost per unit in each month plotted not against how many were produced *that* month but instead, against how many units had been produced in all of the *previous* months. The downward sloping line is evidence of what is called "learning by doing": costs were cut in half as the entire Ford team – engineers, managers and workers alike – learned from their accumulating experiences. Remember from Chapter 6 that learning by doing is one of the sources of decreasing costs as output increases.

In panel b., points farther to the right must be for later months (because they represent larger amounts of previous production) so it is possible that new technologies became available for the production process,accounting for the decline in costs. But this seems unlikely in light of the two-year time period and how long it takes to conceive, develop and install new technologies.[1]

An even more remarkable fall in the costs of photovoltaic modules that make up solar panels occurred over the period 1975-2015. Figure 8.3 records both the cost and the price of photovoltaics (in 2015 dollars to take account of inflation) using data from the U.S., China, and other countries. The costs in 2015 were an astounding 1 percent of what they had been just 40 years earlier.

The curve in this figure shows a rapid percentage decline. In contrast to the decline in the costs of producing the M20 armored truck, which tapered off in the second year, the cost declines in PVs appear to have accelerated towards the end of the period.

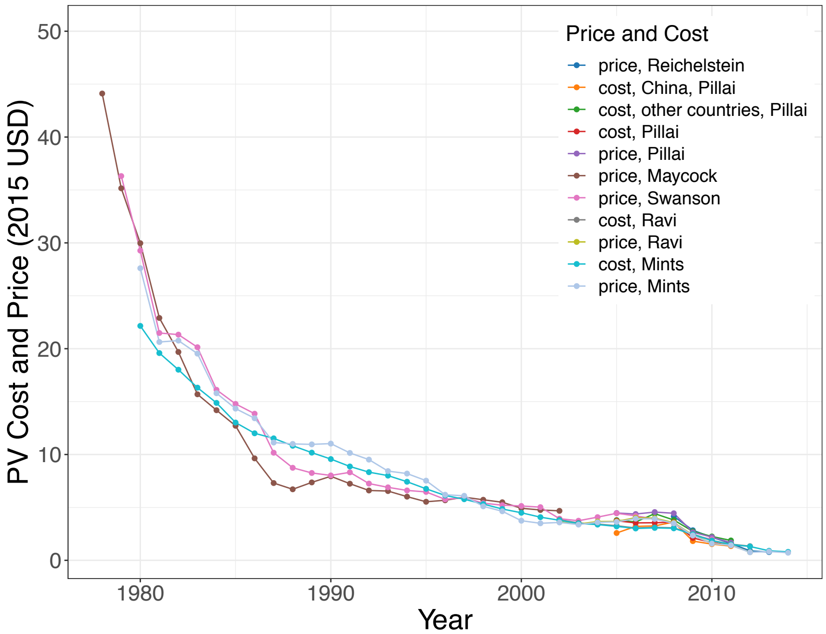


Figure 8.3: **The Declining Cost of Photovoltaic Modules used to generate Solar Power.** We show the results of the available studies of either prices or costs of the modules. Source: Kavlak, McNerney, and Trancik (2018).

Also in contrast to Ford's M20, the major contributor to the decline in the costs of PVs was research and development (both private and government funded in about equal measure). Learning by doing accounts for about ten percent of the cost reduction. Despite a substantial increase in the scale of production of the PVs, economies of scale account for about a quarter of the cost reduction, though this contribution has recently increased.

Figure 8.4: **Solar photovoltaic modules at work**

Reducing cost is an essential objective of owners of firms. They seek to expand the value of their ownership of the firm, which they do by making the firm be as profitable as possible. The means by which they obtain profit is by employing people and using machines to produce goods to sell to customers.

In this they face the same kind of "doing the best you can" problem that you studied earlier as the individual's problem of constrained utility maximization. To solve the problem, the owners (or the managers they have hired) decide such things as how to produce – at the least possible cost – the good they wish to sell and how much to pay the employees they hire.

In Chapter 6 you learned how to represent the relationship between physical inputs and outputs by production techniques and production isoquants. The production functions summarizing these relationships apply just as much to household subsistence production such as Alex choosing between fish and shirts, or a centrally planned such as existed in the Soviet Union under Communist Party rule, as to production organized by in modern, privately owned firms.

8.1 Costs of production: An owner's eye view

If we want to understand why a particular amount of a some good is produced and put on the market, a critical piece of information is the cost of producing the good. To understand costs we take the "owner's eye view," a perspective that will always ask "if I did not invest my money producing this particular product, what else could I do with my funds, and what profits would I make in this alternative use?" The owner's eye view, in other words, takes account of the *opportunity costs* of how the owner's funds are used.

Keep in mind that economists sometimes say things like "firms decide how much to produce." Or the firm does this or that. But a firm is not a single person or a machine, it is is a conglomeration of people. Most of the people do not have the authority to make decisions about such things. Among the people making up the firm, there may be conflicts of interest about the decisions being made. What we mean when we say "the firm does" something is that the *owners* of the firm or their *managers* decide on what is to be done and then exercise their authority over others in the firm to carry out their decisions.

For simplicity, we consider a vertically integrated firm that does not purchase any inputs other than labor and the capital goods (say, machines) owned by a single owner. And to simplify the analysis even further we assume that at the beginning of each period – say a year – he purchases the machines and they run for a year and then have to be replaced. With these machines he has a technology that specifies the number of units of both machines and labor time required to produce a unit of the firm's output.

M-CHECK The characterization of costs here is *different* to the one we used in Chapter 6. Here, we explicitly recognize the role of the **opportunity cost of capital**. In Chapter 6 we avoided considering opportunity costs by assuming that capital goods are rented, not owned. The insights of Chapter 6 remain true as the same tools can be applied once the opportunity costs of capital are included.

A firm's costs for a unit of output are then determined by the following:

- *wages* (w) paid for the amount of labor required to produce one unit (a_L)
- the *price of capital goods* (p_K) times the amount of capital goods required to produce one unit (a_K) called the accounting cost and
- the *opportunity cost* of that capital (ρ).

Putting these ideas together, we obtain the cost to produce one unit of a good. The cost to produce one unit equals the labor cost plus the accounting cost of the capital goods plus the opportunity cost of devoting the value of the capital goods to the particular production process, or:

$$\text{cost of 1 unit}, c = wa_L + p_k a_k + \rho p_k a_k \tag{8.1}$$

$$= wa_L + p_k a_k (1 + \rho) \tag{8.2}$$

The third term on the right hand side of the top equation is the opportunity cost of the firm's capital stock. The owner cares about their opportunity costs because had he not bought the machines, he could have made some other investment that would have yielded him some profits, perhaps purchasing a very safe financial instrument like a government bond.

OPPORTUNITY COST OF CAPITAL *The opportunity cost of capital* is the rate of profit that the owner would make on his next best alternative investment, which could be a low risk government bond, or an investment in some other firm.

Let the total cost of a single unit of the capital good be c_K.

$$c_k = p_k + \rho p_k = (1+\rho)p_k \tag{8.3}$$

c_k is therefore the sum of the price of one unit of the capital good (p_k) and the opportunity cost of devoting the owner's financial resources to this particular use (ρp_K). Now, including the cost of capital goods c_k into the cost function, we have:

$$c \quad = \quad wa_L + c_k a_k \tag{8.4}$$

Equation 8.4 tells us that the cost (c) to produce a single unit of output depends on the factors of production (labor and capital goods) required, the price of those factors, and the opportunity cost of capital goods.

The difference between the owner's eye view of costs and the viewpoint of an accountant or an engineer is this: the engineer or the accountant would ignore the opportunity cost of the use of the owners funds and simply count the actually outlays – the payments to workers, and the cost of the machinery "used up" in producing the good. They would think about what is called the "accounting cost" or:

$$c^A \quad = \quad wa_L + p_k a_k \tag{8.5}$$

Because the owner and the accountant look at costs from a different perspective, they also have a different idea of how to count profits.

8.2 Accounting profits and economic profits

ACCOUNTING PROFIT AND ECONOMIC PROFIT *Accounting profit* is the difference between sales revenue and the direct cost of the inputs required to produce output, excluding the opportunity cost of the funds tied up in financing long-lived plant and equipment. *Economic profit* is accounting profit less the opportunity cost of funds tied up in long-lived plant and equipment evaluated at the average rate of return of the economy, ρ.

A firm's economic profits are given by its total revenues $r(x)$ which is just total output times the price at which it is sold, $p(x)$, minus total costs $c(x)$, which we can divide by the total output (x) to find the economic profit per unit:

$$\begin{aligned} \pi(x) &= r(x) - c(x) \\ &= px - cx \\ \therefore \text{Profit per unit} \quad \frac{\pi}{x} &= \frac{px - cx}{x} = p - c \end{aligned} \tag{8.6}$$

We can substitute c from Equation 8.2 into Equation 8.6 for the per-unit profit as follows:

$$\frac{\pi}{x} \quad = \quad p - wa_L - p_k a_k(1+\rho) \tag{8.7}$$

Equation 8.7 tells us that the per unit profit is determined by the price at which goods are sold (itself determined by the demand curve and how many goods are on the market), the costs of the factors of production per unit of output, and the opportunity cost of capital.

But from the owner's perspective it is not the profits per unit, or even the total profits that matter: what he cares about is the total profits relative to how much money he invested in this firm that he could have invested elsewhere, making some profits in this alternative use of his funds. What he cares about is the rate of economic profit, which is the ratio of the firm's economic profits per unit to the value of its capital stock. The value of the capital stock is the price of capital goods times the total capital goods used by the firm.

RATE OF ECONOMIC PROFIT is the ratio of the firm's profits per unit to the value of its capital stock. The value of the capital stock is the price of capital goods times the total capital goods used by the firm.

We can express the **rate of economic profit** (r^E) either as total profits divided by total value of the capital stock or as the profit per unit of output ($\frac{\pi}{x}$) divided by the value of capital per unit of output ($p_K a_K$), as shown in Equation 8.8:

$$r^E = \frac{\pi}{x p_k a_k} \quad = \quad \frac{\frac{\pi}{x}}{p_k a_k} \tag{8.8}$$

Accounting profits are given by total revenues minus the direct cost of production without taking account of the opportunity cost of capital. Or, what is the same thing (we show this in M-Note 8.1), the rate of accounting profits (r^A) equals the rate of economic profit (r^E) plus the opportunity cost of capital:

$$r^A \quad = \quad r^E + \rho \tag{8.9}$$

So, the rate of economic profit is zero when the rate of accounting profit equals the opportunity cost of capital:

$$\text{Zero (economic) profit condition} \qquad r^E \quad = \quad r^A - \rho = 0 \tag{8.10}$$

Equation 8.10 is called the *zero profit condition*. (It could be more accurately named: the zero *economic* profit condition.) In the owner's eye view the condition is important. When the zero profit condition is satisfied, the owner makes the same profits on his investment in the firm that he would make by investing it in his next best alternative. If economic profits are positive, he is happy with his investment and will perhaps increase it. But if the opportunity cost of capital exceeds the accounting profits he is making, then economic profits are negative and he will be better off investing elsewhere.

REMINDER In Chapter 6 we explained that the owners of a firm seek to maximize their profits in choosing how much to produce by:

- first determining the least-cost way to produce any level of output they might consider producing, and then
- using that information, along with information on the demand for their product and the degree of competition among firms selling similar products, to determine the level of output that would yield the greatest profits.

We explained the first step in Chapter 6. Here we summarize the information from the first step in a cost function, and using this we explain the second.

M-Note 8.1: Economic profits and accounting profit

To see the relationship between economic profits and accounting profits, we can substitute the value of $\frac{\pi}{x}$ given by Equation 8.7 into Equation 8.8 to have the following expression for the economic rate of profit :

$$r^E = \frac{p - (wa_L + p_k a_k(1+\rho))}{p_K a_K}$$

Then we can rearrange this to separate out the opportunity cost of the capital goods used

$$\begin{aligned} r^E &= \frac{p - wa_L - p_k a_k - \rho p_k a_k}{p_K a_K} \\ &= \frac{p - wa_L - p_k a_k}{p_K a_K} - \frac{\rho p_k a_k}{p_K a_K} \\ &= r^A - \rho \end{aligned}$$

The last step is true because

$$r^A = \frac{p - wa_L - p_k a_k}{p_k a_k}$$

So the rate of economic profit is equal to the accounting rate of profit minus the opportunity cost of capital.

8.3 Cost functions: Decreasing and increasing average costs

In Chapter 6 we explained how the owners or managers of a firm will select a technology and a combination of inputs of labor (l) and capital goods (k) to minimize the cost of producing a particular level of output x given the wage rate w and the cost of the capital goods (including the opportunity cost) c_k. The solutions of this constrained optimization problem, applied to every level of output the firm might produce then constitutes the firm's total cost function $c(x)$ which represents the minimum total cost of producing output in the quantity x at the given costs of inputs.

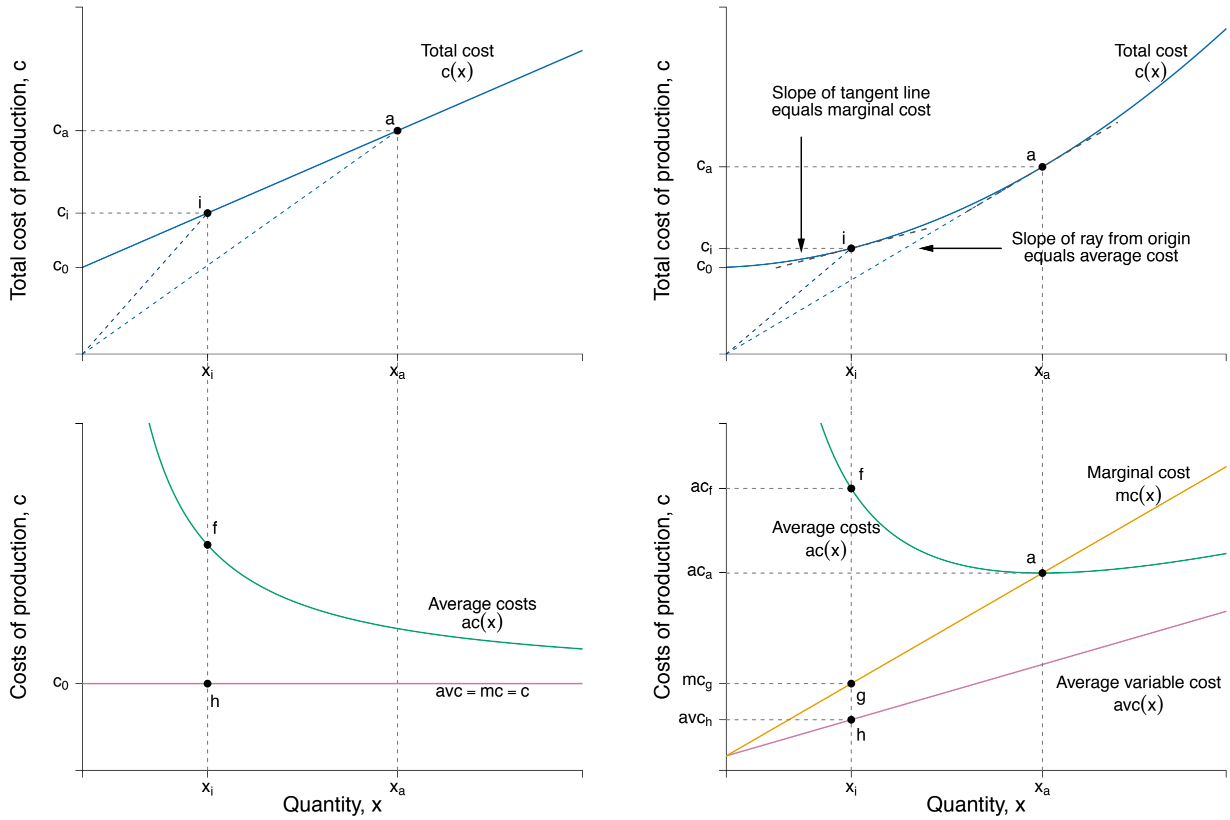
The total cost curves in the two top panels of Figure 8.5 therefore show the minimum cost required to produce each level of output x, for a given wage w and cost of capital goods c_k. Total cost always increases as output increases.

In both panels there is a *fixed cost*, c_0, the firm must pay even at a zero level of output. A cost is said to be *fixed* if the firm cannot avoid it even if it produces nothing. Examples are the opportunity cost of the firm's capital goods which would be very costly to de-install and sell if the firm chooses to produce less. Also included are any patents or other intellectual property required in the production process.

Intellectual property such as patents, trademarks and copyrights even if owned by the firm represent a fixed cost because the opportunity cost of the firm using them exclusively is the sum they could collect in fees by selling

TOTAL, AVERAGE, AND MARGINAL COST *Total cost* is the minimum cost of producing output level x, $c(x)$ at given costs of inputs; *average cost* is the cost per unit produced, $ac(x) = \frac{c(x)}{x}$; and *marginal cost* is the effect on total cost of producing an additional unit of output , $mc(x) = \frac{\Delta c(x)}{\Delta x}$.

EXAMPLE Think about producing this book. Writing it has incurred large fixed costs (the research and writing, decades of experience in the classroom, the publisher's advertising and editorial advice) called first copy costs, namely the cost of producing just one copy of the book. But the marginal costs which are approximately equal to the average variable costs, that is the costs of producing additional copies of the physical book or e-book, are very limited. For some texts, like *The Economy* by The CORE Team an open access introduction to the field, the fixed costs of producing the content was in excess of 2 million U.S. dollars, but the marginal cost is literally zero and it is



(a) Linear total cost with declining average costs and constant marginal costs

(b) Increasing total cost with U-shaped average costs and rising marginal costs

Figure 8.5: **Total, average, marginal and variable costs: increasing and constant marginal costs**. Panel **a.** and Panel **b.** represent different cases: rising marginal costs on the right and constant marginal costs on the left. The upper and lower graphs show two ways of looking at costs for these two cases. All of the information used to produce the lower graphs is contained (but in different visual form) in the upper graphs. In the figure we used the following cost function with $c_0 > 0$ in both panels, $c_1 >$ and $c_2 = 0$ in the right panel and $c_2 > 0$ in the left panel. Total Cost: $c(x) = c_0 + c_1 x + \frac{c_2}{2} x^2$ Marginal Cost: $mc(x) = c_1 + c_2 x$ Average Cost: $ac(x) = \frac{c_0}{x} + c_1 + \frac{c_2}{2} x$ Average Variable Cost: $avc(x) = c_1 + \frac{c_2}{2} x$.

M-CHECK Marginal cost is the change in total cost, Δc, associated with a small change in total output, Δx, expressed as the ratio of the former to the latter, or $mc = \frac{\Delta c}{\Delta x}$, where Δx is small. Marginal cost is not the cost of producing the last unit (the units produced are identical, and the cost of each is the average cost, or $\frac{c}{x}$.) If the cost function is $c(x) = c_0 + c_1 x + \frac{c_2}{2} x^2$, then the marginal cost is the derivative of c(x) with respect to x or $\frac{dc(x)}{dx} = c_1 + c_2 x$.

or leasing them. Other fixed costs are the wages and salaries of people who will be employed even if the firm is producing nothing, and the costs of any licenses, lobbying or advertising in which the firm invests independently of its output.

Some costs are fixed in the short run but not in the longer run. Buildings or equipment not being used because a firm has downsized can eventually be sold or rented to others.

Marginal cost is the change in total costs associated with a small change in output or $\frac{\Delta c(x)}{\Delta x}$, which is the slope of the cost curve at any given point such as x_1.

Average cost is the ratio of total cost to output, which is the slope of the line from the origin to the point on the total cost curve corresponding to a given output, x_1.

Average variable cost is the slope of the line from the intercept of the total cost curve to the point on the total cost curve. Examples of variable costs are the wages and salaries paid to employees engaged in production, the costs

Fixed Cost (c_0)	*Linear Cost* (c_1)	*Quadratic Cost* (c_2)	*Description*	*Example*
0	>0	0	Constant ac = mc	Many firms in the long run (no fixed costs)
>0	>0	>0	U-shaped ac curve; rising mc	Many firms in the short run, (right panel in Figure 8.5)
>0	0	0	Declining ac; mc=0	Digital production with "first copy costs"
>0	>0	0	Declining ac, Constant mc	Many firms (left panel in Figure 8.5)
0	>0	<0	Declining avc	Many firms

Table 8.1: **Cost functions, Average Cost (ac), Marginal Cost (mc) and Average Variable Cost (avc)** The cost function is $c(x) = c_0 + c_1 x + \frac{c_2}{2} x^2$. Evidence on cost functions is presented below. See Table 8.2

of inputs and energy used in the production process, and the wear and tear on the equipment used in production. corresponding to the given output (not shown in the figure).

SHORT RUN, LONG RUN These terms do not refer to the passage of time. Instead they differ in what is assumed to be "held constant" (termed "exogenous") when analysing a problem (in the short run) that may become variable ("endogenous") in the long run. About a firm's costs, for example, we assume that its stock of capital goods and technology is exogenous (constant) in the short run, but may be varied in the long run. So the firm's fixed costs (the opportunity cost of the capital goods it owns) will not be fixed in the long run. A reason why the short run average cost curve may be U-shaped is that with a fixed stock of capital goods and technology increased production incurs higher unit costs, which could be avoided in the long run if the amount of capital goods and the technology in use could be changed.

In the bottom part of Panel b. in Figure 8.5 we show these average, marginal and average variable costs. For a firm producing an amount x_1 units of its output:

- Point **f** shows the average cost, that is, the slope of the ray from the origin to point **i** on the total cost function in the top panel.
- Point **g** is the marginal cost of production, namely the slope of the total cost function at point **i** in the top panel.
- Point **h** is the average variable cost if the firm is producing x_1.

Notice in the top right graph that for a point like **a** where the ray from the origin has the same slope as a tangent to the curve, it will be the case that is $mc = ac$, which, as you can see in the lower left graph (also point a) the average cost is at a minimum.

Checkpoint 8.1

For each of the cases (the rows) in Table 8.2 draw the average, average variable and marginal costs curves. Hint: Whenever you draw average and marginal cost curves, if the marginal cost curve intersects the average cost curve, make sure this happens at the point where average cost is lowest! At zero output, the average variable cost is equal to the marginal cost, and is increasing when it lies below marginal cost.

8.4 Application: Evidence about cost functions

As you will see later in this chapter and in Chapter 9, how the process of competition works and how we should model firms' owners' profit making strategies hinge on facts about cost curves and in particular, does the average cost curve of a firm slope upwards, so that costs per unit rise with increased total output.

There are two sets of influences on the shape of the cost curve:

- whether the production function used by the firm is characterized by economies of scale, diseconomies of scale or constant returns to scale; and
- whether the cost of acquiring inputs – c_k and w in our example above, but including other inputs – is greater, less, or the same for differing levels of output.

Here are two examples of how input prices will depend on the amount the firm is producing. Hiring a large amount of labor in a small labor market may require higher wages than would be necessary if the firm were hiring fewer (we will model such a case in Section 11.12). A large buyer – like Walmart, for example – may be able to bargain for lower wholesale prices of the merchandise it sells than a smaller firm.

REMINDER If doubling all inputs more than doubles output then the production process exhibits economies of scale; if the increase in output is less than double, we have diseconomies of scale. And if output doubles we have constant returns to scale.

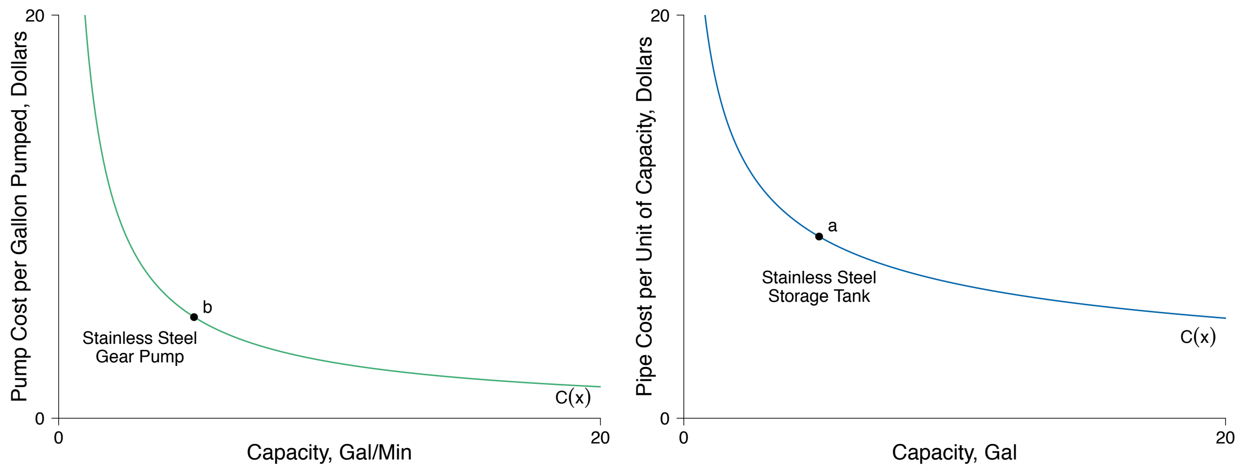
We want to find out, then, what empirical research has been able to determine about the shape of the average cost curve in different industries. There are many ways to investigate this including:

- *Engineering data:* Engineering evidence about the relationship between physical inputs and output, to determine if larger outputs can be produced with less than proportional increases in inputs.
- *Statistics on costs:* Statistical studies of how costs vary with the amount produced.
- *Estimates of production functions allowing inferences about economies or diseconomies of scale or constant returns to scale:* Statistical estimates of production functions, for example to determine if the sum of the exponents for the firms' inputs in a Cobb-Douglas production function are less than or greater than one, or about one.
- *Surveys of managers and owners:* Direct evidence from firm managers on their understanding of the extent of economies of scale and the shape of their cost curves.

FACT CHECK Here is an example of how an engineer might use knowledge of the technology to estimate costs. The cost of piping for conveying liquids in a plant is proportional to the circumference of the the pipe, but its capacity (in gallons) is proportional to the area of a cross section of the pipe. Because doubling the circumference of a circle (the cost of the pipe) more than doubles the area of the circle (the capacity of the pipe) the costs per gallon conveyed by the pipe is less the larger is the pipe's capacity.

Engineering evidence: Physical inputs and outputs

Engineers design production processes and entire production systems to find the least cost ways of producing given units of output. They also seek to determine the least cost size of a production unit such as a plant, asking, for example, would the costs of producing some amount be less using one large



(a) Cost of a pump as capacity increases

(b) Cost of a steel pipe as capacity increases

Figure 8.6: **Cost of equipment involved in chemical production processes.** The cost of a 304 Steel Pipe per gallon (1990 USD). As the size of the pipe increases, its capacity increases more than proportionally. The cost *per gallon* of capacity decreases as the number of gallons of capacity increases.

plants or two small plants? And unlike economists, they start from input-output relationships measured in physical units.

When translated into costs using the prices of the physical inputs required the advantages of large firms can be substantial. A handbook for chemical engineers tells us that, to store dry chemicals in a fiber drum, the cost per square foot of storage would be 28 percent lower for a plant needing a 61 gallon drum than for a small plant needing only a 15 gallon drum.

In Figure 8.6 we plot similar costs per unit of capacity from the same engineering handbook. The cost counted here is the price of the equipment in question, that is the capital good involved in the production process. The capacity is in physical units – gallons per minute for the pump and gallons for the storage tank. Economies of scale are quite common in this kind of manufacturing process.

Econometric evidence on economies of scale and average costs

Table 8.2 summarizes some econometric studies, showing that economies of scale or constant returns to scale and declining or flat average cost curves are common. There is little evidence that average costs rise with higher levels of output.

Managers' assessments of their firm's cost curves

Alan Blinder and his collaborators adopted a research strategy rarely used by economists: they simply asked managers and CEOs what they thought their cost curves looked like. Specifically, they were asked if "their variable costs per unit are roughly constant when production rises" or if instead variable costs were rising or falling with increased production. Of the 190 firms that

Source	*Industry*	*Estimated slope of average cost (AC) curve and/or economies of scale*
Nerlove 1963	Electricity supply	"Marked" economies of scale especially for smaller firms
Griliches and Ringstad 1971	Norwegian manufacturing	Economies of scale
Christensen et al. 1976	Electric power generation	Declining AC for 1955; flat for 1970
Bittlingmayer 1982	Iron pipe	Declining AC (including declining mc)
Friedlaender et al. 1983	Top 3 U.S. auto firms	Declining AC for GM and Chrysler, rising for Ford
Caves et al. 1984	Airline industry	Declining AC for a given market, constant AC for expanding market
Dawson and Hubbard 1987	Dairy	Substantial declines in AC at modest size, slight increases at larger sizes
Hall 1988	U.S. manufacturing	Flat (17 of 21 industries); economies of scale (2) and diseconomies of scale (2)
Klette 1999	Norwegian manufacturing	Flat AC in "most industries." "Moderate diseconomies" of scale in a few industries
Koshal and Koshal 1999; Degroot et al. 1991; Laband and Lentz 2003	Higher education	Economies of scale
Cockburn and Henderson 2001; Henderson 2000	Pharmaceutical research	Economies of scale for firms smaller than the largest, flat for the largest 3 firms
Ashton 2003	Privately owned water services (U.K)	"Slight" diseconomies of scale. Rising average variable costs
Carlos and Voltes-Dorta 2011	Airports	Economies of scale
Aytekin et al. 2018	US firms	Flat AC curves
Loecker et al. 2018	Publicly traded U.S. firms 1955-2016	Economies of scale (increased between 1980 and 2016)

Table 8.2: **Evidence on economies of scale and decreasing average costs**

responded only eleven percent reported that average variable costs increased with additional output. The rest reported having either downward-sloping or flat average variable cost curves, meaning, even in the absence of fixed costs, declining or constant average costs.

FACT CHECK In the Blinder study, managers at 10 of the firms were unable to answer the question even after the idea of average variable costs was explained in a number of different ways, suggesting that the models above about how firms choose output levels and prices may not apply to all firms.

The conclusion most consistent with the engineering, statistical and survey evidence is that average costs do not rise with increased output. This means that marginal cost is either equal to or less than average cost, which is what we will assume in the models to be developed.

For reasons of simplicity we frequently assume that marginal and average costs are equal. But this ignores the important role of fixed costs which if substantial will lead to declining average costs.

8.5 A monopolistic competitor selects an output level

The firm's cost function is part of the information needed to answer the question: how do the firms owners determine the level of output that the firm will produce and the price at which the output will be sold?

REMINDER The *price setting power* that you studied in Chapter 4 describes the situation faced by the monopolistically competitive firm. They do not have "take it or leave it" (TIOLI) power.

We set aside many decisions that firms' owners routinely make, about how much to spend on advertising, research, influencing public policy and so on. Here we model the firm's owners:

- seeking an output level and a price of their product to maximize profits
- using a given technology and
- facing a given set of prices of their inputs and
- subject to the demand curve for their product.

We provide an answer based on a simple cost function in which average cost is a constant, so average cost $= ac = c = mc =$ marginal cost.

We model a single firm selling a unique product – for example, Honda selling the Accord model of the many cars they produce. As the only producer of this particular product – other businesses can sell other cars, but no other firm can sell a Honda Accord – the firm is a monopolist, that is the single seller of what is termed a differentiated product.

HISTORY The seemingly contradictory term "monopolistic competition" is due to Edward Hastings Chamberlin whose book of that title along with Joan Robinson's *Imperfect Competition* (both published in 1933) provided an alternative to the model of perfect competition. We use Chamberlin's term rather than Robinson's because we see competition among large firms with substantial market power as an alternative benchmark model, not as an "imperfection" of the conventional perfect competition approach.

Monopolistic competition and product differentiation

But potential buyers have alternatives: for example, a Subaru Forester is a substitute for a Honda Accord. Even if there is no substitute product, being a monopolist does not mean that the firm can sell any amount that it produces as any price it chooses. It is constrained by a downward sloping demand curve that effectively tells the firm: "if you produce more cars or offer more lessons you will have to lower your price to sell them all." Or: "if you charge a higher price you will not be able to sell as many goods."

These firms engage in what is termed **monopolistic competition**: they are the *single seller of their own product* – Subaru cannot produce and sell an Accord – but they *compete* with other firms selling products that are close substitutes for their own product. A monopolistically competitive firm will face less competition and make higher profits the more different its product is from what other firms are selling.

MONOPOLISTIC COMPETITION Monopolistically competitive firms have a monopoly on the particular good they produce, but compete with other firms that sell similar products. A monopolistically competitive firm faces the *constraint* of a downward-sloping demand curve.

Economists say that these firms produce a *differentiated product*, meaning that a firm's product does not just "happen to be" different from its competitors products but that firms actively try to make its product be or seem to be as different as possible from the products of other firms. Product differentiation strategies include distinctive design and advertising to promote brand loyalty. Another is the use of trademarks, which are privately owned (they are intellectual property) and cannot be used by firms other than their owner. These prevent other firms from competing directly (only Nike can sell Nike shoes).

REMINDER In Chapter 7 you learned that the demand curve for a product will be less elastic if there are fewer close substitutes for the product. Inelastic demand means that the reduction in sales associated with higher prices is modest, so firms seek to design and advertise products so that they will have or appear to have few close substitutes. Later, you will see that the maximum profit per unit of output sold is greater the more inelastic the demand for the product.

Active product differentiation is just one of a great many strategies that owners of firms deploy in order to maximize their profits. Others include innovation to reduce costs (that you studied in Chapter 6), preventing other firms from producing identical or similar products and lobbying government bodies for favorable tax or regulatory treatment. But here we will study just one dimension of this profit-maximizing process: deciding on a quantity of goods to produce and the price at which to sell them.

HISTORY Joseph Schumpeter is considered the father of the field of innovation economics. In 1950 he wrote: "in capitalist reality, as distinguished from its textbook picture, it is not [price] competition which counts but the competition from the new commodity, the new technology, the new source of supply, the new type of organization ... – competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits ... but at their foundations and at their very lives"

M-CHECK Here we use linear demand functions, such as those we derived from quadratic, quasi-linear utility functions.

Isoprofit curves and the feasible set of prices and sales

Monopolistically competitive firms set prices and levels of output to maximize profit. To understand a firm's choice of a particular combination of a price and a quantity, we will use the same architecture for constrained maximization that we have used throughout the book with its two basic elements:

- The *objectives* of the decision maker represented as a new type of *indifference curves* called *isoprofit* ("equal-profit") *curves*, giving the price-quantity combinations that, if implemented, would yield a given level of profit.
- The *feasible set* of choices that the decision maker can implement, that is, all of the price-quantity combinations that the firm could actually implement given the demand curve for its product.

PROFIT *Profit* A firm's economic profit is the difference between sales revenue and the total cost of producing the output, $\pi(x) = r(x) - c(x) = p(x)x - c(x)$. Remember that the firms costs include the opportunity cost of the capital goods it uses. When we use the term "profit" without an adjective ("economic" or "accounting") we mean economic profit.

REMINDER As we showed in Chapter 7, each consumer has a marginal rate of substitution of money, y, for the good, x, which is their *willingness to pay* for the good. A buyer's willingness to pay for all possible levels of purchase is their individual *inverse demand function*. The market demand function for x is the horizontal sum of all buyers' demand functions for the good.

The **isoprofit curves**, along which profit is equal to some constant, are just another kind of indifference curve because the owners of the firm are indifferent between making a particular level of profit by selling a lot of their product for a relatively low price, or alternatively, by selling a few units at a high price.

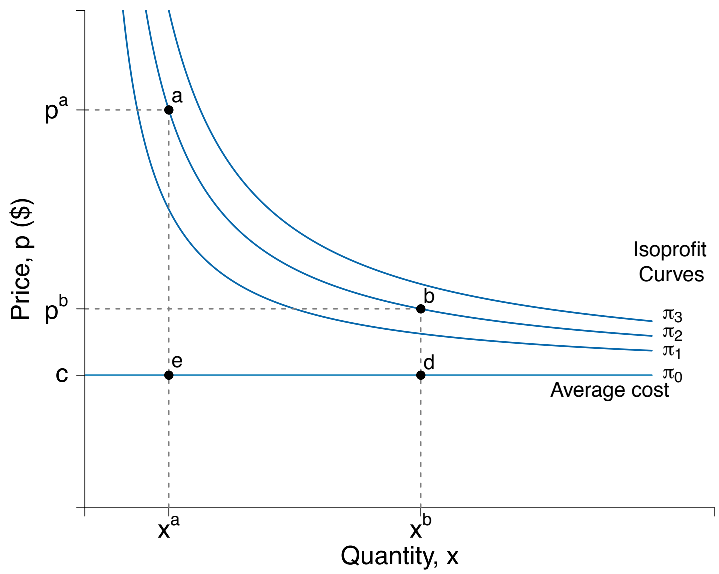


Figure 8.7: **A firm's isoprofit curves**. The labels to the right of the isoprofit curves indicate the level of profits associated with every point on the curve, and $\pi_3 > \pi_2 > \pi_1 > \pi_0 = 0$. Selling x^a units of output at price p^a (point **a**) yields the same profits as selling x^b units at price p^b (point **b**). If the price is equal to the average cost, no profits are made no matter how many units are sold (for example, points **d** and **e**).

Some of the firm's isoprofit curves are shown in Figure 8.7.

ISOPROFIT CURVE An isoprofit curve shows combinations of prices and quantities for which profit is equal to some constant.

Every point in the figure is some price-quantity combination that the firm might think about selecting. Not all of these are possible; what is feasible will depend on the demand curve. So think about the isoprofit curves as hypothetical statements. Point **a** and point **b**, for example, on the same isoprofit curve indicate that *if* the quantity indicated on the horizontal axis could be sold at the price indicated on the vertical axis for both of these points, then the total profits of the firm would be the same. Higher profits are above and to the right (indicating a higher price and or higher quantity sold.)

HISTORY: JOAN ROBINSON (1903-1983) taught economics at the University of Cambridge and developed one of the first models of what she termed "imperfect competition". The term "monopsony" is also due to her, as were many colorful turns of phrase. Concerning debates between critics and defenders of capitalism she wrote: "No one is conscious of his own ideology any more than he can tell the smell of his own breath" And "the misery of being exploited by capitalists is nothing compared to not being exploited at all" (that is, being without a job). Her heated "capital controversy" with American economists Paul Samuelson, Robert Solow and others raised doubts about the general equilibrium model of perfect competition. (We will see in Chapter **??** that subsequently Robinson's critique was substantially vindicated)

The slope of any one of the isoprofit curves is the marginal rate of substitution between selling more and charging a higher price.

Think about the isoprofit curve that is the horizontal line coinciding with the firm's (constant) average cost curve. The information it conveys is that at a price equal to the average cost the profits per unit sold are zero, so selling a lot (point **d**) or a little (point **e**) yields same amount of profit for the firm, namely zero.

Now that the owners of the firm have ranked every possible price-quantity combination as more or less profitable (or exactly as profitable, that is on the same isoprofit curve) we ask: which of these combinations could the firm actually implement?

M-Note 8.2: Slope of an isoprofit curve: Two methods

To find the slope of the isoprofit curve we write the equation for profits as follows:

$$\pi = \pi(p,x) = px - cx \tag{8.11}$$

$$\pi = x(p-c) \tag{8.12}$$

For any two points on an isoprofit curve, the profit difference associated with the difference in price dp is exactly compensated by the (opposite signed) utility difference associated with the difference in quantity dx, so that taking account of both effects, the difference in utility between the two points is zero. We want to find the differences in x and p that are consistent with no difference in profits, that is, being on the same isoprofit curve. To do this we totally differentiate Equation 8.12 with respect to dx and dp and we set the result equal to zero.

$$\begin{aligned} d\pi &= dx\frac{\partial \pi}{\partial x} + dp\frac{\partial \pi}{\partial p} = 0 \\ &= dx(p-c) + dpx = 0 \end{aligned} \tag{8.13}$$

We can re-arrange Equation 8.13 to find $\frac{dp}{dx}$, which is the slope of the isoprofit curve:

$$\begin{aligned} \frac{dp}{dx} &= -\frac{p-c}{x} \qquad (8.14) \\ &= -\frac{d\pi/dx}{d\pi/dp} \\ &= -\frac{\Delta\pi \text{ due to change in } x}{\Delta\pi \text{ due to change in } p} \end{aligned}$$

.

We found the slope of the iso-profit curve by total differentiating Equation 8.12. Alternatively, we can calculate the slope of the iso-profit curve directly. First, by rearranging Equation 8.11 we have the expression of the iso-profit curve

$$p(x;\pi) = c + \frac{\pi}{x} \tag{8.15}$$

where π is a constant and p as a function of x.

- As shown in Figure 8.7, when $\pi = \pi_0 = 0$ the iso-profit curve is just $p = c$, the horizontal line.
- When $\pi = \pi_1 > 0$, the graph of the iso-profit curve $p = c + \frac{\pi_1}{x}$ is the hyperbola $\frac{\pi_1}{x}$ shifted upward by c units.
- Given the constants $\pi_3 > \pi_2 > 0$, we have:

  $$p(x;\pi_3) = c + \frac{\pi_3}{x} > c + \frac{\pi_2}{x} = p(x;\pi_2) \text{ for all } x$$

  That is, for any given quantify x, to achieve a higher profit, the firm has to charge a higher price. Therefore, the iso-profit curve π_3 is above π_2.

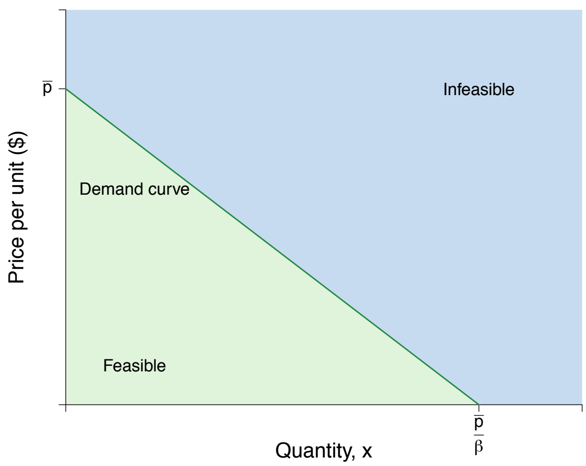
We can calculate the slope of the iso-profit curve 8.15 by taking the first order derivative with respect to x:

$$\frac{dp}{dx} = -\frac{\pi}{x^2}$$

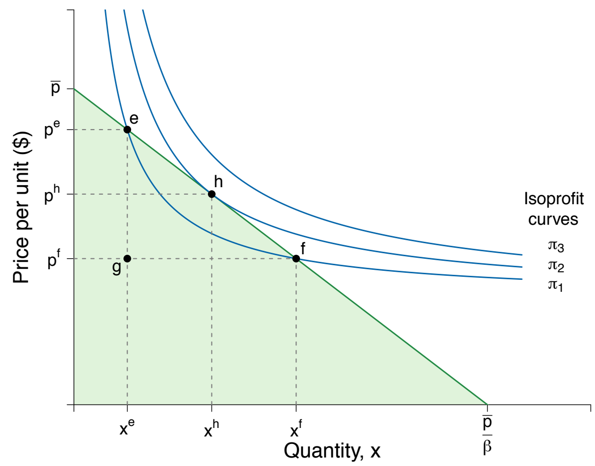
Recall that $\pi = (p-c)x$, so we can rewrite the above expression as

$$\frac{dp}{dx} = -\frac{(p-c)x}{x^2} = -\frac{p-c}{x}$$

which is the same as Equation 8.14.



(a) Feasible combinations of price and quantity



(b) Isoprofits and demand

Figure 8.8: **Feasible price and quantity combinations and the isoprofit curve**. A demand curve for the firm's products determines the feasible combinations of prices and quantity that it can choose in its maximization problem, as shown in panel a. In panel b. we include the firm's isoprofit curves. Along an isoprofit curve, profit is constant and the firm wants to choose the highest isoprofit curve given what is feasible. That is, the firm will choose the isoprofit line that is tangent to the demand curve. At the point of tangency, the firm will choose its quantity of production and the corresponding price on the demand curve.

Checkpoint 8.2: Average costs and the zero economic profits isoprofit curve

Explain why the average cost curve (the line c, e, d in Figure 8.7) is also the isoprofit curve for zero economic profits.

Demand and profit maximization

To choose the profit-maximizing price and quantity combination, we introduce the demand curve facing the firm, shown in panel a. of Figure 8.8. The demand function shown is:

$$\text{Inverse demand} \qquad p(x) = (\bar{p} - \beta x) \tag{8.16}$$

and its slope $\frac{\Delta p}{\Delta x} = -\beta$.

The figure divides the entire space into the *feasible* set of price and quantity combinations and the set of *infeasible* combinations. The firm now has to decide among the feasible points.

You can see at once that while both **h** and **g** are feasible, **h** has both higher prices and a larger quantity of sales, so the firm will surely not choose point **g**. So the choice is narrowed down to points on the boundary of the feasible set, which, as you can see is the demand curve.

The negative of the slope of the demand curve, $-\frac{\Delta p}{\Delta x}$, is the marginal rate of transformation of a larger quantity sold into a lower price charged. The price-quantity combination that maximizes profits – point **h** in the right panel of Figure 8.8 – is at the tangency between the demand curve and the highest feasible isoprofit curve. Because the slopes of the two curves are equal when they are tangent, the price quantity combination that maximizes the owners' profits is that where the marginal rate of substitution is equal to the marginal rate of transformation and is shown in panel b. of Figure 8.8 and in M-Note 8.3:

$$\text{slope of isoprofit curve} = \text{slope of demand curve}$$

The negatives of these slopes are the marginal rate of substitution and the marginal rate of transformation.

$$\begin{aligned} \Rightarrow \text{marginal rate of substitution} &= \text{marginal rate of transformation} \\ \frac{p-c}{x} &= \beta \end{aligned} \tag{8.17}$$

Another way to understand the rule for finding the profit maximizing level of sales is to rearrange Equation 8.17 so that it reads as follows:

$$p = \beta x + c \tag{8.18}$$

On the left of Equation 8.18 is the benefit of selling more (namely the price itself), which is also the average revenue. On the right of Equation 8.18 are both the cost of production of the additional goods (c) and the "cost" to the firm in lost revenue due to the fact that to sell more the firm will have to reduce its price (βx).

We can see in Figure 8.8 b, that when the owners of the firm can choose among points **e**, **f**, and **h** on the demand curve (the feasible frontier) all three points are feasible. At point **e**, the marginal rate of substitution is greater than the marginal rate of transformation (the isoprofit curve is steeper than the demand curve or $mrs(x_e, p_e) > mrt(x_e, p_e)$) and the owners of the firm can make more profits by selling more output at a lower price. At point **f**, the marginal rate of substitution is less than the marginal rate of transformation ($mrs(x_e, p_e) < mrt(x_e, p_e)$) and the owners of the firm can make more profits by selling less output at a higher price. Though points **e** and **f** are both feasible, point **h** results in higher profits on isoprofit curve π_2 rather than points **e** and **f** on isoprofit π_1 ($\pi_2 > \pi_1$). At point **h** $mrs(x_h, p_h) = mrt(x_h, p_h)$ and the firm is on the highest isoprofit curve it can attain subject to the constraint of the demand curve.

M-Note 8.3: Isoprofit curves, the feasible set and maximum profits

The slope of the isoprofit curves. In the previous M-Note, you saw that the slope of the isoprofit curve is

$$\frac{dp}{dx} = -\frac{p-c}{x} \tag{8.19}$$

The marginal rate of substitution, is the negative of the slope of the isoprofit curve.

$$-\frac{dp}{dx} = \frac{p-c}{x} = mrs(x, p) \tag{8.20}$$

Equation 8.20, the marginal rate of substitution between the quantity sold and the profits per unit sold, is profits per unit (the numerator) divided by the number of units sold (the denominator).

The slope of the demand curve. The inverse demand curve is the frontier of the feasible set and using the demand curve shown in Equation 8.16:

$$p(x) \quad = \quad (\bar{p} - \beta x)$$

Differentiating the price with respect to the total amount sold its slope is $-\beta$. The negative of the slope of the demand curve is the marginal rate of transformation of more goods sold into lower prices at which they sell, that is, $mrt(x,p) = \beta$.

The condition for maximum profits. The price and quantity that maximize the firm's profits are the combination of p and x such that the slope of the isoprofit curve is equal to the slope of the demand curve, or $mrs(x,p) = mrt(x,p)$:

$$mrs(x,p) = \frac{p-c}{x} = \beta = mrt(x,p)$$

8.6 *Profit maximization: marginal revenues and marginal costs*

There is a second – equivalent – way to determine the price and output level that will maximize a firm's profits. Recall that the firm's revenue is the product of the price at which it can sell the output, $p(x)$, and the output sold (x):

$$\text{Total revenue} \qquad r(x) = p(x)x \tag{8.21}$$

TOTAL, AVERAGE, AND MARGINAL REVENUE *Total revenue* is the value of total sales, x, at the price p where we express the price as $p(x)$ to show that it depends on the amount sold $r(x) = p(x)x$; *average revenue* is the revenue per unit of output, which is the price, $ar(x) = \frac{r(x)}{x} = p(x)$; *marginal revenue* is the effect on total revenue of a small increase in sales or $mr(x) = \frac{\Delta r}{\Delta x}$.

Like the cost function, the total revenue function has corresponding average and marginal values:

- **Average revenue** is **total revenue** divided by total output (or the ratio of revenue to output): $ar(x) = \frac{r(x)}{x}$. Notice that because $r(x) = p(x)x$, average revenue is also the price the firm can charge when it is selling output x: $ar(x) = p(x)$. The amount of revenue a firm gains *per unit sold* is simply the price of the good. In this chapter, we assume that all customers pay the same price. In Chapter 9 we consider the case of price discrimination, when a firm charges different customers different prices.

- **Marginal revenue** is the change in total revenue associated with a small change in sales, that is, $\frac{\Delta r}{\Delta x}$. Marginal revenue is therefore the slope of the total revenue function at a given output x.

The firm's profit, π, is the difference between its *revenue* from sales of its output, $r(x)$, and the cost of producing output, $c(x)$:

$$\text{Profit} \qquad \pi(x) = r(x) - c(x) \tag{8.22}$$

The owners of the firm would like to find a level of production x that maximizes $\pi(x)$ given the inverse demand function $p(x)$ and the cost function $c(x)$.

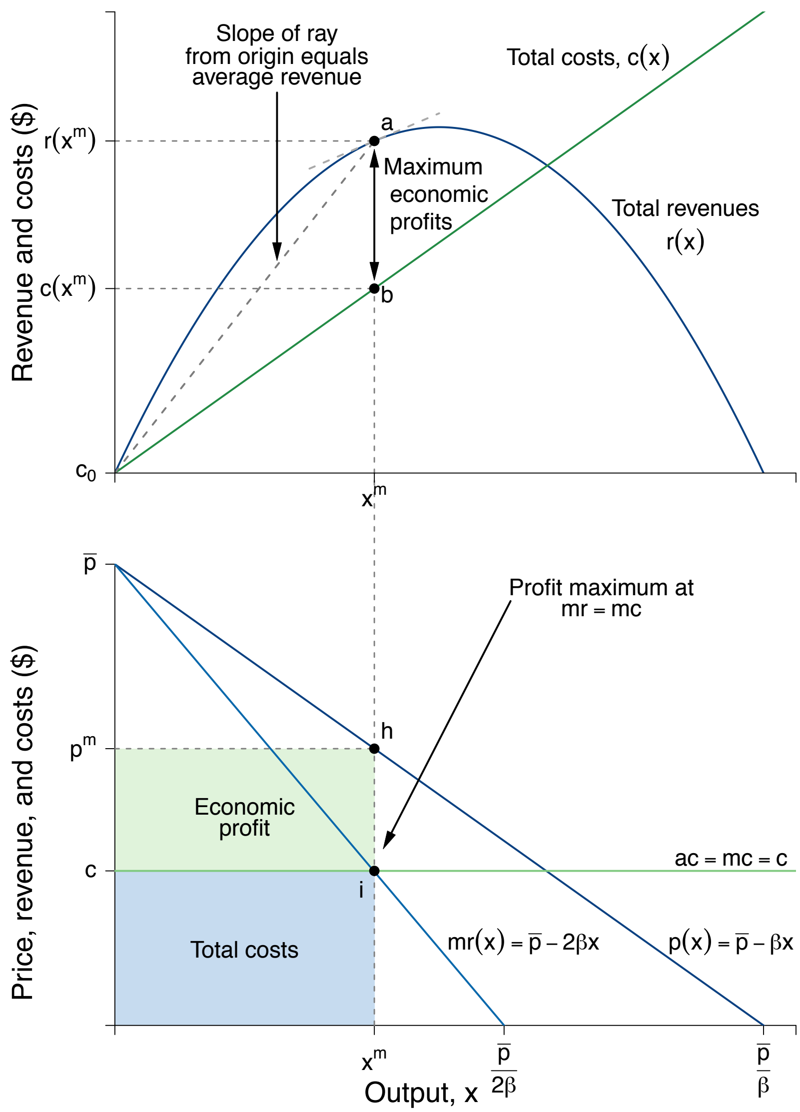


Figure 8.9: **The level of output that maximizes economic profit** Economic profit at output x is the difference between total revenue at x, $r(x) = p(x)x$, and the total cost of producing x, $c(x)$. Economic profit is maximized when marginal revenue, the slope of the tangent to the revenue curve, is equal to marginal cost, the slope of the total cost curve. The maximum profit is the difference between revenue and total cost where the slopes are equal, that is, at points **a** and **b**. The distance **ab** in the upper panel is shown as the green shaded rectangle in the lower panel. The *slopes* of the total cost and total revenue curves in the top panel are the *levels* of the marginal cost and marginal revenue curves in the lower panel. The price p^m in the lower panel is (in the top panel) the slope of a ray from the origin to point **a**, which is the average revenue.

We have already seen how this can be done using the firm's isoprofit curves and demand curve. An equivalent method is to plot the total revenue curve $r(x) = p(x)x$ together with the total cost curve $c(x)$ as in Figure 8.9.

The vertical distance between the two curves at any point, is the total amount of profit. The figure shows how starting at $x = 0$, this first increases up to a maximum, then decreases as revenues begin to increase less than costs. Eventually, at high levels of output, costs exceed revenues.

PROFIT-MAXIMIZING OUTPUT The *profit-maximizing output* of the firm is that for which the vertical distance between the revenue and cost curves is greatest. It is either the output level where marginal revenue equals marginal cost, or zero, if profit when marginal revenue equals marginal cost is less than zero and owners of the firm would do better by shutting down altogether.

The maximum profit of the firm occurs at a level of output, x^m, where the slope of the total cost curve is equal to the slope of the revenue curve at x, as shown in Figure 8.9:

$$\begin{aligned} \frac{\Delta r}{\Delta x} &= \frac{\Delta c}{\Delta x} \\ \text{marginal revenue} &= \text{marginal cost} \end{aligned} \tag{8.23}$$

The firm will then find the corresponding price for its good by substituting the profit-maximizing quantity back into its demand curve. Figure 8.9 illustrates this way of finding the output level (and price that maximizes the firm's profits.

M-Note 8.4: The monopolistic competitor chooses a price and output level: general case

The information relevant to the problem is:

- The firm's total cost $= c(x) = cx$
- The *inverse demand curve* for the firm's product is downward sloping and is $p = p(x)$, and
- The firm's *total revenue* is $r(x) = p(x)x$
- so, the firm's total *profit* = $r(x) - c(x) = p(x)x - cx$

To find the rule that determines the profit maximizing level of output, we differentiate the profit function with respect to x, and set the result equal to zero. So we have:

$$\frac{\partial \pi}{\partial x} = \underbrace{\overbrace{\frac{\partial p}{\partial x}}^{\text{Lost revenue due to fall in price}} x + \overbrace{p(x)}^{\text{Revenue gained through increased sales}}}_{\text{Marginal revenue}} - \underbrace{c}_{\text{Marginal cost}} = 0 \tag{8.24}$$

Equation 8.24 tells the owner of the firm to expand production as long as the additional revenues resulting from a small increment in sales – the marginal revenue – exceed the addition to total cost associated with the increment in output. The first term on the right-hand side is the revenue lost due to the negative effect of selling more on the price at a given level of output (because the demand curve is downward-sloping). The second term on the right-hand side is the positive effect of selling more on revenues (at a given price). So the rule (or first order condition) determining the profit maximizing level of output of the firm is to choose the level of x that equates marginal revenue to marginal cost.

M-Note 8.5: Output, price, and profit of a monopolistic competitor: The example of a linear demand curve.

The relevant information for this case is:

$$\begin{aligned} \text{Inverse demand} \quad p(x) &= (\bar{p} - \beta x) && (8.25) \\ \text{Total revenue} \quad r(x) &= p(x)x = (\bar{p} - \beta x)x = \bar{p}x - \beta x^2 && (8.26) \\ \text{Economic profit} \quad \pi^m(x) &= r(x) - c(x) \\ &= \bar{p}x - \beta x^2 - cx && (8.27) \end{aligned}$$

The marginal revenue is the first derivative of total revenue with respect to output (the slope of its revenue function):

$$\text{Marginal Revenue} \quad mr(x) = \frac{dr(x)}{dx} = \bar{p} - 2\beta x \tag{8.28}$$

The owners maximize profit by choosing the output level that equalizes marginal revenue and marginal cost and then selling that output at the highest price possible, as given by the inverse demand curve:

$$\bar{p} - 2\beta x = c \tag{8.29}$$

Solving Equation 8.29 for x, we have:

$$x^m = \frac{\bar{p} - c}{2\beta} \tag{8.30}$$

To find the price, we substitute the output (Equation 8.30) into the inverse demand curve (Equation 8.25):

$$\begin{aligned} p^m &= \bar{p} - \beta x^m \\ \text{Substitute in } x^m \quad &= \bar{p} - \beta\left(\frac{\bar{p} - c}{2\beta}\right) \\ &= \frac{1}{2}\bar{p} + \frac{1}{2}c \\ \text{Add and subtract } \frac{1}{2}c \quad p^m &= c + \frac{1}{2}(\bar{p} - c) \end{aligned} \tag{8.31}$$

To find economic profit we substitute the monopolist's quantity into equation 8.27.

$$\pi^m = (p(x^m) - c)x^m = \frac{(\bar{p} - c)^2}{4\beta} \tag{8.32}$$

The firm charges a price that is *greater* than its marginal costs by one half the difference of the maximum price and its marginal costs. Price exceeds marginal cost. As a result, the firm makes economic profit, equal to π^m.

Checkpoint 8.3: Numerical example

Assume new values for the parameters in which we are interested: let $\bar{p} = 100$ and $\beta = 1$ and the firm has constant marginal costs $c = 1$.

a. Find the formula for its profit.
b. Find the profit-maximizing quantity.
c. How much profit do the owners make? What is the price and how much greater is the price than marginal costs (that is, what is the markup)?

M-Note 8.6: Marginal revenue and price elasticity of firm demand

As we saw in Chapter 7, the price elasticity of firm demand is the ratio of the percentage change in quantity sold to the percentage change in price. In Chapter 7, $\frac{\Delta x}{\Delta p}$ which is equivalent to finding the first derivative of the demand curve $x(p)$ with respect to p, i.e. $\frac{dx}{dp}$. This allows us to re-write the equation for price elasticity of firm demand as follows:

$$\eta = \frac{\Delta x}{\Delta p}\frac{p}{x} = \frac{dx}{dp}\frac{p}{x}$$

The owners of the firm recognize that a firm's revenues can be affected by the responsiveness of consumers to changes in prices. Therefore, we can examine the connection

REMINDER Remember that the *price elasticity of demand* is the ratio of the percentage change in quantity demanded to the percentage change in price, $\eta = \frac{\Delta x}{\Delta p}\frac{p}{x}$. We will often write this as η and not use the subscript "xp" unless necessary to separate the price elasticity of demand from other kinds of elasticity.

between marginal revenue and price elasticity of demand. Marginal revenue can be written in terms of the price elasticity of the firm's demand curve:

$$\begin{aligned} \frac{dr(x)}{dx} &= p(x) + \frac{dp(x)}{dx}x \\ &= p(x)(1 + \frac{dp(x)}{dx}\frac{x}{p}) \\ mr(x) &= p(1 + \frac{1}{\eta}) \end{aligned} \tag{8.33}$$

Recall that elasticity is a *negative* number because the demand curve is downward-sloping, and therefore $\frac{dp}{dx} < 0$. Consequently, we can re-write the equation above using the absolute value of η, that is, $|\eta|$:

$$mr(x) = p\left(1 - \frac{1}{|\eta|}\right) \tag{8.34}$$

If we equate Equation 8.34 for marginal revenue to marginal cost (c) to satisfy the principle of profit maximization, we can identify a relationship between price elasticity and costs as follows:

$$\begin{aligned} mr(x) = p\left(1 - \frac{1}{|\eta|}\right) &= c = mc(x) \\ \frac{p}{c}\left(1 - \frac{1}{|\eta|}\right) &= 1 \\ \frac{p}{c} &= \frac{1}{1 - \frac{1}{|\eta|}} \\ \Rightarrow \frac{p}{c} &= \frac{1}{\frac{|\eta|-1}{|\eta|}} \\ \Rightarrow \frac{p}{c} &= \frac{|\eta|}{|\eta|-1} \end{aligned} \tag{8.35}$$

Two methods of choosing a profit-maximizing price and quantity compared

We have introduced two methods of determining the profit maximizing price and quantity defined by :

- the tangency of the feasible frontier given by the demand curve and the isoprofit curves and
- equating marginal revenue and marginal costs.

The two methods are illustrated in a single graph in Figure 8.10 where point **h** is the point of tangency between the demand curve and the firm's highest feasible isoprofit curve, and point **i** shows the marginal revenue equated to the marginal cost. As you can see whether determined by the firm selecting point *h*, or point *i* the price (p^m) and quantity sold (x^m) will be the same. M-Note 8.7 confirms that the methods give the same result.

M-Note 8.7: The two methods give the same result

To determine the combination of price and quantity sold that maximizes the profits of the firm we used two methods. This M-Note shows that the methods are identical (also shown in Figure 8.10.)

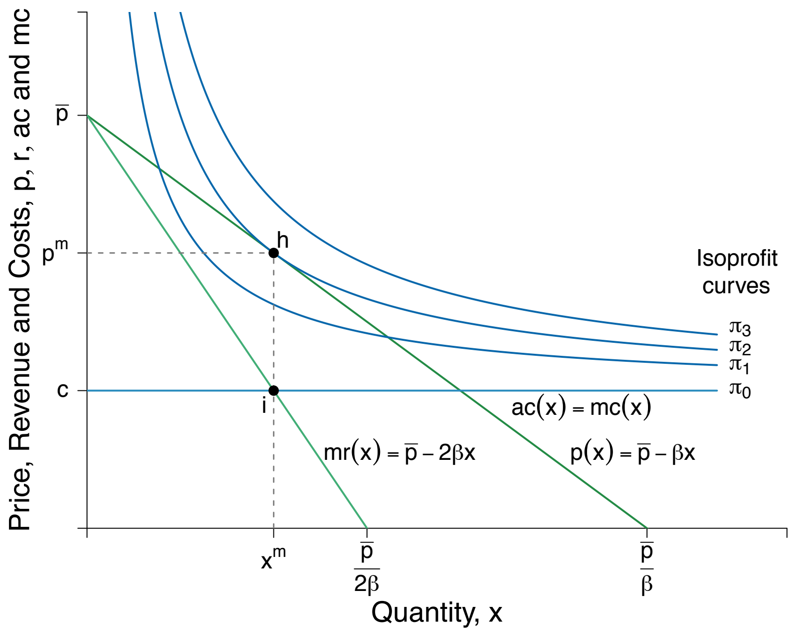


Figure 8.10: **Isoprofits, demand, marginal revenue and marginal costs.** The firm chooses the quantity at which marginal revenues equal marginal costs. This quantity coincides with the quantity the firm would have chosen when its isoprofit curves were tangent to its demand curve. Along π_0, economic profits are zero as the price equals marginal cost.

The first was based on the demand curve as the frontier of the feasible set and the family of isoprofit curves as the basis for choosing the most profitable point on the demand curve. Using the inverse demand curve in Equation 8.16 this gave us the following condition:

$$\begin{aligned} \text{marginal rate of substitution} &= \text{marginal rate of transformation} \\ \frac{p-c}{x} &= \beta \end{aligned}$$

The second method showed that the maximum profit is the level of output such that

$$\begin{aligned} \text{marginal revenue} &= \text{marginal cost} \\ \overline{p} - 2\beta x &= c \end{aligned} \tag{8.36}$$

We can rearrange Equation 8.36 to show that it is the same as Equation We first replace p by the equation for the inverse demand curve which gives the value of p for each value of x:

$$\begin{aligned} \text{marginal rate of substitution} &= \text{marginal rate of transformation} \\ \frac{(\overline{p} - \beta x) - c}{x} &= \beta \\ \frac{\overline{p} - c}{x} - \beta &= \beta \\ \frac{\overline{p} - c}{x} &= 2\beta \\ \overline{p} - c &= 2\beta x \\ \overline{p} - 2\beta x &= c \\ \text{marginal revenue} &= \text{marginal cost} \end{aligned}$$

8.7 The markup, the price elasticity of demand, and entry barriers

When a firm selects the profit maximizing level of output and price, the amount of economic profit its owners will make depends on the extent of competition that the firm faces, that is, the number of firms selling the identical product or close substitutes. The reason is that for some given cost c, the price that maximizes the owners' profits will be higher if buyers do not reduce their purchases very much when the firm raises its price. This means that prices will be higher relative to costs the more inelastic is the demand curve for its product (the smaller is $|\eta|$.)

To see why this is true remember that in M-Note 8.6 we showed that:

$$\text{marginal revenue} = mr(x) \quad = \quad p\left(1 - \frac{1}{|\eta|}\right) \tag{8.37}$$

If the firm chooses a level of sales such that marginal revenue equals marginal cost, then we can represent the ratio of price to costs as follows:

$$\frac{p}{c} \quad = \quad \frac{|\eta|}{|\eta| - 1} \tag{8.38}$$

M-CHECK Equation 8.38 follows from Equation 8.37 by equating marginal revenue in 8.37 to marginal cost c:

$$p\left(1 - \frac{1}{|\eta|}\right) = c \tag{8.39}$$

Then, dividing both sides by c and by $\left(1 - \frac{1}{|\eta|}\right)$, we find:

$$\frac{p}{c} = \frac{1}{\left(1 - \frac{1}{|\eta|}\right)} \tag{8.40}$$

And finally multiplying both the numerator and denominator of the right hand side term by $|\eta|$ we get Equation 8.38 as shown in the text.

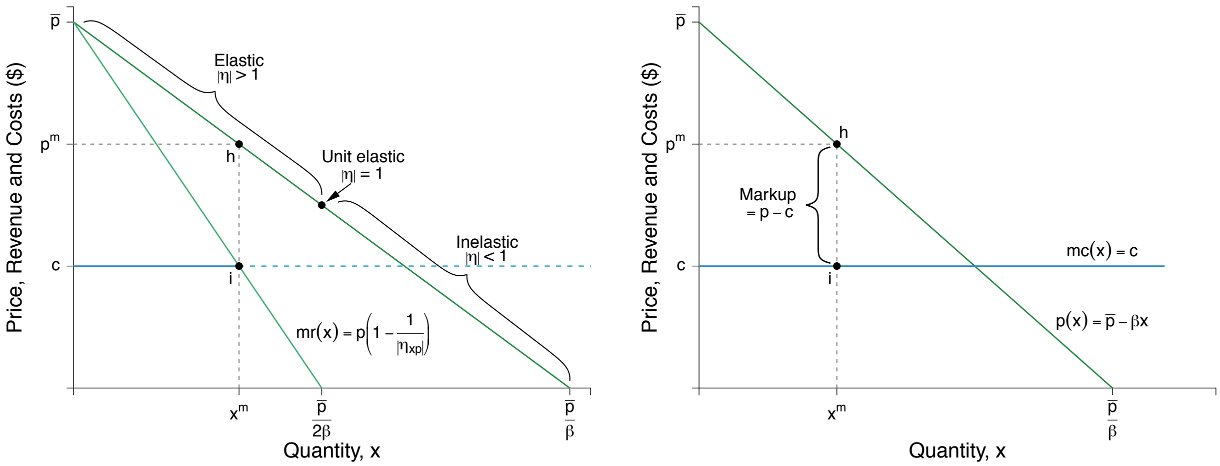
Equations 8.37 and 8.38 shows us two things:

- a point on the demand curve where $|\eta| < 1$ cannot be a profit-maximizing price and output because in that case marginal revenue would be negative as you can see from Equation 8.37. This means that at an output level x such that the price elasticity of demand is less than one, i.e. $|\eta| < 1$, a decrease in the firm's sales (reducing x) would raise its revenues and also lower its total costs, so the output level where $|\eta| < 1|$ could not be a profit maximum.
- Equation 8.38 tells us that as long as the price elasticity of demand the profit maximizing point on the demand curve greater than one (in absolute value i.e. $|\eta| > 1$), the lower it is (the less elastic the demand) the greater will be the price compared to the marginal cost ($\frac{p}{c}$).

Equation 8.38 also shows that if a firm is facing a very elastic demand curve, the best its owners can do is to set prices very close to costs (and make limited profits. This is because $\frac{p}{c}$ approaches 1 (meaning price no larger than marginal cost) as $|\eta|$ becomes very large.

M-Note 8.8: The markup and the price elasticity of demand

Given the relationship derived between $\frac{p}{c}$ and elasticity in Equation 8.45 we have the



(a) Marginal revenue and elasticity

(b) The markup

Figure 8.11: **Marginal revenue, elasticity, and the markup**. In panel a. we show the relationship between elasticity of firm demand (η) and marginal revenue. In panel b. we show the markup over marginal costs that the firm makes as profit per unit. We use the idea of the markup for each unit to find the markup ratio, μ, which is the ratio of the markup to marginal costs, or $\mu = \frac{p-c}{c}$.

following relationship:

$$\begin{aligned}\mu &= \frac{p-c}{c}\\ \mu &= \frac{p}{c}-1\end{aligned} \tag{8.41}$$

We already know that $\frac{p}{c} = \frac{|\eta|}{|\eta|-1}$, which we can substitute into Equation 8.41:

$$\mu = \frac{|\eta|}{|\eta|-1} - 1 \tag{8.42}$$

Now, remembering that a ratio of any number to itself equals 1, we can substitute $1 = \frac{|\eta|-1}{|\eta|-1}$ into 8.42:

$$\begin{aligned}\mu &= \frac{|\eta|}{|\eta|-1} - \frac{|\eta|-1}{|\eta|-1}\\ \mu &= \frac{|\eta|-|\eta|+1}{|\eta|-1}\\ \mu &= \frac{1}{|\eta|-1}\end{aligned}$$

The markup ratio can therefore be understood based on its definition of profit per unit divided by cost per unit and based on a relationship to price elasticity of demand.

The markup and the markup ratio

We can summarize the relationship between a firm's price and costs by the markup and the markup ratio.

The **markup** is the firm's profit per unit or the price it charges per unit minus its costs per unit, or $p-c$. Remember that the firm's profits are given as follows:

$$\begin{aligned}\text{Profit} \quad \pi(x) &= p(x)x - cx\\ \text{markup or profit per unit} \quad \frac{\pi}{x} &= p(x) - c\end{aligned} \tag{8.43}$$

Equation 8.43 is called the markup, because it measures the degree to

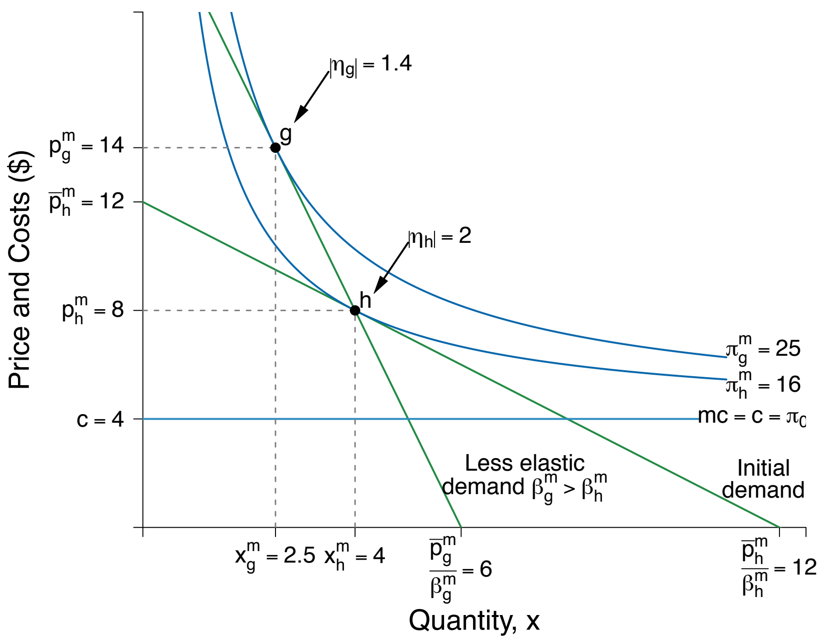


Figure 8.12: **Increasing profits by making a demand curve less elastic**. Initially, the firm faces a more elastic demand curve, so the profit maximum is at point **h** and the firm's total revenues are 8x4 = 32 and its costs are 4x4 = 16, giving it a profit of 16 as shown for the isoprofit curve through point h.

which the owners of the firm "mark up" prices above the cost to produce the good.

The **markup ratio** (μ) is the ratio of the markup to the cost:

$$\begin{aligned} \text{Markup Ratio} &= \frac{\text{Profit per unit}}{\text{Cost per unit}} \\ \mu &= \frac{p-c}{c} \end{aligned} \tag{8.44}$$

The markup ratio measures the degree to which prices are greater than costs. If prices were equal to costs, i.e. p = c, then the markup ratio would be zero ($\mu = 0$). As price becomes greater than marginal cost, the markup ratio measures the extent of that difference as μ gets larger and larger. Given the relationship derived between $\frac{p}{c}$ and the price elasticity of demand in Equation 8.35, we can derive a relationship between the markup ratio and elasticity as follows:

THE MARKUP AND THE MARKUP RATIO
The firm's markup is its profit per unit or how much greater the price it charges is relative to its costs, i.e. $\frac{\pi}{x} = p - c$. The markup ratio, is the profit per unit divided by unit costs, or $\mu = \frac{p-c}{c}$.

$$\mu = \frac{p-c}{c} = \frac{1}{|\eta| - 1} \tag{8.45}$$

If $|\eta|$ is large as will be the case when the firm has many competitors selling close substitutes for the firm's product, the markup ratio is close to zero (meaning that economic profits fall as competition increases). If, on the contrary, η is close to 1, then μ gets larger and larger (μ tends towards infinity as $|\eta| -> 1$).

Price elasticity of demand and profits

You can see how this affects the firm's profits in Figure 8.12 when faced with the "Initial demand" curve. Point **h**, the tangency of the demand initial demand curve and the highest feasible isoprofit curve, indicates the firm's profit maximizing price (8) and quantity sold (4).

If the firm could devise a way to make the demand curve less elastic, like the steeper demand curve in the figure, it would then have a different – and more profitable – constrained optimization problem to solve. It could raise its price (to 14), sell less (2.5) and increase its profits (from 16 to 25).

This shows that the economic profit made by the firm's owners will be greater the less elastic the demand, as long as the price elasticity of demand is greater than one ($|\eta| > 1$). This is why firm's profit-maximizing strategies include product differentiation. Advertising, trademarking, and design can make the firm's own product seem "more different" from other firms' products. As a result there are fewer close substitutes for the firm's product, which as you know from Chapter 7 makes the firm's demand curve less elastic.

INCUMBENT FIRMS are those already selling in a market that may wish to limit the market entry of new firms.

Another way to make the demand curve less elastic is to limit the number of firms producing the same product. For example, incumbent firms may pursue *predatory pricing* – firms temporarily charging a price less than their marginal costs – and other practices to inflict losses on a new firm attempting to enter a market.

Barriers to entry

BARRIERS TO ENTRY The term barriers to entry refers to anything making it difficult for new firms to enter a market, including intellectual property rights that give incumbent firms a monopoly on particular technologies or, economies of scale in production, and predatory pricing.

Because the markup ratio and profits will be greater if the demand curve is less elastic, and this will be the case if there are fewer competing firms in a market, profits will be greater if it is difficult for new firms to enter a market to compete with incumbent firms. A barrier to entry is anything that makes it difficult for a new firm to enter a market to compete when the owners of incumbent firm are making economic profits. Sometimes barriers to entry are called "moats" because, like the protective water barriers around old castles, they protect the incumbent firms from intrusion by "outsiders."

Barriers to entry include

PREDATORY PRICING Predatory pricing occurs when an incumbent firm that can sustain short or medium term losses in revenue, charges a price lower than its (marginal) costs and often therefore lower than its competitors to drive its competitors out of business.

- Predatory pricing, just mentioned above.
- *Economies of scale in production and learning by doing:* A new firm with initially limited output and experience, will have a disadvantage until it can grow larger or accumulate experience in producing the good.

- *Intellectual property and licensing*: Exclusive use of trademarks, production processes, or knowledge protected by *intellectual property rights* and other government-enforced monopolies (e.g. licenses).
- *Naturally occurring barriers* to entry such as limited access to natural resources or advantageous spatial locations owned by incumbent firms.
- *Network economies of scale in demand*. For many goods or services, the the value to the consumer depends on the number of others purchasing the good. An entering ride hail service could not compete with Uber and Lyft unless it already had a large number of drivers, which it could not get unless it had a large number of consumers.

Some barriers to entry are not the result of deliberate strategies by firms, they just exist by the nature of the product being produced. Economies of scale and naturally occurring barriers are examples.

But some barriers are be constructed or at least heightened by the one or more incumbent firms. Predatory pricing, establishing monopolies on essential knowledge through intellectual property rights and licensing, and advertising a company owned trademark are examples.

Checkpoint 8.4: Price elasticity of demand and profits

Use the numerical values in Figure 8.12 to

- confirm that the profits at point **g** are 25 as shown;
- show that the elasticity of demand at point **h** is 2 as indicated, using Equation 8.33; and
- calculate the markup ratio in Equation 8.45
- Many similar questions using the data in figure 8.12

8.8 Application: Evidence on the markup in drug prices

Because the profit maximizing markup ratio depends on the extent of competition that a firm faces, it differs substantially among firms, across different industries and over time. In Chapter 9 (Figure 9.21) we provide evidence on the markup ratio for firms in the U.S. economy as a whole, showing a sharp increase from around 0.2 during the mid 1980s to almost 0.6 three decades later.

For some firms the markup ratio is much greater than these economy-wide averages. A reason is the limited competition facing firms selling products in which they hold intellectual property rights such as trademarks, copyrights, or patents.

An example is the pharmaceutical sector, in which patents on drugs or ingre-

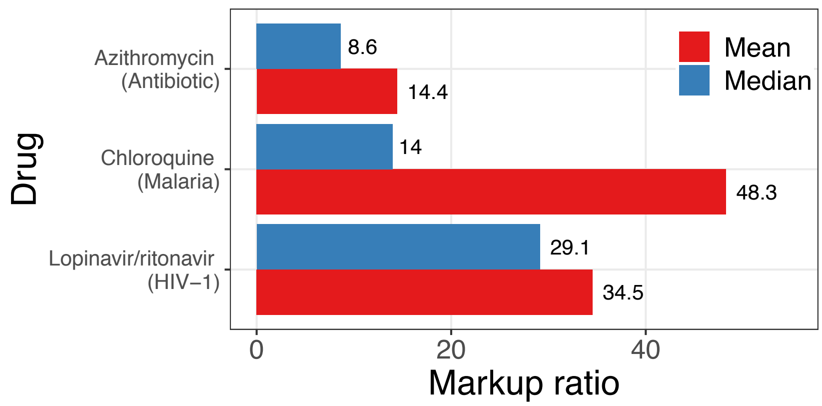


Figure 8.13: **Estimated markup ratios for drugs** The data shown are measures of $\mu = (p - c)/c$ based on a single minimum cost estimate (c) along with prices charged for the drugs shown in the 7 to 11 countries for which this information was available. The blue bar for chloroquine, for example, indicates that the profits per treatment sold (that is $p - c$) were 14 times the minimum cost of the treatment (c) We show both the mean and the median because extraordinarily high prices in the U.S contribute to a higher mean markup than is representative of the countries taken as a whole.

dients of drugs constitute barriers to entry by competing firms. Only Bristol-Myers-Squibb (BMS) or those companies that it licenses can sell the hepatitis C drug declatasvir. But BMS still competes with Gilead Sciences, the only company that can sell another hepatitis C drug, sofosbuvir.

One set of estimates of markups on drugs is in Figure 8.13. Seeking some guidance on what a vaccine for COVID-19 might cost to produce as the pandemic swept the world in early 2020, a team of pharmaceutical scientists devised a measure of the "minimum cost" of producing what they considered as possibly similar drugs. (There was no vaccine for COVID-19 at the time.) For example they used an online data base for tracking actual shipments of the necessary chemicals ("active pharmaceutical ingredients") and their prices to find the least cost inputs.

The costs estimates include these ingredients (the major part of the cost), and the costs of production (called "formulation and tabletting") and packaging. They also include a ten percent markup over accounting costs that represents the opportunity costs of the capital goods used, and the cost of a profit tax. They assumed that production was at at a level to exploit the possible economies of scale in production.

Their estimate is a measure of the marginal cost (and the average variable cost). It does not include costs unrelated to the production of the particular drug such as advertising, lobbying, legal, and other expenditures attempting to promote a favorable legal and regulatory environment for the companies concerned, research and the costs of trials required for regulatory approval.

The price data that the researchers used to calculate the markup ratios in Figure 8.13 came from national drug price databases or online pharmacy sites. Where more than one price was available, the lowest price was used.

The middle set of bars in Figure 8.13, for example, is for a treatment of malaria, which the pharmaceutical scientists estimated has a cost of $\$0.30$ for a 14-day course of the drug. The lowest price at which the drug was being sold was less than that, in Bangladesh, and considerably more than that in India, Malaysia, Sweden, South Africa, China, and the U.K, where the same 14-week course of the drug was sold for $\$8$. In the U.S. the same course of the drug sold for $\$93$. The estimated markups are based on these prices, which for the other drugs shown also include data from France, Brazil, Malaysia, Sweden, Turkey and other countries.

We do not know if the firms in question actually achieved markup ratios of the amounts shown. For example they may have been using production methods that are *not* the least cost solution to their cost minimization exercise. Remember, the cost curve indicates the *minimum* cost of producing a particular amount, and firms may not have implemented this in combining the "active pharmaceutical ingredients" to produce their "final finished product."

Limited competition provides firms' owners with opportunities to increase their profits not only in the prices at which they sell their products, but also in the prices which they pay of inputs into their production.

8.9 Willingness to sell: Capacity constraints and market supply

Recall that for a particular good x, the demand curves introduced in Chapter 7 provided answers to a hypothetical "what if" question: given the prices of all other goods and an individual's budget, how much of good x would she purchase if she could buy any amount she wished consistent with her budget when offered various prices for the good?

The supply curve for a firm provides an answer to a similar hypothetical question: if the price at which the firm could sell any amount it wished were p, what amount (x) of output would the owners of the firm choose to produce and sell? Answers to this question for all possible values of p give us the supply curve, that expresses the amount produced as a function of the hypothetical price, or $x(p)$

FIRM SUPPLY CURVE The supply curve for a firm provides an answer to a hypothetical question: if the price at which the firm could sell any amount it wished were p what amount (x) of output would the owners of the firm choose to produce and sell? Answers to this question for all possible values of p give us the supply curve, that expresses the amount produced as a function of the hypothetical price, or $x(p)$.

The hypothetical question is a thought experiment by which we construct the supply curve as economic concept. It is not something that owners of firms ask themselves. How much a firm places on the market will typically affect the price at which it can sell the goods. So firms do not take the price at which they will sell the good as a given. They choose both price to charge and quantity to market based on their costs and the demand for the product (using the constrained optimization methods described in Sections 8.5 and 8.6).

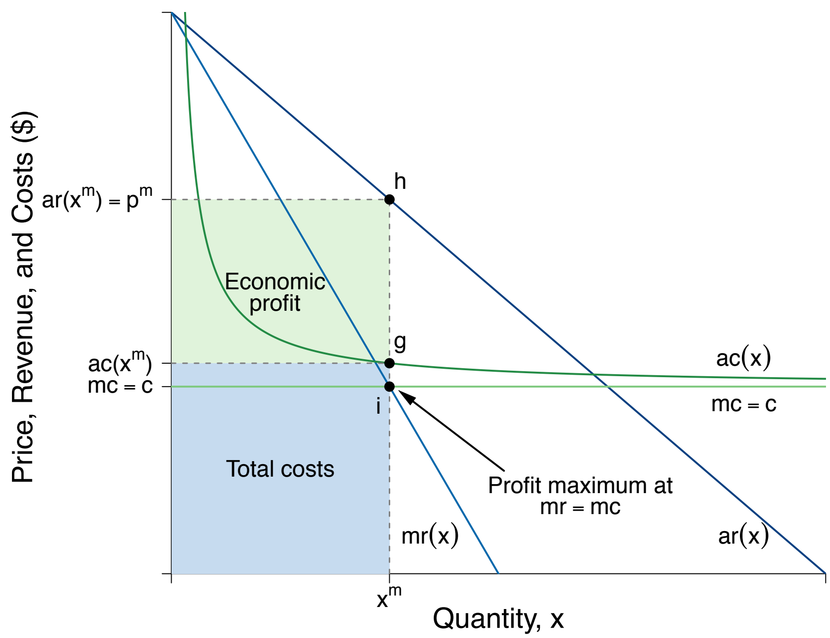


Figure 8.14: **Profit maximization with declining average costs**. Like the firm shown in Figure 8.9, this firm has fixed costs ($c_0 > 0$) and linear variable costs ($c_1 > 0, c_2 = 0$) and therefore its total costs are $c(x) = c_0 + c_1 x$, its average costs are $ac(x) = \frac{c_0}{x} + c_1$ and the firm's marginal costs are $mc(x) = c_1$. The firm with declining average costs chooses a price and quantity to produce. The firm chooses its quantity where marginal revenue equals marginal cost, satisfying the principle of profit maximization. At that output, x^m, the firm's costs will be given on the average cost curve at point **g** as shown by the blue area. The firm's profits are its revenues minus its costs for producing quantity x^m. The revenue per unit (average revenue) is given by the price shown at point **h** for the output x^m. Therefore the economic profits are shown as the green shaded area.

With declining costs there is no firm supply curve

There is another reason why the owners of firms do not ask themselves the "what if" question above about how much the firm would produce: in many cases the answer would be either "nothing" or "an infinite amount". To see why consider the firm in Figure 8.14.

The firm's average costs are declining with increased output, while marginal costs are constant. Suppose we asked the owner of this firm the question: "How much would you produce if you could sell any amount you wished at a some price p?" The owner would think the question odd, because she would of course know that she would not be able to sell any amount she wanted at any positive price.

But if she was willing to answer the question anyway, she would notice as long as she could produce at a level such that average costs were less than the price specified, she would be making economic profits. The more she produced, the more profits she would make.

So if she took the question literally, she would say that she would produce an infinite amount of goods. Or more practically, as much as she possibly could, that is, until her firm ran into some constraint on its capacity to produce (what is called a *supply constraint* not shown in the cost function). If she were asked how much she would direct her firm to produce if the price were $p < mc$ her answer would be "nothing" because no matter how much she produced her economic profits would be negative.

From this example we see that for a firm whose average costs fall as more output is produced unless there is something that eventually limits how much the firm can produce, the firm supply curve does not exist . The same is true for a firm with constant average costs (equal to marginal costs): for a $p > ac$ the owners would want to produce an infinite amount and for $p < ac$ they would in the long run produce nothing; they would go out of business.

Firms' supply curves when they face capacity constraints

Practically speaking, firms do face supply constraints (also called capacity constraints) that limit their production. A supply constraint is some level of production $\underline{x}$ such that it is not possible (at any cost) to produce more than the constraint. Sources of supply constraints include:

- *Natural (long-run) limits on expanding one or more input*: if arable farmland is limited in supply, for example, or the density or zoning of urban locations prevents expansion.
- *Short-run limits on increasing capital goods* or some other input .
- *Intrinsic limits on administrative capacity*: it may be simply impossible to coordinate the activities of a very large organization

These supply constraints may be particularly binding in the short run.

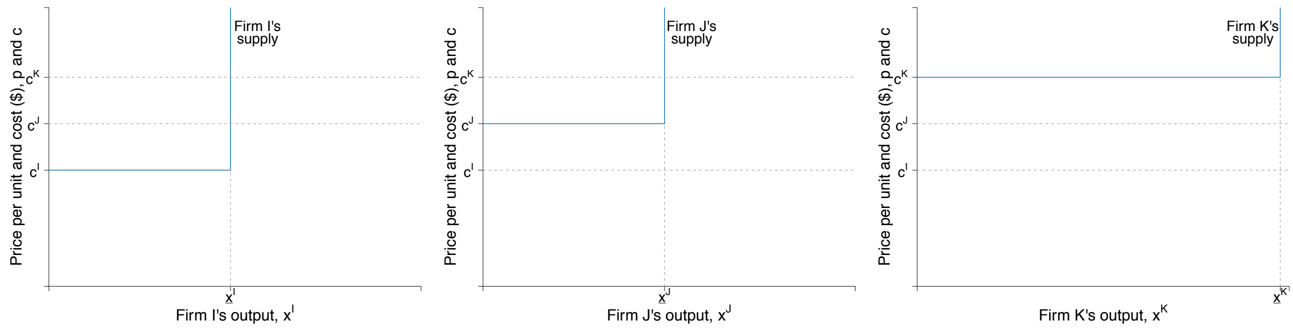
To derive a market supply curve we will simplify by assuming the following:

- Firms have constant average and marginal costs up to some supply constraint, that is, $mc = ac = c$.
- Firms cannot produce beyond the supply constraint, $\underline{x}$.
- Though firms produce an identical sugary drink product, they differ in their marginal and average costs.

Differences in their costs may arise, for example, from their machinery being more or less up to date or their management being more or less competent.

This means that an individual firm's supply function is L-shaped like those shown in Figure 8.15. Firms I, J and K are all firms in the sugary drinks industry, producing liters of sugary drinks that they would like to sell in the sugary drinks market.

Their costs curves are flat and equal to their marginal and average costs – c^I, c^J, and c^K – up to their capacity constraints ($\underline{x}^I$, $\underline{x}^J$, and $\underline{x}^K$), and then vertical. In this instance, the firms' capacity constraints could be determined by sizes of their plants, the availability of ingredients, or other natural limitations that



(a) Firm I's costs and supply (b) Firm J's costs and supply (c) Firm K's costs and supply

Figure 8.15: **Costs, capacity constraints, and supply curves of three firms.** The marginal (and average) costs and capacity constraints of three firms are shown: Firm I, Firm J, and Firm K. They are arranged from lowest cost (Firm I) to highest cost (Firm K). Firm I and Firm J have the same capacity constraint ($\underline{x}^I = \underline{x}^J$), but I's marginal and average costs are lower than J's ($c^I < c^J$). Firm K has a larger capacity than Firms I and J, that is, $\underline{x}^K > \underline{x}^I = \underline{x}^J$, but Firm K's costs are greater than I's and J's, $c^K > c^J > c^I$.

limit the liters of sugary drinks they are capable of producing in the short run.

Though they appear very different, the capacity constrained L-shaped supply curves of the individual firms have an interpretation similar to the individual demand curve. Remember, the demand curve indicates the buyer's *maximum* willingness to *pay* to acquire an additional unit of the good. Similarly, the height of the supply curve is the *least* amount that one could *pay* to a seller – as a take-it-or-leave-it offer – to purchase an additional unit.

Therefore you can think of the supply curve as a minimum *willingness to sell* curve. The minimum level at which Firm I is willing to sell is a price, p, that is at least as great as its costs, c^I. For any price greater than c^I, Firm I is willing to sell a quantity up to its capacity constraint, $\underline{x}^I$. The other firms behave similarly: they are willing to sell goods for any price greater than their minimum willingness to sell or willingness to accept up to their capacity constraints.

8.10 Economic profits and the market supply curve

Just as we constructed a market demand curve by horizontally summing individual willingness to pay demand curves in Chapter 7, the willingness to sell step function in Figure 8.16 is the horizontal sum of the cost curves for firms I, J, and K shown in Figure 8.15, along with the costs of three additional higher cost firms L, M, and N with their corresponding cost curves and capacity constraints.

Remember, the supply curve provides answers to the hypothetical question: at some particular price, p how much output will be supplied? Figure 8.17 shows two different hypothetical prices: p^A (a low price) and p^B (a higher price). The figure shows that at the lower price, p^A, only three firms have a willingness to accept or willingness to sell (considering their costs) that is less than the given market price.

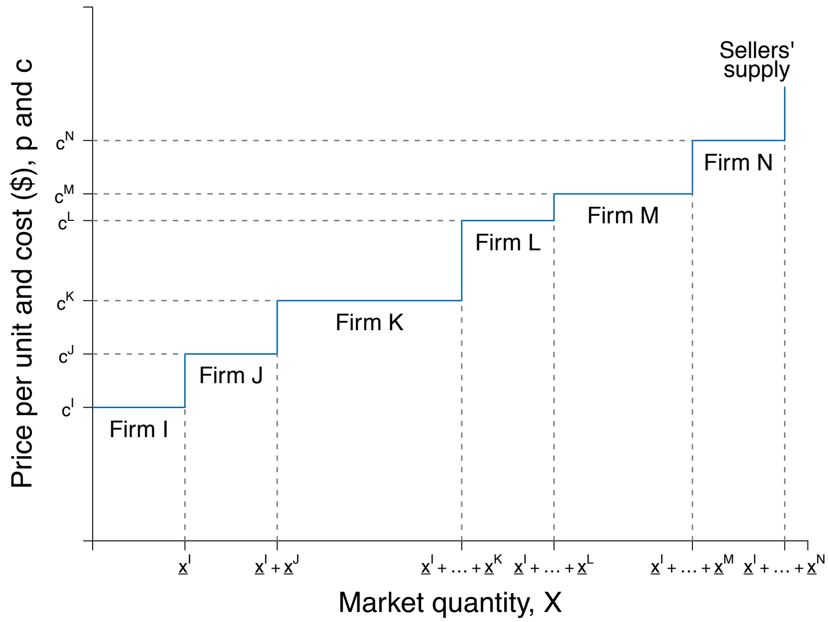


Figure 8.16: **A step-wise willingness to sell function .** The step function is the horizontal sum of the supply curves of the firms : I, J, K, L, M and N over the ranges over their given capacity constraints. The firms are ordered from least cost (Firm I with c^I) to highest cost (Firm N with c^N). The cost curves for firms I, J, and K are shown in Figure 8.15. The costs of Firms L, M, and N are all higher than I, J, and K. The firms' capacity constraints differ, as shown by the differing lengths of the outputs they produce on the horizontal axis.

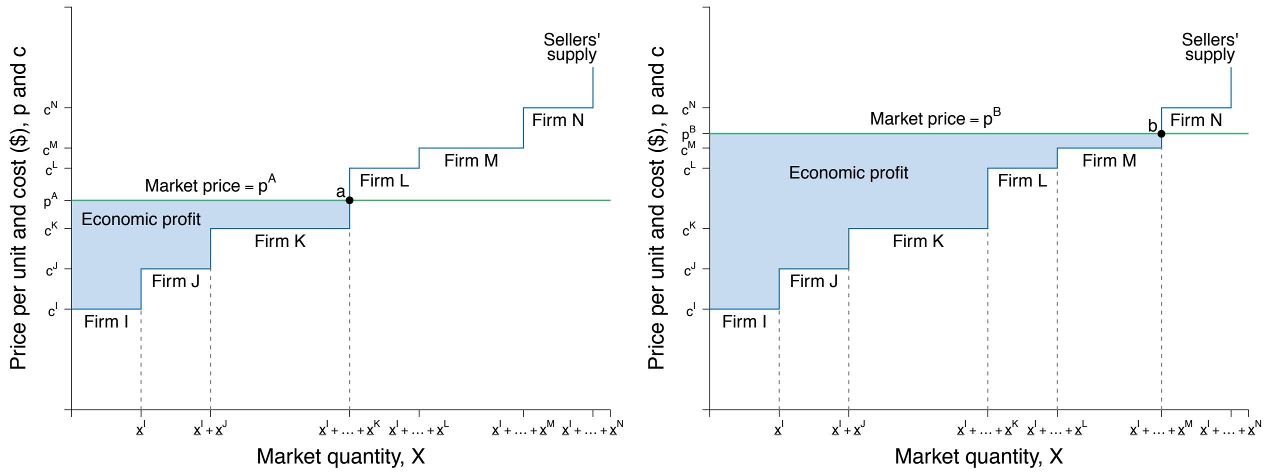
Presuming there is sufficient market demand at this price (remember, the price is just an "as if" thought experiment), each firm will sell a quantity equal to its capacity constraint and will sell a total number of units equal to $x^I + x^J + x^K = X$. This makes the price and quantity sold at **a** in Panel **a.** one point on the market supply function: $X = X(p)$. The other firms L, M, and N will not produce anything.

MARKET SUPPLY FUNCTION The market supply function is the horizontal summation of the individual supply functions of each firm in the industry, which is given by each firm's costs. It is a *willingness to sell* function for the market as a whole.

The higher price p^B, exceeds the minimum willingness to sell of two additional firms: Firm L and Firm M with costs c^L and c^M. So they enter the market, and along with other lower cost firms they produce at their capacity constraints ($\underline{x}^L$ and $\underline{x}^M$). This means that in Panel b., point **b** is another point on the market supply curve. The supply function for the entire market is constructed by repeating the same exercise with other hypothetical prices. So the step function giving the willingness to supply is itself the market supply function.

Economic profits, market supply and firm entry or exit

For each unit that a firm sells at a price above its cost per unit, it receives an economic profit, namely the difference between the price at which it sells the good and its costs. Therefore for Firm I, when it sells at the price p the largest output it can produce is $x^I = \underline{x}^I$, so its profits are $\pi^I(x^I) = x^I(p - c^I)$. In similar fashion, the other firms choose an output level and make economic profits based on their costs, supply constraints, and the price at which they sell their output.



(a) A low price and low economic profits

(b) A higher price and high economic profits

Figure 8.17: **Economic profits with the step-wise supply function.** In Figure a. the price is p^A and three of the firms – I, J, and K – produce at that price because the price is greater than each of their costs, $p^A > c^I$, $p^A > c^J$, and $p^A > c^I$ meaning that each of them makes economic profits per unit equal to the price minus their costs. Therefore, $\pi^I = x^I(p^A - c^I)$, $\pi^J = x^J(p^A - c^J)$ and $\pi^K = x^K(p^A - c^K)$. These profits are shown in the shaded blue area above the costs and below the price. In Figure b. the price is p^B and, following similar reasoning to the above, Firm L and Firm M enter the market as the price is now greater than their marginal costs. They make corresponding economic profits and the profits of the incumbent firms increase. Firm N's cost remains above the price and therefore it does not produce any output.

The total profits at the two different prices are shown by the shaded blue areas in Figure 8.17. Firm N does not enter the market even at the higher price, because its economic profits would be negative, i.e. accounting profits less than the opportunity cost of the capital goods used. Similarly, if the price were at p^B with all but Firm N supplying to the market, but then it fell to p^A, Firms M and L would cease production, as they would be making negative economic profits.

Economic profits and losses therefore regulate whether firms enter the market (or not, like Firm N) or cease production (like Firms M and L at the lower price), a topic to which we return in Chapter 9.

In most we have many more than 5 firms supplying the market so that showing each of them separately as in Figure 8.17 would be impractical. In Figure 8.18 we present a smoothed version of the same market supply curve. The smoothed supply curve we show there based on the previous step-wise functions is linear, but it can have any shape as long as it is upward rising.

Checkpoint 8.5: Supply curves

Why must the smoothed supply curve shown in Figure 8.18 be upward sloped?

8.11 Perfect competition among price-taking buyers and sellers: Shared gains from exchange

By far the most widely used economic model of competition is based on the intersection of the supply and demand curves as illustrated in Figure 8.19, using the sugary drinks market as an example. In Panel a. we show an upward-

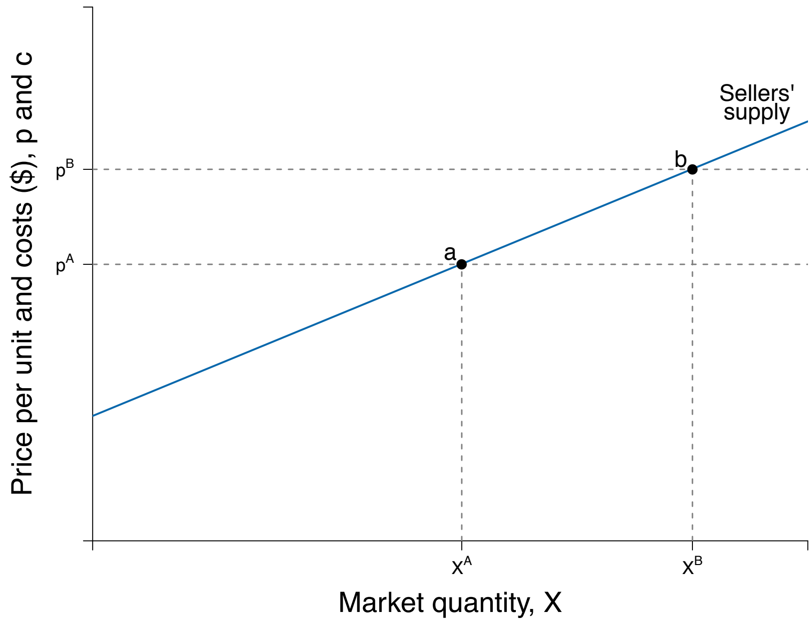


Figure 8.18: **The smoothed market supply curve.** The figure shows a smoothed version of the step-wise function in Figure 8.16. Points **a** and **b** here correspond to points **a** and **b** in that figure.

rising supply curve (like the smoothed variant of the one just derived) and a downward-sloping demand curve consistent with those derived in Chapter 7.

A Nash equilibrium of price taking buyers and sellers

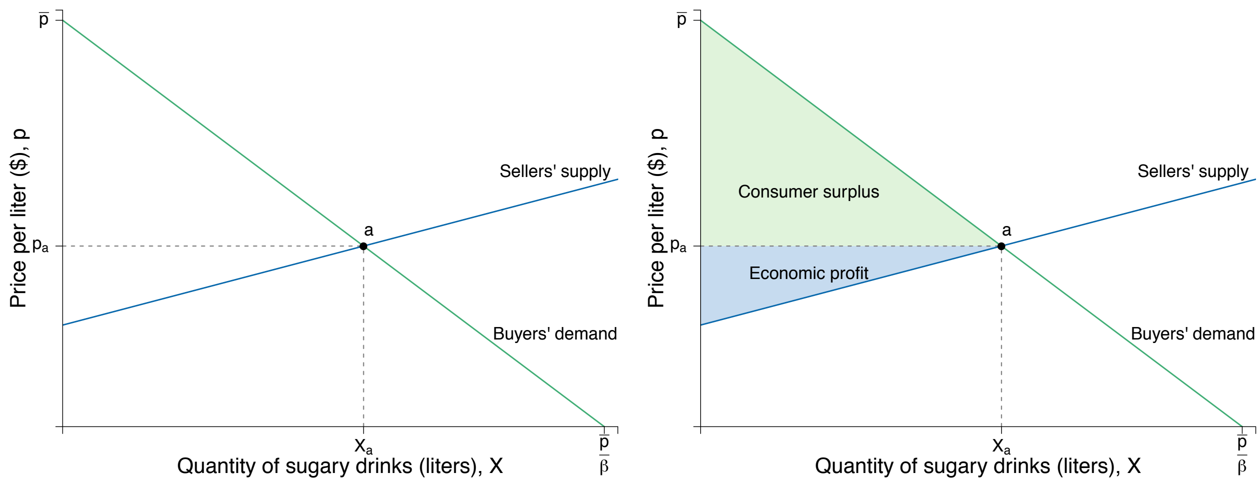
The price and quantity given by the intersection of the two curves is the equilibrium of what is termed the perfectly competitive equilibrium model. This is because the hypothetical questions that we used to construct the supply and demand curves correspond to a way that economists have commonly represented perfect competition. In both cases we asked the buyer or seller to hypothetically take the prices as given, and then say how much they would buy or sell.

PRICE TAKING, PRICE MAKING Price taking and price making are alternative strategies that firms and individuals may follow. Price taking means taking as given the prices at which one might buy or sell. A price maker is a buyer or seller who considers altering the price at which they offer to buy or sell, or altering the level of output in ways that change the price at which they can transact. Monopolistic competitors are price makers, as are the monopolists, duopolists, and oligopolists that we model in the next chapter.

Taking a price as given is termed *price taking*. It means not considering different prices at which one could buy or sell.

Two things to note about the term:

- *Price taking is an assumption about strategies:* Price taking does not mean that a buyer or seller cannot change the price (of course they can, a buyer or seller can post any price they wish at which they are ready to buy or sell); price taking is just an assumption about the strategies that a player in a game – a buyer or seller – considers when deciding what to do.
- *Price taking is a behavior:* Price taking may make sense in some situations



(a) The demand and supply market equilibrium

(b) Consumer surplus and economic profit at market equilibrium

Figure 8.19: **Competitive equilibrium in the market for sugary drinks.** In Panel **a.**, the intersection of the buyers' demand curve and the sellers' supply curve gives the competitive equilibrium in the sugary drinks market. The intersection of the curve provides the market-clearing price: the price at which quantity supplied by the sellers equals quantity demanded by the buyers. The intersection is shown by point **a** with quantity demanded X_a and market price p_a. In Panel b., the consumer surplus is the area shaded in green beneath the demand curve and above the market price, p_a. The economic profits are shown by the shaded area in blue above the supply curve (the minimum willingness to sell) and the market price, p_a.

but not in others. Price taking is not an attribute of a firm or other seller or buyer, it is a behavior that buyers and sellers may adopt as a best response to a particular situation. So when you use the term price taking you should be thinking about the situation, not the person.

From the way we constructed the market supply curve we know that if the price is p_a and firms are price takers, then firms will supply a total of X_a. From the way we constructed the market demand curve, we know that if the price is p_a and buyers are price takers the total demand will be X_a.

If this is the case then the intersection of the supply and demand curves (point **a** in Figure 8.19) is a Nash equilibrium: both buyers and sellers are best responding to the price p_a that is being charged or offered by the other side of the market.

Because supply equals demand, there are no sellers wishing to sell at that price and unable to find buyers, and there are no buyers wishing to buy at that price and not finding a seller. The only relevant actors, therefore, are those transacting at the price p_a. And they are all doing the best they can, given what everyone else is doing.

This shows that if the buyers and sellers are acting as price takers, then the intersection of the supply and demand curves – both constructed on the basis of a hypothetical price taking question – is a Nash equilibrium. We know that monopsonistic or monopolistically competitive firms will not act as price takers: they will affect the prices at which they sell their outputs and purchase their inputs by selecting a level of production.

Shared gains from trade and Pareto efficiency

In Panel b. of Figure 8.19, we show the consumer surplus and economic profits – the gains from exchange or rents – which result if the price is p_A.

Consumer surplus and economic profits are similar in two ways:

- They are both rents, that is utility (for consumers) and income (for firms' owners) in excess of their next best alternative (not buying and not producing, respectively).
- They are similar in the dimensions in which they are measured, namely, money.

  This means that as long as we assume that an extra Euro is worth the same to all consumers and all owners, we can do more than add up the surpluses received by consumers or owners. We can add the two classes of rent together. So the area between the supply and demand curves from the origin to however many goods are transacted is the total rent made possible by the transactions.

There are three important characteristics about the price and quantity at the intersection of the supply and demand curves:

- *Market clearing*: Because the outcome (p_a, x_a) is a point on both curves, the amount supplied is equal to the amount demanded, so there is no excess demand (demand greater than supply) or excess supply (supply greater than demand).
- *Price is (approximately) equal to marginal cost*: To see this, remember that costs differ among firms, but for a given firm costs are constant up to some supply constraint. So over this range of the firm's output, average and marginal costs are equal. Figure 8.17 shows that at any point on the supply function – like **a** or **b** – the price is either equal to the average and marginal cost of the highest cost firm producing for the market, or (as is the case with **a** and **b**) it is between that cost and the marginal and average cost of the lowest cost firm that is not producing for the market.
- *Total rents (consumer surplus plus economic profits) are maximized*: A consequence of the above result – that price is approximately equal to the marginal cost of the highest cost producer in the market – is that the there is no other price and quantity that could feasibly be transacted for which the sum of consumer surplus and economic profits would be larger.

MARKET CLEARING A market clears when the amount supplied is equal to the amount demanded, so there is no excess demand (demand greater than supply) or excess supply (supply greater than demand).

The last point means that the Nash equilibrium of this interaction of price taking buyers and sellers is Pareto-efficient. Starting from the (p_a, x_a) there is no alternative technically feasible price and quantity transacted under which consumers could be made better off without making owners worse off and conversely. The buyers could be made better off, of course, if the price were lower, but this would make the sellers worse off. And vice versa.

8.12 The effects of a tax: Consumer surplus, profits, tax revenues and deadweight loss

We can use the perfectly competitive model to analyse the effects that changes in the supply curve or the demand curve have on prices and amounts transacted, and the resulting changes in consumer surplus and economic profits. We will illustrate this application of the model by returning to the tax on sugary drinks.

As we did in Chapters 5 and 7 we will use comparative statics: compare the Nash equilibrium after the shift in demand or supply with the status quo Nash equilibrium or the Nash equilibrium *before* the change.

- The word *static* refers to the Nash equilibrium because at a Nash equilibrium there are no reasons for the actors to change what they are doing.
- The process is *comparative* because we compare two or more states before and after a change.

We consider a tax that is imposed on firms producing sugary drinks at the rate τ per liter of the drink. Figure 8.20 shows the demand curve for sugary drinks and the supply curve prior to the tax with the Nash equilibrium price p_a and the quantity sold X_a indicated by point **a**, along with the respective consumer surplus and economic profits.

When a tax is imposed, it increases the costs for each firm by the tax per unit, τ. So if Firm F's cost per unit before the tax was c^F then the cost after the imposition of the tax is $c^F + \tau$. Because the supply curve is the horizontal summation of the cost curves of each of the firms and the cost for every firm has risen by the amount of the tax, the market supply curve shifts upwards by that amount too, as is shown in Figure 8.20. You can see that imposing the tax therefore reduces the amount of sugary drinks that firms wish to produce at any given price.

The new Nash equilibrium is point **b** with price p_b and quantity sold X_b. The intended effect of the tax – a reduction in the quantity consumed – is achieved: consumption falls by the amount $X_a - X_b$.

From the figure we can identify the following changes:

- *Consumer surplus lost*. The consumer surplus which before the tax was the areas $A + C + E$ is now just A because fewer units are purchased and at a higher price. In Chapter 7 we noted that the consumers of sugary drinks tend to be poorer than average (in the U.S.) and raised the concern that the tax might be thought to unfairly burden them.

HISTORY In his 1890 *Principles of Economics* Alfred Marshall labeled the area between the market-clearing price and the supply curve as a "producer's surplus" analogous to his consumer's surplus. Others have called this the "seller's surplus." Like Marshall starting from the fact that firms differ in their costs, we have derived a "willingness to sell" based market supply curve for which at either the firm level or for the market, the area between the price and the supply curve is a rent. This is because it is a payment above the seller's next best alternative which is to not produce and sell the good at all, or to sell it at cost. But this is just economic profit, so we do not introduce Marshall's term.

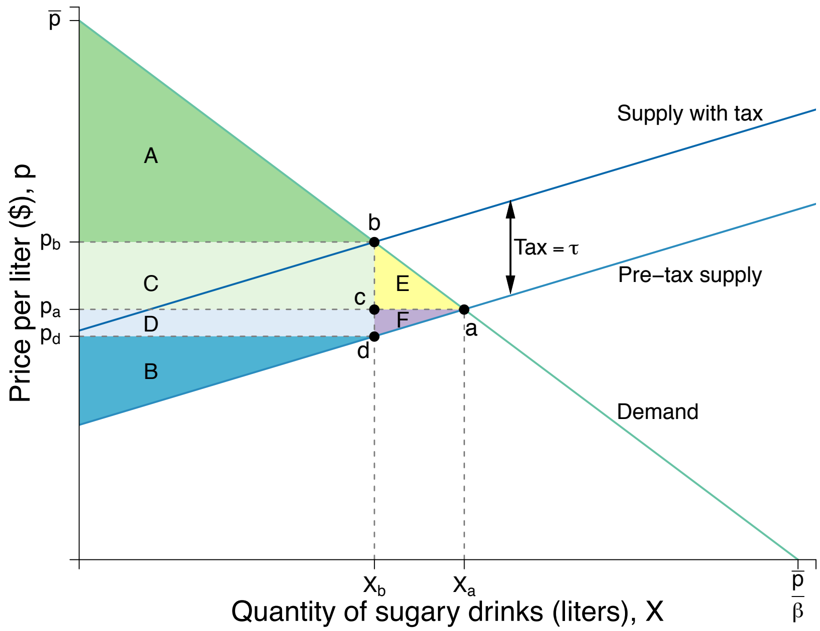


Figure 8.20: **The effects of a tax on sugary drinks.** Before the tax, the total quantity of sugary drinks sold is X_a with price p_a. After the tax, the supply curve shifts upwards because of the increased cost imposed by the tax, with a corresponding decrease in quantity demanded to X_b and a higher price p_b. Consumers and firm owners both bear some cost of the tax. With the tax, only those firms with lower marginal costs, indicated by point **d** will produce and sell the sugary drinks.

- *Economic profits lost.* But the consumers are not the only ones who have suffered losses. The total economic profit, too, has been reduced. Before the tax it was the area $D+F+B$ but because the amount sold is less and the price minus the tax paid per unit is now lower, economic profits are just B after the tax.

- *Tax revenues and government services gained.* There is now a substantial amount of tax revenue equal to the tax times the number of units sold, or areas $C+D$ in the figure. These revenues support governmental programs – education, public safety, income transfer programs for the less well off – that provide benefits for both consumers and owners.

- *Dead weight loss.* The area of the triangle $E+F$ is the *deadweight loss* associated with the tax. The top part of it (E) is consumer surplus lost by consumers (you already know this from Chapter 7; the lower part (F) is the firm's owners lost economic profits.

To understand why the tax creases a "deadweight loss," think about two contrasting effects on the total rents enjoyed in the form of consumer surplus and economic profit:

- *Redistribution of the rents* from consumers and owners on the one hand to the government (and to those who benefit from government financed services); here, it is the case that what the consumers or producers lost the government gained.

- *Reduction of total rents* This is the deadweight loss triangle; and it is not a transfer from one group to another, it is a quantity of benefits that existed prior to the tax that is lost.

Of course we cannot say that the owners or consumers are worse off as a result of the tax. Even setting aside the health benefits and associated reduced medical expenditures, the tax revenues resulting from the policy may finance public policies that confer benefits on both groups more than offsetting the surpluses that they lost either in transfers to the government or as deadweight losses. And there are others, not in the model, that are affected by taxes of this kind.

8.13 Competition among price takers: An assessment

Even before Adam Smith's *Wealth of Nations*, economists developed models of how market competition works. Today, the most widely taught of these in introductory economics – the model of perfect competition – represents buyers and sellers as *price takers*.

HISTORY In his 1948 introductory economics textbook Nobel Laureate Paul Samuelson, called the "founder of modern economics," wrote that monopolistic competition "includes most firms and industries" while the perfectly competitive firm "includes a few agricultural industries."

As is the case with any model, the usefulness of the perfectly competitive Nash equilibrium of price-taking buyers and sellers depends on the insights it can generate despite not being an exact representation of the empirical problem at hand. For example, while consumers of sugary drinks are price takers, that assumption does not apply to some of the producers. Just two firms – Coke and Pepsi – account for well over two thirds of the carbonated soft drinks sold in the United States. They are effectively price making duopolists or at best oligopolists, not price takers.

HISTORY Today, when people think about economics, the first thing that comes to mind is "supply and demand." But in the middle of the last century, in the most famous economics textbook ever, the topic "Determination of price by supply and demand" was put off until p. 447. Exactly ten pages later, the author, Paul Samuelson, wrote: "This is all there is to the doctrine of supply and demand. All that is left to do is to point out some of the cases to which it can be applied and some to which it cannot."

But contrasting two Nash equilibria of a market with price taking buyers and sellers using the comparative static method provided a simple way of identifying and adding up some of the gains and losses associated with the tax on sugary drinks, even if it was far from a complete picture. The model, in other words, captured enough of the reality of the market for sugary drinks to be a useful tool of analysis.

Stepping back from the sugary drinks market to the analysis of market competition in general, there are four conditions under which the Nash equilibrium of price taking buyers and sellers would be a plausible model:

Standardized products: firms produce a product that is indistinguishable or barely so from that of its competitors.

Limited barriers to entry so that no firm or other actor can affect the price in its favor by altering the amount that it buys or sells.

Rising or supply constrained average cost curves: remember that if the average cost curve is decreasing or constant, the supply curve does not exist.

Market equilibrium with market clearing so that the intersection of the supply and demand curves approximates the real situation we are studying.

The first two conditions – standardized products and price taking by buyers and sellers - define what competition means in the "perfect" sense. But the third and fourth conditions – rising cost curves and market clearing equilibria – are equally essential to make the model work.

To evaluate this theory we can use two types of standards:

- Is it *coherent*: does it make sense internally? Or, for example, are its assumptions contradictory?
- Does the model and the predictions based on it *correspond to the reality* it is designed to describe?

Is the perfectly competitive model coherent? We have shown that if buyers and sellers are price takers then the intersection of the supply and demand curves is a Nash equilibrium. But we now have to ask: is acting as a price taker a Nash equilibrium? Does the perfectly competitive model arbitrarily limit the strategies that buyers and sellers can follow? Is price taking the best a buyer or seller can do?

FACT CHECK The ingredients and recipe by which Coke is made is a legally protected trade secret; and the company also has a government enforced monopoly on the trademark Coca Cola (and Coke).

We have already seen that to make the price elasticity of demand less elastic and to thereby raise profits, firms will seek to differentiate their products and to limit entry of firms to the markets in which they sell. If they are successful in doing this they will be able to profit from being price *makers*. In this case the assumption of profit maximization and price taking are contradictory. So price taking is not "doing the best they can" and cannot be part of the Nash equilibrium. In cases where this is true, the perfectly competitive model is incoherent.

Does the model and its predictions correspond to the reality that it claims to describe? Models are useful because they do not correspond to all of the complexities of real economic problems. So this is a question about where the model applies sufficiently to be useful and where it does not. While we gain important insights about the sugary drinks market from the model of perfect competition – despite the huge market shares of Coke and Pepsi – there are other markets in which the model will be substantially misleading.

We will see in Chapters 12 and 11 that neither the assumption of price taking nor the condition that markets clear correspond to the reality of some of the most important markets of a modern capitalist economy. Examples are the credit market (where banks set interest rates and the demand for loans typically exceeds the supply) and the labor market (where employers set wages and the demand for labor typically falls short of the supply, meaning that people are unemployed). In Chapter 9 we will see that barriers to entry can be sufficient so that price making strategies for selling goods and services best characterize a substantial part of modern economies.

There are other limitations of the perfectly competitive model. We have shown that the Nash equilibrium is Pareto efficient by restricting our analysis to consumers and their surplus and owners and their profits. But there will be others affected by the production and consumption of the goods or services in question who are not considered in showing that the price taking competitive equilibrium is Pareto efficient. These external effects concern:

- Those who produced the goods
- Those affected by the external environmental effects of the goods' production
- Others affected by the external effects of the good's consumption (think of the health impacts of sugary drinks and their effects on family members and tax payers.)

So our statement that the perfectly competitive equilibrium is Pareto efficient is based on a useful but incomplete model. We will consider these and other external effects in Chapter 14.

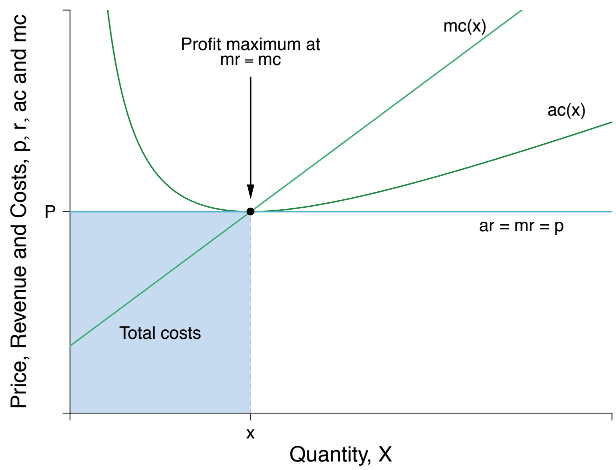
None of these limitations means that the model is "wrong" or cannot be insightful in answering some questions. But they do suggest that the domain over which the model of perfect competition applies in the real economy is somewhat restricted.

8.14 Two benchmark models of the profit-maximizing firm: Price takers and price makers.

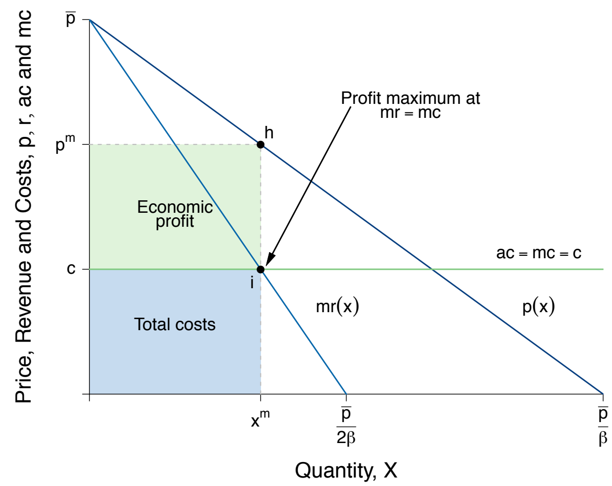
The domains for which the price taking model is of limited applicability – markets with substantial barriers to entry, labor markets, credit markets for example – are all interactions in which some form of price making figures prominently in the profit-maximization strategies of firms. Here we bring together what you have learned about price taking and price making firms.

The price taking owners of firms take the market price as given in choosing an output level to maximize profits. The size of firms in this model is limited by the fact that marginal costs are assumed to rise with increasing levels of output and will exceed average costs (because average costs are rising with increased output.) As a result, for any given price, firms will expand up to the point that price equals marginal cost, but no further. Higher levels of output would reduce profits.

As you have learned, the model does not apply to cases in which average costs decline with increased output or are even constant as output expands.



(a) Perfectly competitive firm



(b) Monopolisticly competitive firm

Figure 8.21: **The perfectly competitive firm and the monopolistically competitive firm: two benchmark models**

In these cases, if instead the price exceeds average costs then the firm will want to grow without limit.

A downward-sloping average cost curve (like the cost curves in Figure 8.14) raises two problems for the theory:

- If the firm grows without limit, then we have to question the assumption that all firms are small relative to the size of the market so that the decisions the owners of firms make will not affect the market price. A consequence is that the process of competition itself may destroy the conditions under which a large number of small firms could compete.
- Price cannot be equal to marginal cost because average cost (at any level of output) exceeds marginal cost, so firms that set $p = mc$ would eventually go out of business. A consequence, we will see in the next chapter, is that the resulting prices and quantities will not be Pareto-efficient.

An alternative benchmark model accepts the evidence that long run average cost curves are flat or even downward sloping and explains why firms do not grow without limit by the fact that firm sales are constrained by a downward sloping demand curve. Demand curves slope downward because in practice, many firms have a limited number of competitors selling exactly the same product that they are producing and the absence of close substitutes – a Honda is not a Ford.

The two benchmark models – perfect competition and monopolistic competition – are illustrated in Figure 9.19. For the monopolistically competitive firm, the downward sloping demand curve replaces the upward rising marginal cost curve as the limit on firm growth. And the flat cost curve replaces the flat demand curve as a part of the model facilitating firm growth.

Under these conditions, firms will not grow without limit. This is true even though the owners of the monopolistic firm, unlike the perfectly competitive firm, will be receiving economic profits. So the downward-sloping demand curve allows us to reconcile flat or downward-sloping cost curves with a limit on firm size.

But by accepting the monopolistic competition model as a benchmark, one has to give up the result that in equilibrium the price will be equal to marginal cost. This will be important when we consider the relationship between market competition and the efficiency of the resulting allocations.

Which of the two models is best to use depends on the question that you would like the model to help you answer. For example, the model of supply and demand with perfectly competitive firms may not be the best representation of the sugary drink industry.

There are two reasons:

- We have already seen that in the U.S. carbonated drinks are produced and wholesaled not by large numbers of firms constrained to sell at the market price, but by just a few large price setting firms facing downward sloping demand curves. This is the case for most sugary drinks in most of the world's markets.
- Moreover, the cost of producing sugary drinks is probably either constant or declining with increased sales, so the supply curve essential to the model of perfect competition does not exist.

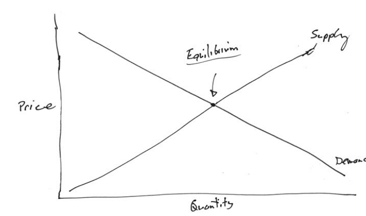


Figure 8.22: **Supply and demand according to Bill Gates.** Bill Gates of Microsoft tweeted this figure with the heading "At the time I was in college this was basically how the global economy worked" adding ". . . cost . . . increases as supply increases." He went on "But software doesn't work like this. Microsoft might spend a lot of money to develop the first unit of a new program, but every unit after that is virtually free to produce. . . . And it's not the only example: data, ebooks, even movies work in similar ways."

For both reasons, the model of monopolistic competition would seem to be a better representation of the sugary drinks market. We will study the process of competition among large firms in the next chapter.

8.15 Application: Dynamics – The growth of firms and the survival of competition

The downward-sloping demand curve can place a limit on firm growth, but other factors may work in favor of continuing growth in size, especially of large firms. These include:

- *Radically declining average costs* as occur in the production and sale of knowledge intensive products such as software and pharmaceuticals where "first copy costs" (c_0 in our cost function) are substantial while marginal costs are effectively zero;
- *Demand-side increasing returns* introduced in Chapter 7, in which the willingness to pay (or to endure advertising) is greater for the one millionth member of a network than it is the first hundred and

- *Learning by doing* which gives advantages to incumbent firms with a larger amount of cumulative sales as the example of the M20 armored truck during the Second World War showed.

This brings us to the question: If successful firms grow, taking over a larger market share, then why doesn't the number of firms in a given market shrink?

FACT CHECK The last few decades have witnessed spectacular growth in some firms. In 2020 for example, Walmart employed over 2 million people worldwide, and in the U.S. about half of all employees work in firms with more than 500 workers. Since 1988, the average firm size in the U.S. has remained roughly constant (around 20 employees).

To provide a possible answer, a completely different set of models of firm size and competition has been proposed, inspired by biological models of competition for fitness in the natural world. In these "life history" models, firms are born, grow and die, so the size of any firm in existence is typically growing. But they do not grow without limit (they die) due to influences such as bad luck or mismanagement that are not directly related to prices and average costs.

Think about a similar but simpler question. Why is the average age of many populations constant despite the fact that every member of the population is getting older? The answer is that those leaving the population are older than those being born into it.

A similar possible explanation for why the average firm size could stay constant even if some firms grow perpetually may be that when firms die, they are replaced on average by one or more smaller firms. So, constant growth of surviving firms can be consistent with an constant average firm size.

To see how this could work, let us say that there is a firm i whose size, by some measure, say sales or employment, is s_{it} at time t, and it grows in size at the percentage rate, $g^s > 0$. Firm i's size at time $t+1$ would, therefore be:

$$s_{it+1} = (1+g^s)s_{it} \tag{8.46}$$

Let us assume that there are a constant number, n , of firms in the industry and that the average size of a firm at time t is $\bar{s}_t$. Next, suppose that in every time period, each of the n firms dies with a probability f. Once dead, it is replaced by a firm of size $\underline{s} < \bar{s}$. In other words, the new firm is smaller than the average firm size in the industry.

We can then find values of the growth rate of firms g^s, the size of new firms $\underline{s}$ and the "mortality rate" of firms f such that there is some average firm size that does not change even if the surviving firms grow at the rate of g^s. In other words, we can find a set of values $(g^s, \bar{s}, f)$ for which $\bar{s}_{t+1} = \bar{s}_t$.

To see this, we write the weighted average firm size in $t+1$, as the increased size of all surviving firms, multiplied by the probability that the firms survived, plus the probability that a firm did not survive, times the size of the average replacement firm.

$$\bar{s}_{t+1} = (1-f)(1+g^s)\bar{s}_t + f\underline{s} \tag{8.47}$$

To find the conditions under which firm size could be unchanging from period to period, we equate the average firm size in time t and $t+1$. That is, we find the common value of $\bar{s}_t = \bar{s}_{t+1} = \bar{s}$ satisfying the equations:

$$\bar{s}_{t+1} = (1-f)(1+g^s)\bar{s}_t + f\underline{s} = \bar{s}_t \tag{8.48}$$

We can simplify the above to get:

$$g^s\bar{s}(1-f) = f(\bar{s}-\underline{s}) \tag{8.49}$$

The above equation gives us a condition such that the average firm size will not change despite the growth of surviving firms. The average firm size will be constant if its growth caused by firms surviving with probability $(1-f)$ and growing at the percentage rate g^s (the left side of the equation) is offset by the shrinkage in firm size caused by firms dying with probability f and being replaced by firms that are on average smaller by the amount $\bar{s}-\underline{s}$ (the right side of the equation). This will be possible (for positive firm size) if the positive effect of increased average size due to firm growth exceeds the negative effect of greater average firm size associated with replacement firms being smaller than failed firms, or $g^s(1-f) < f$.

The resulting constant firm size can then be found by solving equation 8.49 for $\bar{s}_t$ as is illustrated in Figure 8.23.

$$s^* = \frac{\underline{s}f}{f-g^s(1-f)} \tag{8.50}$$

If we take f as a measure of the degree of competition, then the average size of firms will be greater when competition is less. The average size of the firm will also be greater if firm growth is greater, as is shown by the dashed line in the Figure 8.23, or if the size of the replacement firm is greater. (The latter would be the case if barriers to entry are greater for smaller firms.)

The model could be extended to take account of the possibility that firm growth relates positively to firm size, and that firm failure probability relates negatively to firm size. Additionally, if firms fail by merging with a larger firm, then there is no smaller "replacement firm." All of these possibilities make the puzzle of constant firm size more difficult to reconcile with surviving firms growing.

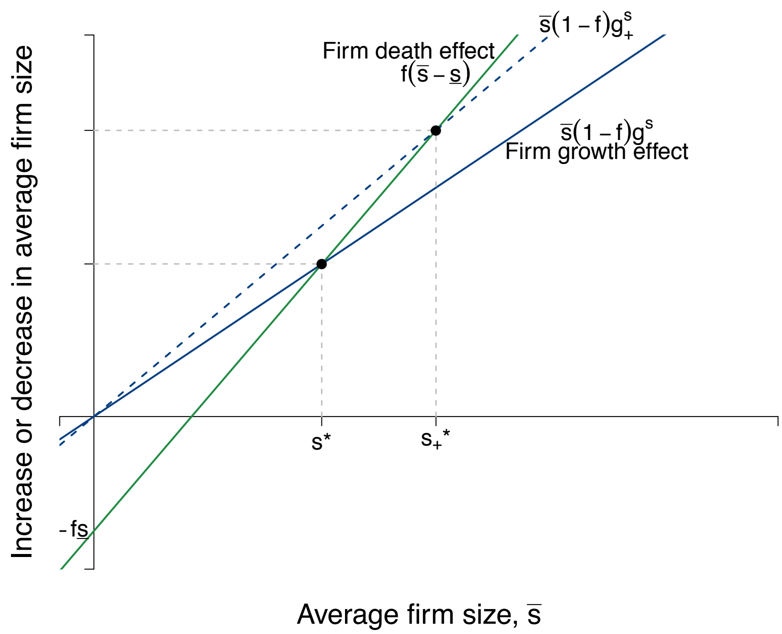


Figure 8.23: **Constant average firm size with every firm growing in size.** The constant firm size is s^*. This occurs when the firm death effect $f(\bar{s}-\underline{s})$ is just equal to the firm growth effect $\bar{s}(1-f)g^s$. With a more rapid growth of existing firms (a higher g^s, say g^s_+), the constant firm size would be higher, at s^*_+.

Economists will continue seeking models in which both competition and the advantages of large scale production play an important role. The biologically inspired "life history" models of the firm allow for both economies of scale and the enduring importance of competition. Whether they will prove insightful for the analysis of firm behavior in other respects remains to be seen.

8.16 Conclusion

We began this section of this book in Chapter 6 with an explanation of how economies of scale (along with learning by doing) lead to specialization by task and product, and how the resulting division of labor requires a method of distributing goods from the specialist producers to the end generalist user. Markets play an essential role in coordinating this distribution process.

Economics has demonstrated that for many goods and services where conditions approximating perfect competition obtain, markets can perform this task reasonably well, at least by comparison to alternatives such as centralized allocation of goods and services by governments.

But it may be that the factors that make specialization so beneficial – economies of scale and learning by doing – in the long run also undermine the process of competition itself. It appears therefore that the economies of scale and learning by doing are among the reasons why markets have come to play such an important role in the economy and may be inconsistent with the price taking model of competition that is the basis for economists thinking that markets do a good job.

We turn, then, to models of how firms compete that are more consistent with what is known empirically about modern economies.

Making connections

Profit maximization using feasible sets and indifference curves The firm's choice of price and output level is another constrained optimization problem, with the demand curve as the frontier of the feasible set and the firm's objectives summarized by a set of isoprofit curves.

Consumer surplus, economic profits and mutual gains from exchange Both consumers and firms' owners will typically receive rents arising from the the production and sale of goods.

Restricting output and hiring: A larger slice of a smaller pie The firm facing a downward sloping demand curve or an upward-rising cost of labor (or other input) curve will increase its profits by restricting hiring and sales so as to sustain higher prices and lower wages; the firm's customers and employees do less well as a result.

Economics as an empirical science Our models are informed by evidence; for example that average costs are not greater (and maybe less) at higher levels of output, and that intellectual property rights and other barriers to entry can result in substantial price markups over cost.

Two contrasting benchmark models are a) firms that face a flat demand curve and are constrained by rising costs and b) monopolistic competition or other settings with limited competition in which firms with flat cost functions are constrained by a downward sloping demand curve.

Public Policy The supply and demand model of price taking buyers and sellers on competitive markets illuminates the effects on consumers, owners, and others of a tax on sugary drinks.

Learning by doing, the survival of competition, and dynamic analysis The declining cost of the M20 armored car during the Second World War and of solar panels recently along with the the model of how average firm size might stabilize even if all firms are growing are examples of a dynamic analysis, looking at how something changes over time rather than comparing the equilibrium before and after some exogenous change, as in comparative static analysis.

Important ideas

total cost	average cost	marginal cost
fixed cost	falling average cost	production set
accounting cost and profit	economic cost and profit	average profitability
total revenue	average revenue	marginal revenue
profit maximization	equalization of marginal cost and marginal revenue	firm shutdown
real world cost curves	isocost curve	birth and death of firms

Mathematical notation

Notation	**Definition**
x	total output of a firm
X	market output of a good
p	price of a unit of output
a_l, a_k	labor and capital inputs required to produce a unit of output
c	total costs of production
w	wages
l	amount of labor
p_k	price of capital goods
ρ	opportunity cost of capital
c_k	total cost of a unit of capital
$r(x)$	total revenues function
π	economic profits
r	profit rate
β	slope of the demand function
η	elasticity of market demand with respect to price
μ	mark-up ratio over costs of production
τ	amount of tax
s	firm's size
g^s	growth rate of a firm
$\underline{s}$	size of a firm entering to the market
f	probability of failure of each firm

Note on super- and subscripts: A: accounting; E: economic; m: monopoly; C: competitive market.

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

See reference list.

9

Competition, Rent-seeking & Market Equilibration

. . . according to Dr. Johnson, competition is "the action of endeavouring to gain what another endeavours to gain at the same time." Now, how many of the devices adopted in ordinary life to that end would still be open to a seller in a market in which so-called "perfect competition" prevails? I believe that the answer is exactly none. Advertising, undercutting, and improving ("differentiating") the goods or services produced are all excluded by definition–"perfect" competition means indeed the absence of all competitive activities.

Friedrich Hayek "The Meaning of Competition," 1948

DOING ECONOMICS This chapter will enable you to do the following:

- Understand competition as a strategic process among firms, actively rent seeking through price-making, choice of output, advertising, innovation, product differentiation, and more.
- See why (except under special conditions) the "price equals marginal cost" condition for Pareto-efficiency of the level of output will not be a Nash equilibrium.
- Explain how perfect price discrimination, like perfect competition is a an abstract ideal model illustrating conditions under which price would equal marginal cost.
- Show how the prices, quantities sold, the price mark up over costs, economic profits and consumer surplus vary with the extent of competition in a market.
- Explain how barriers to entry restrict competition among firms, increase deadweight inefficiency , and raise the owners' profits owners while reducing customers' consumer surplus.
- Understand the economics of non-clearing markets and how rent seeking by buyers and sellers may equilibrate supply and demand.
- See how the forces of supply and demand work by altering the fallback positions of buyers and sellers.
- Contrast "modern monopoly" (e.g. Microsoft) with the kinds of concentrated market power that were more common in the past (e.g General Motors).

The year 1890 was a big one for computers. It was the year that Herman Hollerith, frustrated with the eight years it took to tabulate the 1880 US census, decided to introduce punch cards into data collection and entry processes for the census. It took 1 year rather than 8 to tabulate the 1890 census, a massive improvement in productive efficiency, that heralded a revolution that would echo through the twentieth and twenty-first centuries.

Charles Flint, recognizing the genius of Hollerith's system, bought out Hollerith's company and, merging it with several of his own companies, created the company that became, in 1924, the International Business Machines company, more commonly known as IBM. For decades, IBM was at the forefront of computing innovation, providing large machines to companies, governments, and universities; and raking in huge profits as a result.

But IBM, like its competitors, didn't see the coming wave of the personal computing **industry** and the dominance of software companies. It took other innovators – Bill Gates and Paul Allen of Microsoft, Steve Jobs and Stephen Wozniak of Apple, among others – to see the potential that a personal computing machine would have in the everyday lives of people around the world.

Even Steve Jobs – famous for advising Stanford University's graduating class in 2005 to "stay young, stay foolish"– could not keep up with the pace of

Figure 9.1: **An Apple Macintosh Classic introduced in 1990.** Courtesy Alexander Schaelss, 2004, CC 3.0

things. It took him and others a couple of decades after the invention of the personal computer to see the potential for personal musical devices – the iPod – and to design a smart phone – the iPhone. The iPhone would go on to dislodge the giants of the cell phone industry, Nokia and Research in Motion (the manufacturer of the Blackberry).

Big, near-monopolistic, profitable, and established firms did not foresee the coming changes to their industries. The story of IBM echoes within and across industries. Kodak a once-dominant firm in the photography industry failed to develop its own digital camera even though it had people within the company who invented one. Their invention had been discarded because of the damage it might do to their core business, selling film and film-based cameras.

Microsoft, which had won the personal computing battle of the early nineties through its innovative software, failed with its smartphones (losing to Apple), search engine (losing to Google, now Alphabet), web browser (losing to a variety of Chrome, Firefox and other non-Internet Explorer packages), or networking systems (losing to the Linux-based Apache system).

INDUSTRY An industry is a set of firms producing a similar product. "Similar" products are close substitutes for each other: tea and coffee, not cookies and coffee. In the Cournot model the firms in an industry produce identical products. Because they are identical, they are perfect *substitutes* for each other.

When we think of sports, politics, the job market, or social status climbing, we usually think of competition as a process in which competitors actively seek to gain advantage over others. In the economy, too, competition is the equivalent of warfare where there are big winners, and, as IBM and Kodak found out, big losers. Depending on the nature of the competition, consumers too can be big winners, or big losers.

But as Friedrich Hayek pointed out in the head quote, what economists call perfect competition as taught in introductory economics courses differs from competition in politics and war. Buyers and sellers are passive "price takers" (who take the price as given). In this chapter, we present a more empirically grounded view of competition. Firms are not price takers, they are price makers. Firms set prices and innovate to capture a larger market share, and even create and then dominate entirely new markets.

HISTORY Auguste Cournot (1801-1877) was a French mathematician and economist who studied the process of competition among firms. He is considered by many to be the founder of the field of mathematical economics.

To do this we study how competition works whether there are a substantial number of firms producing similar products or there is a single firm selling a unique product. We will introduce the perfect competition model as a special case.

9.1 Modelling the continuum of competition: From one firm to many

To understand how the number of firms in a market affects the outcomes for firms and consumers, we start with what is called the Cournot ("Coor-NO") model of competition.

REMINDER Remember from Chapter 7 that the inverse demand curve ($p(X)$) gets its name from the fact that it is the mathematical inverse of the conventional demand function, which makes quantity sold a function of the price ($X(p)$).

To determine the outputs the two or more firms will produce and the price at which the goods will sell we cannot consider the firms in isolation as we did when we modeled monopolistic competition in Chapter 8. The firms are engaged in a strategic interaction: their owners know that their profits depend on not only their own firm's actions but the actions taken by the other. This is why we use game theory to understand the process of competition. Cournot competition is represented by a game with the following characteristics:

REMINDER When we refer to a "firm" seeking to maximize profits or making some other decision we mean the owners of the firm (who will receive the profits and directly or indirectly through management make the relevant decisions.)

- *Players:* The owners of firms in an industry (if there is more than one firm) sell an *identical* product;
- *Strategies*: Each firm simultaneously selects a level of output to produce and sells that entire output at the highest price possible given by the other firms' sales and the industry inverse demand curve.
- *Payoffs:* For every a set of outputs for each of the firms (called the *industry output profile)* there is a level of economic profit (possibly zero or negative) received by each firm's owners.
- A *Nash equilibrium:* is an *industry output profile* and a price such that each firm's output level is a (profit maximizing) best response given the other firms' output levels when the *single price* at which all of the firms outputs are sold is determined by the industry demand curve and the total output produced by the firms.

There are two important consequences of this set-up.

First, firms in the Cournot model are not price takers, they are price makers. As in the models of monopolistic competition and monopsony in Chapter 8 the owners of the firm know that the level of output the put on the market will affect the price at which they can sell it. They deliberately "make" prices by choosing how much to produce, taking account of the fact that they are constrained by a downward-sloping demand curve.

REMINDER The Nash equilibrium describes a situation in which, were it to occur, none of the firms' owners would have an incentive to alter their strategy. It does not describe the *process* by which the equilibrium would come about.

LAW OF ONE PRICE The Law of One Price states that in equilibrium identical goods or services will transact at the same price.

Second, competition in the Cournot model conforms to the **Law of One Price**. The Law of One Price states that in equilibrium identical goods or services will transact at the same price. The reason why prices do not differ among the firms if one firm's output were selling at a higher price than other firms' outputs, then buyers would switch to the lower priced firms.

Identical prices for identical goods may seem a truism too obvious to warrant a law of its own, but it is not always true: airlines, for example regularly charge different prices to different categories of customers – the elderly for example – for exactly the same seats on the same flights. We introduce this case – price discrimination later in this chapter, and other violations of the Law of One Price in subsequent chapters.

THE LAW OF ONE PRICE states that in a competitive equilibrium identical goods or services will transact at the same price.

Cournot's model allows us to consider a continuum of competition that we illustrate by three cases differentiated by the number of firms in the industry, n:

- *Monopoly*, $n=1$: there is only one firm in an industry.
- *Duopoly* $n=2$: a second firm shares the industry demand.
- *Oligopoly* and *"unlimited competition"* $n>2$: several firms (oligopoly) or many firms (unlimited competition) share the demand.

REMINDER In Chapter 8 we introduced the model of monopolistic competition (due to Joan Robinson and Edward Chamberlin) that originated during the Great Depression, showing how the owners of a firm decide how much to produce when there are a small number of firms producing products that are differentiated, but are similar enough so that raising a price will lead some potential buyers to switch to a competitor. Here we study a model in which a single product is produced by *n* firms with *n* ranging from 1 (monopoly) to many (similar to perfect competition.)

"Monopoly" and "monopolistic competition"

The outputs and prices that maximize profits in the monopoly case $n=1$ are identical to the case of monopolistic competition introduced in Chapter 8 as long as the firm's product is differentiated in some way (by trade mark for example) so that no competitor can sell an identical product.

The term "monopoly" suggests a sole seller of a product with few substitutes. Examples would be a single local seller of electricity or the drug Daraprim (life saving treatment of an HIV AIDS related illness) which in 2015 was priced at 750 dollars a pill, having previously had been sold (presumably at a profit) for seven U.S. dollars a pill.

By contrast, the term "monopolistic competition" stresses that many single sellers of a differentiated product (literally monopolies) do have to compete with firms selling close substitutes (unlike Turing Pharmaceuticals, the seller of Daraprim). Examples are the sugary drinks discussed in Chapter 8 with very price elastic demand (Coca cola at $|\eta| = 3.79$ and Mountain Dew at $|\eta| = 4.39$). These highly price elastic demand curves are a indication of the competitive nature of the sugary drinks industry. The price elasticity of demand for sugary drinks as a whole, however, is much lower at $|\eta| = 1.4$. So if that entire sugary drinks market were served by a single firm, the term "monopoly" would be appropriate.

The economic environment: Demand, revenue, costs and profits

268

For each of these cases (where the number of firms, $n = 1, 2$, few, many) we will use the *inverse demand function* for which price depends on the quantity sold (we assume that firms sell everything they produce, so we use "sales" and "output" interchangeability. For simplicity we assume the demand curve is linear where p is price, X is total output and sales in the market, $\bar{p}$ is the maximum price when output is zero, and $-\beta$ is the slope of the industry inverse demand curve, $\frac{\Delta p}{\Delta x}$:

$$\begin{aligned} \text{Inverse demand function} \quad p(X) &= \bar{p} - \beta X \\ \text{where} \quad X &= x^1 + x^2 + ... x^n \end{aligned}$$

FACT CHECKWe use this cost function for simplicity, not for realism. As we saw in the Chapter 8, firms typically have some fixed costs and large firms may enjoy either economies of scale in production or increased bargaining power in purchasing inputs from suppliers (meaning that larger firms pay less for inputs and have lower costs).

In the model, all firms have an identical cost function with *no fixed costs* and a *constant marginal cost*, c.

$$\text{Cost function} \qquad c(x) = cx \tag{9.1}$$

The opportunity costs of the machinery, intellectual property, and other capital goods used in production are included as a cost of production. In this case, the firm's marginal and average cost are equal and do not vary with the firm's output.

When a firm sells its product at a price greater than its costs it makes economic profits, meaning that accounting profits exceed the opportunity cost of capital.

REMINDER If costs are $c(x) = c$, then average costs (the total costs divided by the total output) are $ac(x) = \frac{c(x)}{x} = c$ and marginal costs (the change in total costs associated with a small increase in output Δx) are $\frac{\Delta c(x)}{\Delta x} = c$.

To calculate the firm's economic profits, we use its revenue and costs as follows:

$$\begin{aligned}
\text{Firm i's revenue} \quad r^i &= p(X)x^i \\
\text{Firm i's profit} &= \text{Revenues} - \text{Costs} \\
\pi^i &= r(x^i) - cx^i
\end{aligned} \tag{9.2}$$

We can substitute the inverse demand curve $p(X)$ into Firm i's profit function to find their profits:

$$\begin{aligned}
\pi^i &= p(X)x^i - cx^i \\
&= (\bar{p} - \beta X)x^i - cx^i && (9.3) \\
&= (\bar{p} - \beta(x^1 + \cdots + x^n)x^i - cx^i && (9.4)
\end{aligned}$$

A firm's profits depend both on their own production and on the production on of other firms (included in X). Each firm has a negative external effect on the revenues of other firms by producing more output.

Checkpoint 9.1: Constant Marginal Costs

Consider these idea of constant marginal costs and fixed costs:

1. Think about examples of firms that you think likely have constant marginal costs for certain products. Many of these may even have constant marginal costs that are close to zero. Explain your example and why it has constant marginal costs.
2. Many firms have large and significant fixed costs, for example, Amazon's distribution centers or Microsoft's intellectual property assets. But, many firms these days have very low fixed costs, needing merely some laptops and free software to get started. Which of these do you think is more likely to correspond to a more competitive market? Why?

A coordination problem: Over-harvesting fish and crowding the market.

With a downward-sloping demand curve and more than one firm, firms face a coordination problem like over harvesting fish from the lake in Chapter 5. We can think of the potential buyers of the product of the firms as analogous to the fish in the lake: the more customers that one firm "harvests" the fewer are left for the other firms.

This means that like the lake, the market is a common property resource: like catching fish, selling to customers is *rival* because the customer one firm sells to will not buy goods from another firm. But the firms *cannot be excluded* from competing on the market. The analogy of market competition to the over-fishing problem is summarized in Table 9.1. Comparing the objectives of the

	Over-harvesting fish	*Overselling*
Objective	utility $u^i = y^i - \frac{1}{2}(h^i)^2$	profit $\pi^i = p^i x^i - cx^i$
Production and demand	Fish caught (y^i) $y^i = h^i(\alpha - \beta H)$ $H := h^1 + h^2 + ... + h^n$	Revenue ($p^i x^i$) $p^i x^i = (\overline{p} - \beta X)x^i$ $X := x^1 + x^2 + ... + x^n$
External effect, $\mathbf{n = 2}$ (i and j)	β: j's fishing reduces i's catch	β: j's sales reduce i's revenue

Table 9.1: *Comparison between over-harvesting fish and "over-harvesting" customers* The lake and the market are both common property resources: additional fishermen and additional firms cannot be excluded, and they compete with incumbent fishermen or firms for fish stocks and customers.

players – the utility function of the fishers and the profit function of the firms (in the table) – you can see the following:

- the *firm's level of output* x^i is analogous to the fisherman's fishing time h^i – it is necessary to the firms objectives (profit) and it is also a cost.
- *other* firm's level of output (contributing to the total output sold on the market, X) is similar to other fishermen's fishing hours contributing to H because it has a negative effect on the firm's objective.
- this negative external effect in each case is represented by β.

We shall see that as long as firms do not coordinate (forming what is called a cartel) they end up producing more than would be Pareto-efficient, their owners make lower profits than if if they each had produced less or there were fewer firms. There is one difference between the over-harvesting fish model and the over-harvesting customers problem modeled below. In Chapter 5 the only players in the game were the fishermen themselves (they did not sell the fish they caught, they just ate them).

When we model markets, there are not only the owners of the firms but also customers. We will see that when firms fail to coordinate so as to limit their sales firm owners make lower profits, but customers benefit from lower prices. Correspondingly, when firm owners do cooperate, agreeing jointly on a level of output and a price, they can all improve their profits, but at a cost to consumers who then would face higher prices.

Before we see study the coordination failure among firms we look at what happens with one firm in the industry. This optimization problem and its solution are identical to that for the monopolistically competitive firm shown in Figure 8.9.

9.2 Reviewing the monopoly case, $n = 1$

In the monopoly case the firm's output and sales x is the same thing as the industry output and sales X (i.e. $x = X$) so the firm's revenue is $r(x) = p(x)x$. The firm costs are $c(x) = cx$, and so its profit is:

$$\text{Profit} \qquad \pi(x) \quad = \quad p(x)x - cx$$

Which, using the inverse demand function to represent the price, is:

$$\text{Profit} \qquad \pi(x) \quad = \quad (\bar{p} - \beta x)x - cx$$

As with was the case for the monopolistically competitive firm in Chapter 8, the firm chooses the profit maximizing level of output and sales finding the value of x that equates marginal revenue to marginal cost or

$$\begin{aligned} \text{Marginal revenue} \quad &= \quad \text{Marginal cost} \\ \frac{\Delta r(x)}{\Delta x} = \frac{\Delta p}{\Delta x}x + p = \bar{p} - 2\beta x \quad &= \quad c \end{aligned} \tag{9.5}$$

Solving Equation 9.5 for x gives the profit-maximizing level of output and sales for the Cournot model with $n = 1$, that is, the monopoly case:

$$x = \frac{\bar{p} - c}{2\beta} \tag{9.6}$$

The top panel of Figure 9.2 depicts a monopoly firm, Firm A, choosing to produce output x^{Am} which it then sells at the price p^m given by the inverse demand function. The superscript m is for monopoly. The firm's total revenue is $x^{Am}p^m$ and its total profits are $x^{Am}(p^m - c)$, the green-shaded area. The area of the triangle $p^m h \bar{p}$ is the amount of consumer surplus that this firm's sales generate by selling x^{Am} and the price p^m.

REMINDER Consumer surplus is a measure of consumer welfare that is the difference between their willingness-to-pay and what they actually pay for each unit of the good that they consume, generally summed over all purchasers of the good.

9.3 Duopoly: Two firms' best responses and the Nash equilibrium

If there are two firms, they compete for a share of the market demand for the identical product they are selling. This is called a **duopoly**. The strategic analysis of a duopoly and competition among more than two firms is based on

DUOPOLY When there are two firms selling the same output, we call the industry a *duopoly* and we call each firm a *duopolist*.

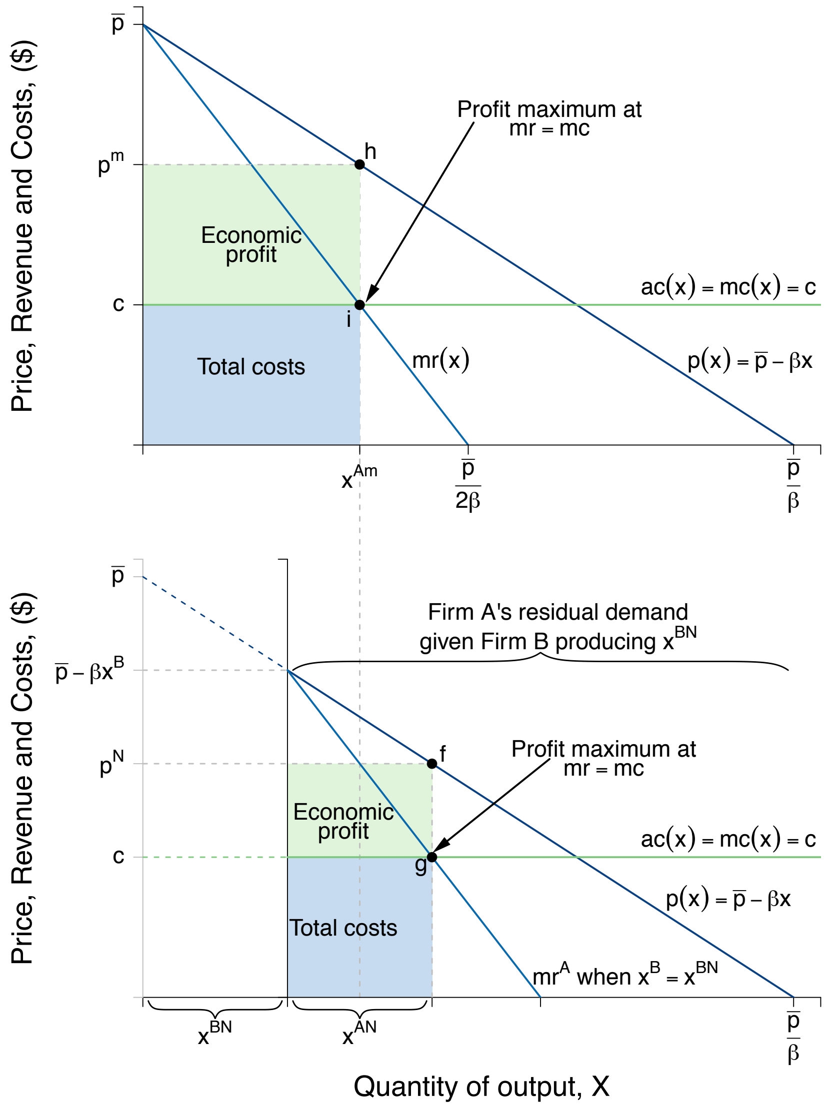


Figure 9.2: **Comparison of monopoly and duopoly.** The top panel shows a single firm, A, choosing the profit maximizing output level. In the bottom panel a second Firm B has entered the market with sales of X^{BN}, so Firm A's residual demand curve is now given by the maximum price at which amounts greater that X^{BN} can be sold.

each firm's best response (meaning profit maximizing level of output) to each of the other firms' choices of outputs. We derive each firm's best-response function by first considering the profit maximizing decision of a single firm and how this depends on the output level of the other firms.

Duopoly profit, revenue, and costs

The industry or market output (X) is the sum of the firms' outputs. The market price will be affected by both firms' outputs. We assume that firms sell everything they produce, so sales and production are the same quantity, namely x. Therefore when there are two firms, A and B, the market output is as follows:

$$\text{Duopoly market output} \qquad X = x^A + x^B \tag{9.7}$$

We show how the duopoly case differs from monopoly in the lower panel of Figure 9.2. If the new firm, B, produced nothing, Firm A would be a monopoly, as before. So, consider the case of interest when Firm B produces and will sell some amount X^{BN}. These are sales that Firm A will not be able to make: customers who have been removed from the market. The result is that Firm A now faces a reduced demand curve, shown in Figure 9.2 as a right-ward shift of the vertical axis by the amount of sales that will be implemented by Firm B.

Firm A then faces the problem depicted in the lower panel. The solid portion of the demand curve is the portion of the market "left over" for Firm A. This is called the **residual demand curve** expressed as $x^A(p) = X(p) - x^B$

RESIDUAL DEMAND CURVE The residual demand curve for Firm A if there is one other firm in the market (Firm B) is $x^A(p) = X(p) - x^B$.

The profit maximum for the firm will be to produce and sell an output such that the marginal cost equals the marginal revenue (now based on the residual demand curve), or x^{AN}, and then to sell this amount at the price p^N.

The consumers are identical and the firm's products are also identical so this price is also the price at which Firm B sells its output. We can represent the duopoly case mathematically, starting with each firm's revenue and profit function. :

$$\begin{aligned}
\text{A's Revenue} \quad r(x^A, x^B) &= p(X)x^A \\
\text{A's Costs} \quad c(x^A) &= cx^A \\
\text{A's Profit} \quad \pi^A &= r(x^A, x^B) - c(x^A) && (9.8) \\
&= p(X)x^A - cx^A && (9.9)
\end{aligned}$$

Because the firms are identical, Firm B's revenue, costs and profit would be *mirror images* of these.

MATH NOTE When we say that the functions are *mirror images* – sometimes also called symmetrical – it means that if we were to repeat the process for B, each time you encounter x^A, substitute in x^B, each time you encounter x^B, substitute in x^A. For example, B's profit function is $\pi^B = p(X)x^B - cx^B$.

The market demand curve is given by the same function we used for the monopoly scenario.

$$\begin{aligned}
\text{Market Demand} \quad p(X) &= \bar{p} - \beta X \\
&= \bar{p} - \beta(x^A + x^B) && (9.10)
\end{aligned}$$

We can substitute equation 9.10 into equation 9.9 to obtain Firm A's economic profit:

Firm B's economic profit would similarly be:

$$\pi^B = (\bar{p} - \beta x^A - \beta x^B)x^B - cx^B \tag{9.11}$$

Because each firm's profit depends on the output of the other firm, we can see that the profits the firms make are *interdependent*. In effect, B's sales shift the demand curve of Firm A to the left, meaning that for any given price the amount of that can sell is less, the greater are the sales of Firm B. This is a type of *negative external effect* of one firm's production on the other firm's demand curve and profitability. As a result, like the Fishermen's Dilemma in Chapter 5, the two firms are engaged in a *strategic social interaction*.

Isoprofit curves and best responses

Acting independently, each firm needs to choose a strategy, the quantity of output that it will produce, as a *best response* to the strategy – the quantity of output – of the other firm. Each firm will have a *best-response function*, its profit-maximizing output for each potential level of output of the other firm.

We know from the analysis of Figure 9.2 that for any given output and sales of the other firm, each firm will produce the amount that equates its marginal revenue and its marginal cost. For the duopoly case M-Note 9.1 shows that this gives us, for Firm A, the following rule to follow:

$$\text{A's best response} \qquad x^A(x^B) = \frac{\bar{p}-c}{2\beta} - \frac{1}{2}x^B \qquad (9.12)$$

Comparing Equation 9.15 to the profit-maximizing rule for the monopoly in Equation 9.1 we see that Firm A's best-response function says the following:

- *First term*: produce what you would have produced had you been a monopoly, $\frac{\bar{p}-c}{2\beta}$ $minus$ the
- *Second term*: one half what the other firm produces, $\frac{1}{2}x^B$.

ISOPROFIT CURVE An isoprofit curve for some given level of profit is composed of those combinations of outputs (x^A, x^B) that result in the given level of profit.

To understand the strategic relationship between the two duopolists we introduce the **isoprofit curve** that shows those combinations of outputs (x^A, x^B) that result in the given level of profit. Figure 9.3 presents three isoprofit curves for Firm A, each corresponds to a different level of profit: π_1^A, π_2^A and π_3^A where $\pi_3^A > \pi_2^A > \pi_1^A$. The level Firm A's profits is greater for isoprofit curves closer to the horizontal axis. The reason is that the less Firm B produces (that is the closer to the horizontal axis) the higher will be Firm A's profit for any level of output that it chooses.

REMINDER Remember that when Aram and Bina engaged in the Fishermen's Dilemma, the same was true of their indifference curves in (h^A, h^B) coordinates: they had higher utility the closer their indifference curve was to their own effort's axis. We also used isoprofit curves in Chapter 8, but in that chapter the isoprofit curve was determined by the price of the good and the quantity of output rather than by the two firms' quantities.

M-Note 9.1: Best responses in a duopoly

Firm A's profits are:

$$\pi^A = (\bar{p} - \beta x^A - \beta x^B)x^A - cx^A$$

Given the output of Firm B, we find the output of Firm A that maximizes A's profits by differentiating the above equation with respect to x^A and setting the result equal to zero.

$$\partial\pi^A/\partial x^A = (\bar{p} - \beta x^B) - 2\beta x^A - c = 0 \quad (9.13)$$

Equation 9.13 requires that Firm A maximizes profit by choosing the output level that makes marginal revenue equal to marginal cost:

$$mr^A(x^A, x^B) = (\bar{p} - \beta x^B) - 2\beta x^A \quad = \quad c = mc(x^A) \quad (9.14)$$

Rearranging Equation 9.14 to isolate x^A, we find:

$$x^A(x^B) \quad = \quad \frac{\bar{p} - c - \beta x^B}{2\beta}$$

we get A's best response $$x^A(x^B) \quad = \quad \frac{\bar{p} - c}{2\beta} - \frac{1}{2}x^B \quad (9.15)$$

We can repeat the derivation for Firm B, to find Firm B's best-response function:

$$x^B(x^A) \quad = \quad \frac{\bar{p} - c}{2\beta} - \frac{1}{2}x^A \quad (9.16)$$

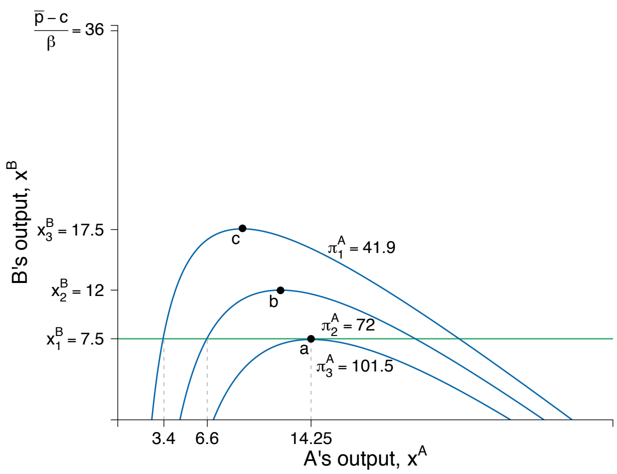
The two best-response functions show that as Firm A produces more, Firm B produces less, and vice versa.

The family of Firm A's isoprofit curves is constructed exactly as were the indifference curves of the two fishers in Chapter 5. To see how this is done, suppose Firm B is producing and selling 7.5 units, and Firm A considers producing some small amount placing in on the upper most isoprofit curve π_1^A in Figure 9.3 a (which is the lowest level of profits shown). The owners of Firm A could do better if they proceeded to the right (producing more) until they encountered the middle isoprofit curve π_2^A in Figure 9.3 a. But they could do still better if they increased output to 14.25.

Producing more than 14.25 would bring them down to the middle isoprofit curve with lower profits. So, $x^A = 14.25$ is Firm A's best response to Firm B producing 7.5.

We can generalize this example. A horizontal line from Firm B's axis representing a given level of B's output is the feasible frontier for Firm A's constrained profit maximization, and Firm A's best response is the level of output where the isoprofit curve is tangent to the horizontal line. We show three dashed lines in Figure 9.3 panel b corresponding to three levels of output by Firm B: x_1^B, x_2^B and x_3^B. For each of these *given* levels of output by Firm B, Firm A's best response is on the highest isoprofit curve that is *tangent* to that output. A maximizes profit at the output level where Firm A's isoprofit curve is tangent to the horizontal line representing Firm B's output.

Therefore Firm A's best-response function is made up of points on the isoprofit curves like points **a**, **b**, and **c** in the figure where the isoprofit curve is *horizontal*, because the best-response function shows A's *profit-maximizing* output at a *given* level of Firm B's output. Firm A's best-response function is made up of all points at which A's isoprofit curves are horizontal. We could draw an equivalent figure for Firm B where each of B's isoprofit curves are tangent to a vertical line representing a given level of output by Firm A.



(a) The feasible set and choosing a best-response isoprofit

(b) The duopolist's best-response function

Figure 9.3: **Isoprofit and best-response curves for one duopolist.** In Panel a. three of Firm A's isoprofit curves are shown along with a horizontal green line indicating some hypothetical level of output that Firm B might choose, namely 7.5 units of output. Firm A's best response to $x^B = 7.5$ is the point on that horizontal line at which Firm A's profits are greatest, i.e. point **a** where Firm A produces 14.25 units. So, $x^A = 14.25$ by Firm A is a best response to $x^B = 7.5$ by Firm B. Panel b. shows the derivation of A's best-response function plotted in dark green Firm A's isoprofit curves are the same as in Panel a. Points b and c in Panel b on A's best-response function are derived in the same way that we derived point a in Panel a. Notice that these three points are all at the horizontal point on A's isoprofit curves. The numerical values used to produce this figure are: $\bar{p} = 20, \beta = 0.5, c = 2$. A's best-response function is given by the Equation 9.15 in M-Note 9.1.

M-Note 9.2: The slope of an isoprofit curve

Here we show why the best response of Firm A to any level of output of Firm B is where Firm A's isoprofit curve is horizontal.

The equation for a particular isoprofit curve of Firm A – one with $\pi^A = k$ where k is some constant – has the form:

$$\pi^A(x^A, x^B) = (\bar{p} - \beta(x^A + x^B))x^A - cx^A = k$$

The isoprofit curve for this particular level of profit (k) is made up of points with differing levels of x^A and x^B but with the same level of profit, namely $\pi^A = k$. These points satisfy the following equation:

$$d\pi^A = dx^A \cdot \frac{\partial \pi^A}{\partial x^A} + dx^B \cdot \frac{\partial \pi^A}{\partial x^B} = 0 \tag{9.17}$$

Which we can rearrange as follows:

$$dx^B \cdot \frac{\partial \pi^A}{\partial x^B} = -dx^A \cdot \frac{\partial \pi^A}{\partial x^A} \tag{9.18}$$

To find the slope of the isoprofit we need to find $\frac{dx^B}{dx^A}$:

$$slope = \frac{dx^B}{dx^A} = -\frac{\partial \pi^A / \partial x^A}{\partial \pi^A / \partial x^B} \tag{9.19}$$

The denominator is always negative because the more one firm sells, the lower will be the residual demand and hence the profits of the other firm.

Where

$$\partial \pi^A / \partial x^A = 0 \tag{9.20}$$

the slope is zero so the isoprofit is flat. This point is also the best response of Firm A to the level of output of Firm B, because Equation 9.20 is the condition defining A's profit maximizing output, as you can see from M-Note 9.1.

Checkpoint 9.2: Re-interpreted BRFs

- Redraw Figure 9.3, but instead draw Firm B's best-response function and isoprofit curves.
- Use the reasoning in M-Note 9.1 to determine the slope of B's isoprofit curves.

Nash equilibrium in Cournot Duopoly

In Figure 9.4 the duopolists' best responses are plotted together. Because the firms are identical, the best response functions are symmetrical. So, for example the vertical axis intercept of A's best response function (36) is the same as the horizontal axis intercept of B's best response function. The symmetrical best response functions together with the fact that neither firm has any particular bargaining advantage – such as being first mover or having the power to make a take it or leave it offer – also means that at the Nash equilibrium, the firms will produce the same output. So the Nash equilibrium will lie on the 45-degree line in the figure where the firms'outputs are equal.

Each best-response function is negatively sloped. The Nash equilibrium occurs at the intersection of the firms' best-response functions: where each firm best responds to the strategy of the other player. At the equilibrium, the firms produce x^{AN} and x^{BN} and both sell their output at the following price:

$$p^N = \bar{p} - \beta(x^{AN} + x^{BN}) \tag{9.21}$$

HISTORY Seeing that Antoine Cournot developed this model almost a century before John Nash (1928-2015) was born, it may seem strange that we do not refer to the Cournot equilibrium, and economists sometimes do. But to limit confusion we will stick to the more familiar term *Nash* equilibrium. The best-response functions that we use here (and in the case of the interacting fishermen earlier) are also Cournot's idea.

Are we certain that this is a Nash equilibrium? We constructed it as a mutual best response, so it should be. But in our construction we assumed that each firm took the other's output and sales as given, their customers already 'extracted' from the market, like fish taken from the lake by another fisherman and no longer "available" to be caught.

Have we overlooked any opportunity for profit that the actors might adopt? The owner of one of them, Firm A for example, might reason that seeing that the products of the two firms are identical, he could capture the entire market just by offering a price somewhat lower than p^N.

But suppose he tried this. Remember the other firm has already produced X^{BN} and is going to sell that amount at the highest possible price (whatever that price is). So Firm B would match or beat any price Firm A selected. So the price-cutting strategy would be self-defeating. This confirms that the

Figure 9.4: **Best-response functions for the two duopolists.** Because the Nash equilibrium is a mutual best response it must be a point on both best-response functions. There is only one such point, namely, the intersection. The numerical values used to produce this figure are: $\bar{p} = 20, \beta = 0.5, c = 2$. The best-response functions are given by Equations 9.15 and 9.16 in M-Note 9.1

intersection of the two best-response functions is indeed a Nash equilibrium.

How does the duopoly outcome contrast with monopoly case?

- The presence of the second firm (B) dilutes the monopoly power of the first firm (A)
- We say that B *crowds* the market, leading to a larger total output than in the monopoly case;
- Therefore, there is a lower market price closer to marginal cost; and
- There is lower total economic profit of the two firms compared to the single monopolist.

M-Note 9.3: Nash equilibrium output with two firms

The output levels given by the intersection of the duopolists' best-response curves is a Nash equilibrium. To find these equilibrium outputs, we substitute B's best-response function (Equation 9.16) into A's best-response function (equation 9.15):

$$\begin{aligned}
& x^A &=& \frac{\bar{p}-c}{2\beta} - \frac{1}{2}\left(\frac{\bar{p}-c}{2\beta} - \frac{1}{2}x^A\right) \\
& x^A &=& \frac{\bar{p}-c}{2\beta} - \frac{\bar{p}-c}{4\beta} + \frac{1}{4}x^A \\
\text{subtract } \frac{1}{4}x^A \quad & \frac{3}{4}x^A &=& \frac{\bar{p}-c}{4\beta} \\
\text{multiply through by } \frac{4}{3} \quad & x^A &=& \frac{4}{3}\left(\frac{\bar{p}-c}{4\beta}\right) \\
\text{Cournot (Nash) equilibrium output} \quad & \therefore x^{AN} &=& \frac{\bar{p}-c}{3\beta} = x^{BN}
\end{aligned} \tag{9.22}$$

If you derive B's best-response function as an exercise, you will see that B's (Cournot) Nash equilibrium output is mirror image of A's as shown in Equation 9.22.

Below we provide an alternative way to solve the Nash equilibrium output in this symmetric setting. At the symmetric Nash equilibrium, both firms will choose the same output $x^{AN} = x^{BN} = x^N$. Therefore, x^N should best reply to itself, i.e.,

$$\begin{aligned}
x^N &= \frac{\bar{p}-c}{2\beta} - \frac{1}{2}x^N \\
\frac{3}{2}x^N &= \frac{\bar{p}-c}{2\beta} \\
x^N &= \frac{\bar{p}-c}{3}
\end{aligned}$$

Graphically, the symmetric Nash equilibrium is the intersection of the best response function and the 45-degree line $x^B = x^A$, as shown in Figure 9.4.

M-Note 9.4: Prices, total output and profits with two firms

To find the prices and profits associated with the Nash equilibrium levels of output, we sum x^{AN} and x^{BN} to get total market output use the inverse demand function to find the equilibrium price. From the equilibrium price, we can find each firm's profit and sum them to find the total economic profit in the duopoly. Remember, the superscript N denotes a Nash equilibrium.

$$
\begin{array}{rrcl}
\text{Total output} & X^N := x^{AN} + x^{BN} & = & \dfrac{2(\bar{p}-c)}{3\beta} \\
\text{Using the inverse demand function} & p^N & = & \bar{p} - \beta\left(x^{AN} + x^{BN}\right) \\
 & & = & \bar{p} - \beta\left(\dfrac{2}{3}\dfrac{\bar{p}-c}{\beta}\right) \\
 & & = & \dfrac{1}{3}\bar{p} + \dfrac{2}{3}c \\
\text{Add and subtract } \dfrac{1}{3}c & \therefore p(X) & = & c + \dfrac{1}{3}(\bar{p}-c) \\
\text{Firm profits} & \pi^{AN} & = & \dfrac{(\bar{p}-c)^2}{9\beta} = \pi^{BN} \\
\text{Total profits} & \Pi^N & = & (p^N - c)(x^{AN} + x^{BN}) = \dfrac{2(\bar{p}-c)^2}{9\beta}
\end{array}
$$

Relative to the monopoly outcome, market output is higher, market price is lower, and total profits are lower in the duopoly.

Checkpoint 9.3: Willingness to pay, slope of the demand and duopoly output

Let us assume new values for the parameters in which we are interested: let $\bar{p} = 100$ and $\beta = 1$ and the firm has constant marginal costs $c = 1$.

a. Find each firm's formula for its best-response function.
b. Find the Nash equilibrium quantity for each firm.
c. How much profit does each firm make? What is the price and how much greater is the price than marginal costs (that is, what is the markup)?

9.4 Oligopoly and "unlimited competition": From a few firms to many firms

A feature of Cournot's approach is that it allows us to use a single model to study the entire range of competition from monopoly to an industry with very many firms. If more identical firms enter the market, they all have to share the demand reducing the residual demand curve of each firm. We include a summary of the results that we derive in Table 9.10.

Oligopoly

We can represent process of competition and its outcome graphically for *n* firms, taking account of two characteristics of the Nash equilibrium:

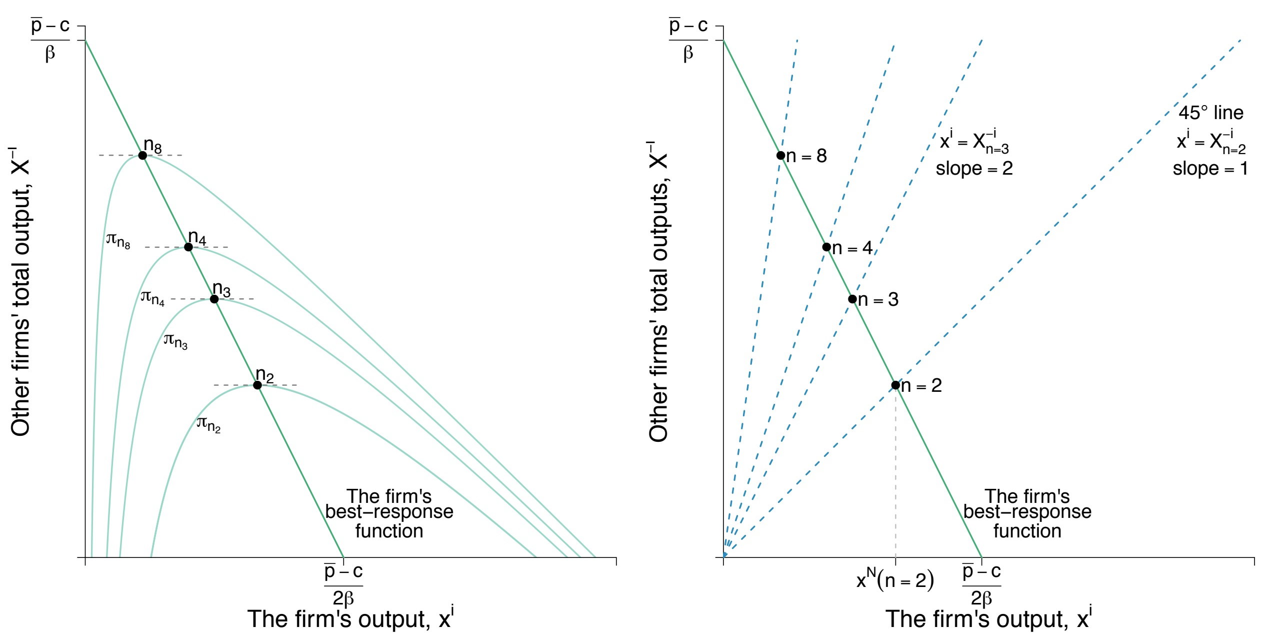
- Because the firms are identical, in equilibrium they produce the same level of output.

- For each firm, the Nash equilibrium output must be a best response to the total output being produced by the $n-1$ *other* firms.

REMINDER M-Note 9.1 and Figure 9.3, explained why the best response of Firm A is at a point where A's isoprofit curves are horizontal.

To see how this works we arbitrarily pick one firm ("the firm") and study its choice of an output level given what all of the other firms are producing. It does not matter which firm we pick as because they are identical. If "the firm" best responds to the outputs chosen by the other firms and those firms (like "the firm") also best respond to what the other firms do, then the result is a Nash equilibrium. We can visualize the equilibrium by plotting the *total output* of the other firms, X^{-i}, on the vertical axis and the best response of "the firm", x^i, on the horizontal axis, as we do in Figure 9.5.

M CHECK The superscript $-i$ means "not Firm i" and uppercase letters refer to totals, so X^{-i} is the total output of the n-1 firms other than firm i.



(a) The firm's isoprofits along their best-response function

(b) The firm's Nash equilibrium output for different outputs of the other firms

Figure 9.5: **Nash equilibrium output with *n* firms** In both panels the arbitrarily selected firm's output is on the horizontal axis, and the total output of the remaining $n-1$ firms is on the vertical axis. Panel a shows the construction of the best response function of "the firm". Each of the other firms has an identical best response function. In Panel b, each dashed ray from the origin has a slope of $n-1$ and shows for different values of n, the total output levels of the other firms that are n-1 times the output level of "the firm". We know that this condition must be true in equilibrium because the firms are identical.

In Figure 9.5, the firm's best-response function is plotted against the total output of the other firms in the industry. The best response shows each point at which the firm's isoprofit curve is tangent to a horizontal line (shown by the dashed gray lines) indicating some hypothetical given output of the other firms, such as at points $\mathbf{n}_2$ and $\mathbf{n}_3$ (each of which have corresponding outputs for the firm).

The best-response function is made up of points where the the isoprofit curves are *horizontal*.

In the Nash equilibrium, two conditions must be met:

- the output of the firm must be on its best-response function; and
- the output of each of the $n-1$ other firms must be equal to the best-response output of "the firm", so the total output of the other firms has to satisfy $X^{-i} = (n-1)x^i$.

The second condition means, as you have already seen, that the if there is just one other firm then equilibrium lies on the 45-degree ray from the origin, which is $X^{-i} = x^i$ which has a slope of 1, namely the number of other firms.

As a result the Nash equilibrium is the point on both the firm's best-response function and the line $X^{-i} = (n-1)x^i$, namely their intersection at each of the Nash equilibrium points: $\mathbf{n}_2$ for $n = 2$ firms, $\mathbf{n}_3$ for $n = 3$ firms, and so on.

We show in M-Note 9.5 that the firm's best response when it faces $n-1$ competitors is as follows:

$$x^i(X^{-i}) = \frac{\bar{p}-c}{2\beta} - \frac{1}{2}X^{-i} \tag{9.23}$$

Equation 9.23 is the monopoly firm's level of output ($\frac{\bar{p}-c}{2\beta}$) minus $\frac{1}{2}X^{-i}$, which is the same that we found (in Equation 9.15) for Firm A in the duopoly case where $n = 2$ and $X^{-i} = x^B$. And it replicates the output choice for the monopoly (given in equation 9.6), too, because in that case, $n = 1$ so the second term drops out (the output of the other firm is zero because there is no other firm).

Figure 9.5 shows that though with 8 firms in the industry each of them is producing less (6 rather than 9), the total output is greater ($8 \times 6 = 48$ rather than $3 \times 9 = 27$.) Because the inverse demand curve is downward-sloping, the market price will therefore be lower with many firms competing.

M-Note 9.5: Cournot competition with many firms: Best responses

To study the case with many firms we focus on a single firm i (called "the firm") whose output is x^i and all of the $n - 1$ other firms taken together, whose total output is X^{-i}. So total production and sales is $X = x^i + X^{-i}$ and the demand curve seen by the firm when there are $n-1$ other firms is:

$$\text{Demand faced by the firm} \quad p(x^i, X^{-i}) = \bar{p} - \beta(x^i + X^{-i}) \tag{9.24}$$

From this we have we have the firm's revenue and profit:

$$\text{The firm's revenue} \quad r(x^i, X^{-i}) = p(x^i, X^{-i})x = (\bar{p} - \beta X^{-i})x^i - \beta(x^i)^2 \tag{9.25}$$

$$\text{The firm's profit} \quad \pi(x^i, X^{-i}) = (\bar{p} - \beta(n-1)X^{-i})x^i - \beta(x^i)^2 \tag{9.26}$$

Therefore, given X^{-i}, partially differentiating the total revenue function with respect to x^i, we have:

$$\text{The firm's marginal revenue} \quad mr(x^i, X^{-i}) = (\bar{p} - \beta X^{-i}) - 2\beta x^i \tag{9.27}$$

The firm maximizes profit by choosing the output level that equates marginal revenue to marginal cost:

$$mr(x^i, X^{-i}) = (\bar{p} - \beta X^{-i}) - 2\beta x^i = c$$

solving this equation for x^i

$$x^i(X^{-i}) = \frac{\bar{p} - c - \beta X^{-i}}{2\beta} \tag{9.28}$$

(9.29)

Re-arranging Equation **??** we can find the firm's best-response function:

The firm's best response

$$x^i(X^{-i}) = \frac{\bar{p} - c}{2\beta} - \frac{1}{2}X^{-i} \tag{9.30}$$

M-Note 9.6: Nash equilibrium output with many firms

The intersection of the firm's best-response curve with the line $X^{-i} = (n-1)x^i$ is the Nash equilibrium, which we find by substituting it into the firm's best-response function, Equation 9.30:

$$\begin{aligned} x^i &= \frac{\bar{p} - c}{2\beta} - \frac{1}{2}(n-1)x^i \\ x^i + \frac{(n-1)}{2}x^i &= \frac{\bar{p} - c}{2\beta} \\ \frac{2x^i + (n-1)x^i}{2} &= \frac{\bar{p} - c}{2\beta} \\ \frac{(n+1)x^i}{2} &= \frac{\bar{p} - c}{2\beta} \end{aligned}$$

$\therefore$ multiplying by $\frac{2}{(n+1)}$

$$x^N(n) = \frac{\bar{p} - c}{(n+1)\beta} \tag{9.31}$$

M-Note 9.7: Equilibrium price n firms

We can substitute the equilibrium outputs from Equation 9.31 into the inverse demand curve to find the price and each firm's profit:

$$\begin{aligned} p^N &= \bar{p} - \beta n x^N \\ &= \bar{p} - \beta n \left(\frac{\bar{p} - c}{\beta(n+1)} \right) \qquad (9.32) \\ &= \frac{1}{(n+1)}\bar{p} + \frac{n}{(n+1)}c \end{aligned}$$

Add and subtract $\frac{1}{(n+1)}c$

$$\therefore p^N(n) = c + \frac{1}{(n+1)}(\bar{p} - c) \tag{9.33}$$

As equation 9.33 shows, the price in Cournot competition is equal to the marginal cost, c, plus $\frac{1}{n+1}$ times the difference between the maximum price, $\bar{p}$, and marginal cost.

Notice that for any given number of firms in the market, the price does not depend on β, the (negative of the) slope of the inverse demand curve. With the demand curve we are using a steeper (more inelastic) demand curve (which would induce firms to raise prices) is also associated with a reduction in demand (which has the opposite effect). This is why the two βs cancel out in Equation 9.32.

Unlimited competition

Additional competing firms dilute the market power of the incumbent firms (those already in the industry). Adding firms to the markets leads to each firm producing less but with a larger total output. As a result of the greater output the market price decreases and becomes closer to marginal cost and the firms make lower total profits. Like Cournot, we call this very large n case "unlimited competition."

M-Note 9.7 shows that the price at which each of the firms sell their product is:

$$p^N = c + \frac{1}{n+1}(\bar{p} - c) \tag{9.34}$$

Equation 9.34 says that as n becomes very large the following happens:

- The second term in Equation 9.34, namely $p^N - c$, approaches zero.
- The price therefore falls to just above the cost per unit (c), and
- the firms' economic profits all but disappear because profit per unit $(p - c)$ approaches zero.

Checkpoint 9.4: Higher willingness to pay and competition

Let us assume the following values: $\bar{p} = 100$, $\beta = 1$ and the firm has constant marginal costs $c = 1$. Let $n = 43$ firms.

a. How much output will each firm produce with the parameters as described above?
b. What will the market price be, how much profit will each firm make, and what is the markup over marginal costs?

9.5 The extent of competition and the markup over costs

We can now show that the markup ratio introduced in Chapter 7 – profits divided by costs – falls as the number of firms – that is the extent of competition – increases. As shown in Equation 9.34, the *price* (revenue per unit) in the Nash equilibrium with *n* firms is $p^N(X^N(n))$.

In Figure 9.6 we plot the resulting $p^N(n)$ along with the cost c per unit which does not vary with the number of firms. The difference between the two is the markup.

REMINDER The markup ratio is: $\mu(n) = \frac{p(n)-c}{c}$.

M-Note 9.8: The markup ratio, $\mu(n)$ and the degree of competition

We know from M-Note 9.7 that in the Cournot model that the price with n firms is given by Equation 9.33, which is:

$$p^N(n) = c + \frac{1}{n+1}(\bar{p} - c) \tag{9.35}$$

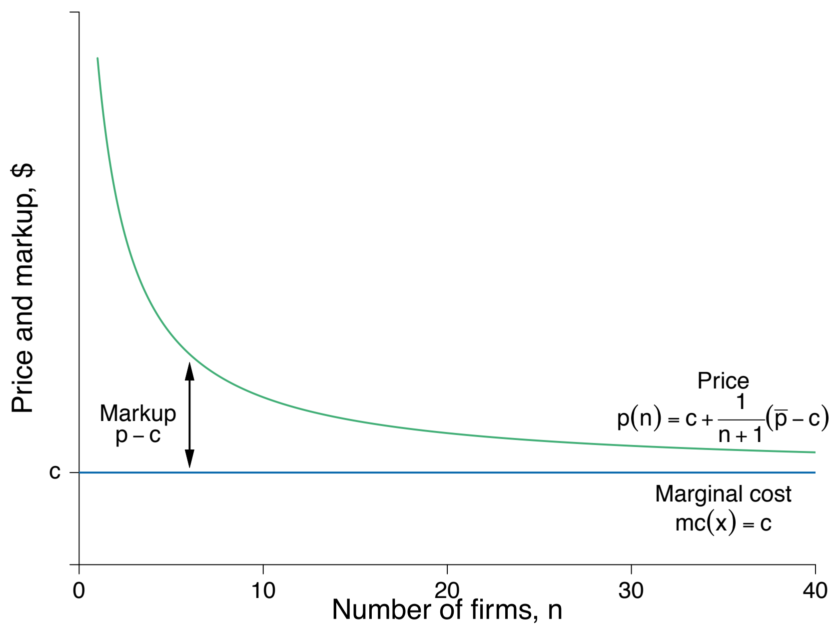


Figure 9.6: **Price and markup over costs** In M-Note 9.7 we showed that the Nash equilibrium price with n firms is given by Equation 9.33, which is the green curve in the figure. It slopes downward because while each firm produces less the more firms are in the industry, the total effect of more firms is to increase industry output. The numerical values used to produce this figure are: $\bar{p} = 20, \beta = 0.5, c = 2$.

Let us re-arrange the equation and find the ratio :

$$\text{Subtract } c: \quad p^N - c = \frac{1}{n+1}(\bar{p} - c)$$

$$\text{Divide by } c: \quad \frac{p^N - c}{c} = \frac{1}{n+1}\left(\frac{\bar{p} - c}{c}\right) = \mu(n) \tag{9.36}$$

Equation 9.36 shows that the markup ratio for an industry with a degree of competition given by n firms is the maximum possible markup ratio times the inverse of $n + 1$. An increase in competition (increase in n will reduce the markup as can be seen from equation 9.36.

The results with large n are similar to what is called the "perfect competition model": price is (approximately) equal to marginal costs, and economic profits are (approximately) zero. Moreover the model gives us reason to expect n to be large. Equation 9.34 means that in the Nash equilibrium the price will always exceed marginal (and average) costs even if there is a very large number of firms. This being the case positive economic profits will persist, providing an incentive for more firms to enter, increasing the level of competition, but never eliminating the difference between price and cost.

The Cournot model includes results approximating "perfect competition" as a limiting case. We will also see that it also can incorporate the effects impediments to firms entering the market, so that "perfect competition" is never realized.

At end of Chapter 8 we asked: why does the size of firms not grow forever, eventually eliminating competition? Here we ask the opposite question: Why do firms not continue to enter the market until there are so many firm competing that it approximates "perfect competition?"

Checkpoint 9.5: The markup ratio in monopoly, duopoly and oligopoly

Assume the parameters of $\bar{p} = 20$, a slope of $\beta = 0.5$, and a marginal cost, $c = 2$.

a. Calculate the markup ratio in each of the following situations: monopoly, duopoly, and oligopoly with $n = 2$ and $n = 43$ respectively.

b. Draw a figure with the number of firms (n) on the horizontal axis and them markup ratio on the vertical axis. Label four points corresponding to your answers to a. and sketch the line connecting these points.

9.6 Barriers to entry and the equilibrium number of firms

We have so far illustrated our three cases by assuming particular values of n the number of firms: $n = 1, n = 2$, and $n > 2$ (either few, or many). We now consider what determines the number of firms, meaning: what determines the extent of competition.

To understand the factors affecting the extent of competition, we ask: what determines the *equilibrium* number of firms, that is, the number of firms that does not change over time? We asked a similar question in Chapter 5 when we studied the equilibrium number of people fishing on a lake. This was the number of people fishing such that the utility from fishing is equal to the utility of the fallback option that people would experience if they did not fish.

The equilibrium number of firms in an industry is determined in the same way. For a firm to attempt entry to the market it must be the case that the expected profits to be had by attempting to enter the market are greater than or equal to the opportunity cost of entering fallback option of the potential entrant. So the Nash equilibrium number of firms in the industry, n^N, will be the largest whole number such that the expected profits exceed the opportunity cost of attempting to enter.

We will see that the equilibrium number of firms will be smaller:

- The more the equilibrium price falls as the number of firms in the market increases;
- The more profitable are the alternative uses of the funds that a firm's owners might commit to entering the market (the opportunity cost of capital).
- The extent of barriers to entry of new firms.

REMINDER The *opportunity cost of capital* is the rate of return on an *alternative* use of a firm owner's funds where the risk is negligible as is the case for U.S. Treasury bills or other government bonds. These investments pay a given amount of money at some future date unless the government that issued them defaults (which for many governments is very unlikely). We call this "riskless" rate of return ρ, which therefore is the opportunity cost of capital.

REMINDER In Chapter 5 the Nash equilibrium number of people fishing on the lake is the largest whole number n^N such that the utility of those fishing $u(n^N)$ is not less than the utility that they would receive at their fallback option, or $u(n^N) \geq u^Z$.

M-CHECK The equilibrium number of firms must be a whole number because the entry of a "fractional firm" would be meaningless, just as with the case of the equilibrium number of people fishing on the lake it did not make sense to talk about a "fractional fisherman" taking up fishing.

REMINDER The term barriers to entry refers to anything making it difficult for new firms to enter a market, including intellectual property rights that give incumbent firms a monopoly on particular technologies or, economies of scale in production or in demand, and predatory pricing.

The decision to enter an industry and barriers to entry

We model barriers to entries as a probability that a firm attempting to enter will fail. The greater are the barriers to entry the higher is the probability of failure for a firm attempting to enter. The new firm attempting to enter a market is identical to the *incumbent* firms, so if it succeeds in entering, it becomes just one of $n+1$ identical firms. To determine the conditions under which a new firm might enter, we have to return to the "owner's eye view of costs" introduced in the Chapter 8. In deciding whether to enter a market, the owners of the firm consider the value of the capital goods they will devote to that project. They then compare the profits they expect to make if they attempt to enter the industry with the profits they could make if they devoted those same funds to an alternative use.

EXAMPLE Consider a person choosing whether to purchase capital goods worth \$1000 in starting a firm, where the risk-free annual rate of return, ρ = 2% which is the rate of returns he could have by purchasing, instead, government bonds. Then the opportunity cost of that owner's capital (what they would earn if they did not purchase the capital goods for the start up firm) is \$20 (\$1000 $\times$ 1.02 - \$1000) per year. Economic profits are accounting profits in excess of this opportunity cost. For more information, review these ideas in Chapter 8.

REMINDER An *incumbent firm* is a firm that is already producing in the industry.

EXPECTED PROFIT The expected profits are the profits if successful multiplied by the probability of success plus the profits if unsuccessful (that, is, the losses) multiplied by the probability of failure .

But entering a new industry is *risky*. Because the owners commit some funds to the project (entering the industry) but they do not know what the outcome of their attempt will be. To consider just the extremes it could be either:

- *Failure*: There is a probability (b) that attempting to enter will fail so that resulting in no revenues to offset the costs of attempted entry.
- *Success*: Alternatively, there's a probability $(1-b)$ that the firm succeeds in entering the market and be able to sell its products at the same price as the other firms, and thereby offsetting the costs of the initial investment.

M-CHECK The expected price is just the weighted average of the price if entry fails ($p=0$) and the price if entry succeeds ($p(n+1)$) with the weights being the probability of these two events occurring, that is b and $1-b$ respectively.

The key idea for modeling the risk faced by an entering firm is that it incurs its costs with certainty regardless of whether it succeeds or fails, but earns profits only if it succeeds. The entering firm intending to produce an amount x with certainty pays the cost cx. The risk arises because with some probability b (for *barriers to entry*) the firm will fail to sell the product.

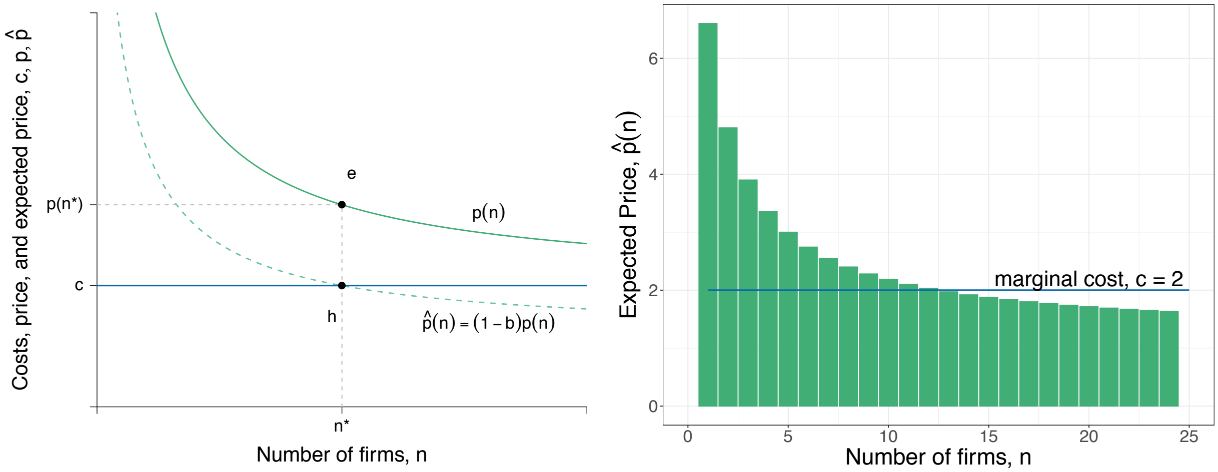
A firm considering entering a market with n firms, if successful, will be in a market with a total of $(n+1)$ firms selling its product at the price that results when there is an additional firm in the market. Therefore it calculates its **expected profits** per unit of output produced as follows:.

$$\text{Expected profit per unit} = \underbrace{-cb}_{\text{Failed entry}} + \underbrace{(1-b)(p(n+1)-c)}_{\text{Successful entry}} \tag{9.37}$$

$$= \underbrace{c}_{\text{Cost}} + \underbrace{(1-b)p(n+1)}_{\text{Expected price}} \tag{9.38}$$

Determinants of the equilibrium number of firms

Figure 9.7 shows how the equilibrium number of firms in the industry is determined. The expected price as a function of the number of firms ($\hat{p}(n)$) slopes downwards because as more firms enter the total amount to be sold increases



(a) Computing the equilibrium number of firms

(b) Equilibrium number of firms (numerical example)

Figure 9.7: **The equilibrium number of firms with barriers to entry.** The number of firms in an industry is plotted on the horizontal axis, and the price and costs of the the firm are plotted on the vertical axis. It is in a firm's interest to attempt to enter as long as $\hat{p} \geq c$. In panel **b** the barriers to entry (b = 0.4) and marginal costs $c = 2$, with corresponding equilibrium number of firms $n = 12$. With $n = 12$, the expected price is 2.03; with $n = 13$ the expected price falls to $\hat{p} = 1.97 < 2 = c =$ marginal cost. So the 13th firm will not enter.

and the price falls (due to the downward-sloping market demand curve). The cost curve (including the opportunity cost of capital) is horizontal and does not depend on the degree of risk in entering, because the firm pays the cost whether or not it succeeds in selling its output.

The rule governing firm entry and exit is the following:

- Entry: If with n firms already in the market, the prospective entering firm's expected price (taking account of the probability b that the price will be zero) exceeds the cost (including the opportunity cost of investing capital in this firm), then the owners will decide to enter.
- Exit: If at the existing n some firm's profits fall short of what they could make in an alternative investment then it will *exit* the market.

M-Note 9.9: Barriers to entry and the equilibrium number of firms

From Equation **??**, we know that potential entrants consider an expected price in the market, $\hat{p}$, when deciding whether to enter. We can now substitute in the value for price as a function of the number of firms in the industry (Equation 9.33) into the condition given by

Equation **??**:

$$
\begin{aligned}
(1-b)p(n) &= c \\
(1-b)\left(c+\frac{1}{n+1}(\bar{p}-c)\right) &= c \\
(1-b)c+(1-b)\left(\frac{1}{n+1}(\bar{p}-c)\right) &= c \\
(1-b)\left(\frac{1}{n+1}(\bar{p}-c)\right) &= bc \\
\frac{1-b}{n+1}(\bar{p}-c) &= bc \\
\text{Multiply by } (n+1) \qquad (1-b)(\bar{p}-c) &= nbc+bc \\
\text{Subtract } bc \qquad nbc &= (1-b)(\bar{p}-c)-bc \\
\text{Simplify RHS} \qquad nbc &= \bar{p}(1-b)-c \\
\text{Divide by } bc \qquad n^N &= \frac{\bar{p}(1-b)-c}{bc} \qquad (9.39)
\end{aligned}
$$

The Nash equilibrium number of firms n^N is therefore determined by:

- consumers' maximum willingness to pay ($\bar{p}$)
- the size of barriers to entry (b)
- the firm's marginal costs (c), which include the opportunity cost of capital (ρ).

Taken together these two rules suggest a level of n at which firms will be neither entering nor leaving, what we call n^N (N for Nash equilibrium). This is the equilibrium number of firms.

In Figure 9.7 the price at the Nash equilibrium for each number of firms $p(n)$ shown in green. All firms in the industry will sell at this price. With barriers to entry the probability that a firm attempting entry will fail and not be able to sell its output at all is b. A firm that fails to enter will face a price $p = 0$ (it can't sell its goods). So the expected price of the firm considering entering is the dashed line, $\hat{p} = (1-b)p(n)$. As $\hat{p}$ exceeds the cost as long as there are fewer than n^N firms in the industry, entry of new firms will occur until $n = n^N$.

We can now consider some of the economic factors that might affect the level of competition in an industry by increasing or decreasing the equilibrium number of firms.

- An *increase in barriers to entry* (b): Increasing b will shift down the expected price function as shown in Figure 9.8, resulting in a smaller number of firms, less competition, and a higher price markup over costs.

- An innovation that reduces the cost of production: *Decreasing c shifts down the cost line* (not shown in the figure) increasing the equilibrium number of firms and hence the degree of competition.

- An increase in the opportunity cost of capital: This could occur if profitability in some other economy increased, improving firm owner's next best alternative use of their funds, or if the central bank's monetary policy increased the cost of borrowing. Because an increase in the opportunity cost of capital increases c, the increase will reduce the equilibrium number of firms in the industry and support a less competitive environment.

The figure shows that with barriers to entry, in the Nash equilibrium each of the firms in the industry will be selling goods at a price above the marginal cost. As a result, the owners of each firm would prefer to have more buyers, if they could figure out some method, such as advertising or the other modes of competition, to sell more output without lowering the price.

We bring together all of the results on Cournot competition as n varies from 1 to many in M-Note 9.10.

M-Note 9.10: Summary: Nash equilibrium results for the Cournot model as the number of competing firms vary

You can generate the results for monopoly ($n = 1$), duopoly ($n = 2$), oligopoly ($n > 2$, but not large), and unlimited competition ($n \to \infty$) using the equations below. In the equations below we use the inverse demand function $p = \overline{p} - \beta X$ and the measure of barriers to entry, b. We use the superscript N to denote a Nash equilibrium value, so that, for example $x^N(n)$ "x super N of n" means the Nash equilibrium level of output of a firm when there are n firms in the market.

Firm output (Equation 9.31): $x^N(n) = \frac{1}{n+1}\frac{\overline{p}-c}{\beta}$

Industry output: $X^N(n) = nx = \frac{n}{n+1}\frac{\overline{p}-c}{\beta}$

Market price: $p^N(n) = \overline{p} - \beta X = \overline{p} - \beta\left(\frac{n}{n+1}\frac{\overline{p}-c}{\beta}\right) = c + \frac{1}{n+1}(\overline{p}-c)$

Markup ratio: $\mu^N(n) \equiv \frac{p-c}{c} = \frac{1}{c}\left(c + \frac{1}{n+1}(\overline{p}-c) - c\right) = \frac{\overline{p}-c}{(n+1)c}$

Firm profits: $\pi^N(n) = (p-c)x = \left(\frac{1}{n+1}(\overline{p}-c)\right)\left(\frac{1}{n+1}\frac{\overline{p}-c}{\beta}\right) = \frac{1}{(n+1)^2}\frac{(\overline{p}-c)^2}{\beta}$

Industry profits: $\Pi^N(n) = n\pi = \frac{n}{(n+1)^2}\frac{(\overline{p}-c)^2}{\beta}$

Equilibrium condition: Expected revenue per unit $=$ cost per unit $\hat{p} = (1-b)p = c$

Equilibrium number of firms: $n = \frac{p(1-b)-c}{bc}$

9.7 A conflict of interest: Profits, consumer surplus, and the degree of competition.

Another strategy that owners might pursue is to find ways of increasing *b* – barriers to entry – so as to reduce the number of firms in the industry. You can see from Figure 9.6 that the profits per unit produced – the markup – is greater the fewer firms there are competing.

REMINDER Consumer surplus, a measure of consumer welfare made possible by an exchange, is the difference between each consumer's willingness-to-pay and what they actually pay for each unit of the good that they consume, generally summed over all purchasers of the good.

While barriers to entry benefit the owners of the incumbent firms, they reduce

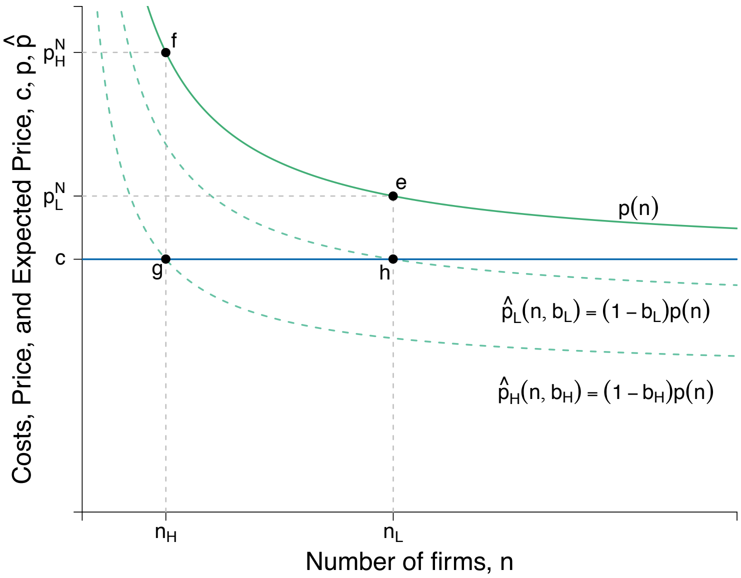


Figure 9.8: **How barriers to entry affect equilibrium number of firms.** The extent of entry barriers affects the equilibrium number of firms. There are two expected price lines: with low barriers to entry (b_L) the expected price line is $\hat{p}(n, b_L)$ and with high barriers to entry (b_H) the expected price line is $\hat{p}(n, b_H)$. Notice that the high barriers to entry line is *lower* than the low barriers to entry line. With *low* barriers to entry, the number of firms is given by the intersection of costs and $\hat{p}(n, b_L)$ (point **h**) with a relatively higher number of firms n_L^N. With *high* barriers to entry, the number of firms is given by the intersection of costs and $\hat{p}(n, b_H)$ (point **g**) with number of firms n_H^N. With greater barriers to entry, there are fewer firms and lower expected profits because of how hard it is for firms to enter the market.

the economic benefits of consumers, measured by consumer surplus. You can see this in Figure 9.9. For the monopoly case where $n = 1$ the red and blue dots show that profits (the blue dot = 162) is twice the red dot (consumer surplus = 81). You have already seen this result in the top panel of Figure 9.2 where the profit rectangle is twice the size of the the consumer surplus triangle.

Adding even just a single competing firm – the duopoly or $n = 2$ case – increases competition and lowers the price sufficiently to bring profits down somewhat and to substantially increase consumer surplus, so that profits and consumer surplus are equal at 144.

Notice something important: with n = 2 profits plus consumer surplus, namely 288 is much greater than with n = 1 where profits plus consumer surplus is 243. This means that the conflict of interest between owners and consumers is not a zero sum game: when the number of firms goes from 1 to 2 the sum of the gains to consumers and losses to owners does not sum to zero, it sums to 45. The same pattern persists as n goes to 3 and higher numbers: the consumers gain more in consumer surplus than the owners lose in reduced profits. Where did the extra benefits come from?

9.8 *Limited competition and inefficiency: Deadweight loss*

The answer is that, in addition to redistributing income towards owners' profits and away from consumers, limited competition is also a source of inefficiency. As the number of firms competing increases, the inefficiency is reduced so the total benefits – profits plus consumer surplus – increases.

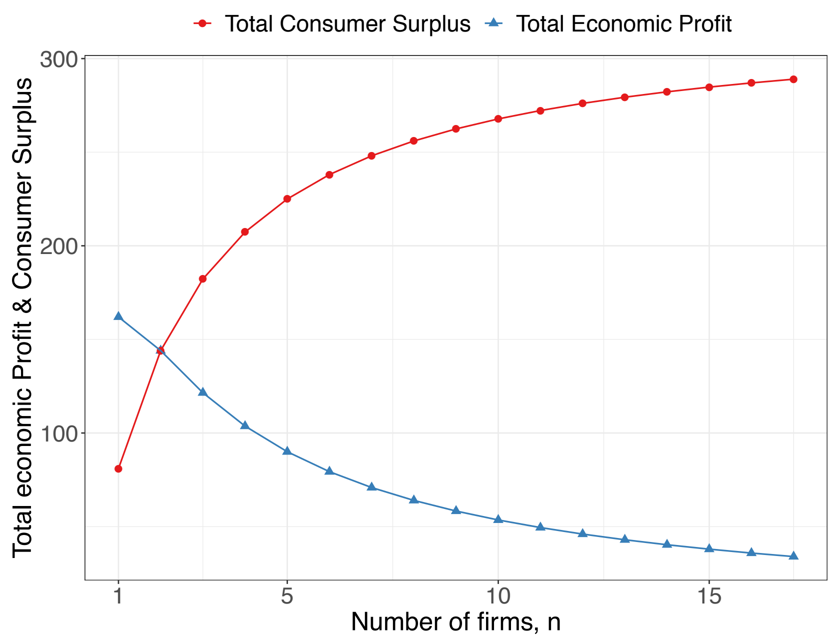


Figure 9.9: **Conflicts of interest between firms' owners and consumers: consumer surplus and economic profit** Total consumer surplus increases and total economic profit decreases as the number of firms competing for the same market increases. The parameters used are the same as in the other figures: $\bar{p} = 20$, $\beta = 0.5$, $c = 2$.

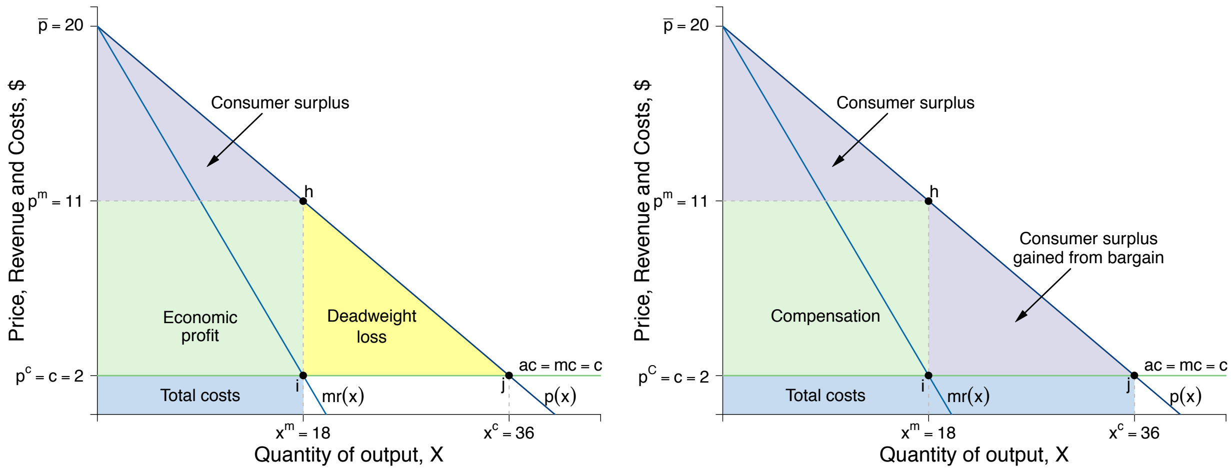
Limited competition and deadweight loss

We measure the extent of the inefficiency by a quantity called the dead weight loss which you also encountered in Chapter 8 (e.g. in Figure 8.20). The dead weight loss represents the quantity of either economic profits or consumer surplus that could have been realized if the firm or firm had produced more. This is shown in Panel **a** of Figure 9.10 as the area of the yellow shaded triangle for the case of a single firm – a monopoly like the one in Figure 9.2. Recall that the firm will produce the quantity of output at which its marginal revenue equals its marginal cost. It will then sell that output at the maximum price possible, indicated by point **h** in the figure. The total revenue of the firm is composed of two parts: costs (including the opportunity cost of the capital goods used) and economic profits, their quantities indicated by the area of the green and blue rectangles.

REMINDER Remember from Chapter 8 that *deadweight loss* is the feasible consumer surplus or economic profits that are not realized because price is above marginal cost.

To see why the monopoly producing x^m units of output and selling them at the price p^m is inefficient, think of an alternative. Suppose that the firm produced x^c and sold that amount at the maximum price it could, namely p^c which is also equal to the marginal and average cost. (Don't ask *why* the firm would do this, just imagine that it did). Then there would be no profits and consumer surplus would be the entire area under the demand curve an above the cost price line, that is, the two purple triangles and the light blue rectangle in Panel **b** of Figure 9.10.

If this were to occur in reality, consumers would benefit and owners lose. So



(a) Monopolist and deadweight loss

(b) Hypothetical bargain

Figure 9.10: **Monopoly and deadweight loss: Why limited competition is Pareto-inefficient.** In Panel **a** is the profit maximizing output and price of a monopolist. Panel **b** illustrates the hypothetical case of the same firm producing twice as much and selling the resulting output (x^c) at the highest price feasible given the demand curve. In the absence of compensation to the owners the consumer surplus would be the entire area above the cost curve and below the demand curve. The area of rectangle "Compensation" is identical to the Economic profit rectangle in Panel **a** showing that the firm's owners are as well off in Panel **b** as in Panel **a**. The two purple triangles in Panel **b** are the consumer surplus, double the amount in Panel **a**.

Consumer surplus: $cs(x) = \frac{1}{2}(\bar{p} - p^m)x^m$

Deadweight loss: $dwl(x) = \frac{1}{2}(p^m - c)(x^c - x^m)$

Marginal revenue: $mr(x) = \bar{p} - 2\beta x$

Inverse demand: $p(x) = \bar{p} - \beta x$

the hypothetical doubling of output is *not* a Pareto improvement. But it *could* result in a Pareto improvement if the consumers were to give up an amount of their increased consumer surplus sufficient to compensate the owners for their lost economic profits. This is shown in Panel **b** of Figure 9.10. The consumers would have doubled their consumer surplus and the owners would have been exactly compensated for their lost economic profits, so they would be no worse off.

With this compensation for the lost economic profits of the owners, the hypothetical increase in production to x^c is therefore a Pareto improvement. This means that the monopoly output and price is *not* a Pareto efficient allocation. The fact that we did not explain how the hypothetical increase in production to x^c could occur does not matter; we will see how that might be actually implemented shortly. All we require to show that some allocation is Pareto inefficient is that there exist some *other* allocation that is technically feasible, meaning does not violate the basic facts given by the demand function and cost function.

A Pareto-efficient allocation with price = marginal cost

Is the new hypothetical allocation itself Pareto efficient? To see that it is, think about why the monopoly allocation was inefficient: there was a deadweight loss arising from the fact that the price p^m exceeded marginal cost. If the firm were to produce x^c the price would equal the marginal cost, so there would be no deadweight loss. And so we could not find a technically feasible allocation that is a Pareto improvement over the allocation with x^c and p^c. Therefore x^c and p^c are a Pareto efficient allocation.

Are there any conditions in this model under which the Pareto-efficient allocation would actually occur, not hypothetically as a thought experiment but as a Nash equilibrium, given the relevant players' objectives and constraints?

There are, at least approximately. It will be a Nash equilibrium in which the price is approximately equal to marginal costs, or what is the same thing, where deadweight loss is approximately zero. We know from Equation 9.34 that

$$p^N(n) \quad = \quad c + \frac{1}{n+1}(\bar{p} - c)$$

which means that the Nash equilibrium price p^N approximates marginal cost when the number of firms n is very large. This will occur when the barriers to competition **b** is close to zero because (as we know from Equation 9.39), the Nash equilibrium number of firms is given by the following:

$$n^N \quad = \quad \frac{\bar{p}(1-b) - c}{bc}$$

Figure 9.11 illustrates how deadweight losses diminish and virtually disappear as the number of firms competing becomes very large.

In Chapter 8 we showed that the supply and demand model with price taking buyers and sellers – often referred to as *perfect competition* – also implements a Pareto efficient Nash equilibrium. And it is for the same reason: that under those conditions the price would approximately equal the marginal costs of the highest cost producer.

The result is that perfect competition in the price taking supply and demand model yields the same results as "unlimited competition" – that is large n – in the Cournot model of price making competition.

Efficiency, competition and inequality

Because owners of firms tend to be wealthier than are consumers on average, the extent of competition affects not only the degree of deadweight loss but also the extent of inequality in the economy. We have seen that by restricting its sales to x^m the owners of the monopoly gained a larger slice of a smaller pie, to the disadvantage of consumers.

A solution is to reduce the market power of the monopolist. Because firm owners tend to be more wealthy than are consumers on average, reducing barriers to entry and thereby increasing the number of firms competing would by this standard make the economy both more efficient and more equal.

Notice from Figures 9.11 and 9.9 that the number of firms competing need not be very large to substantially raise the level of consumer surplus relative to profits, and reduce the extent of deadweight losses. If the number of firms is *very* large, this is similar to what many economists have called *perfect*

Figure 9.11: **Inverse relationship between number of firms and deadweight loss**. As a market becomes more competitive and the number of firms increases, the size of the deadweight loss decreases. As the number of firms increases, the markup decrease and price moves closer to marginal cost, which results in lower deadweight loss. Deadweight loss is indexed to being 0.5 under a monopoly when $n = 1$.

competition, which we explain later in this chapter. Cournot refers to the case of very many competitors as *unlimited competition* and we will use his term, because his model is quite different from the model associated with the terms perfect competition in which firms act as price takers.

We will see in Section 9.10 that another way to implement an efficient outcome is to give the monopolist more power, not less, allowing it to make take it or leave it (TIOLI) offers to each consumer, in what is termed *perfect price discrimination*. You have already encountered similar cases in Chapters 5 and 10 where the actor with TIOLI power implemented a Pareto efficient but highly unequal outcome.

9.9 Coordination among firms: Duopoly and cartels

CARTEL A *cartel* is an industry structure in which firms producing separately jointly agree on what the output and price should be in an industry, and also agree on how to share the resulting market demand.

EXAMPLE The most famous contemporary example of a cartel is OPEC (the Organization of Petroleum Exporting Countries), which has regularly agreed on the number of barrels of oil each country would produce therefore affecting the world price for oil and gasoline or petroleum.

When there is more than firm, the conflict of interest is not just between the owners of the monopoly and the consumers. There are two conflicts:

- between the owners of the firms on the one hand and the consumers on the other; and
- among the owners of the firms who are competing to sell their goods.

In Figure 9.12 we represent both dimensions of conflict for the case where we have $n = 2$. At the Nash equilibrium, point **n** (identical to the case shown in Figure 9.4) the two firms each produce 12, so total output is 24, larger than the monopolist produced, namely 18.

Recall (from Figure 9.10) that consumers' surplus is maximized when a total output of 36 is produced. One possible allocation that would maximize consumer surplus is point **j** in the Figure 9.12 with each of the duopolists producing 18. The arrow from **n** to **j** shows that both firms producing *more* is better for consumers.

Along the line, the outputs of two firms sum to $\frac{(\bar{p}-c)}{\beta}$, which is equal to 36 in this case. Consumers would value points **k** and **l** as much as they do point **j**. Consumers do not care which of the firms produces more output as long as the total sum of output equals $\frac{(\bar{p}-c)}{\beta}$. At point **j** or any other point on this line the the following are true :

- Price is equal marginal cost, $p(X) = c$
- Consumers enjoy the maximum consumer surplus
- Deadweight loss is zero,
- Economic profit is zero, and

- The allocation is Pareto efficient.

Somewhere along that line is where consumers would like the allocation to be.

But not the duopolists. The orange arrow shows that both producing less (than the Nash equilibrium) can raise profits to each. Just as the fishermen Chapter 1 could do better if they restricted their over-fishing the lake, duopolists could receive higher profits if they could cooperate so as to not "over-harvest" their potential market by producing too many goods.

We have superimposed on the figure the isoprofit curves of the duopolists. Contrary to the interests of the consumers, the duopolists would do better if they could agree to *reduce* output. The yellow lens shows all of the combinations of the outputs of the two firms that are Pareto superior to the Nash equilibrium (**n**) *if we forget about the consumers and consider only the interests of the owners of the firms.*

Restricting output of the two firms to just 9 each (point *i*) would bring total output down to 18, the level the monopoly chose to maximize its profits. While consumers would like the outcome to move from **n** towards **j** the duopolists would like the outcome to move from **n** towards **i**.

figures/CompetitionMarkets/cournot_brfs_profits_social.pdf

Figure 9.12: **The duopolists' coordination problem and the maximum consumer surplus.** If point **n**, the Nash equilibrium is the status quo, owners of firms will seek to restrict output by implementing a point in the yellow Pareto-improving lens, ideally (for them) a Pareto-efficient point such as **i**. Consumers, in contrast, will be better off if output is increased, lowering prices, reducing deadweight loss and increasing consumer surplus. For consumers the ideal level of production is given by the purple line; they are indifferent between points on the line, but if both firms expand their output equally to a point (**j**) where price equals marginal cost and consumer surplus is maximized (eliminating both deadweight loss and economic profit for the firm's owners).

The duopolists' dilemma: A coordination failure among owners

If they could agree to each produce an output of $x^{B*} = 9 = x^{A*}$. This would maximize their joint profits. It would also be Pareto efficient as you can see from Figure 9.14.

But how could they enforce such an agreement (they could get the government to enforce it because in most countries it would be illegal). If Firm A knows that Firm B will produce output x^{B*}, Firm A's best response is to *increase* its output to the output x^A_O. Firm A behaves *opportunistically* by taking advantage of Firm B reducing its output (point **d**). When Firm A behaves opportunistically, it increases its profit to π^A_O which is substantially *higher* than the Nash equilibrium profit. A's opportunism results in Firm B obtaining much *lower* profit than at the equilibrium, π^B_V where it is the *victim* of A's opportunism. The same reasoning applies to the owners of Firm B. Their best response is to violate the agreement.

You can see from Figure 9.15 that the two duopolists face a prisoners'

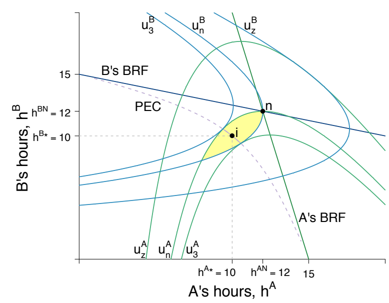
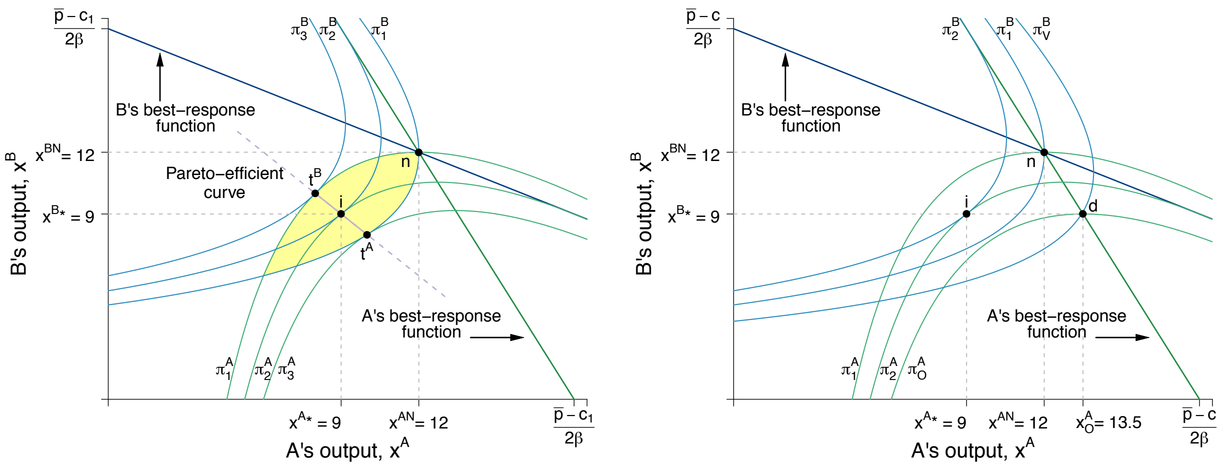


Figure 9.13: **The over-fishing coordination problem.** Figure 5.12 from Chapter 5 shown here is about how many hours two fishermen will spend fishing on a lake, which is a common pool resource. They could both do better if they coordinated and fished less than they do at the Nash equilibrium (point **n**). As in Figure 9.12, the yellow lens shows all of the allocations that are Pareto superior to **n**.

REMINDER An outcome is Pareto-efficient when the firms' isoprofit curves are tangent to each other, meaning that the marginal rates of substitution are equal. When the isoprofits are tangent, no firm's owners can do better without making another firm's owners worse off and therefore there are no alternative outcomes that are Pareto-superior to the tangencies of the isoprofit curve.



(a) Nash equilibrium and cartel choice for two firms

(b) Firm A behaves opportunistically

Figure 9.14: **Duopolists' isoprofit curves and best-response functions.** Firm A's output is on the horizontal axis and firm B's output is on the vertical axis. Looking at each firm's isoprofit curves, we can see that the Cournot (Nash) equilibrium is Pareto-inefficient with potential Pareto-improvements. At the Nash equilibrium, each firm is on its equilibrium isoprofit line: π_1^A for Firm A and π_1^B for Firm B. If both firms could simultaneously *decrease* their output, then at least one firm would obtain higher profits, that is, at least one firm could move to a higher isoprofit curve (π_2 or π_3) while the other remained at their Nash equilibrium isoprofit line (π_1), or both firms could make higher profits, π_1 to π_2. π_O^A and π_V^B correspond to the extreme payoffs that the firms would receive when Firm A behaves opportunistically and Firm B is the victim of A's opportunism, as shown in Figure 9.15.

dilemma. A similar Prisoners Dilemma type coordination problem occurs when there are many oligopolistic firms competing. On the basis of this reasoning the Cournot model reasons that firms will not coordinate and form cartels.

How owners address their coordination failure (consumers beware!)

A coordination mechanism that owners of firms use to solve their coordination problem is a *cartel* that negotiates output and price levels by dividing up markets among potentially competing firms. The firms would produce the total market output equivalent of the monopoly and each would obtain a share of the monopoly profit by selling at the monopoly price. They would still have a conflict, because having agreed on a total output (namely 18) each firm would prefer to produce more of it.

Cartel-type behavior was outlawed in the United States by the Sherman and later anti-trust acts, although U.S. firms in some industries have strong incentives to limit competition if they can get away with it. In other societies cartels have sometimes become the normal, expected way of organizing production. But, because of the availability of opportunistic behavior, cartels that are not protected by law are often very fragile, as the prisoner's dilemma game above would suggest.

But, behaving as a cartel is not the only institution that firms adopt to try to obtain greater market power and increase their share of the rents on a market. Firms regularly *merge* with other firms (that is, to form one larger firm), in effect pooling their resources and becoming more like a monopoly. Firms also seek to *acquire* other firms either by buying them at a mutually agreed price or by setting up *hostile takeovers* through buying enough shares to name directors favorably disposed to the acquisition.

EXAMPLE For the duopolist, a cartel performs the function that the Impartial Spectator hypothetically, or government policies in reality might perform for the two fishers seeking to overcome the over-fishing problem: the cartel seeks to maximize the total profits of the two firms, just as the Spectator maximized the total utility of the two fishermen.

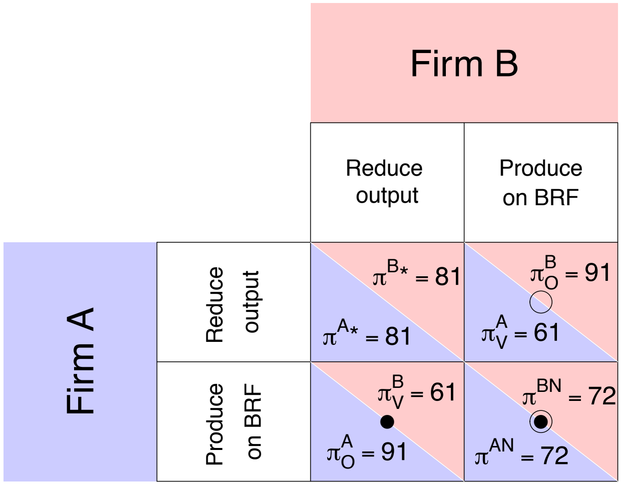


Figure 9.15: **Cournot Production: A Prisoner's Dilemma Game.** A's payoffs are in the left-bottom corner of each cell; B's payoffs are in the top-right corner of each cell. Firm A's payoffs are ranked: $\pi_O^A > \pi^{A*} > \pi^{AN} > \pi_V^A$. Similarly, for B: $\pi_O^B > \pi^{B*} > \pi^{BN} > \pi_V^B$. The subscript "O" corresponds to "Opportunistic" and the subscript "V" corresponds to 'Victim of Opportunism." For each player, their *opportunistic* profit is greater than their cooperative, reduced output payoff. Neither firm wants to receive the profit they would obtain when they are *victims* of opportunism by the other firm. As a result, for each firm, to choose the output on their best-response function (Produce on the BRF) strictly dominates Reduced output, the strategy that would bring about the Pareto-efficient outcome.

To sum up, a Nash equilibrium in the Cournot model with a limited number of firms is a Prisoners' Dilemma from the point of view of the firms resulting in their failure to maximize their joint profits by enforcing the price and quantity that a monopolist would choose.

Dynamic inefficiency

There is another possible source of inefficiency in monopoly pricing. Think about a firm that is considering developing a new product that will have few close substitutes and hence will face cost and demand conditions similar to those shown in Figure 9.10.

From the firm's standpoint, the incentive to do this is the area labeled economic profit. But the profits the firm would gain does not fully measure the benefits that introducing the new product would yield. Consumers buying the produce would also benefit in the form of consumer surplus. Thus, the profits accruing to the innovating firm fall short of the entire benefits of the introduction of the new product (consumer surplus plus producer surplus (economic profit)).

The economic profit alone may not be sufficient to offset the development costs of the new product. This is called *dynamic inefficiency* because it arises from the firm failing to make an innovation that would have increased the sum of consumer surplus and economic profit.

Checkpoint 9.6: Price discrimination and dynamic efficiency

Explain why the example of dynamic inefficiency – the failure to introduce a new product – would not occur if the innovating firm could perfectly price discriminate.

9.10 Perfect price discrimination: Eliminating deadweight loss at a cost to consumers

PRICE DISCRIMINATION means selling the same product to at different prices, for example, charging more to buyers with a greater willingness to pay.

Surprisingly, the deadweight losses associated with a monopoly's market power can be eliminated if we let the monopolist have a little *more* power so that the monopolist can charge different prices to different individual buyers. This is called **price discrimination**. Price discrimination is an example of a situation in which the Law of One Price does not hold.

Some aspects of racial, religious, sexual preference and gender discrimination fall into the broad category of price discrimination. When an automobile dealer sells a car to a woman at a higher price than would have been charged to an otherwise identical man, we have an instance of price discrimination. But here, we consider cases where price discrimination is based on the buyers willingness to pay, not her gender or some other characteristic.

Perfect price discrimination and Pareto-efficiency

To see why how price discrimination might address the inefficiencies associated with monopoly or other forms of limited competition, take another look at Figure 9.10. Deadweight losses exist because the monopolist restricts how much he sells to maximize profits. Therefore, the price the monopolist charges exceeds marginal cost. Consequently, many consumers are willing to pay more for the good than the marginal cost that the firm would incur to produce it, which we can identify in Figure 9.16 because the demand curve represents the maximum price a buyer is willing to pay for each unit of the product sold.

FACT CHECK A study of the markets for rice, kidney beans, sugar, tomatoes and tortillas in Mexican villages found that those buying larger amounts paid substantially lower prices. For the tortilla market, for example, a ten percent increase in the amount of tortillas purchased was associated with a 4.5 percent decrease in the price. This may have occurred because those buying small amounts were generally poor, and had fewer alternative places they could purchase the commodities, and fewer substitute goods for these staples. Richer buyers, purchasing larger quantities had better alternatives for shopping around, and also could consume substitute goods if faced with too high a price for the goods studied.

But suppose the firm could make a private bargain with each of these consumers who are not buying the good at the price p^m. As first mover the monopolist would make the customer a take-it-or-leave-it offer charging a price equal to the consumer's maximum willingness to pay.

If this were possible then the firm would produce more. The firm would produce and sell the good even to the consumer whose willingness to pay just barely exceeded the marginal cost. In this case, the price charged to the marginal consumer (the one with the least willingness to pay that exceeds the marginal cost) would (virtually) equal the marginal cost and there would be no dead weight loss.

Of course if the firm could make a private deal with each of the consumers whose willingness pay fell short of the price p^m in Figure 9.16 then it would also want to make a similar deal with all the rest of the consumers. But in these cases the firm would charge *more* than p^m because, as you can see from the demand curve, the remaining consumers have a willingness to pay that is greater than that price.

REMINDER The participation constraint of a buyer represents the requirement that in order to voluntarily participate in a social interaction, such as choosing to purchase a good from a firm, an individual must receive a utility at least as great as the next best alternative, which constitutes the buyer's fallback in the transaction. If a seller makes an offer that violates the participation constraint, the buyer will refuse it.

EXAMPLE The perfectly discriminating firm

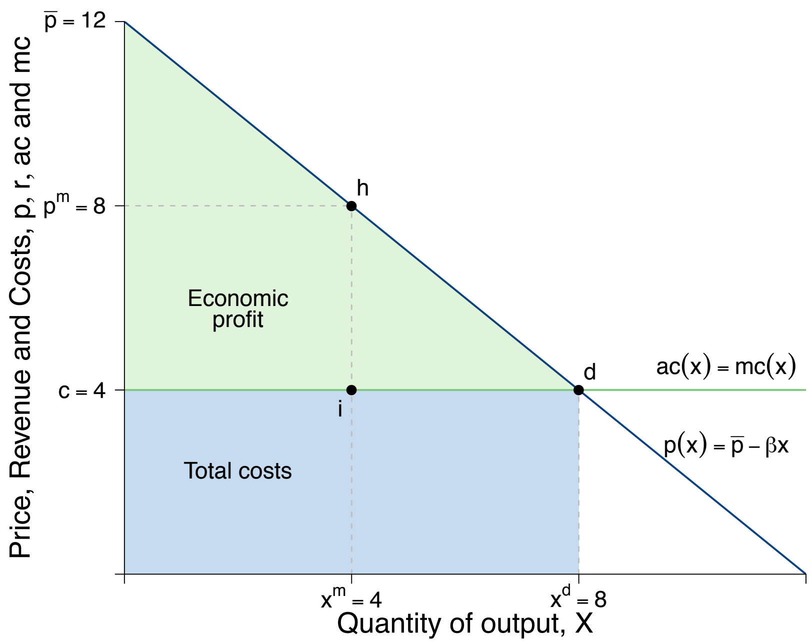


Figure 9.16: **Cost curve, demand curve, and the perfectly discriminating firm.** The perfectly discriminating firm charges each customer a price equal to the customer's reservation price. For example, the customer who buys the x'th unit of the good is charged p'. Because no buyer pays less than their maximum willingness to pay for the good, there is no consumer surplus; the perfectly discriminating firm appropriates the total economic surplus from the transaction.

We can visualize the policy of the perfectly discriminating firm in terms of cost curves and the demand curve, in Figure 9.16. The area under the demand curve represents the maximum sales revenue that the firm could get if it could somehow charge each consumer a personally-chosen price for each unit (a different price for each consumer for every unit which that consumer buys) designed that the consumer would be indifferent between buying or not buying. **Perfect price discrimination** allows the firm to capture all of the economic surplus of the firm's production of a good through price-setting.

There are no benefits conferred on consumers in the form of consumer surplus. And there is no deadweight loss either.

Is the outcome when the monopoly practices perfect price discrimination a Nash equilibrium? You can check that given that the monopolist can make a take-it-or-leave-it price offer and sets a price just slightly below the buyer's maximum willingness to pay, the buyer's best response to this price offer is to buy. Given the range of different people's different willingness to pay, the monopolist cannot do better than a perfect price discrimination strategy.

PERFECT PRICE DISCRIMINATION *Perfect price discrimination* occurs when a firm can make a separate take-it-or-leave-it offer to each individual buyer for each unit sold at a price equal to the consumer's maximum willingness to pay.

Perfect price discrimination therefore results in Pareto-efficient level of output even by a monopoly because the price charged to the marginal consumer is equal to the marginal cost of producing that good. There is also another way to see this.

Remember that an outcome must be Pareto-efficient if it is the result of one party maximizing his or her utility subject to a participation constraint constraint requires that the other person's utility must not be less than some given amount. "Participate" here means "buy the good", and the consumer's fallback is not to buy the good, and have more money for other purposes. Since, the firm cannot secure buyers at prices above the demand curve; the demand function is the participation constraint for the perfectly discriminating monopolist.

The result seems paradoxical: the inefficiencies resulting from the market power of a monopoly can be eliminated by giving the monopoly even more power, the ability to charge different prices to each consumer. The reason why this works is that when the firm charges a single price to all buyers, the owner faces a trade off: he could sell more, but in order to do so he will have to lower the price; or he could to charge more, but then fewer consumers will buy.

Given the trade-off posed by the demand curve and the requirement to sell all goods at a single price, the monopolist adopts a strategy to get a larger surplus by reducing the total surplus produced in the transaction. Therefore, the monopolist restricts production and sales to allow a higher price. In other words, because he cannot claim all the surplus, the monopolist finds a way to get a larger slice from a smaller pie.

Granting the monopoly price-discriminating power liberates the owner from the trade-off posed by the demand curve. It does this because the firm can now produce and sell to any consumer from whom the firm might extract a profit, without lowering the price at which the firm sells to other buyers.

When you think about it, it makes sense that perfect price discrimination is efficient given what we have already seen in Chapters 4 and 5. For example, when Ayanda and Bongani bargained and Ayana had take-it-or-leave-it power, she was acting exactly as a perfect price discriminator appropriating for herself the entire surplus, constrained only by Bongani's participation constraint. The outcome was Pareto-efficient. The same was true when we considered one of the fishermen having take-it-or-leave-it power as owner of the lake in Chapter 5.

To sum up, perfect price discrimination and potential buyer's responses to it:

- leads to a Pareto-efficient allocation even in the case of monopoly, since the firm will sell to every buyer whose reservation price is above the marginal cost of the goods produced; and
- constitutes a Nash equilibrium.

REMINDER The word *distribution* refers to *how much* of the benefits or costs of any interaction each person gets. Perfect price discrimination results in greater inequality in the distribution of economic surplus from exchange because the firm gets the whole economic surplus, and buyers get none.

Perfect price discrimination and perfect competition

So far we have presented price discrimination as an almost magical solution to the problem of monopoly, but it is far from that.

It has important effects on *distribution*: the owners of the firm appropriate all the available economic surplus in the form of a monopoly rent. Compare this to the case in Figure of the non-price-discriminating monopolist. He produces the amount that equates the marginal revenue with marginal cost and sells this amount at the maximum price possible. Figure 9.10 a. shows that buyers receive some consumer surplus. But with perfect price discrimination (shown in Figure 9.16) buyers are worse off, since they have lost all the consumer's surplus, which now takes the form of economic profits for the monopolist.

Perfect competition (introduced in Chapter 8) is another abstract idea which, like perfect price discrimination, under some not very realistic modeling assumptions ensures that a Nash equilibrium allocation will be Pareto-efficient. But there is an important difference: as Figure 8.19 shows, at the equilibrium of competition with price-taking buyers and sellers the gains from exchange are *shared* between consumers and owners.

Perfect competition does this works not by enhancing the market power of sellers to the disadvantage of buyers (as in the case of price discrimination) but instead by improving the fall-back positions of both buyers.

Because buyers in the perfectly competitive model have the option of buying exactly the same good from another seller at the same price, their fallback position is exactly what they are getting in equilibrium.

Rents are shared in the perfect competition model because buyers have access to something even better than a close substitute: an identical good at the same price from another supplier. A result is that instead of the entire surplus being captured by the discriminating monopolist, both consumers and producers enjoy a share of the surplus.

Checkpoint 9.7: The PC vs. the ICC

a. What is the difference between the participation constraint and the incentive compatibility constraint?
b. How are each of these constraints relevant to understanding price discrimination and the distribution of the economic surplus between consumers and owners?

9.11 Application: Price discrimination in action

If price discrimination can eliminate the gains from the exchange enjoyed by buyers by distributing their the entire consumer surplus to owners economic profits, then is it called "perfect"?

In science and other scholarly disciplines the word "perfect" does not necessarily mean "flawless." It often refers to some abstract idea, such as a "perfect gas" in physics which does not exist but which is a helpful simplification for understanding the important aspects of some processes. Perfect price discrimination is an abstract idea designed to clarify why monopolies are *in*efficient, not a value judgment about how good it is.

Although price discrimination benefits owners, firms do not typically act as perfect price discriminators. Many firms do price discriminate to some extent, but perfect price discrimination (like perfect competition) is not a business practice that we expect to see often. There are three reasons why few, if any, businesses practice price discrimination in its pure form:

FACT CHECK Price discrimination is often viewed as "price gouging" and it is generally condemned on ethical grounds. For example, eighty percent of Americans in a study by psychologist Daniel Kahneman disapproved of stores charging more for snow shovels after a storm (when presumably people's willingness-to-pay is higher than usual).

- *Information*: it requires information on potential buyers' maximum willingness to pay that is costly or even impossible to obtain;
- *No re-sale*: the monopoly firm would have to find a way to ensure that those who purchased the good at a low price could not resell it to those with a higher willingness to pay, thereby undermining the monopoly's high price sales; and
- *Consumer lash-out*: buyers often react negatively to price discrimination, thinking that the firms should not be able to profit by charging different prices to different people, or to people living in different circumstances.

But even with these impediments, there is a price discrimination strategy that is feasible in many cases. When a firm has information that allows it to distinguish groups of potential buyers who have different willingness to pay, but not enough information to find out every individual buyer's reservation price, it can engage in group price discrimination.

The discriminating firm will charge a different price to members of each of the groups (or *sub-markets*) depending on the firm's estimate of the willingness to pay of the members of each group.

Suppose the firm considers raising the price at which it sells to a particular group. If buyers in the group can meet their needs by purchasing a close substitute for the firm's good, then their willingness to pay for this particular product will be limited, and so a small increase in price will drive them away. The same will be true of members of the group have little income their marginal utility of money left over for other purchases will be substantial, so that their willingness to pay is limited.

If, on the other hand, buyers in the group have no good alternatives to buying from the firm, or if they are wealthy their willingness to pay will be greater and the firm can raise price a lot and lose only a few customers. In this case, we would expect the price offered to the group to be high.

Competitive strategies like advertising the distinctive features of the product and consumer loyalty programs such as airline miles or discount cards are designed to deter customers from switching to there next best alternative. Advertising by the firm is often designed to attract a group of buyers who are particularly attached to the advertised features of its product, or particularly taken by the glamour of using it, rather than some close substitute. Advertising also targets higher income people, whose willingness to pay will be greater because their marginal utility of money left over is less.

Mac computers tend to be a lot more expensive than PCs, so the travel site Orbitz steered Mac users to more expensive hotel sites. Here the price-setter exploits the fact that rich people have higher willingness to pay. So lower income buyers *might* see a lower price. But not necessarily. Discounts for the elderly (on average lower income) follow this logic. For the identical stapler the Staples.com website charged a price of $15.79 to one buyer and $14.29 to another based on where they lived. They only lived a few miles apart.

Office Depot told the Wall Street Journal (WSJ) that they use "customers' browsing history and geolocation" to personalize price making. In the WSJ's study the areas that tended to get *lower* prices were *higher* income areas where there was more competition from rival sellers. In other words, the rich paid less and the poor paid more.

FACT CHECK Find out more at the following website: msu.edu/ conlinmi/teaching/MBA814/WSJpricediscrimination.pdf.

Airline companies charge fliers who stay in a city over Saturday night a lower fare for the same seat than fliers who return without staying over Saturday night. Stayover requirements for discount airline tickets are an example of *statistical price discrimination*. These consumers have a lower willingness-to-pay, and greater price elasticity of demand than other fliers. They have a higher elasticity of demand at any price because non-business travelers typically have more flexibility in their travel schedules and routes, and more alternatives than business flyers. A business flyer who wants to stay over Saturday night will benefit from the lower fare, showing the airlines' ability to discriminate is limited by the information their customers reveal.

Checkpoint 9.8: Perfect price discrimination and dynamic inefficiency

Explain why the dynamic inefficiency described above – a socially valuable innovation not being introduced – would not occur if the monopolist considering the innovation were capable of perfect price discrimination.

9.12 Rent-seeking, price-making, and market equilibration

The model of perfect competition introduced in Section 8.11 describes a Nash equilibrium, but it does not explain *why* we might expect a market to be at or near the intersection of the supply and demand curves.

Why would we give special attention to the market-clearing equilibrium point? The answer must be that other points are not Nash equilibria so that they will be disrupted by buyers or sellers (or both) changing the prices or quantities at which they are willing to transact.

REMINDER A Nash equilibrium is an allocation such that none of the actors could do better by altering their strategies, given what the strategies adopted by others. This explains why if actors are implementing a Nash equilibrium they have incentives to *stay there*. The Nash equilibrium concept says nothing about how they might *get there* in the first place.

This does not mean that we can ignore all of the graph except the intersection of the two curves. There are two reasons:

- the economy may be out of equilibrium for long periods of time either because adjusting to a change in demand or supply can take time or because public policies prevent prices from adjusting.
- what happens when the market is not at a Nash equilibrium is essential to how a market might equilibrate, that is, get to an equilibrium, and why sometimes they may not get to an equilibrium. To see this we study the cases illustrated in Figure 9.17.

Non-equilibrium prices and the short side of the market

Suppose that for some reason the price is p^H, a price higher than the price given by the intersection of the supply and demand curves. To understand what this means remember that "supplied" does not mean "sold." It means produced and brought to market. Whether it is sold will depend on demand.

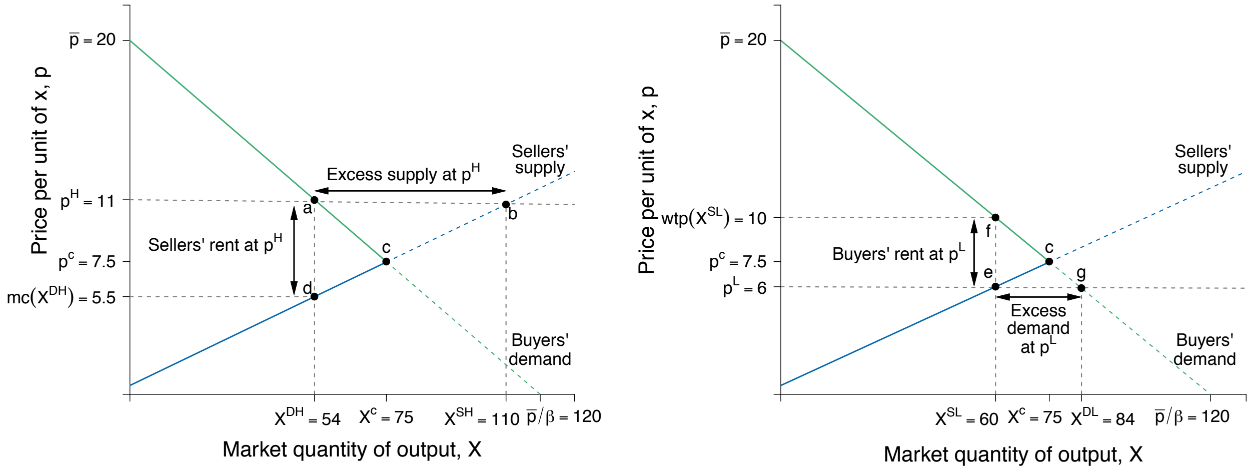
Figure 9.17 shows that at the price p^H the amount sold will be the amount demanded X^{DH}, which means buyers will not be willing to buy more than this quantity. The quantity supplied is X^{SH} which exceeds the quantity demanded X^{DH}. Sellers therefore want to sell more than buyers want to buy at p^H. This case is called **excess supply**, referring to the excess of goods produced and brought to market *but not sold*, remaining in warehouses or on the shelves rather than going into someone's shopping bag.

EXCESS SUPPLY OR DEMAND exists when at the prevailing price, the amount supplied exceeds the amount demanded, or the amount demanded exceeds the amount supplied.

By contrast, when the price is $p^L < p^c$ buyers want more goods than the sellers want to sell, so there is **excess demand**, which might show up with people standing in lines to get limited numbers of goods or heading home with empty shopping bags. When either excess demand or excess supply exist, we say the market does not *clear*. To see how what happens in this case we need to introduce some new terms about the economics of *non-clearing markets.* When someone cannot buy or sell the amount that they would like at the going price we say that they are *quantity-constrained.*

We refer to two "sides" of the market with respect to demand and supply:

- *Demand side*: the buyers; and



(a) Excess Supply

(b) Excess Demand

Figure 9.17: **Excess supply and excess demand: Rents in a non-clearing market.** When $p^H > p^c$, (Panel a) sellers would like to supply X^{SH}, but are only able to transact with buyers at the quantity *they* demand at p^H, which is X^{DH}. As a result, those sellers who sell to buyers gain an economic rent equal to the distance **ad**, which is the difference between their marginal costs at X^{DH}, $mc(X^{DH})$, and the price they obtain p^H. When $p^L < p^c$, (Panel b) buyers would like to buy X^{DL}, but are able to transact with sellers only at the quantity sellers are willing to sell at p^L, which is X^{SL}. As a result, those buyers who buy from sellers at p^L gain an economic rent equal to the distance **ef**, which is the difference between their willingness-to-pay at X^{SL}, $wtp(X^{DL})$ and the price they pay p^L.

- *Supply side*: the sellers.

When the market does not clear, there is a long side of the market and a **short side of the market**.

- *Short side*: The short side of the market refers to the actors – buyers or sellers – as short siders, who are able to make all of the transactions they wish. They are on the side of the market where the number of *desired transactions* is lowest.

- *Long side*: Some of the long-siders will be able to make the transactions they wish, but others will not. Some long-siders are quantity constrained. They are on the side of the market where the number of *desired transactions* is greater.

THE SHORT SIDE OF THE MARKET The *short side of the market* is the side – either supply or demand – on which the number of desired transactions is least, given the price.

Either buyers or sellers can be on the short side of the market, depending on the price being above or below the market-clearing price. If the price is $p^L < p^c$ (Figure 9.17 b.) so that there is excess demand then sellers on the short side of the market. When $p^H > p^c$ (Figure 9.17 a.) by contrast, sellers want to sell more goods than buyers want to buy, so buyers, who can buy all they wish, are on the short side of the market.

Table 9.2 summarizes these results about non-clearing markets.

Notice that in Figure 9.17 the supply curve above p^c is a dashed line, and the same thing is true of the demand curve below p^c. We do this to emphasize that these portions of the supply and demand curves are irrelevant to our analysis: we will never observe an outcome on the dashed segments.

This is because the exchange is voluntary: there is no way that those who would like to buy or sell more at the going price can force others to exchange with them. As a result, it is the short side of the market that determines the

Figure	*Price*	*Excess supply or demand*	*Short side*	*Rents*	*Unable to transact*
9.17 a	$p^H > p^c$	Excess supply	Demand (buyers)	Some sellers get economic profits	Other sellers get nothing
9.17 b	$p^L < p^c$	Excess demand	Supply (sellers)	Some buyers get consumer surplus	Other buyers get nothing

Table 9.2: **Excess demand and supply, rents and quantity constraints** The two rows of the table refer to the case in which the price is higher (top row) or lower (bottom row) than the price that would clear the market.

quantity transacted: that is, buyers (the demand side) when $p^H > p^c$ and sellers (the supply side) when $p^L < p^c$.

Rents in non-clearing markets

In Figure 9.17 the fact that the market does not clear is indicated by either *excess supply* or *excess demand*, the extent of which is measured in terms of quantity not bought or not sold. The quantity not sold is a *horizontal* distance in the figure.

The extent to which the market does not clear can also be measured *vertically* in the same figure, that is, by the difference between the price at which the good is transacted and the buyers' willingness to pay or the sellers' willingness to sell.

This means that in a non-clearing market there must also be *economic rents*, in the form of either consumer surplus or or economic profits. These measure the extent of excess demand or supply in the vertical dimension, that is, in terms of the price.

At the price p^H the marginal cost of the last unit sold ($mc(X^{DH})$) is less than the price, so we have both:

- *excess supply* because $X^{SH} > X^{DH}$ (the horizontal dimension) and
- some (but not all) sellers who succeed in selling their output gain *economic profits* on their last unit sold equal to the line ad that is the vertical distance between p^H and $mc(X^{DH}$ in Figure 9.17 Panel a.

In the other case, at a price $p^L < p^c$, we have both:

- *excess demand* because $X^{DL} > X^{SL}$ and
- some (but not all) buyers who are able to purchase what they want gain an *consumer surplus* (another form of economic rent) on the last unit purchased that is equal to the line ef in Figure 9.17 Panel b, because they paid p^L for a good for which their willingness-to-pay (the height of the demand curve at X^{DL} or $wtp(X^{DL})$, is greater than the price.

NON-CLEARING MARKET If there is either excess supply or excess demand the market does not clear. Market clearing occurs when demand equals supply.

Notice the pattern in the two examples above:

- *Short side*: Those on the short side of the market are able to transact all that they wish at the given price and
- *Long side*: Those on the long side of the market fall into two different groups with decidedly contrasting outcomes: Those who succeed in making a transaction gain an economic rent while the rest of the so called "longsiders" get nothing, they are excluded from the market.

EXAMPLE In colloquial English, to "get the short end of the stick" means to get a bad deal, to come up short in some bargain. But being on the short side of a market can be an advantage, as we will see when we study the market in labor in Chapter 11. In the equilibrium of the labor market there is an excess supply of workers so employers are on the short side of the market and workers both employed and unemployed are on the long side. This is also true in the credit market, where there is excess demand for loans, so banks and other lenders are on the short side and borrowers are on the long side, some of whom would like to borrow at the going rate of interest but cannot.

Rent seeking and market equilibration

How does a competitive market adjust when the going price is such that there is excess demand or excess supply? For any kind of adjustment to occur the buyers and sellers cannot act as they might in the equilibrium of the perfect competition model, that is, as price takers. As a thought experiment, if all of the buyers or sellers in the market described in Figure 9.17 were price takers (and therefore do not change the price), then a price like p^H or p^L would persist forever.

But in a situation with excess demand or supply being a price taker is not the best you can do. In a competitive market for goods like that described in the figure, a non-clearing market is not a Nash equilibrium. Why?

- **Rent-seeking**: If economic rents exist when the market does not clear, then there will be opportunities to gain either consumer surplus or economic rent by changing the amount demanded or supplied, or offering a different price. These activities are called rent seeking.
- **Equilibration**: Under most conditions, how the rents are obtained will result in prices and quantities *changing* so that the market eventually *equilibrates* and *clears*.

EQUILIBRATION is the process of getting to an equilibrium from a non-equilibrium point. Equilibration explains the dynamics or step-by-step process of how to reach an equilibrium.

To see why, look at Figure 9.17 and imagine the price is p^H as in the left panel and you are selling the *last* unit you produce at *more* than its marginal cost.

What would you do?

You would want to sell more goods. But you are *constrained* by the demand curve: you cannot sell more if you remain a price-taker and do not lower your price. So you become a price maker and acting as a rent-seeker you would lower your price a little bit and the following will happen:

- *Greater quantity sold*: You sell more as buyers switch to you because of the lower price you are offering (your goods are substitutes for the goods other firms produce).

- *Expanded sales*: You will be able to expand your production and sales of the good.
- *Lower profit per unit*: You would make a slightly smaller rent on each unit (a lower profit per unit) because the price is lower (and therefore $p - c$ is lower).
- *Best-responses by others*: Other firms losing business to you would do the same and lower their prices to best-respond to your change in strategy.

The process would go on until the price fell to p^c.

Something analogous happens if the price is lower than p^c. At price p^L it is the buyers who are receiving rents (their willingness to pay exceeds the price they are paying). Put yourself in one of the buyers' shoes: you would like to buy more but cannot without paying more. So you would act as a rent-seeker and price-maker, and would offer to pay a higher price than p^L. Sellers would flock to you, and other prospective buyers would find that they too would have to increase their prices. Again the process will go on until the economic rents disappear, that is, until the market clears at p^c.

RENT-SEEKERS AND RENT-SEEKING Buyers and sellers are *rent-seekers* because they wish to obtain the maximum rent possible. They each wish to obtain the highest rent – either consumer surplus for buyers or economic profit or producer surplus for sellers – given their fallback positions (their next best alternatives).

This is why the dashed portion of the supply and demand curves in the figure will not have any bearing on quantity transacted at the given price. All of the action must be on or between the solid lines, which encompass the *short side of the market*.

Checkpoint 9.9: Short side vs. long side

Analyze the market outcome when $p = p_L$.

1. Explain what the initial economic rent would be when $p = p_L$.
2. What would happen to the number of firms if $p < p^*$?
3. Would $p = p^L$ persist? Why or why not?

9.13 *Application: When rent-seeking does not equilibrate a market – A housing bubble*

There are two reasons why we study how the equilibration process works through the rent-seeking activities of buyers and sellers when the market is not in equilibrium.

The first is that this process can take a very long time. To see why, think about a labor market in which wages for a given kind of labor are substantially higher in one region or city than in another. The rent-seeking that might equilibrate this market requires workers and their families to leave their communities and families. Equilibration, therefore, often takes place over decades if not a generation.

FACT CHECK The dramatic increase in U.S. imports of manufactured goods from China resulted in plant closings in many industries, like furniture making. This so called "China shock" destroyed the old equilibrium, and economic reasoning would predict that unemployed furniture workers would move to other occupations and localities. This has occurred, but very slowly over a decade and a half.

As a result, the equilibrium of a market may be a poor guide to what we will find when we observe the economy empirically. But the idea of equilibrium remains informative. Equilibrium often provides predictions about what changes we should observe in data: for example, workers moving to areas with higher wages or less unemployment.

The second reason to study the equilibration process is that sometimes it does not work. Rent-seeking may drive the market *away* from equilibrium over a least a long enough period to create a substantial amount of insecurity.

To see how equilibration fails, lets consider a particular market, housing. Let us assume that the housing market starts in equilibrium. For some reason the price of housing rises, so that there now is excess supply of housing. If the rent-seeking process described above were to work, then owners of the unsold houses could do better by offering to sell at a lower price. Eventually the market would return to the earlier equilibrium.

But **rent-seekers** might reason differently. Imagine you have just gotten your first well-paying job. You want to have two choices:

- Buy either a larger home and don't acquire other assets (such as stocks or bonds)
- Buy a smaller or and also buy other assets

The increase in prices of houses could make you choose the smaller home. This is how the equilibration process is supposed to work.

But if you think ahead, you could reason as follows:

1. Housing prices rise,
2. and they may be even higher in the future
3. which means that putting a lot of money into my home rather than stocks will result in my being wealthier in the future.
4. Therefore buy the large house.

If this kind of economic reasoning were common – and there is no reason why it should not be – the increase in the price of housing would result in a shift *upwards* in the demand curve for houses. The shift in demand – if sufficiently great – would transform the market situation from one of excess supply to excess demand. This would result in additional increases in housing prices, thereby fulfilling your expectation that your home value would go up. The result would be a sustained rise in house prices. This is called a *housing bubble*. The term housing bubble suggests that prices do not increase forever. The bubble will *burst* if something happens that leads you or others to doubt that housing prices will continue rising. Then the reasoning runs in reverse.

FACT CHECK A housing bubble in the U.S., Spain, and other countries contributed to the 2007-2009 global financial crisis. Read about it in Atif Mian and Amir Sufi, "How they (and you) caused the Great Recession," University of Chicago Press, 2014.

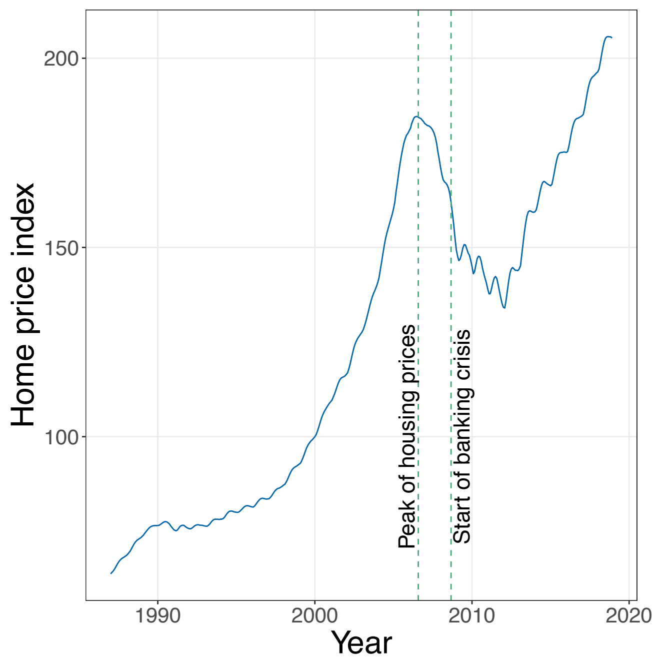


Figure 9.18: **The Case-Shiller index of housing prices in the United States from 1987 to 2018.** The housing mortgage crisis started after the high price for houses was reached in August 2006. The subsequent banking crisis began in September of 2008 with the closure of Lehman brothers, an investment banking firm.

A fall in the price of housing indicates that owning housing is no longer a good investment. People will decide to give up the home improvement or summer home in favor of alternative forms of wealth. Housing prices then crash. Figure 9.18 shows the pattern of a bubble, then a crash for US houses: housing prices increased until August 2006, then started to drop a little bit, then crashed. They began to recover again in January of 2012 and have continued to increase since then.

Bubbles and the instability they cause are often attributed to people's "irrationality." But notice there was nothing inconsistent (and therefore irrational) about your calculation above. You were using the same kind of "doing the best you can" in light of the available information that characterizes the constrained optimization process that is sometimes termed *economic rationality*.

DURABLE ASSET A durable asset is an asset that lasts for more than one period of time and often for many periods. Contrast a slice of pizza, which you quickly eat, with a home. When you buy a home you are "consuming" housing, but you are also buying an asset which can experience increases or decreases in its price. The future price of the asset is important to you when you buy the good.

What made the bubble possible was not stupidity or a gambling temperament, but the fact people (reasonably) took an increase in house prices as a signal of things to come, and that housing is a **durable asset**, whose future value must be a critical concern when buying or selling. Bubbles do not occur in markets for goods that are non-durable or do not last. You do not see people flocking to buy up as much fish as possible when the price rises, hoping to sell it for a profit later! (With improvements in refrigeration and storage, however, even this cannot be ruled out).

9.14 How competition works: The forces of supply and demand

We have also seen how competition among rent-seeking buyers and sellers in a non-clearing market can equilibrate supply and demand. The fact that most goods are not durable (and are costly to store) means that bubbles are the exception, and movement towards a market equilibrium is more common. This shows that the supply and demand framework provides a useful conceptual tool, even if the perfectly competitive model of supply and demand does not explain how price-taking buyers and sellers would ever *get to* a Nash equilibrium. . The rent-seeking approach, teaches an important lesson. Locating the intersection of a supply and a demand curve is not sufficient to understand how competition works and how buyers and sellers best respond to changes in supply, demand, and the degree of competition.

BILATERAL MONOPOLY In a *bilateral monopoly* transaction there is one monopoly transactor on each side (hence *bilateral*) of the market – one potential buyer and one potential seller.

Changes in supply and demand affect the prices at which goods are exchanged by altering the fallback options of buyers and sellers. A seller's fallback (next best alternative) in an interaction with a potential buyer is to sell to an alternative buyer or not to sell at all. A buyer's fallback is to buy from an alternative seller or not to buy at all.

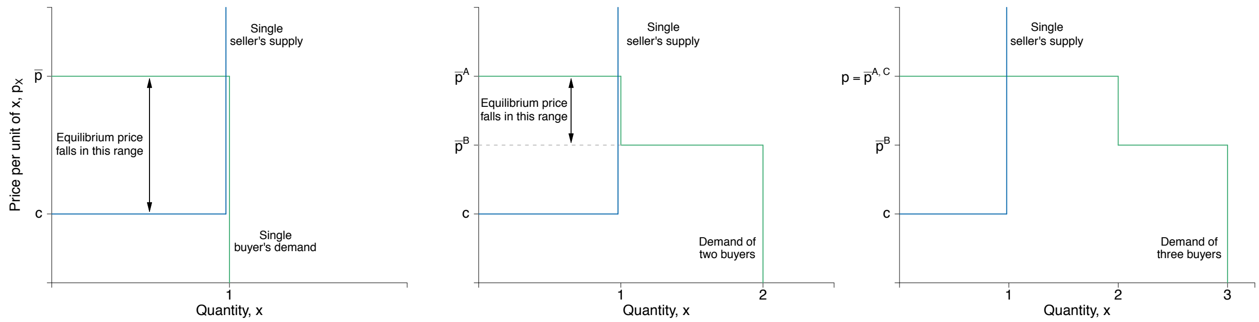
The price at which a seller can sell her good will depend on her own fallback and the fallback options of those to whom she could potentially sell the good. Similarly, the price at which a buyer buys a good depends on his fallback option and the fallback options of those from whom he could potentially buy the good.

REMINDER A fallback is each buyer or seller's next best alternative to the current exchange or social interaction. The fallback may involve not exchanging or not participating in the interaction, or it may involve an alternative, such as buying from a different store, or selling to a different customer.

Consider the three markets in Figure 9.19 where there is a single seller, Zenji, who can produce one unit of a good at a cost c. This is his minimum *willingness-to-sell*: he will not sell the good for less that it would cost him to produce it. (Because there is only one unit of the good produced we omit the "marginal" and "average" in describing its cost). The good is not divisible so Zenji must sell one unit to just one buyer.

We will consider three potential buyers: Amos (A), Bella (B), and Carlos (C). Panel a in Figure 9.19 is a **bilateral monopoly**, which we have encountered in the bargaining between Aisha and Betty in Chapter 4: there is one buyer, Amos (who could be called a monopsonist) and one seller, Zenji. In Panel b Zenji has two prospective buyers, Amos and Bella (A and B) with different maximum willingness-to-pay for the good ($\bar{p}^A > \bar{p}^B$). In Panel (c) Carlos has joined the market, so we have three prospective buyers: A and C have the same willingness to pay for the good, and B's willingness to pay is lower than theirs ($\bar{p}^A = \bar{p}^C > \bar{p}^B$).

We start with market (a) the bilateral monopoly. If the buyer and seller exchange the good, then the total economic rent that results will be sum of Amos's consumer surplus and Zenji's economic profit. This is the difference



(a) Bilateral Monopoly

(b) One seller, two buyers

(c) One seller, three buyers

Figure 9.19: **Supply and demand for one seller and the cases of one, two, and three buyers.** Figure (a) is a bilateral monopoly with one buyer and one seller. We assume buyer A has the highest willingness to pay for the good. The transaction can occur at a price anywhere in the range between the seller's cost and the buyer's willingness to pay, $[c, \bar{p}^A]$. Figure (b) has one seller and two buyers, where buyer B's willingness to pay is lower than buyer A's willingness to pay, but higher than the seller's cost, $c < \bar{p}^B < \bar{p}^A$. The availability of B as an alternative buyer narrows the range of prices over which A and the seller will negotiate, since the transaction price p will now be in the range $[\bar{p}^B, \bar{p}^A]$. Figure (c) shows what occurs when a third buyer, C, with the same willingness to pay as A ($\bar{p}^C = \bar{p}^A$) enters the market.

REMINDER Each of these scenarios is like an Ultimatum Game or a Dictator Game from Chapter 2 with different institutions determining whose actions determine the distribution of the economic surplus.

between Amos's willingness to pay and Zenji's willingness to sell: $\bar{p}^A - c$. Any price in the range $[c, \bar{p}^A]$ would satisfy both Zenji's and Amos's participation constraints because both would do better than their fallback option. The lower the price, the more consumer surplus Amos would get, and the higher the price, the more economic profit Zenji would get.

Which of the seller or the buyer obtains the larger share of the economic rent will depend on the rules of the game governing the interaction. For example if Zenji is first mover and can commit to a particular price, then he will charge $\bar{p}^A$, Amos' maximum willingness to pay, or just a bit less. Amos will accept and the total economic rent made possible by the transaction $\bar{p}^A - c$ (or virtually all of it) will go to Zenji in the form of economic profit. Amos has barely improved on his fallback position, which in this case was to not buy the good, gaining a consumer surplus of zero.

Of course had Amos been the first mover the results would have been exactly the opposite. He would have offered to buy the good at the price c, which is Zenji's fallback option, or a bit higher, capturing all of the economic rent in the form of consumer surplus.

How does an increase in demand affect the outcome of the interaction? This could take one of two forms. First, if Amos for some reason came to value Zenji's good more, then his maximum willingness to pay would be greater than $\bar{p}^A$. Now supposing that the rules of the game had not changed (Zenji is still first mover) then the result would be that Zenji would charge a higher price, and receive larger economic profit.

There is a second way the demand could increase: the entry of a second prospective buyer, Bella. Bella's willingness to pay changes Zenji's willingness to sell. Because Bella will purchase the good for any price less than or equal to $\bar{p}^B$, Zenji can credibly refuse to sell the good to Amos for any price lower than $\bar{p}^B$. Zenji's minimal willingness to sell before was c the cost of the good but now it has increased to $\bar{p}^B$ because his fallback option in dealing with Amos is no longer simply not selling the good, but instead selling it to Bella. Whatever the rules of the game are, as long as the exchange is voluntary, the result is to narrow the potential prices to the range $[\bar{p}^B, \bar{p}^A]$, and ensures that Zenji now takes at least $\bar{p}^B - c$ as economic rent from the exchange.

If Amos is first mover, the best he can do is offer to buy the good from Zenji at the price $\bar{p}^B$, a higher price than when Bella was not in the market. The increase in demand resulting from Bella's entrance to the market changed the price because it improved Zenji's fallback option. Zenji could sell to Bella if he did not sell to Amos.

Finally, Figure c. shows what happens when a third buyer, Carlos, who has the same willingness to pay as Amos, $\bar{p}^C = \bar{p}^A$ joins the market. The seller, Zenji, now has the fallback option of selling the one unit either to Amos or to Carlos at their common willingness to pay $p = \bar{p}^A = \bar{p}^C$.

Bella will not participate in the exchange once Carlos enters the market, because her willingness to pay is lower than either Amos's or Carlos's. As a result Zenji would not agree to sell her the good at a price she would accept, even though his cost of producing the good is less than the minimum she would be willing to pay to acquire the good.

The arrival of Carlos means that Zenji's willingness to sell is no longer c when his fallback option was to not sell the good (and gain zero economic profit), and it is no longer Bella's willingness to pay $\bar{p}^B > c$ when his fallback option in dealing with Amos was to sell to Bella. It is now $\bar{p}^A = \bar{p}^C$ because his fallback option in selling to Amos is instead to sell to Carlos.

The presence of the fallback buyer for the seller – either Amos or Carlos – ensures that Zenji will capture all of the economic rents in the form of economic profit.

The interactions illustrated in Figure 9.19 shows that changes in demand (including an increase in the degree of competition among the buyers) alters the price by changing the fallback option of the seller.

EXAMPLE In Chapter 11 we will see that an increase in the demand for labor will affect the Nash equilibrium age by altering the fallback option of employees. Employers will pay them more because workers' fall back options – finding another job – will be better, the greater is the demand for labor.

Checkpoint 9.10: More Sellers of the good

Imagine that instead of Amos being joined by two other buyers of the good (Bella and Carlos), two other sellers (Yao and Xiao) joined the market. Explain the role of fallback options and what the distribution of consumer surplus and economic rents would be in each of the following cases. Be sure to sketch the buyers' demand and sellers' supply in each case.

a. A second seller, Yao, joins the market with costs, c^Y, which are greater than Zenji's costs, c^Z. Amos is the only buyer.
b. A third seller, Xiao, joins the market with costs, c^X, which are equal to Zenji's costs, c^Z. Amos is the only buyer.
c. Three sellers and three buyers participate in the market. The following is true of their willingness to pay and their costs: For buyers, their willingness

to pay is $\bar{p}^A > \bar{p}^B > \bar{p^C}$. For the sellers, $c^Z < c^Y < c^X$. Also, $c^X = p^C$. Notice that Carlos (C) and Xiao (X) have different willingness to pay and costs compared to the earlier examples.

9.15 *The "perfect competitor:" Rent-seeking firms competing in and for markets*

The Cournot model of the continuum of competition from monopoly to many firms and the theory of price discrimination suggest a view of competition quite different from the model of perfect competition. In this alternative model of competition, buyers and sellers are not price takers, rather they are price makers and rent seekers of the kind that Hayek envisioned in his view of competition in the headquote of this chapter.

HISTORY Friedrich Hayek was a leading critic of centralized economic planning and an advocate of limited government (see Chapter 14). But as you know from the headquote to this chapter, he was an equally vociferous critique of the model of perfect competition.

You have already seen an example of these active rather than passive traders in the case of the seller who can perfectly discriminate in the prices they charge capturing all of the consumer surplus, and eliminating the dead weight loss associated with monopoly, resulting in a Pareto-efficient outcome. The new insight here is that monopolies implement inefficient outcomes not only because of limited competition, but because of the limits of price discrimination which prevent sellers from acting as perfect competitors. The fact that monopolies limit output in such a way that some consumer surpluses that are technically feasible are not realized is a consequence of the assumption that every unit must be sold at the same price, that is, a limitation on price discrimination.

If perfect price discrimination were possible, and in other cases like this, where active rent-seeking buyers and sellers lead to in a Pareto-efficient outcome, we refer to the actor as the Perfect Competitor, substituting this term for "perfect competition." Of course, active rent-seeking buyers and sellers are not *trying* to implement a Pareto-efficient outcome any more than the passive price-taking actors of the perfectly competitive model are.

To see how this active view of competition works, we need to broaden our view of how economic competition takes place. So far we have looked at two different ways to compete. In the Cournot model, firms compete by setting output levels to maximize their profits taking account of the effect of their output choice on the prices at which they can sell the product, given the output level of other firms. In the model of market equilibration by rent-seeking and price-setting we have just examined, firms compete by setting prices so as to capture rents that are available when markets do not clear.

HISTORY Joseph Schumpeter is considered the father of the field of innovation economics. In 1950 he wrote: “in capitalist reality, as distinguished from its textbook picture, it is not [price] competition which counts but the competition from the new commodity, the new technology, the new source of supply, the new type of organization ... – competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits ... but at their foundations and at their very lives.”

But setting prices or quantities are just two of the competitive strategies adopted by firms.

Reducing the level of competition

Odd as the phrase sounds, an important way firms compete is by attempting to reduce the number of competitors. Recall that a firm's profit increases the smaller the number of competitors there are in a market. We have already mentioned one method of reducing competition: forming a cartel (a group of firms that decide jointly on their level of output and pricing). A second competition reducing strategy is deterring entry by competitors either by buying the would be competitor or practicing what is called predatory pricing, that (recall from Chapter 8) is setting prices below cost for a period of time to discourage entry or drive out a competitor.

EXAMPLE In 2008, Amazon expressed interest in buying a small by rapidly growing e-commerce firm, Quidsi. Quidsi was a firm that specialized in baby care, household goods and beauty products. In 2009, Quidsi declined the offer by Amazon.com to buy them. Amazon immediately cut prices on diapers and baby products by 30 percent and shortly thereafter launched Amazon Mom offering free two day shipping for baby and household products. Quidsi's sales fell and their investors looked for ways to sell the business. In 2011, Quidsi was purchased by Amazon. In 2017, Amazon shut down Diapers.com, their site that had replaced Quidsi.

Entry deterrence by a monopolist in the Cournot model could be accomplished if the monopolist, say Firm A, when facing a possible entry by a competitor, Firm B, were to temporarily produce not at the monopolists profit maximizing level of output but at the level of output for which the competitor's best response is to produce nothing.

To see how this would work return to Figure 9.4. Because Firm B's best response to Firm A's selling more is to sell less, if it is willing to temporarily forego making profits, Firm A can induce Firm A not produce at all. In the figure this would require Firm A to produce 36 units. With the parameters used to generate that figure ($\bar{p} = 20$ and $\beta = 0.5$) the highest price at which this level of output could be sold is $p = 2$ which is equal to the cost of marginal and average production. This "entry deterring level of output" is the output level of the entire industry in the case of unlimited competition.

Advertising and product differentiation

A monopolistically competitive firm – the sole seller of a particular good – can devote resources on advertising to shift the demand curve for its product up (increasing $\bar{p}$). It can also increase brand loyalty so that consumers consider other similar product to be less good substitutes (less similar), thereby making the demand for the product less price elastic. The effect will be to allow the firm to raise prices with a lesser cutback in sales. Firms also invest in product design features that are primarily aimed at **product differentiation** rather than increasing the functionality of the product.

PRODUCT DIFFERENTIATION *Product differentiation* is a business practice aimed at making the firm's product appear more distinct, less similar to substitute products, and hence making the demand for the firm's product less elastic.

Consider, for example, the array of smartphones that exist on the market. Apple's iPhone uses iOS (a proprietary operating system) and the phone is sold as a high-end product, with competition from firms like Samsung with phones like the Samsung Galaxy S series using the Android operating system as a competitor. Apple adopts a variety of strategies (combined with advertising) to *differentiate* the iPhone from its competitors. (An indication of how important Apple's differentiation is can be seen in other firms' attempts to *copy* the iPhone). As with other strategies, each firm will have a best-response product differentiation reaction to other firms. In some cases, firms will differentiate their products so dramatically that they form entirely new markets. Rent-seeking product differentiation leads to *market-making*.

Lobbying

Firms engage in political activities to ensure an environment – taxes, property rights, interest rates, import tariffs, immigration rules, and other policies – that favors the owners' profits. Lobbying in some countries takes the form of contributions to political campaigns, hiring people to persuade elected officials of the firm's point of view, public advertising, and even vacation retreats for judges who might affect legal decisions affecting the firm.

Innovation

HISTORY Joseph Schumpeter (1883-1950) held that innovation occurred through a process that he termed *creative destruction*. Disruption, not equilibrium, is his central idea. The "creation" of the innovating firm "destroys" firms that do not keep up, giving the innovators at least temporary economic profits called *Schumpeterian rents*. These are competed away when other rent seeking firms come up with new ideas or duplicate the ideas of innovating firms. Notice: the process of diffusion of innovations works because innovating firms do not have intellectual property rights sufficient to make them sole monopolists of the new idea.

Improvement in technology, we saw in the Chapter 8, can create innovation rents for those who adopt the new methods first. These rents are temporary until competitors manage to adopt the same new technologies, or find equivalent alternatives. Another competitive strategy is to combine innovations with attempts to prevent other firms from following. Intellectual property rights (patents, copyrights, or trademarks) establish a firm's monopoly over a production method, an idea or a name, and prolong the period over which the innovation rents will flow to the firm. Like product differentiation, intellectual property rights are a competitive strategy that works by limiting competition.

Innovations need not be in technology; a new form of organization or institution may reduce costs and allow a firm to capture economic rents until others copy it. An example is outsourcing by firms, purchasing from other firms inputs previously produced within the firm (such as the components of an iPhone, or a car's engine). This strategy allows firms located in high wage countries to acquire inputs at low cost from countries with lower wages. During the Industrial Revolution, the factory itself was an organizational innovation. Spinning yarn, weaving cloth, stitching cloths, and other kinds of work had previously been done mostly in people's homes under what was called the "putting out" system. The industrial factory brought hundreds of workers together under the direct supervision of a single management, creating new opportunities for cooperation and control in the labor process.

Market making: Competing for markets not in markets

Lastly, consider *market-making*, the creation of new markets to generate new rents.

We look at two examples to illustrate the idea of market-making.

Figure 9.20: **A Mandarin Red Regency TR-1 Texas Instruments transistor radio. Courtesy Joe Haupt, CC 2.0 Sharealike**

Let's go back to the mid-1950s and relate a story about how consumers got to have transistor radios. Pat Haggerty of Texas Instruments, having found out about the invention of transistors, managed to negotiate a license for the patent for transistors from Bell Labs. When he purchased the license for $25,000 dollars, transistors were being sold to the military for $16 each. Haggerty told his engineers at Texas Instruments that they would have to produce a transistor for $3 to make it possible to sell to consumers. Haggerty made this declaration in June 1954, telling his engineers they were going to market by November 1954 – just in time for Christmas. They made it. Texas Instruments started selling a transistor radio – The Regency TR-1 – for $49.95. The Regency TR-1 came in four colours (black, ivory, Mandarin Red, and Cloud Gray) and it could fit in your pocket. At the time, the other radios on sale were all large, clunky pieces that took up space on the kitchen counter or living room side table.

HISTORY Hayek stressed that competition is a social relationship among people. "Especially remarkable in this connection is the explicit and complete exclusion from the theory of perfect competition of all personal relationships existing between the parties. The fact that our incomplete knowledge of the available commodities or services is made up for by our experience with the persons or firms supplying them – that competition is in a large measure competition for reputation or good will – is one of the most important facts which enables us to solve our daily problems." Hayek "The Meaning of Competition," 1948.

Few businessmen at the time thought people wanted a pocket-sized radio, but Haggerty knew better and engaged in remarkable market-making. By the end of 1955, over 100,000 units had been sold. Haggerty was also lucky that Elvis Presley's song "That's All Right" began blaring over the airwaves in 1954, and everyone wanted to hear it on their own transistor radio.

The Regency TR-1 was the iPod of the 1950s, taking advantage of the advent of Rock and Roll, and personalizing access to music. No one knew they wanted one, until they did.

HISTORY Our thinking is influenced by the work of Makowski and Ostroy "Perfect competition and the creativity of the market" (2001), a modern day restatement of ideas of Hayek. Ronald Coase, and Joseph Schumpeter. The term "perfect competitor" is from Makowski and Ostroy.

Let's track back from the iPod of the 1950s to the iPad of 2010. Apple tried to sell a tablet once before in 1993 - the Newton, but it tanked. Microsoft produced a tablet in 2000 called the Microsoft Tablet PC, but it didn't catch on either. Though many firms tried to produce a tablet that would satisfy customers, the market for tablets did not take off until the Apple iPad.

Apple's iPad was released in 2010 and many people wondered why consumers would want "a bigger iPhone that can't make phone calls." But these skeptics were wrong. Having released the iPad in April of 2010, by 2011 Apple had sold almost 15 million iPads. No one outside of Apple had predicted such success.

Apple's *market-making* with the iPad paved the way for competition in a new market, with Microsoft re-entering with the Surface Pro, Google selling the Google Nexus, Samsung the Galaxy tab, and Amazon the Amazon Fire Tablet (among others). Once Apple had made the market, other firms entered it in order to obtain a share of the economic rents on the tablet market.

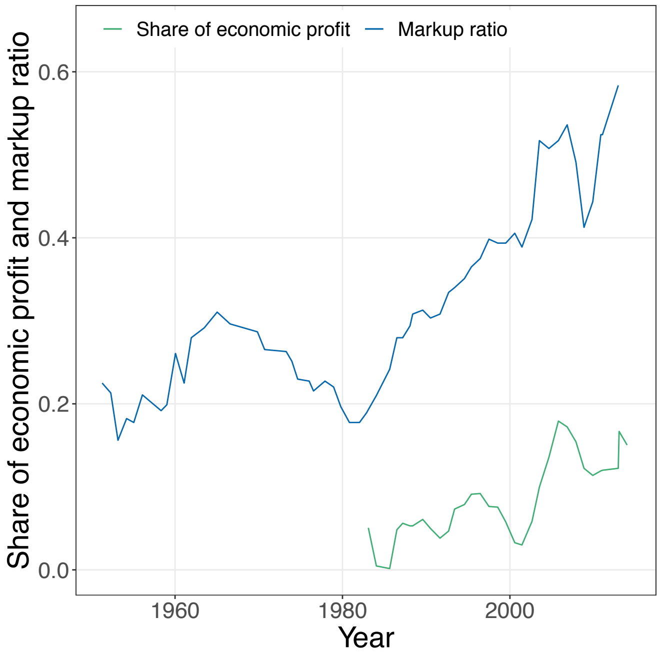


Figure 9.21: **Rising markup ratio and economic profit as a share of total income**. The markup ratio is $\mu = \frac{p-c}{c}$ as defined earlier. It equals 0 when $p = c$ and increases beyond that and is greater than zero when $p > c$. Data are from Loecker et al (2017). The markup ratio is for the entire U.S. private economy with estimates for each sector weighted by their share of total sales. The share of economic profits is defined as the share of accounting profits minus the share representing the opportunity cost of capital. Due primarily to falling interest rates, the latter has declined substantially over this period. The share going to wages and salaries has also declined. Data are from Barkai (2018). As the figure shows, the markup ratio and the profit rate move in the same way, suggesting a strong relationship between decreasing competition and increased profitability of firms.

9.16 Application: Declining competition and public policy

The many forms of competitive rent seeking just described may lead to a reduction in the degree of competition. In the four decades following 1980 this appears to have been the case in the United States.

There are many ways that these trends can be measured. Focusing on the U.S., the revenue of the top 50 firms as a share of the industry total revenue has risen in many sectors including retail trade, finance and insurance, real estate and leasing, wholesale trade, transportation and warehousing. Over this period, the number of new firms entering industries has fallen and as a result the age of firms has risen. Since the early 1980s, the fraction of employment in firms less than five years old has been cut almost in half.

A better overall measure of the decline in the degree of competition in the U.S. economy is the the extent to which price exceeds marginal costs, which we have measured by the markup ratio, namely the excess of price over costs, divided by costs. The blue line in Figure 9.21 shows an estimate of the markup ratio across the U.S. economy since the middle of the last century.

It rose during the period of strong demand from mid-century to the mid 1960s and then declined as the growth of the economy slowed during the period called stagflation. But from 1980 onward, it has risen by over 40 percent, consistent with a dramatic reduction in competition.

FACT CHECK At the turn of the century economists regarded the European economy as less competitive than the American. But this relationship has dramatically reversed. In the telecommunications sector, for example, measures of competition decreased substantially in the U.S. while rising in Europe. Consumers have felt the difference. The average monthly cost of broadband in the U.S. in 2018 was double that in France and Italy.

As the markup measures the excess of a firm's sales revenues over its costs including the opportunity cost of capital, it is no surprise that economic profit as a share of national income has also greatly increased. This is shown in the green line in the figure.

Trends towards a less competitive economy in Europe are less pronounced, in large measure because the European Union has much stronger policies to ensure competition.

The Cournot model and other models of markets demonstrate the deadweight inefficiency and losses in consumer surplus that occur when a firm with a limited number of competitors faces a downward-sloping demand curve and therefore act as a price maker.

To address these consequences monopoly, duopoly, oligopoly and other cases of limited competition governments have pursued **"competition policies"** often breaking up large firms into smaller firms. By breaking up firms, competition policies increase n, the number of firms. As a result, they lower the price markup over costs, reduce profits, and increase the consumer surplus.

COMPETITION POLICIES deal with market concentration and large firms in industries by breaking up large firms into smaller firms to increase competition. Alternatively, they also exist to *prevent* mergers or acquisitions of firms to ensure that markets do not become *more* concentrated with less competition.

Models and policies of this type have motivated anti-trust legislation such as the Sherman Act (1890) in the United States, and the breakup of the giant Standard Oil company by a decision of the U.S. Supreme Court in 1911. Similar policies have been pursued by the European Commission, and Competition Commissions in South Africa, India, and many other countries. The interest of consumers is served by the increase in the number of firms as long as the smaller firm size is not associated with significantly higher cost.

HISTORY Senator John Sherman called the act in his name "a bill of rights, a charter of liberty" and in 1890 on the floor of the U.S. Senate said: "If we will not endure a king as a political power, we should not endure a king over the production, transportation, and sale of any of the necessities of life. If we would not submit to an emperor, we should not submit to an autocrat of trade, with power to prevent competition and to fix the price of any commodity."

We will return to the question of competition policy in Chapter 15 and the challenge of what are called the "modern monopolies" such as Amazon, Microsoft, and Apple.

Checkpoint 9.11: Cournot model with decreasing costs

Consider the firm with a fixed cost of c_0 and no variable cost.

1. Use Cournot analysis of the monopolistic firm to determine the profit maximizing level of the firm's output.
2. Redraw the left panel of Figure 9.12 for duopolisic firms A and B assuming that they have a fixed cost of c_0 and a variable cost of zero.
3. Explain the economic reasons (not just the math) why the Nash equilibrium levels of output change and are different from those in Figure 9.12.

9.17 Conclusion

Making connections

Mutual gains and conflicts of interest over their distribution: Here, conflicts of interest occurred a) among firms' owners over the distribution of economic profit and b) between owners as a group and consumers over the distribution of the total rents (gains from exchange) in the form of economic profit as opposed to consumer surplus.

Institutions policies and rules of the game: How these conflicts are resolved depend on barriers to entry and other influences on the degree of competition that are affected by public policies such as intellectual property rights and competition policies.

External effects and coordination problems: The coordination problem facing the owners of competing firms is identical to the "fishermen's dilemma" that you encountered earlier. In both cases the production by each of the actors (fishers, firm owners) reduces the benefits enjoyed by the others. They could all be better off if they could coordinate and all produce less. A difference with the fishermen's dilemma setting is that in the Cournot model we consider the interests of consumers who would be better off if the firms produced *more.*

Rent-seeking, price-making and strategic behavior: Profit seeking is an active and strategic process of wage and price making (including discrimination), product differentiation, innovation, securing a favorable legal and policy environment and much more. This contrasts with the passive view of price taking competitors, whose decisions are confined to the amount to be bought or sold.

Important ideas

competition	the firm	industry
profit-maximizing best response	average firm output	short-run Cournot/Nash equilibrium
monopoly	duopoly	oligopoly
price	marginal cost	price and marginal cost in equilibrium
principle of profit maximization	marginal revenue	profit
best-response function	cartel	merger
opportunism	Prisoner's Dilemma	Pareto-efficiency
price-taking	market demand and supply curves	price-taking equilibrium
perfect price discrimination	price gouging	statistical discrimination
price above marginal cost	coupons, miles, and advertising	customer loyalty programs
rent-seeking & rent-seeker	appropriation	fallback
positive feedbacks		multiple equilibria
free-entry and exit	long-run industry equilibrium	opportunity cost of capital
perfect competition	perfect competitor	full appropriator
modern monopoly	first-copy cost	network economies of scale
demand-side economies of scale		

Mathematical Notation

Notation	**Definition**
x	output of an individual firm
X	industry output
p	price
$\hat{p}$	expected price with barriers to entry
β	slope of the industry demand curve
c	marginal cost
$r()$	revenue function of a firm
π	profits of a firm
Π	industry profits
n	number of competing firms in industry
b	probability of failure of a new firm
μ	markup ratio over costs of production
ρ	opportunity cost of capital

Note on super- and subscripts: A, B: firms; N: Nash equilibria; m: monopoly; C: Pareto efficient; *: agreed allocation; O: opportunistic behavior; V: victim of opportunistic behavior; S: supplied; D: demanded; H: high; L: low.

Discussion questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

See reference list.

Part III

Markets with Incomplete Contracting

> Not everything in the contract is contractual.... the contract is not sufficient in itself but is possible only thanks to a regulation of the contact, which is social in origin.

Emile Durkheim, *De la Division du Travail Social (The Division of Labor)* 1967 [1893]:189, 193.

In most economic interactions there are some important bits of information that are known to some of the actors but not to the others. The employee knows how hard she worked yesterday, her employer may not. The borrower knows how the loan will actually be used – prudently or wrecklessly, for example – the banker may not. These are examples (as you already know) of asymmetric information.

Other bits of information may be known to a buyer or seller, but not admissible as evidence in a court of law (that is, non-verifiable). In most legal systems, for example, the employer's account that he found the worker asleep at her desk, unless substantiated by witnesses, would not be considered to be verifiable and could not be used as evidence against the employee for example to recover the wages paid to the employee for her nap time.

Where critical information is either lacking or cannot be used in court, contracts can not cover everything that matters in an exchange, or they may simply be unenforceable. A result is that some things that a buyer or seller cares about in a transaction are not subject to contract.

This explains Emile Durkheim's statement: contracts do not cover everything and they do not enforce themselves. Durkheim (1858-1917), considered (along with Max Weber , 1864-1920) a founder of sociology, stressed the importance of social norms (rather than the law) in making mutually beneficial exchanges possible.

The information problems and the incomplete contracts to which they give rise are illustrated by the difficulty of contracting for quality in Chapter 10. You purchase a made to order software package capable of executing a series of calculations and creating visual representations of the results. But how good, really, is the resulting code? You and the software developer will often disagree. And there is no way that the exact capabilities of the code could have been specified in an enforceable contract (how cool the visuals had to be, how quickly the algorithms involved would compute which kinds of results, and so on).

In situations like this the contract will necessarily be incomplete, and you will see that this fundamentally changes how markets work. This is why we devote separate sections of our book to markets with complete contracts (the more conventional subject matter of undergraduate economics texts) and those with incomplete contracts.

The employer and the employee in the labor market illustrate a similar problem in Chapter 11. The employer effectively rents your time, but he cannot purchase and contractually enforce exactly the tasks he needs you to do in order to make a profit. In Chapter 12, we show that banks and other lending institutions face similar problems of incomplete or unenforceable contracts in the credit market.

In these and other markets where enforceable contracts do not cover everything that matters in a transaction, mutually beneficial exchanges can nonetheless take place. Social norms may step in to facilitate mutually beneficial exchanges where contracts are incomplete. If the employee's work ethic commits her to doing the best she can even when her employer is not looking, then the fact that hard work cannot be specified in a contract need not stand in the way of an exchange. Similarly, if the borrower's truthfulness in describing the project which the loan will finance, commits him to not undertaking excessively risky options, the prudent use of loaned funds will occur, even if it cannot be enforced in a contract.

In these and other important markets, mutually beneficial exchanges are made possible by some combination of contract, social norms, along with a third element. This is the exercise of power by principals (lenders, employers) over agents (borrowers, employees). In these markets, principals typically set prices (wages, interest rates and prices) so that agents receive a payment above their next best alternative, that is a rent. Because the agent will lose the rent if the principal terminates the relationship, the employee, lender or other agent has a good reason to work hard, use borrowed funds prudently and othewise to behave in the principle's interest, even if it is not required by contract. The threat of losing the rent is the basis of the principal's power.

The critical role of the agents' social norms (work ethic, truthfulness) and the exercise of power by principals in facilitating exchanges where contracts are incomplete are features that do not appear in markets with complete contracts. Another key difference is that where contracts are incomplete, markets typically do not clear in competitive equilibrium. A consequence is that unemployment – that when labor supply exceeds labor demand – is to be expected even under ideal competitive conditions and need not be explained by "sticky wages" or other "market frictions."

Taking account of the fact that contracts are often incomplete leads to new insights about how markets really work. And so, looking back to the epitaph for the previous part of this book, to Voltaire's amazement at how markets facilitate exchanges even among people who in other contexts might be at war, we add a caveat. This is that where contracts are incomplete exchanges are facilitated if one of the parties (the principal) has power over the other party.

And we can underline Voltaire's observation: buyers and sellers care very much about the social norms of the people they are trading with. Trust matters. As Voltaire put it: "There the Presbyterian confides in the Anabaptist, and the Churchman depends on the Quaker's word." And so the exchange process is far from anonymous, it is often very personal. This is why when looking for a used car to buy, we typically go to somebody we know or a seller with a good reputation, not to a total stranger.

Because the exercise of power by principals over agents and social norms are essential to how markets work when contracts are incomplete, modern economists draws on the insights of political science, sociology, psychology and the other social sciences. This is why we make occasional reference to these disciplines.

10

Information: Contracts, Norms & Power

> In an economic theory which assumes that transaction costs are non-existent, markets have no function to perform and it seems perfectly reasonable to develop the theory of exchange by an elaborate analysis of individuals exchanging nuts for apples in the edge of the forest or some similar fanciful example . . .
>
> Ronald Coase, *The Firm, The Market, and the Law*, 1988: 7-8.

DOING ECONOMICS

This chapter will enable you to:

- Understand the difference between complete and incomplete contracts, and why, due to limited information, many contracts are incomplete.
- Model an interaction of a principal and an agent in which their interests conflict concerning some aspect of the the exchange that is not subject to a complete contract.
- Distinguish between hidden actions and hidden attributes as a source of the contractual incompleteness
- Show in the hidden actions case that principals will set prices so that agents receive a rent, the market does not clear, and the outcome is Pareto-inefficient due to the external effects of the agent's actions that are not subject to contract.
- Explain why, when contracts are incomplete, social norms and the exercise of power by principals facilitate mutually beneficial exchanges.
- Understand the conflict of interest between the principal and agent, and how the information available to the principal affects the distribution of the gains from exchange between them.
- See how the nature of the contract affects social aspects of exchange such as trust and repeated interactions.

UNITED COLORS OF BENETTON.

Figure 10.1: **Trade marks can reduce uncertainty about the quality of goods.** Our "Benetton model" illustrates this case.

10.1 Introduction

When we talk about the market for, say, wheat, or toothpaste, what are we talking about? To a farmer, there is really no such thing as "wheat." There are literally hundreds of different species called "wheat," and until recently a farmer's crop might be a mixture of quite a few of them. This makes buying and selling "wheat" difficult because while the farmer knows what he is selling, the buyer does not.

But we do have markets for a limited number of species and grades of wheat. These markets came about not because nature conveniently produced a standardized product called "hard red winter wheat #2" but because an economic institution – the Chicago Board of Trade – in the mid-nineteenth century adopted classifications and policies to create homogeneous categories of grain so as to facilitate such transactions.

Wheat once referred to a diverse collection of products, with size, genetic strain, and quality differing from one sack of wheat to another. For two farmers – Jones and Svenson – the supply, demand, and price for farmer Jones's wheat differed from the corresponding supply, demand, and price for farmer Svenson's wheat. The *markets* for the two farmers differed even though they both sold wheat.

But, thanks to the Chicago Board of Trade, different grades of white winter wheat, red winter wheat, spring wheat, and many other standardized categories came to be of such uniform quality that the ownership of grain no longer referred to any specific sack or particular lot of wheat, but to a contract entitling the owner to the delivery of specified amount of some particular grade of wheat. As a result the biodiversity of wheat on the American great plains fell, but a limited number of new markets, resembling what one sees in an economics textbook, came into being.

Each type of grain has become a homogeneous good, for which traders can write *enforceable* and **complete contracts** – simply for an *amount* without worrying about the quality. The categories of grain are now like electricity: you purchase any amount of it by the Kilowatt hour without caring or knowing which power plant generated it. You do not worry about the quality of the electricity you buy. What you buy is what you get. And if you paid for electricity that you did not receive, you can get your money back.

Grain is not unusual in this respect. Goods like Sugar Number 11, Corn Number 2 Yellow, or Light LA Sweet (that's crude oil) are not gifts of nature. Rather, they are created by a deliberate *process of standardization* to eliminate difficult-to-monitor differences in quality.

HISTORY Remarkably, this standardization of grain trading was accomplished by an entirely private body, the Chicago Board of Trade. Memberships in this body would themselves become marketable commodities before the nineteenth century ended.

10.2 Incomplete contracts: "... not everything is in the contract"

COMPLETE CONTRACT A contract is *complete* if it a) covers all of the aspects of the exchange in which any party to the exchange has an interest, and b) is enforceable (by the courts) at zero cost to the parties. A contract is *incomplete* if it lacks either of these two features.

But unlike electricity and red winter wheat #2 much of what is transacted in a modern economy are not homogeneous goods or services. Instead, in many markets, contracts are *incomplete*: traders cannot easily verify quality, or they can observe the quality but such quality cannot be verified in a court of law or by some other third party to enforce the contract. Think of hiring someone to care for an elderly relative or to take care of your children while you are at work. How can you know the quality of the care they are giving? And even if you later found out that they had been careless, could you go to court to get back the wages you paid them?

Contract enforcement by courts or other third parties is called **exogenous enforcement** because the enforcers are not parties to the exchange (they are *outside* the exchange). But for many exchanges this is not the case: the contract is incomplete.

ENDOGENOUS ENFORCEMENT OF CONTRACT *Exogenous enforcement* of the terms of an exchange is done by courts or another third party – not the parties to an exchange themselves – and is the defining characteristic of a complete contract. (The prefix "exo" means outside or external as in exotic, exodus, or exogamy, the practice of restricting marriage to members outside of one's own group). When the parties to an exchange – employers and employees, buyers and sellers, borrowers and lenders – themselves adopt strategies to ensure favorable terms of an exchange for aspects of it not covered by a contract enforcement is *endogenous*.(The prefix "endo" means within or internal; for example endogamy is the practice of restricting marriage to members of one's own group.)

Complete and incomplete contracts share the two features:

- that a transaction is based on the mutual expectation of gain by all the participants and
- that there is a conflict of interest over how these gains will be divided among the parties to the exchange.

The key difference is that when contracts are incomplete the division of the gains is not enforced entirely by an external body – the courts – based on terms specified in the contract. The terms of the contract matter and courts may be involved, but the outcome of the exchange is also the result of strategic interactions among the participants involving rewards, punishments, and the exercise of power. These include threats, promises, and the creation of incentives through offering repeated interactions. The preferences of the parties to an exchange also matter, for example a commitment to telling the truth about the condition of the used car you are selling, or a worker's intrinsic motivation to do high quality work.

This what Durkheim meant – quoted in the introduction to this section of the book – when he said that "not everything in the contract is contractual" and that market exchanges are socially regulated. We call this the **endogenous enforcement** of the terms of an exchange.

Often, contracts are incomplete and the terms of the exchange are enforced endogenously because information is asymmetrical and non-verifiable. Information is **verifiable** if it can be used in court to enforce a contract. Non-verifiable information such as hearsay, or even direct but uncorroborated eyewitness observation, generally cannot be used to enforce contracts. Information is asymmetric if something is known by one party but not by another. This affects the kinds of contracts that can be enforced because a party's information about their own attributes or actions may be *private information*.

VERIFIABLE INFORMATION Information is *verifiable* if it can be used in court to enforce a contract. Non-verifiable information such as hearsay, or even direct but uncorroborated eyewitness observation, cannot be used to enforce contracts.

ASYMMETRIC INFORMATION Information is asymmetric if something is known by one party but not by another. This affects the kinds of contracts that can be enforced because a party's information about their own attributes or actions may be *private information*.

Students are experts on incomplete contracts. Suppose as a condition of your employment at a consulting firm following graduation you contracted to "learn microeconomics, "How would it be determined that you had fulfilled your obligation? The contract might have been written "pass this particular course in microeconomics" but the employer would hardly be satisfied that this would guarantee that you were able to do the kinds of work they need.Or suppose you did poorly on the exam. Had you failed to learn economics? Or did the exam not test your knowledge of microeconomics? Or did you have a particularly bad allergy attack the day of the exam? How could your employer ever enforce this contract?

Here are some other examples of exchanges under incomplete contracts.

- Owners of firms want managers to maximize the value of the owners' assets, but managers have their own objectives (first class air travel, lavish offices) and managerial contracts fall far short of having an enforceable requirement to maximize the owner's wealth.
- Families devote a sizable fraction of their budgets to purchasing educational and health services, the quality of which is rarely specified in a contract (and would be unenforceable if it were).

Three of the most important examples of incomplete contracts are the subject of this chapter and the next two chapters. Here we consider the case where information on the quality a good is known to the seller but not to the buyer

– whether it is a used car or a piece of clothing provided to Benetton by a subcontractor. In Chapters 11 and 11 we study incomplete contracts in the labor market where information on the effort you put into your job is not readily available to your employer, and the credit market in which your promise to repay the money you borrowed may be unenforceable if you are broke.

As these examples suggest contractual incompleteness is the rule rather than the exception in economic transactions. Here are five reasons why.

1. *Asymmetric or non-verifiable information*: Third-party enforcement of contracts requires information that is available to both parties and can be *verified* by third-parties such as courts of law. **Verifiable information** is often not available, on, for example, the quality of service provided or whether the poor condition of a rented apartment is due to normal wear and tear or the tenant's negligence. It is often the case that a party to an exchange simply does not have the information (whether verifiable or not) necessary to enforce a contract: the insurance company may not know that you are already ill when you seek to extend your coverage. This is a case of asymmetric information.

2. *Time*: A contract is generally executed over a period of time as when a contract specifies that Party A does X now and Party B does Y later. But what if what B does later depends on other things that cannot now be determined? A complete contract must specify what the parties must do in every possible future situation or contingency or "state of the world." In general, people cannot completely specify these future states, and in any case, it is not ordinarily cost-effective to specify what to do in each contingency.

3. *Measurability*: Many of the services or goods involved in the exchange process are inherently *difficult to measure* or to describe precisely enough to be written into a contract. The restaurant owner would like his serving staff to interact in a pleasant manner with customers, but how can this be observed by the owner, and even if it were to be observed, how could it be measured or considered to be verified for use in a legal proceeding?

4. *Authority*: For some transactions there is no institution – no court or other relevant third party – capable of enforcing a contract. Many international transactions are of this type. For example, if a country defaults on its debt to international creditors, no third party enforces the claims on the debt. This has happened in a variety of countries internationally, such as Lebanon's default on $1.2 billion of Eurobonds in 2020, or Argentina's debt re-structuring that took place repeatedly during the period 2005-2016.

5. *Motivation*: Even where the nature of the goods or services to be exchanged would permit a more complete contract, traders may favor a less complete contract for motivational reasons. Intrusive surveillance of work-

REMINDER In Chapter 2 the imposition of a fine on parents who were late in picking up their children from a day care center is an example of "making the contract as complete as possible". Recall that the result was that more parents were late; completing the contract back-fired.

ers by employers to establish verifiable information on work activities, for example, may backfire if the employer's distrust angers employees, leading to less satisfactory work performance.

As the final reason suggests, how incomplete a contract will be is in some measure a matter of choice. For example, where the parties to an exchange are trusting and trustworthy people committed to reciprocity in their dealings, perhaps having reciprocal or altruistic preferences like those we saw in Chapter 2, they may deliberately leave some important aspects of the exchange unspecified even when the relevant enforceable contractual clauses could be written.

The first and fourth reason for incomplete contracts above – lack of verifiability and authority – make it clear that whether a particular good or service is subject to complete contracting will differ from one legal system to another. The completeness and enforceability of contracts depends on legal institutions in other ways as well. The ability of a lender to enforce a debt contract against a borrower may be greatly influenced by whether legal institutions include bankruptcy or other forms of limited liability that protect some of the borrower's assets from being taken by the lender, or, at the other end of the spectrum, imprison delinquent debtors on behalf of creditors.

Checkpoint 10.1: The Five Reasons

1. Which of these five reasons why contracts are incomplete are involved in the examples given at the beginning of this section? Hint: more than one reason is typically involved.
2. Come up with your own example of an incomplete contract and explain why one or more of the five reasons apply to it.

10.3 Principals and agents: Hidden actions and hidden attributes

A **principal-agent relationship** (also called an *agency problem*) arises when two conditions hold:

- *Conflict of interest*: the actions or attributes of the agent affect the payoffs of the principal in such a way that there is a conflict of interest between the principal and the agent. The employer, for example would like the employee to work harder; the employee would like to go home a little less exhausted at the end of the day.
- *Incomplete contract*: the agent's actions or attributes are not known to the principal (or, if known, they are not verifiable) and so cannot be subject to enforceable contract. How hard the worker works cannot be specified in an enforceable contract.

PRINCIPAL-AGENT RELATIONSHIP A principal-agent relationship (also called an *agency problem*) arises when two conditions hold:

- *Conflict of interest*: the actions or attributes of the agent affect the payoffs of the principal in such a way that there is a conflict of interest between the principal and the agent.
- *Incomplete contract*: the agent's actions or attributes are not known to the principal (or, if known, are not verifiable) and so cannot be subject to enforceable contract.

EXAMPLE Many of the economic interactions between members of different economic classes – for example, employers and employees, landlords and tenant farmers – are principal-agent relationships.

Good or service	*Principal/agent*	*Non-contractual action or attribute*	*Hidden action or attribute*	*Examples of strategies by principals to get agents to act in ways favorable to them.*
labor services	employer/agent	labor effort care	action	contingent renewal
managerial services	owner/manager	effort, maximizing owners' profits	action	profit sharing, contingent renewal
private debt	lender/borrow	level of risk taken	action	collateral shared control
sovereign debt (between nations)	lending government/borrower government	probability of default	action	trade sanctions other intervention
goods	buyer/seller	product quality	action	contingent renewal by buyer
public policy	citizen/ government citizen	policy choice and implementation	action	contingent renewal referendum
residential tenancy	landlord/tenant	care of residence local amenities	action	security deposit contingent renewal
agricultural tenancy	landlord/tenant	labor effort and quality of land	action	shared residual claimancy
equipment rental	owner/renter	care of the equipment	action	deposit, ownership share in equipment
car insurance	insurer/insured	driving habits and competence	action and attribute	higher rates for younger drivers or those with previous accidents
second hand cars	buyer/seller	quality of car	attribute	purchases only from sellers with good reputations
health insurance	insurer/insured	pre-existing health of insured	attribute	required medical exam as a condition

Table 10.1: **Principals and agents: Hidden actions and attributes** These are a few of the applications of principal agent models. In this chapter we illustrate the model by the problem of variable and difficult to measure quality of goods. In Chapters 11 and 12 we look at labor services and private debt. Political scientists have used similar models to study how citizens (the principals) can hold accountable public officials (their agents).

Both conditions are necessary. If there were not a conflict of interest, then the agent would simply do what the principal desired (both would desire the same thing) without an enforceable contract. It would be as if the principal himself carried out the necessary action.

If a complete contract covered all of the agent's actions that mattered to the principal, then a conflict of interest would not make the relationship special: "purchasing" the agent's action would be no different from the principal buying a shirt. The models of exchange with complete contracts – used earlier in the book – would be perfectly adequate.

REMINDER In Chapters 8 and 9, we explained the how buyers' willingness to pay and sellers' willingness to sell result in consumer surplus and economic profits – the mutual benefits from exchange. These models assumed that contracts were complete, such as for the purchase of wheat, or other commodities for which the quality was readily verifiable. We can use the same concepts for cases in which contracts are incomplete.

We can classify principal-agent problems into two broad categories:

- *Hidden actions*: These are things that an agent does that the principal has some interest in, but does not know (or lacks verifiable information about), such as the effort of an employee or the business practices of a borrower. An example is that a person whose home is fully insured against fire may take less care to avoid fires. Insurance companies call this behavior *moral hazard*, and that term is sometimes used to apply to any hidden action problem.
- *Hidden attributes*: There are characteristics of an agent that the principal has some interest in, but does not know (or lacks verifiable information about) such as which drivers are reckless, or which patients are seriously ill.

EXAMPLE Patronage or patron-client relationships are common in politics and have also been studied in principal agent models. The patron – perhaps a political leader – is a principal who interacts with a subordinate, the client, who is the agent. The client provides difficult to monitor services (loyalty in political conflicts) in return for well-defined compensation (patronage jobs, money transfers or access to public services).

To understand how mutually beneficial exchanges take place without complete contracts and how these benefits are divided among the parties, we need to use the concepts we have already developed – best response, first mover, fallback option, and non-clearing markets – to construct a new set of analytical tools called principal agent models. Especially important among the methods you have already learned is the theory of repeated games introduced in Chapter 5. Most jobs are not one-shot interactions: they go on year after year. So the game between the employer and the employee is *repeated*, and what each party does in one period depends on what they expect to happen as a result next period. Models based on repeated games of this type are used to study transactions between employers and employees, lenders and borrowers, and a wide range of other exchanges as shown in Table 10.1. They range from the landlords and sharecroppers that we mentioned in Chapter 2 to a fundamental problem of democracy: citizens trying to control their governments.

REMINDER In a repeated game what is called the stage game – like a simple prisoners' dilemma - is played more than once with the same players.

In the right-hand column we list some of the strategies followed by principals to get agents to act in their interest. In the first row – about employers as principals and employees as agents – the term *contingent renewal* means that because employment is repeated game, the principal has the option to not *renew* the relationship (fire the employee) and whether this happens is

Day	Number of vehicles	Maximum vehicle value ($)	Total value divided by number of vehicles	Estimated value	Choice of highest value vehicle owner	Choice of lowest value vehicle owner
1	10	9,000	$\frac{45,000}{10}$	4,500	Leave	Remain
2	9	8,000	$\frac{36,000}{9}$	4,000	Leave	Remain
3	8	7,000	$\frac{28,000}{8}$	3,500	Leave	Remain
⋮	⋮	⋮	⋮	⋮	⋮	⋮
9	2	1,000	$\frac{1,000}{2}$	500	Leave	Remain
10	1	0	$\frac{0}{1}$	0	Leave	Leave

Table 10.2: **The market for used cars and the choices of the buyers and sellers**. Buyers are willing to pay up to the full value of the car. Sellers are willing to sell their cars if they can get more than half of the true value of the car. The surplus of a transaction – the mutual benefit from the exchange – is the willingness to pay minus the willingness to sell. The seller of the maximum value vehicle observes the estimated price and as this is below his willingness to sell, he will leave the market. The same is true for the second most valuable car owner, who will also leave, and so on as the days progress, sellers leave, and the market unravels.

contingent on (depends on) whatever information the employer has on the employee's job performance.

Checkpoint 10.2: Asymmetries and hidden-ness

a. Think back to previous chapters. What were the different kinds of asymmetries we saw between players in different games or social interactions? Identify each and consider who was in each role and what the benefits were.
b. What is an example of a hidden **attribute** (characteristic of a person) that a consumer, worker, or borrower might want to keep hidden. Explain why.
c. What is an example of a hidden **action** (choice) that a consumer, worker, or borrower might want to keep hidden. Explain why.

10.4 *Hidden attributes and adverse selection: The Lemons Problem*

A "lemon" (in American slang) is a used car you discover is defective after you buy it. A "peach" is a used car you discover works better and costs less to maintain than you expected after you buy it. The problem the existence of lemons poses to economic markets can be illustrated in a model of a used car market where the principals are the prospective buyers, the agents are the sellers, and the hidden attribute (whether the car is a lemon) is known only by the seller.

Consider the following example (summarized in Table **??**):

- Every day, 10 owners of 10 used cars consider selling.

- The cars differ in quality, which we measure by the monetary value of the car to its owner. Quality ranges from zero to 9,000 in equal steps: there is one worthless car, one worth 1,000, another worth 2,000, and so on. The average value of the cars is therefore 4,500.
- There are many prospective buyers, and each would happily buy a car for a price equal to its true value, but not more, which is their *willingness to pay* for a car.
- Sellers do not expect to receive the full value of their vehicle, but they are willing to sell if they can get even just a little more than half the true value, which is their *willingness to sell* for the car.
- The potential economic surplus – the sum of buyers' and sellers' surpluses – will be half the price of the car.

If prospective buyers can ascertain the quality of each car, and approach each seller to bargain over the price, by the end of the day all of the cars (except for the entirely worthless one) can be sold at a price somewhere between their true value and half the true value. If information were complete, then all mutually beneficial trades would take place.

But if potential buyers cannot ascertain the quality of any particular car that is for sale, the market will not work. Suppose that those who bought cars yesterday find out the true value of their purchase and post it on social media. Then today, the potential buyers will know the true value of the cars sold the previous day. They still do not know the true value of any of the cars for sale today. But they might reasonably adopt the rule that the most they are willing to pay for a car today will be the *average value* of the cars sold yesterday.

HISTORY In 1970, George Akerlof wrote a paper, "The Market for Lemons" in which he developed the model we present here. Before being published in the *Quarterly Journal of Economics*, the paper – now one of the most highly cited in economics and the main reason for Akerlof's Nobel Prize – was rejected by the *American Economic Review* and the *Review of Economic Studies* as "trivial" and by the *Journal of Political Economy* due to one referee's opinion that if Akerlof were correct than much of economics would be wrong! Apparently the quality of an economics paper is at least as difficult to judge as is a used car.

Now suppose that 10 cars had been offered on the market on the first day. We use a proof by contradiction to show that, one by one, the highest quality cars will drop out of the market, until there is no market in used cars. Consider the market on the second day:

- On the first day all the cars (as we assumed at the start) were put on the market and sold at their true value.
- The average value of these cars was $4,500, so the most a buyer is willing to pay today for any car will be $4,500.
- At the beginning of the second day, each prospective seller expects a price of $4,500 at the most. Most of the sellers are happy: $4,500 is more than half the true value of their car.
- But one owner isn't pleased. The owner of the best quality car ($9,000) would not sell unless the price exceeds half the value of his car: more than $4,500.

- Prospective buyers will not pay more than $4,500. On the second day the owner of the best car will not offer it for sale. No one with a car worth $9,000 will be willing to participate in the market on the second day.
- The rest of the cars will sell on the second day: their value averages $4,000.
- On the third day, buyers will know the average value of the cars sold on the second day, and will be willing to pay at most $4,000 for a car.
- The owner of the second day's highest-quality car (the one worth $8,000) will know this, and will not offer her car for sale on the third day.
- As a result, the average quality of cars sold on the third day will be $3,500. The owner of the third-best car will not put his car up for sale on the fourth day.
- And so it goes on, until, after ten days, only the owner of a lemon worth $1,000 and a totally worthless car will remain in the market.
- If cars of these two values sell on the tenth day, then, on the eleventh day, uninformed buyers will be willing to pay at most $500 for a car of any quality.
- Knowing this, the owner of the car worth $1,000 will decide she would rather keep her car than try to sell on the eleventh day.
- The only car on the market on the eleventh day will be worth nothing: the cars that remain on the market are lemons, because only the owner of a worthless car would be prepared to offer that car for sale.

Economists call this process *adverse selection*, because the prevailing price selects which cars will be left in the market. The market of uninformed buyers selects against quality, and is *adverse* to the interests of potential buyers and sellers.

Unlike the lemons used car market based on one-shot interactions, many economic interactions are *repeated*: workers return to the same job and the same boss day after day, borrowers borrow repeatedly from the same bank, shoppers visit the same grocery store day after day, and buy repeatedly from the same online sites. All of these interactions involve principals and agents engaging repeatedly over time and they therefore have more contractual options available to them than they would in one-shot interactions.

Checkpoint 10.3: Lemons problem and online transactions

The lemons problem explains why many used good markets would not exist. But a great many of these markets do exist. Using what you have learned so far about incomplete contracts, explain what prevents used sales markets (e.g. on

E-Bay) from unraveling like the used car market above.

The lemons problem emphasizes the dependence of real-world markets on the nature of the information available to the exchanging parties and to the courts. Unless information is available to sustain the differentiation of products, contracts will converge to the lowest common denominator level for products. Lemons problems are particularly acute in markets for insurance and credit, where information is at a premium, but can occur in a wide variety of other situations.

EXAMPLE Here's an example of the problem of hidden attributes. If you really enjoy speed when you drive your car, you might consider buying more insurance. But if more risky drivers than safe drivers buy insurance, insurance providers have to cover the resulting higher costs to stay in business. Insurers will raise prices so that insurance will not be worth the cost for the safe drivers. This can result in a vicious circle of rising premiums and an increasingly risky pool of those willing to buy insurance. If the insurance company knew which drivers were risky and which ones were safe, they could design high cost policies specifically for the risky drivers, while prudent drivers would be offered insurance at lower prices.

Whether the information available is sufficient to allow market to work well depend on the economic and legal institutions governing the exchange process. In the case of a commodity like #2 red winter wheat, the information required to sustain the market is provided as a kind of *public good* by the agency that sponsors the market and certifies the quality of the product.

There are many factors that could help to ensure that the lemons problem does not persist. Quality assurance bodies, like those for the wheat described above, are one such solution. Another is the introduction of legislation by third parties (states) to create compensation or liability rules such that those who purchase defective products in one-shot exchanges can receive full compensation or exchange their goods. This idea is demonstrated in the 'lemon laws' in most developed countries nationally, or across a variety of US states.

Another method by which parties can exchange to reduce the severity of the lemons problem is to repeat the interaction. When interactions are *repeated*, players can build up and sustain a reputation. Reputations are costly to build and maintain and if a player acts contrary to their reputation, then they may lose the investment they have built up in their reputation. Examples abound. In online markets, second-hand goods are sold by sellers with a star rating saying whether or not it is worthwhile purchasing from them and the rating information is symmetric. Car dealerships build up reputations for honest dealing and good quality vehicles through word of mouth and customer satisfaction reports.

Finally, as we have seen in Chapter 2, people often wish to uphold social norms and to punish those who do not. Among these norms are honestly reporting the nature or quality of a good one is selling, providing no guarantees, but sometimes providing the social regulation of the contract that Durkheim mentioned as essential to buying and selling.

REMINDER Be sure that you are now clear on the links between *hidden actions* with *moral hazard* and *hidden attributes* with *adverse selection*.

When we move from the problem of hidden attributes to hidden actions we encounter a new set of reasons why sustaining mutually beneficial trade may be difficult, even when mutual gains are technically possible.

10.5 Application: Health insurance

The lemons problem, that is, the problem of adverse selection due to hidden attributes, is far more general than the used-car market. To see why, think about health insurance. Imagine hypothetically that you will be born into a population, but do not know whether you would be born with some serious health problem, or might contract such a problem later in life, or perhaps be entirely healthy until you die of old age.

Now imagine that before you were born, or even knew who your parents would be you were asked this question: Would you buy health insurance if the premium – that is, the cost of the insurance to you – which is the same for everyone, is just sufficient to pay for the medical services required across the entire population if everyone agreed to purchase it? This kind of hypothetical decision-making process is called making a decision behind a **veil of ignorance** because it is a thought experiment in which you are invited to think about how you would act or what policies you would favor if did not know the state of of your actual health.

VEIL OF IGNORANCE The veil of ignorance is a thought experiment asking you to imagine making a choice – about an anti-discrimination law, immigration policy, insurance plan – while *ignorant* of your position -your gender, nationality, health status, for example – that will matter for how you fare under the decision you take. In John Rawls' words, we should think about what a just society would be like as if "no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like." The veil of ignorance is similar to the thought experiment proposed by Adam Smith when he wrote that "the natural preference the every man has for his own happiness over that of other people, is what no impartial spectator can go along with." We used the Impartial Spectator experiment in Chapter 10 .

Most people would be willing to purchase health insurance behind the veil of ignorance, because they would rather pay a premium representing the average costs of health care to the whole population rather than be individually responsible for paying for the treatment of a serious illness, which, even if it is very unlikely to strike them, would impose high costs that an most families could not pay. The benefit of protecting oneself and one's family from a financial catastrophe (or the possibility that you can't afford health care when you need it) is worth the insurance premium on average.

But the thought experiment is unrealistic: We cannot sign people up for health insurance before they know how healthy they will be because that means signing them up before they are born. Though most people would buy fairly priced health insurance if they did not know about their future health status, the situation changes dramatically if they can choose whether to buy health insurance knowing something about their current and future health status.

Let's look at the situation from the standpoint of the insurance provider (the principal) paired with a prospective buyer of insurance (the agent):

- People are more likely to purchase insurance if they know that they are ill or likely to become ill. The average health of people buying insurance will be lower than the average health of the population.
- This information is *asymmetric*: The person buying the insurance knows more than the insurance company about how healthy they are .
- Insurance companies selling insurance to people who are sicker than average will be profitable only if they charge higher premiums than if everyone

bought the same insurance.

- This will lead some – those who are reasonably certain that they are healthy and will remain so – to not purchase insurance.
- To remain in business, the insurance companies will have to charge even higher premiums. Eventually the vast majority of the people buying insurance will be those who know they already have or are likely to have a serious health problem.

This is a case of adverse selection. The reasoning above shows why the hidden attributes of initial health status can result in an unraveling of the health-care market. Why? Insurance companies make profits by insuring people who are healthy. Healthy people who want to buy insurance in case they fall ill in the future are priced out of the market, and will not buy insurance. In the extreme case, the health insurance premium will be so high that only people who know they are likely to become seriously ill will buy insurance (people with initial conditions of poor health).

HISTORY One of the most controversial aspects of U.S. President Barack Obama's health care legislation was the requirement that all citizens purchase insurance. Supreme Court Justice Antonin Scalia opposed this requirement as an unwarranted intrusion of the government into people's choices. Reasoning that if you can make people buy health insurance "you can make people buy broccoli." The President's lawyer, Donald Verrilli replied that the market for broccoli was different from the market for health insurance. Is Verrilli right? Are the two markets different in a way that might justify requiring everyone to buy health insurance while not requiring everyone to eat broccoli?

In this case we have what is called a *missing market*. It is a market that could exist, but it would only exist if health information were symmetrical and verifiable (ignoring for the moment the problem of whether everyone would want to share their health data). Under those imaginary conditions, the market could provide benefits to both insurance company owners and people who wanted to insure themselves. Not having such a market is Pareto-inefficient.

To address the problem of adverse selection due to asymmetric information and the resulting missing markets for health insurance, many countries have adopted policies of compulsory enrollment in private insurance programs or universal tax-financed coverage, such as the national health service in the United Kingdom, or similar services in Canada, France, and elsewhere.

Checkpoint 10.4: Applications other than health

Consider applications other than healthcare – think about other kinds of insurance, for example, car insurance or home insurance. What are the characteristics you might want to think about.

a. What are the attributes people would want to keep hidden?
b. What does hiding these attributes do to prices?
c. What does hiding the attributes do to peoples' beliefs about the goods being sold on the markets?
d. Why might the market be missing or what might allow the market to be present?
e. For the latter question, consider that in wealthier countries there exist many markets, especially for insurance, that do not exist in low income or middle

income countries. Why do you think that is?

10.6 Hidden actions and moral hazards: A contingent renewal contract

Insurers, whether private or governmental, face problems other than hidden attributes. There is also the problem of hidden actions: buying the insurance policy may make the buyer more likely to take exactly the risks that have been insured against. A person who has purchased full coverage for his car against damage or theft, may as a result take less care in driving or in securing his vehicle than someone who had not purchased the insurance.

Insurers typically place limits on the insurance. For example, insurance coverage may not apply or be more expensive if someone other than the insured person is driving, or if it is parked on a daily basis in a theft-prone location. These provisions can be written into an insurance contract and are enforceable. But, the insurer cannot enforce a contract about how fast you drive or whether you drive after having had a drink. These are the actions that are hidden from the insurer because of the asymmetric information: you know these facts, but the insurance company does not.

Here is a model of incomplete contracting in the case of hidden actions, inspired by Benetton, the global casual wear marketer. To understand the model, let's first explore some of the history of the Benetton company.

Benetton: Decentralized production

Benetton grew from a small company selling sweaters to shops near the town of Treviso in northern Italy half a century ago to one of the world's largest designers, producers, and sellers of casual wear and other garments. A key to its early success was a highly decentralized system of production: the labor intensive aspects of production – primarily sewing – were carried out by hundreds of small producers, working to designs and schedules and with materials supplied by Benetton. A few processes were done by Benetton itself – notably dying, performed at the last minute so as keep the product in tune with fashion trends.

Importantly, *quality control* and marketing were centralized, performed by Benetton's own staff. Sub-contractors who reliably produced goods of the specified quality benefited from permanent orders, quick payment by Benetton, and other benefits of their long-standing relationship with the company.

As a result of this decentralized subcontracting structure, Benetton saved on the costs of establishing its own production facility. The most valuable asset of the company was not garment-making factories but instead the Benetton

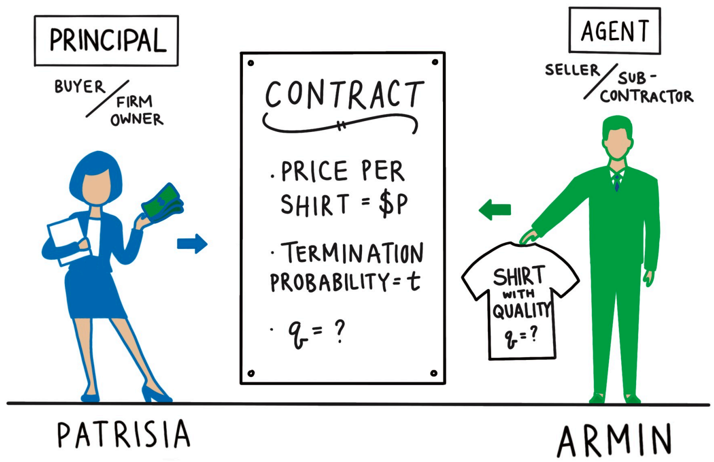


Figure 10.2: Patrisia and Armin interacting in a principal-agent interaction: Armin selling Patrisia a shirt with non-contractible quality (q) and Patrisia specifying a price (p) and termination schedule (t).

name itself (called an *intangible asset*), which, once the garments were acquired from the subcontractors, was of course attached to the product before sale. Many contemporary companies have similar structures, from manufacturing to design to developing code and online applications.

The Benetton model

To model these relationships we introduce buyer (the *principal*, Patrisia) who purchases a good, say, a shirt, from a supplier (the *agent*, Armin, a subcontractor) for a price, p and then sells the good to a consumer (who is not a player in this game). Their interaction is depicted in Figure 10.2.

The good – such as a shirt – has an attribute, its *quality*, q that is important to the principal, and costly or difficult for the agent to provide. Patrisia (who might be the Benetton company in this model) would like to pay a low price for a high-quality shirt, and Armin (one of the subcontracting shirt producers) would like to receive a high price for a shirt of lesser quality, so their interests conflict.

The principal-agent problem arises from this conflict of interest along with the fact that the quality of the shirt cannot be readily determined by Patrisia. She has the problem that shirts are sometimes rejected by customers, and Patrisia then has to refund the price when they return the good.

The challenge that the principal, Patrisia, faces is that she knows that Armin would prefer to produce lower quality goods, but it is not cost effective to determine the actual quality of each unit that she purchases. So she comes up with the following plan:

- *Repeated relationship:* Give Armin the possibility of an ongoing (or repeated) sub-contracting relationship with her, so he can count on her buying his products year after year.

- *Termination probability:* But let him know that if he provides a good that is rejected by a consumer, she will and terminate the relationship.

- *Enforcement rent*: pay him enough for the goods he provides so that he does not want the relationship to be terminated, providing him with a motive to provide quality goods.

In other words she structures her interaction with Armin as a repeated game that she can terminate if she is not satisfied with the quality he provides.

Patrisia interacts with a large number of sub-contractors like Armin. We simplify the problem by assuming that each period she purchases a given number of shirt (it could be one or one thousand) from Armin and each of the other sub-contractors. So she is really choosing the price and the number of sub-contractors from whom to buy.

Here is the structure of the game. The principal is first mover and seeks to maximize her profit by deciding on:

- *Price:* The price to offer the agent (and the other subcontractors) and

- *Quantity*: How many shirts to purchase.

The agent seeks to maximize the expected value of his utility over the duration of his doing business with the principal; he has just one decision:

- *Quality*: The quality of the good to supply.

The principal sets the price knowing the agent's best response function, that is, the quality he will supply for every price she could offer. The agent chooses the quality to provide knowing the price the principal has offered and the probability that the transaction will be terminated for every level of quality he could provide.

To determine what the first mover (the principal) will do we need first to derive the best response function of the second mover (the agent). The level of quality the agent supplies will depend on how valuable it is to the agent to continue the relationship with the principal. This is the rent that the agent will receive, and the reason why he will supply higher quality than he would provide otherwise.

Figure 10.3: **Armin's disutility of providing quality.** The disutility of providing quality increases as the agent provides more quality. The slope of the tangency lines is the marginal disutility of effort at the corresponding level of quality provided. Comparing the slopes at points **j** and **k** shows that a small increase in quality will impose a greater disutility the higher is the quality he provides. In the figure Armin's baseline disutility of providing quality is set to $\underline{u} = 5$.

This is called a contingent renewal model because the principal has made renewal of the contract contingent on(meaning depend on) the agent supplying adequate quality. You will see in the next two chapters that labor and credit markets are also analysed using contingent renewal models.

10.7 The value of the transaction to the agent

Because this is a repeated game, how important it is to the agent that relationship with the principal continue will depend on

- the utility that the agent experiences in each period,
- the number of periods that the principal and agent will interact, and
- the utility the agent will experience if terminated by the principal, that is the agent's fallback option.

The agent's utility in a single period

The seller, Armin, prefers to receive a higher price p and to provide a lower level of quality q (which can take any value from 0 to 1, representing a range from a quality of 0 percent to 100 percent). Armin prefers to provide lower quality because it is costly – in terms of effort or care – for him to produce high quality goods.

For concreteness we will express the general form of his utility function (below on the left) by a specific form (on the right below):

$$\text{Agent's utility} \qquad u(p,q) = p - \frac{\underline{u}}{(1-q)} \tag{10.1}$$

Equation 10.1 says the following:

- Because the marginal utility of the price, $u_p > 0$, Armin considers p to be a "good" something he would like more of because getting more money is valuable to Armin.
- Because the marginal utility of providing quality $u_q < 0$, Armin considers q to be a "bad", so he would rather not exert the effort required to provide $q > 0$.

The second term on the right hand side ($\frac{\underline{u}}{(1-q)}$), is Armin's *disutility of providing higher quality* shown in Figure 10.3. We can see three things:

- If he provides no quality (he just hands over a product with the lowest possible quality, i.e. $q = 0$) his disutility is $\underline{u}$.

M-CHECK The marginal utility of providing quality is the derivative of u with respect to q, denoted u_q. It is negative because the second term in the utility function is preceded by a minus sign. When we refer to the marginal *dis*utility of quality – the slope of the curve in Figure 10.3 – we mean the derivative of $\frac{\underline{u}}{1-q}$ (itself, without the minus sigh) with respect to q which is $-u_q$, so it is positive.

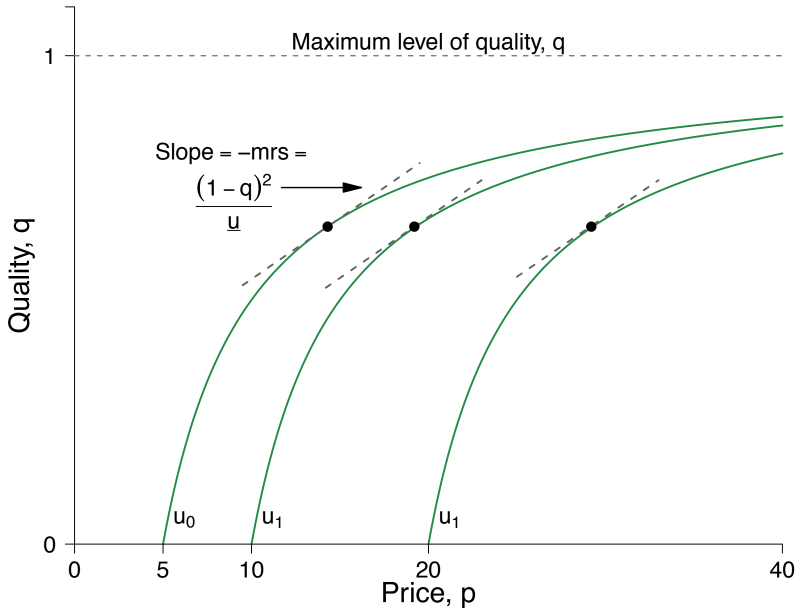


Figure 10.4: **Armin's indifference map for "good" p and "bad" q.** Armin's indifference map for "good" p and "bad" q for three values of u and $\underline{u} = 5$. With this quasi-linear utility function (Equation 10.1) his indifference curves are *horizontal* displacements of each other: their slopes depend only on the level of quality, not on the price. For the indifference curves shown, $u_3 > u_2 > u_1$. Using the notation that $u_q = \frac{\partial u}{\partial q}$

$$-u_q = \frac{\underline{u}}{(1-q)^2}$$

$$-mrs = -\frac{1}{u_q} = \frac{1}{\frac{\underline{u}}{(1-q)^2}} = \frac{(1-q)^2}{\underline{u}}$$

- The higher the quality he provides, the more disutility he experiences: the curve is upward sloping and
- The marginal disutility of quality is increasing: the curve is steeper at point j than at point k.

From Equation 10.1 you can also see that he will never choose produce a perfect good (100% quality or $q = 1$) because at $q = 1$ his disutility is infinite or undefined.

There are two confusions to avoid here, one about slopes and the other about signs. First be clear about the difference between the height of the curve – the disutility of quality – and its slope – the *marginal* disutility of quality. The second is the distinction between the marginal *dis*utility of quality, which is positive, and the marginal utility of quality, which is just the same thing with a negative sign.

M-CHECK You can confirm that his disutility is infinite when $q = 1$ by substituting it into $\frac{\underline{u}}{1-q}$ and confirming that it would result in division by zero.

We can construct indifference curves for Armin based on his utility function: this is another example in which the choices include both a good (receiving the price) and a bad (providing quality). As a result for any quality q on the vertical axis Armin would prefer to obtain a higher price, p, which would place him on a higher indifference curve; and for any price, p, on the horizontal axis: Armin would prefer to provide a lesser quality.

REMINDER As in Chapter 3 when we are considering both a good and a bad, the indifference curves are upward sloping. This is because having more of the bad can be compensated by having more of the good. In this chapter, though, the "bad", quality (q), is on the vertical axis and the "good" (price, p) is on the horizontal axis.

The negative of the slope of the indifference curve is Armin's marginal rate of substitution between more pay (the price he receives) and more quality. We derive this in M-note 10.1

We already know from Figure 10.3 and Equation 10.1 that increasing q reduces Armin's utility (increases his *dis*utility) and does so more and more as q approaches 1. This is the reason why in Figure 10.4 Armin's indifference curves become almost flat when Armin provides high levels of quality: Armin experiences an extremely high marginal disutility of providing any additional quality. At *already high* levels of quality providing additional quality (moving up in the figure) can be compensated only by a very large increase in price (moving to the right), as shown by the flattening of his indifference curves as q gets larger.

M-Note 10.1: The marginal rate of substitution for per-period utility

Recall that the general form of Armin's utility function is:

$$u = u(p,q)$$

We want to find the changes in q and p that are consistent with no changes in Armin's utility – that is, staying on the same indifference curve. To do this, we totally differentiate Armin's utility function with respect to dp and dq and set the result equal to zero:

$$du = dqu_q + dpu_p = 0 \tag{10.2}$$

In Equation 10.2 we are using the notation that $u_q = \frac{\partial u}{\partial q}$ and $u_p = \frac{\partial u}{\partial p}$ Equation 10.2 requires that for any two points on an indifference curve, the utility difference associated with the difference in price (dpu_p) is exactly compensated by the (opposite signed) utility difference associated with the difference in quality (dqu_q), so that taking account of both effects the difference in utility between the two points is zero. We can rearrange Equation 10.2 to find $\frac{dq}{dp}$, which is the slope of the indifference curve:

$$\text{Slope of indifference curve} \quad \frac{dq}{dp} = -\frac{u_p}{u_q}$$

Because the marginal utility of the price (p) is one, we have:

$$\begin{aligned} \text{Slope of Armin's indifference curve} &= -mrs(p,q) \\ &= -\frac{u_p}{u_q} \\ &= -\frac{1}{u_q} \end{aligned}$$

We now consider the specific utility function we introduced as an illustration. To find out the marginal utility of providing quality, u_q, we differentiate Armin's utility function, Equation 10.1, with respect to q to find u_q:

$$\begin{aligned} u(p,q) &= p - \frac{\underline{u}}{1-q} \\ u_q &= (-1)(-1)\left(-\frac{\underline{u}}{(1-q)^2}\right) \\ &= -\frac{\underline{u}}{(1-q)^2} \end{aligned} \tag{10.3}$$

Using Equation 10.3, we can now find the marginal rate of substitution given for the general case by Equation :

$$\begin{aligned} mrs(p,q) &= -\frac{1}{u_q} \\ &= -\frac{1}{\frac{\underline{u}}{(1-q)^2}} \\ &= -\frac{(1-q)^2}{\underline{u}} \end{aligned}$$

The fact that p does not appear in the expression for the $mrs(p,q)$ means that the slope of the indifference curves depends only on the level of quality, not on the price as can be seen in Figure 10.4. This is because we have used a quasi-linear utility function that linear in the price.

Checkpoint 10.5: Goods, bads and axes

a. Consider Figure 10.3, what would happen to the curve if $\underline{u}$ increases or decreases (recall that this means that the disutility from working at all on the contract is changing).

b. Consider Figure 10.4, the indifference curves for p and q look different to the indifference maps for goods and bads we've looked at previously. Why do you think that is? (Think here about what 'causes' what when you think about *why* we might put the variables on a given axis).

c. Consider Figure 10.4 again, what would increasing (or decreasing) $\underline{u}$ do in the context of the indifference curves in $u(p,q)$? Why?

d. Substitute the values $q = 0.4$ and $q = 0.8$ into Armin's marginal rate of substitution, $mrs(p,q)$. Assume $\underline{u} = 5$. What are the values for the mrs? How do you interpret them? (Be clear about what Aram is getting more of and what he is "willing to pay").

The value of the transaction and the enforcement rent

To make her plan work Patrisia has to do two things:

- get at least some information on the quality Armin provides as the basis for terminating the contract if necessary; and
- offer him a high enough price so that he is receiving a rent, so that he will prefer continuing doing business with Patrisia rather than being terminated.

She uses consumer complaints about quality as the basis for her termination decision and as a result she will terminate the relationship with Armin with probability $t = 1 - q$. This means that:

- If Armin provides no quality at all ($q = 0$),there will surely be a complaint, so he will surely be terminated ($t = 1$).
- If Armin were to provide $q = 1$ no consumer would ever complain, so he would be certain that Patrisia would continue buying from him ($t = 0$).

- If Armin provides $q = 0.5$, whether he gets terminated at the end of each period is a coin toss: 50% chance of keeping the contract and 50% chance of losing the contract.

If he provides quality of q then at the end of the first period he will lose his contract with probability $1 - q$. If he is lucky and does not lose the contract, then the first period is just repeated – she offers the same price and so he offers the same quality. As a result, the next period will be his last with the same probability $(1 - q)$ and so on until his luck runs out and he is terminated.

How long will his contract with Patrisia last? Armin does not know for sure: he has to think about probabilities. To do this, imagine a game of flipping a coin, in which you flip it once and if it came up heads you flip it again, and continue doing this until it comes up tails, at which point the game ends. If you did this many times, the game would sometimes end right after the first flip (because it came up tails); sometimes it would continue for many flips.

But on average how many flips would you expect to make, including the first one? The answer (which we show in the Mathematical Appendix) is 2 periods, which is just 1 divided by one-half, the probability that the game will end after each flip. So the expected number of periods that Armin's contract will last, $T(q)$ will be one divided by the probability of termination at the end of each period, or $\frac{1}{t(q)}$.

The value he gets from doing business with Patrisia is the utility he gets in a single period times the number of periods he expects to do business with her.

- The utility per period is $u(p,q)$.
- The expected lifetime of the contract is $T(q)$, which is equal to 1 divided by the probability that he gets terminated in a given period, or $T(q) = \frac{1}{t(q)}$.

So:

$$\begin{aligned} \text{Value} &= \text{Utility per period} \times \text{Expected lifetime of contract} \\ v(p,q) &= u(p,q) \cdot T(q) \\ &= u(p,q) \cdot \frac{1}{t(q)} \end{aligned} \tag{10.4}$$

This is how much he values being Patrisia's subcontractor. But to understand how motivated he will be by the threat of her ending the contract we need to know about his other options. How much he would like to keep his job is the difference between how much it is worth to him, v, and what he would be able to get if his relationship with Patrisia were to be terminated.

This next best opportunity is his fallback option, which is finding another buyer and attempting to sell him the (presumably low quality) good he has made.

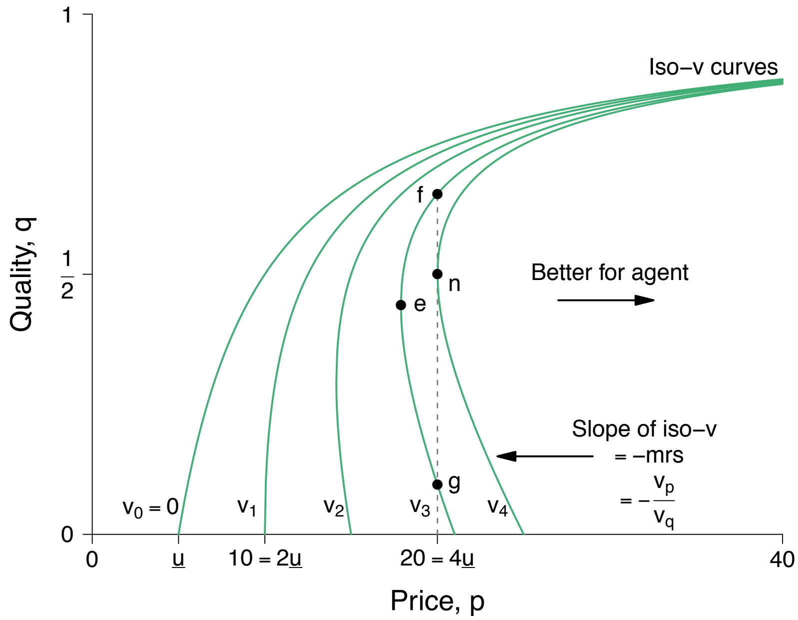


Figure 10.5: **The agent's iso-value curves.** Armin's map of iso-value curves in the "good" p and "bad" q for five values of v using Equation 10.4 and $\underline{u} = 5$.

Later we will discuss his fallback option but for now we assume that his utility is zero if Patrisia terminates their relationship.

So his fallback option is zero, and v is Armin's enforcement rent, the loss of which he would like to prevent by supplying higher quality than he would otherwise do. If $v > 0$, then the following will hold:

ENFORCEMENT RENT In a principal-agent relationship an enforcement rent is the excess of the value of the transaction to the agent over the agent's fallback. A principal offers an enforcement rent to an agent along with the threat that the relationship will be terminated if the agent does not act in the interest of the principal. The possibility that the rent will be terminated motivates the agent to carry out the principal's wishes more than would be the case in the absence of the rent.

- It is a *rent* because it is how much his current situation is preferable to his next best alternative.
- It is an *enforcement rent* because the rent motivates him to provide more quality than he otherwise would (he doesn't want to lose the contract).

In Figure 10.5 we show Armin's iso-value curves. Each curve is made up of all of the combinations of p and q that give the same values of v in Equation 10.4 with $v_4 > v_3 > v_2 > v_1 > v_1 = 0$.

The iso-value curves differ from the single-period utility indifference curves in Figure 10.4: they are not uniformly upward sloping. The upward-sloping portions of the iso-value curves are easy to understand. In Figure 10.5 Armin is indifferent between point **f** – providing high quality and getting paid a high price – and point **e** – providing lesser quality at a lesser price.

But why is Armin indifferent between point **e** and point **g**, where he is providing less quality and getting a higher price, which would seem to be a better deal than point **e**?

The answer is that when Armin takes account of the likelihood of being terminated, providing more quality is not necessarily a "bad." The reason is that providing more quality will *increase* his chance of keeping his contract with Patrisia which (if $v > 0$) he values. At some low levels of quality, providing a little more quality will increase v because it will prolong the duration for which he receives the per-period utility. The iso-value function has a different shape from the single-period utility indifference curves in Figure 10.4 because the game is repeated, and Armin has an interest in continuing his relationship with Patrisia.

Recalling that definition of the marginal rate of substitution is the negative of the slope of the indifference curve, here the slope of an iso-value curve is the negative of the $mrs(p,q)$:

$$\text{Slope of iso-value :} \quad \frac{\Delta q}{\Delta p} = -\frac{v_p}{v_q} = -mrs(p,q) \qquad (10.5)$$

M-CHECK Equation 10.5 is similar to the expression for the slope of the per period utility based indifference curves derived in M-Note 10.1.

Checkpoint 10.6: Armin's iso-value curve

a. Use Armin's iso-value function to check that were he paid $2\underline{u}$ and provided $q = 0.5$, he would have $v = 0$.
b. What would happen to Armin's iso-value curves if his disutility of starting work ($\underline{u}$) were to increase or decrease?
c. In Figure 10.5, along an iso-value curve what would happen as Patrisia offers Armin a higher price?

10.8 The agent's best response: An incentive compatibility constraint

M-CHECK If we have a function like $u(p,q)$ then $u(p,0)$ means the utility if the price is p and the quality level is zero.

For any particular price that Patrisia offers, there are three things that Armin could do:

- refuse the contract and have utility $u = 0$
- accept the contract but deliver $q = 0$ in which case his contract would be terminated with certainty, so he would receive $u(p,0)$ for a single period, and the value of his transacting with Patrisia would be $v(p,0) = u(p,0)$
- accept the contract, deliver some level of quality $q > 0$ and receive $u(p,q)$ for an expected number of periods, $\frac{1}{t(q)}$, and receive $v(p,q)$.

We show how to determine the action taken by the agent based on the price offered by the principal in Figure 10.6. Think about the iso-value function for $v = 0$. The combinations of q and p making up this iso-value function are such give the agent a value of the transaction no better than his reservation option (which is zero). So the v_0 is the agent's participation constraint.

REMINDER In a principal-agent relationship, the incentive compatibility constraint is the agent's best-response function. It gives the combinations of the principal's actions – her choice of a price, a wage, or an interest rate, for example – and the agent's response. It shows the agent's actions that are compatible with the incentives and constraints offered by the principal.

Now think hypothetically about how the agent might respond to different prices offered by the principal. You can see from Equation 10.1 that if the principal were to offer a price less than $\underline{u} = 5$ that is, the disutility of supplying a good of no quality at all – then the best the agent could do even supplying zero quality would result in a utility less than zero, so he would refuse the offer.

At a price grater than $\underline{u}$ the agent will accept the contract. What quality, if any, will the agent supply? This is shown by the best response function – the purple curve – which is also the incentive compatibility constraint.

M-Note 10.2: The agent's best-response function for the general case

We know that Armin (the agent's) value function is per-period utility times the number of periods that he expects the transaction to continue as given by the following equation:

$$\text{A's value} \qquad v(p,q) = u(p,q) \cdot \frac{1}{t(q)} \tag{10.6}$$

As a result, he will choose q – the only variable he controls – to maximize his value given the price that is offered and the inverse relationship between the probability of termination and the quality of the good he provides. So, he will differentiate his value function with respect to q and set the result equal to zero (imposing the first order condition for a maximum). Notice, to find the first-order condition we need to use the quotient rule:

$$\begin{aligned}
v_q &= \frac{u_q \cdot t - u \cdot t_q}{t^2} \\
\text{First order condition:} \quad 0 = v_q &= \frac{u_q \cdot t - u \cdot t_q}{t^2} \\
0 &= u_q \cdot t - u \cdot t_q \\
u_q \cdot t &= u \cdot t_q \\
\text{Divide by: } t \qquad \therefore u_q &= t_q \left(\frac{u}{t}\right)
\end{aligned} \tag{10.7}$$

But, recall that $v(p,q) = \frac{u(p,q)}{t(q)}$ (Equation 10.6), so we can rewrite the equation:

$$u_q = t_q \cdot v \tag{10.8}$$

This is the agent's best-response function, as can be seen if we make clear how u, t, and v depend on p and q.

$$u_q(p,q) = t_(q) \cdot v(p,q) \tag{10.9}$$

For any given value of p, the value of q that satisfies Equation 10.9 is the agent's best response $q(p)$.

Also recall that the slope of Armin's iso-value curve is:

$$\text{Slope of iso-value :} \qquad \frac{dq}{dp} = -\frac{v_p}{v_q} = -mrs(p,q)$$

Given that Armin maximizes by setting $v_q = 0$, we can see why the point on each indifference curve with a vertical slopes makes up the best response function.

Armin will choose quality (q) to maximize his value (v), taking account of the fact the higher q will reduce the probability of termination t (remember at the end of any period, $t = 1 - q$).

He will want to supply more quality as long as the disutility of doing that – the marginal cost of quality – is less than the marginal benefit that he derives from providing additional quality. That is he will want to compare two *negative* quantities:

- *Marginal cost*: u_q, namely the *reduction* in his utility associated with providing more quality (the marginal utility of effort) .
- *Marginal benefit*: $t_q \cdot v$, namely the *reduction* in likelihood of being terminated made possible by providing more quality times the effect of working harder on the probability of keeping the job.

So, as shown in M-Note 10.2 the level of quality that will maximize his value is that at which the marginal costs and marginal benefits are equal:

$$\begin{aligned} u_q &= t_q v \qquad (10.10) \\ \text{Marginal disutility} &= \text{Reduction in termination probability} \times \text{enforcement rent} \\ \text{Marginal cost} &= \text{Marginal benefit} \end{aligned}$$

M-CHECK Equation 10.10 is the agent's best-response function for selecting q in response to the principal's choice of p. Below we use a specific utility function to derive a simple best response function.

To see what Equation 10.10 means, return to Armin's iso-value curves as shown in Figure 10.5. Suppose hypothetically that the the price offered by Patrisia is $p = 20$. Then we can think of Armin's optimizing problem as starting at some low level of quality (point **g**) and noting that as he provides more quality, he reaches higher iso-value functions (not shown) until point **n**, the tangency of the iso-value curve labeled v_4 with the vertical line indicating the hypothetical price. If he proceeds upwards – offering more quality – then he will be crossing ever *lower* iso-value curves, such as the one at point **f**.

M-CHECK Using v_p and v_q respectively for the partial derivative of v with respect to p and q, know that the slope of the iso-value is $-mrs(p,q) = -\frac{v_p}{v_q}$ which is vertical where $v_q = 0$. Because we know that $v_q = 0$ is the first order condition for a maximum for Armin we also know that the point where the iso-value curve is vertical is a point on the agent's best response function.

The quality the agent offers and price indicated by point **n** is one point on Armin's best-response function. Two others , constructed in the same way but at lower prices are points **e** and **d**.

From the figure you can see that the agent would accept the contract and supply zero quality if the price were between $2\underline{u}$ and $\underline{u}$. This is because over that price range, the participation constraint is satisfied (so the agent "participates" in the contract) but prices in that range to not provide sufficient incentives for the agent to raise quality above zero.

We show in M-note 10.3 that if we use the specific utility function in Equation 10.1, then Armin's best-response function (for values of $p \geq 2\underline{u}$) is given by Equation 10.10is:

$$\text{Armin's best-response:} \quad q(p) = 1 - \frac{2\underline{u}}{p} \qquad (10.11)$$

This is the best response function we derived in the figure, using $\underline{u} = 5$. You can see from Equation 10.11 why at prices lower than $p = 10 = 2\underline{u}$ Armin will not provide any quality at all.

Checkpoint 10.7: The best-response function

1. Refer to Figure 10.6. What happens along Armin's best-response function as Patrisia offers him a higher price? What is he willing to do in response?

2. What happens to his best-responses function as $\underline{u}$ increases or decreases? For example, what would happen if Armin had a longer commute (increasing $\underline{u}$) or if he derived self-worth from the the work of producing the good (reducing $\underline{u}$).

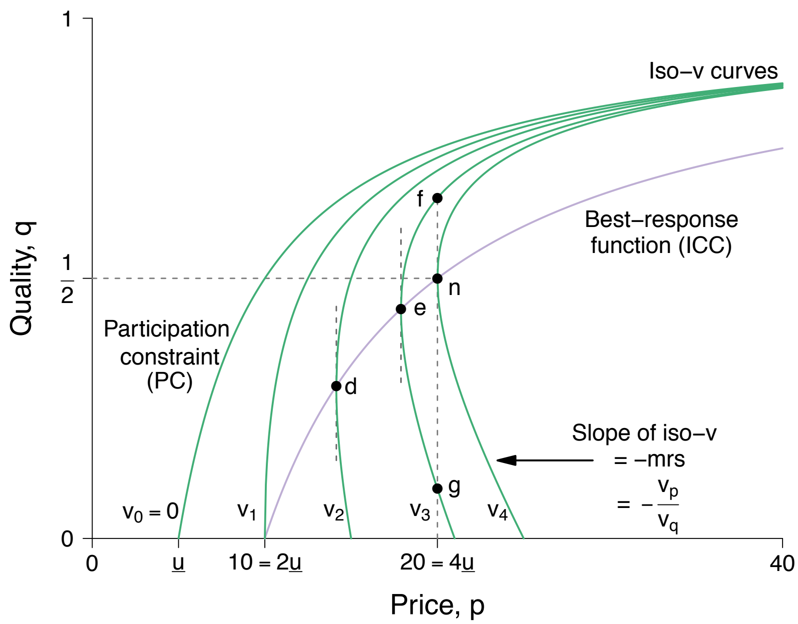


Figure 10.6: **The agent's best-response function**. The agent's best-response function is made up of all the the points on an iso-value curve where the slope of the curve is vertical, as shown by three iso-value curves and corresponding points on the best-response function. Notice that where the best-response function intersects the horizontal axis the iso-value function is vertical. This is why at the price of $2\underline{u} = 10$ the best response of the agent is to provide $q = 0$. The best-response function is the incentive compatibility constraint or ICC. The participation constraint shows all the combinations of p and q that would make the agent indifferent to being terminated because the value of the transaction would be no better than his fallback option, namely zero. The segment of the horizontal axis between $\underline{u}$ and $2\underline{u}$ is the range of prices over which the price is high enough to get the agent to agree to a contract, but not high enough to get him to provide quality greater than zero.

REMINDER Recall from Chapter 4, that when a first-mover has price-setting power, but not take-it-or-leave-it-power, the first-mover maximizes their objective function (utility or profit) given the other person's best response or incentive compatibility constraint. In Chapter 4, this was also Biko's *offer curve*.

M-Note 10.3: The agent's best response with a specific utility function

Recall that the condition for the agent's choice of a level of quality to offer is:

$$\begin{aligned} \text{Marginal cost} &= \text{Marginal benefit} \\ u_q &= t_q v \end{aligned} \tag{10.12}$$

Recall that $t = 1 - q$, so $t_q = -1$, and we can rewrite the condition given by Equation 10.12 using our specific utility function, termination probability, and value function:

$$-\frac{\underline{u}}{(1-q)^2} = -1\left(p - \frac{\underline{u}}{(1-q)}\right)\frac{1}{(1-q)} \tag{10.13}$$

To see how Equation 10.13 gives us Armin's best-response function, multiply both sides of

Equation 10.13 by $-(1-q)$:

$$
\begin{aligned}
\frac{\underline{u}}{1-q} &= p - \frac{\underline{u}}{1-q} \\
\frac{2\underline{u}}{1-q} &= p \\
\text{Divide through by } 2\text{:} \quad \frac{1}{1-q} &= \frac{p}{2\underline{u}} \\
\text{Raise both sides to the power } -1\text{ :} \quad 1-q &= \frac{2\underline{u}}{p} \\
\therefore \text{ Isolate q :} \quad q &= 1 - \frac{2\underline{u}}{p}
\end{aligned}
$$

Thus we have:

$$
\text{Armin's best-response function:} \quad q(p) = 1 - \frac{2\underline{u}}{p} \tag{10.14}
$$

10.9 *The principal's cost minimization and the Nash equilibrium*

The principal makes two decisions: how many units to purchase (which is equivalent to how many subcontractor-agents to engage) and how much to pay them for each unit. We have set up the problem so that we can focus on the second – the principal agent problem. We have already analysed the first, how many shirts she should sell to consumers, in Chapters 8 and 9. To study the price-setting process in the principal-agent relationship, we proceed in three steps:

REMINDER In minimizing p/q the principal is seeking the least cost way of acquiring the goods she is to sell, just as did the cost-minimizing firms choosing technologies and input mixes to minimize their costs of production in Chapter 8. The only difference is that here the principal is buying the goods she will sell rather than producing them. In Chapter 11 we will study a similar problem in which the principal – an employer – is seeking to minimize the costs of production.

- Patrisia, the principal, knowing Armin's best response $q(p)$, determines the price p^N that will minimize the cost of acquiring quality (p/q) (we use the N superscript because this will be the Nash equilibrium price.)
- Armin, the agent, best responds to the price offer by choosing the Nash equilibrium quality, q^N, the quality level that maximizes his value given the price (using his best-response function).
- When p^N is offered and q^N is the response, then the expected number of periods that Armin's relationship with Patrisia will last is $T^N = \frac{1}{t(q^N)} = \frac{1}{1-q^N}$.

The principal minimizes costs

Patrisia will set the price p in order to *minimize* the cost of acquiring quality, that is, minimizing the price she pays for given quality p/q.

To show how the principal will choose among the prices she may set, we show variety of different potential ratios of price to quality to price (p/q) in Figure 10.7 as *isocost rays*. Along a given isocost ray, Patrisia has the same cost of quality, as can be seen by comparing points **b** and **d**. This is why they are called isocost ("equal cost") rays. Patrisia prefers isocost c_1 to c_2 to c_3

figures/information/principal_isoprofit.pdf

Figure 10.7: **Some of the principal's isocost curves.** The principal's isocost curves for different ratios of quality for a given price, q/p.

because $c_1 < c_2 < c_3$. Comparing points **a** on c_3, with **b** on c_2 and **c** on c_1, we can see that at the same price, Patrisia would get a higher quality at point **c**. So, steeper rays are better for the principal (corresponding to lower cost): the slope of any ray is q/p.

The Nash equilibrium price and quality

The isocost rays tell the principal which regions of the graph she prefers, roughly, which way is up. But she is constrained by having to provide the agent with incentives to implement the level of quality she desires. So the q she wants to motivate Armin to provide must lie on his best-response function. This equation – the incentive compatibility constraint – tells her what is feasible. In Figure 10.8 she would prefer to be, for example, at point **c** rather than at point's **d** or **n**, because **c** is on a lower (steeper) isocost ray. But point **c** is infeasible it is above Armin's best response function.

She could be at point **d** on Armin's best response function, but were she there she would be paying a lot for low quality. The resulting allocation will be somewhere on the incentive compatibility constraint, but where?

Figure 10.8: **The principal's cost-minimizing price and the agent's utility-maximizing quality provided** The Nash equilibrium, point **n** (p^N, q^N) is found at the tangency of the isocost ray with slope $\frac{q}{p}$ and the best-response function with slope $\frac{2u}{p^2}$.

Isocost ray: $q = \frac{p}{8\underline{u}}$

Best-response function: $q = 1 - \frac{2\underline{u}}{p}$

It is simpler to find the price that maximizes q/p than that minimizes costs (p/q); the two are equivalent. So, putting her *objectives* together with what is *feasible*, Patrisia wants to find the p that will:

$$\text{maximize } \frac{q}{p}, \text{ where } q \text{ has to be such that } q = 1 - \frac{2\underline{u}}{p}$$

The price that accomplishes is given by the following condition for cost minimization (as shown in M-Note 10.4):

$$\begin{aligned} \text{Condition for cost minimization:} \quad q_p &= \frac{q}{p} \qquad (10.15) \\ \text{slope of agent's BRF} &= \text{slope of isocost ray} \end{aligned}$$

This condition is shown in Figure 10.8: the price is chosen so that the slope of the agent's best-response function q_p is be equal to the slope of isocost ray with the highest ratio of quality to cost, $\frac{q}{p}$.

The allocation (p^N, q^N) is a Nash equilibrium because it is a mutual best response:

- *Patrisia's choice*: given Armin's strategy choice – his best-response function – offering the price (p^N) is the best Patrisia can do; and
- *Armin's choice*: given Patrisia's price offer, providing (q^N) is the best Armin can do.

M-Note 10.4: The principal's cost-minimizing price

Here is her maximization problem: p is the variable over which the principal has control, so she maximizes $\frac{q}{p}$ by varying p.

$$\text{Maximize} \quad \frac{q(p)}{p}$$

To find the price selected by this optimization problem we differentiate this expression with respect to p:

$$\text{Quotient rule} \quad \frac{d}{dp}\frac{q(p)}{p} = \frac{q_p \cdot p - q}{p^2} \qquad (10.16)$$

We then set the resulting Equation 10.16 equal to zero and solve for p:

$$\begin{aligned} \text{Condition for minimum cost} \quad \frac{q_p \cdot p - q}{p^2} &= 0 \\ \text{Multiply through by } p^2: \quad q_p \cdot p - q &= 0 \\ q_p &= \frac{q}{p} \qquad (10.17) \end{aligned}$$

Equation 10.17 is the first order condition for the maximum quality per unit of price for the principal. The left-hand side is the slope of the agent's BRF and the right-hand side is the slope of the isocost ray. We can see that the slope of the agent's best-response function must be equal to slope of the principal's isocost curve.

With the specific functions, we have

- *Slope of BRF*: The slope of the agent's BRF $q_p = \frac{2\underline{u}}{p^2}$.
- *Isocost ray*: We know that $q = 1 - \frac{2\underline{u}}{p}$. So the slope of isocost ray $\frac{q}{p} = \left(1 - \frac{2\underline{u}}{p}\right)\frac{1}{p}$.

Therefore, we can re-write Equation 10.17 as follows:

$$q_p = \frac{2\underline{u}}{p^2} \quad = \quad \left(1 - \frac{2\underline{u}}{p}\right)\frac{1}{p} = \frac{q}{p} \tag{10.18}$$

M-Note 10.5: Nash equilibrium: price, quality and contract duration

To find out what price Patrisia will pay to her sub-contract Armin, we want to isolate p, so we multiply both sides of Equation **??** by p^2, then isolate p to get:

$$\begin{aligned}
& \frac{q}{p} &=& \frac{2\underline{u}}{p^2} \\
\text{Multiply by } p^2 \quad & pq &=& 2\underline{u} \\
\text{Divide by } q \quad & \Rightarrow p &=& \frac{2\underline{u}}{q}
\end{aligned} \tag{10.19}$$

We can then use Armin's best-response function to eliminate q, by substituting in $q = 1 - \frac{2\underline{u}}{p}$ into Equation 10.19:

$$\begin{aligned}
& p &=& \frac{2\underline{u}}{q} \\
& p &=& \frac{2\underline{u}}{1 - \frac{2\underline{u}}{p}} \\
& p(1 - \frac{2\underline{u}}{p}) &=& 2\underline{u} \\
\text{Dividing by } p \quad & (1 - \frac{2\underline{u}}{p}) &=& \frac{2\underline{u}}{p} \\
\text{Collect terms} \quad & 1 &=& \frac{4\underline{u}}{p} \\
\text{Isolate } p \quad & \therefore p^N &=& 4\underline{u}
\end{aligned}$$

To find out how Armin responds, we substitute p^N into his best-response function:

$$\begin{aligned}
& q &=& 1 - \frac{2\underline{u}}{p} \\
\text{Substitute: } p^N = 4\underline{u} \quad & q &=& 1 - \frac{2\underline{u}}{4\underline{u}} \\
& \therefore q^N &=& 0.5
\end{aligned}$$

Finally we use q^N to determine the expected number of periods that Armin will sell to Patrisia and find that it is $T(q^N)$=$\frac{1}{1-q^N}$=2

Figure 10.9: **The possibility of Pareto-improvements over Nash equilibrium with the contingent renewal incomplete contract.** The shaded area shows feasible quality-price combinations such as point **f**, with lower costs (that lie on lower – meaning steeper – cost isocost rays) for the principal and also on higher iso-value curves for the agent, such v_5. The Pareto-improving shaded lens extending upwards and to the right of **n** *must* exist because the agent's iso-value curve at that point is vertical (it is on the agent's best response function) and the principal's iso-cost ray at that point cannot be vertical.

Incomplete contract isocost: $q = \frac{p}{4\underline{u}}$

Best-response function: $q = 1 - \frac{2\underline{u}}{p}$

Checkpoint 10.8: Armin's pride in the quality of his work

Suppose that Armin has come to have pride in the quality of the work he does, so the disutility of the effort it takes to produce high quality goods that Armin experiences is cut in half. Redraw Figure 10.8, showing what happens to Armin's best-response function and the resulting ratio of quality to price, q/p.

Incomplete contracts, external effects and an inefficient Nash equilibrium

The Nash equilibrium allocation (p^N, q^N) is not Pareto efficient. We can see this in Figure 10.9. Remember that the principal prefers allocations that are higher and to the left (more quality for a lower price) and the agent prefers the opposite. So any point that is both above the isocost line thru the Nash equilibrium (better for the principal) and to the right of the iso-value curve v_4 (better for the agent), is a Pareto improvement over **n**.

In Figure 10.9 the shaded lens is the set of allocations that have these two properties and therefore are Pareto-superior to the Nash equilibrium . You can see, for example that point **f** is better for both Armin and for Patrisia than is the Nash equilibrium, **n**. So **f** is Pareto-superior to **n**. It follows that **n** is not Pareto efficient.

You are already familiar with reason why the Nash equilibrium is inefficient. Like the fishermen over-exploiting their resource and diminishing each others catch, the subcontracting agent in deciding on his action – the level of quality to provide – is not taking account of the effect of his choice on another person, in this case the buyer-principal. The reason this occurs is that the contract is incomplete: it does not cover the quality he provides, which is then another example of an external effect.

We will see in Section 10.11 that if the principal could just purchase quality (offering a particular p for some amount of q as if she were buying electricity or red winter number 2 wheat), then the agent would have a reason to take account of her interest in quality because she would be paying him for it. In this case – where a complete contract is available – the result will be Pareto-efficient.

The agent's enforcement rent

A key element in Patrisia's strategy, remember, and the reason why it works is that she pays him enough so that he wants to avoid being terminated. His enforcement rent is shown in Figure 10.10 as the distance between **n** and **c**. To see this imagine Armin were at the Nash equilibrium, **n**, and we asked:

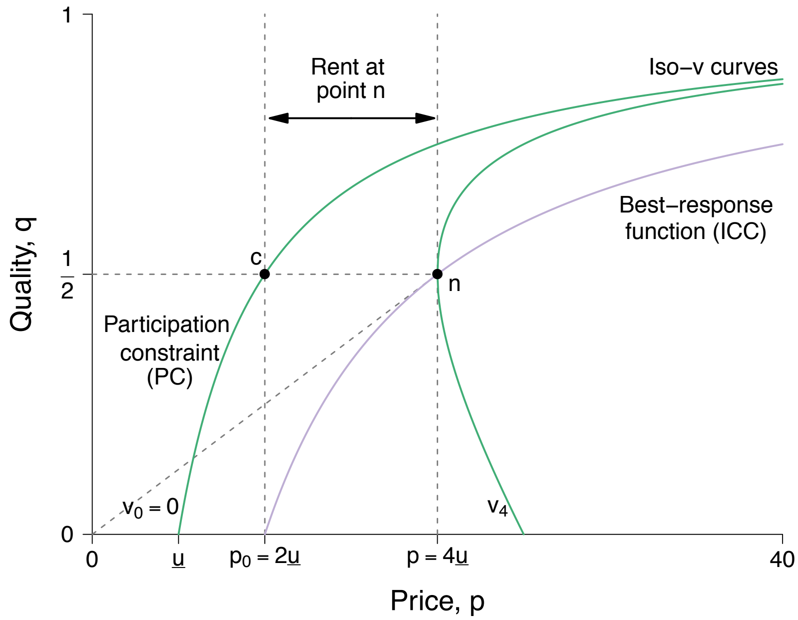


Figure 10.10: **The agent's rent at the the Nash equilibrium.** We can measure the rent that the agent receives at the Nash equilibrium by horizontal (that is, a price difference) between the Nash equilibrium **n** and his participation constraint, $v_0 = 0$.

hypothetically holding constant the level of quality he provides, how much less could he be paid and still be no worse off than at his fallback option, namely zero. To answer the question we imagine moving to the left (lowering the price) from point **n**. We eventually hit the participation constraint $v_0 = 0$. At that point, **c** he receives no rent. So the difference in price between **n** and **c** namely, $2\underline{u} = v_4 - v_0$ is his per period rent.

Checkpoint 10.9: Enforcement rent in another dimension

In Figure 10.10 we measured the rent received by the agent in monetary units, that is the price. We could also measure his rent by answering the following question: at his Nash equilibrium price how much more quality could he provide without being worse of than in his next best alternative. Show this quantity in Figure 10.10.

10.10 Short-side power in principal-agent relationships

The rent that the agent receives each period is the reason why terminating the exchange is an effective threat. If Patrisia cuts him off Armin will not immediately find another principal to transact with on terms as good as his current transaction.

The key feature of these equilibrium contracts is that principals transact with agents who receive economic enforcement rents and *prefer* the current transaction to their next-best alternative. Because some agents receive enforce-

REMINDER: SHORT AND LONG SIDE OF A MARKET THAT DOES NOT CLEAR You learned in Chapter 9 that the *short side* of a market is the side on which at a given price the number of desired transactions is least. The *long side* is the side on which the number of desired transactions is greatest. The short side of a market may be buyers (as in the case of quality) or sellers (as in the case of financial institutions supplying credit, as we will see in Chapter 12.)

ment rents in ongoing contracts, there must exist some other identical agents who are *quantity-constrained*, namely, the suppliers who fail to make a sale. If this were not the case then immediately upon termination the agent could find another principal so the termination would not impose a cost on the agent. It is the fact that the market does not clear that makes the threat of termination effective.

In the Benneton model there are more agents looking for transactions than there are transactions being offered by principals. This is why, if Armin was terminated from his current transaction, it would take him time and a costly search process to find another. (For simplicity, we made the assumption that his next best alternative is to get nothing, but as will see in Section 10.13, this is not an essential feature of the model.)

Patrisia, the employer and other principals in similar principal agent relationships, are on the short side of the market. Armin, the employee and other agents are on the long side; the long side includes agents who would like a transaction but cannot secure one. Armin fears being one of these: unable to sell his product.

We use the term short side power to describe the power that Patrisia is able to exercise over Armin because she is on the short side of a market that does not clear. In Chapter 9 we studied out of equilibrium markets in which supply does not equal demand. In in "Benetton model" of principals and agents, when contracts are incomplete the market does not clear even when it *is* in equilibrium (this will be the case also in of labor and credit markets).

Essential to Patrisia's ability to get Armin to do what she wanted is her ability to threaten to impose a major economic cost on him. This cost – called a sanction, or penalty – is part of our definition of power.

POWER For B to have power over A, it is sufficient that, by imposing or threatening to impose sanctions on A, B is capable of affecting A's actions in ways that further B's interests, while A lacks this capacity with respect to B.

Sufficient condition of power: For B to have power over A, it is sufficient that, by imposing or threatening to impose sanctions on A, B is capable of affecting A's actions in ways that further B's interests, while A lacks this capacity with respect to B.

The definition can be applied to the Benneton interaction. In equilibrium the following conditions hold:

- Armin provides a shirt of higher quality than he would have in the absence of the Patrisia's threatened termination
- Patrisia benefits from this
- Their relationship is asymmeterical: Armin could not get Patisia to act in a way beneficial to him by threatening her with termination.

To see why Armin did not have power similar to Patrisia, imagine that he threatened to end their relationship unless Patrisia raised the price above p^N.

Patrisia would refuse to raise the price, knowing the following:

- because she is on the short side of the market, she could easily find another supplier (remember some of them cannot find a buyer) and
- it would not be in Armin's interest to carry out the threat because he would then have to find another buyer, and as a longsider, this might not be possible.

If Armin were to threaten to sanction Patrisia should she not raise the price (for example, to supply lower quality shirts), his threat would not be credible. This illustrates the idea of a credible threat, namely a threat that will be in the interest of the actor to carry out, if the threat alone did not have its desired effect.

CREDIBLE THREAT A credible threat is one that will be in the interest of the actor to carry out, if the threat alone did not have its desired effect.

The definition of power emphasizes its use to advantage one person over the other. But the exercise of power is also essential to making mutually beneficial exchanges possible when contracts are incomplete.

Imagine, for example, that Patrisia were prevented from threatening to replace Armin with another supplier. Then he would provide only low quality shirts, which she could not sell. As a result she would not purchase shirts from Armin. They would then both be worse off.

The case which we have analyzed – the exercise of power by the buyer (principal) over the producer (agent) – is just one example of power relationships that are sustainable as the Nash equilibrium of a system of voluntary competitive exchanges among private individuals. Other examples include the power that employers wield over employees, or lenders over borrowers, and the other hidden action examples in Table 10.1.

The principal-agent model of hidden actions shows the following:

- that the exercise of power is essential to the principal's strategy
- that a non-clearing market is essential to the principal's ability to exercise power, and
- that, without intending to do so, the result of the principal's actions – along the other actors in the market all acting independently – is a Nash equilibrium in which the market does not clear.

The last point is important and a bit counter-intuitive. In a competitive environment, no principal, acting singly, can create markets that do not clear. But their uncoordinated profit-maximizing actions– setting prices so that agents receive enforcement rents – taken together do just this.

None of this would be possible if markets cleared. To see this we imagine a case in which quality was subject to a complete contract in the Benetton model.

Checkpoint 10.10: Power and contract

- Does Patrisia's contingent renewal contract with Armin satisfy the sufficient contition for Patrisia to be exercising power over Armin?
- Is Patrisia's threat credible?
- Think of a threat that a principal might make that would not be credible.

10.11 A comparison with complete contracts

Suppose that Patrisia has discovered some magical device that can determine exactly the quality offered by Armin and that this information is verifiable. So a complete contract is now possible. She can just name a price and the exact amount of quality that she would like in return. If he does not deliver the goods of the specified quality he does not get paid.

Exchange with a complete contract

Armin will provide that amount, as long as she pays a price that makes Armin even just a bit better off than he would be without the contract. Of course she would not require him to deliver anything close to $q = 1$ because that would be so costly to him to achieve that she would have to pay a very high price in order for him to accept the contract.

REMINDER In a principal-agent relationship, the participation constraint requires the principal to make an offer at least as valuable to the agent as what the agent would get – his fallback – were he to refuse the principal's offer. If the principal's offer is inferior to the agent's fallback, he will not participate in the transaction. Complete contracting in this context is the equivalent to Patrisia having take-it-or-leave-it power over Armin where she controls *both* p and q, whereas with incomplete contracting she only controlled p.

The big difference that the complete contract makes is that Patrisia is no longer constrained by Armin's best-response function – the incentive compatibility constraint – but instead by Armin's participation constraint. This requires that Armin be not worse off agreeing to Patrisia's proposal than he would be were he to walk away, that is, to receive instead his fallback, which as before we assume to be zero.

She no longer has to provide him with a rent – a utility greater than zero, his fallback option – along with a threat to terminate the contract if the goods he has supplied are returned by disgruntled consumers. She can simply refuse to pay for the goods when he delivers them if they are of less quality than she has specified.

Her interaction with the agent is no longer repeated, it is a one-shot game. And it is no longer a principal-agent interaction: the conflict of interest over quality remains, but the second characteristic of a principal agent relationship is now missing: quality is now something that can be enforced by contract.

Nash equilibrium of the take it or leave it complete contracting game

Here is the new complete contracting game:

- Patrisia is first mover and she makes a take it or leave it offer to pay Armin p^C to purchase his product of quality q^C
- Armin either *accepts* the offer and the exchange is carried out or *rejects*.
- In either case, this ends the game.

We use the C superscript to indicate the hypothetical complete contracting Nash equilibrium.

Because, due to the complete contract, she can just purchase quality, we can now interpret Armin's participation constraint as the minimum price at which he is *willing to sell* his products with the quality indicated by the height of the curve. So in Figure 10.11 he would be willing sell his products with quality of $q = 0.5$ if the price were $2\underline{u}$. And the same goes for all of the the price/quality combinations that make up the willingness to sell (participation constraint) curve.

Just as in the case where contracts are incomplete, her maximum will be where the ray from the origin with slope q/p is tangent to the constraint. But this is now the participation constraint, not the incentive compatibility constraint. The outcome of the game is point **c**, with $q^C = 0.5$ and $p^C = 2\underline{u}$. We can confirm that this is a Nash equilibrium:

M-CHECK The fact that the $q^C = q^N$ is just a consequence of the utility function we have selected. The difference between the two contracts that is true in general (irrespective of the choice of functional forms) is that the cost of acquiring effort under the complete contract is less than under the incomplete contract i.e. $c_1 < c_2$.

- Given Armin's utility function and fallback option (described by his participation constraint) Patrisia is doing the best she can by offering the contract: $q^C = 0.5$ and $p^C = 2\underline{u}$.
- Given the her offer and his fallback position Armin is doing the best he can by accepting the contract.

Under the complete contract, Patrisia has paid half as much but received the same quality as when the contract was incomplete; so her miracle machine for detecting quality has cut her costs in half. Were Patrisia to offer the complete contracting price $p^C = 2\underline{u}$ under an incomplete contract, you can see from Figure 10.11 that Armin would provide zero quality (his best response function $q(p) = q(2\underline{u}) = 0$.

Three characteristics of the Nash equilibrium with complete contract are

- *Pareto efficiency*: It is not possible – by choosing some allocation other than $q^C = 0.5$ and $p^C = 2\underline{u}$ – for Patrisia to be better off without Armin being worse off.
- *No rent for the agent*: The supplier's utility under the contract is identical to his next best alternative. Remember we used the distance between point **n** and point **c** to calculate the rent that Armin received under the incomplete contract, so you can see that completing the contract transferred the rent from the agent to the principal.

- *The market clears*: The fact that the agent receives no rent means – by the definition of a rent – that he is no better off with the transaction than he would be at his fallback option, that is, if it were terminated. But this in turn means that his fall back option – what he gets if terminated – must be to immediately secure the same deal — providing q^C for the price p^C from some *other* principal. And for this to be true, it must be that there are no other agents just like Armin who are unable to sell their goods, because if there were he would be among them looking for a buyer, that is, worse of than he was with Patrisia. This is why the absense of rents indicates that markets clear.

Considering the second bullet we know from Chapters 10 and 5 that when one actor optimizes subject to a participation constraint (rather than an incentive compatibility constraint) of another actor, the outcome is Pareto-efficient. This follows from the definition of a Pareto-efficient outcome: If Patrisia has indeed minimized her costs subject to the requirement that Armin have utility of at least zero, then it must be that any *other* allocation would make at least one of them worse off.

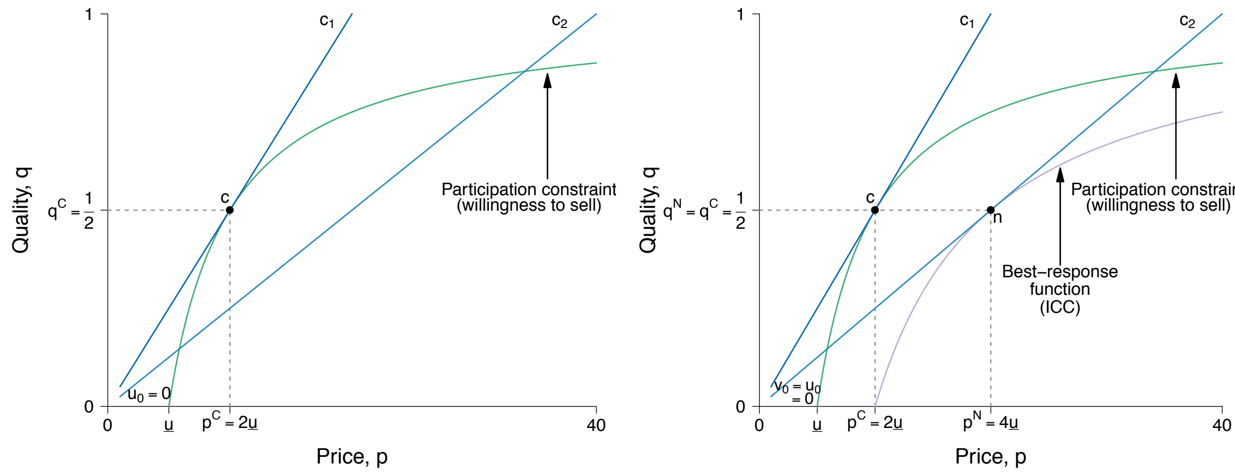
The price-setting game with complete contracts

In order to implement point **c** Patrisia did not need to specify the quality that she required Armin to deliver. She did not have to have take it or leave it power. To see this we modify the complete contacting game.

- Patrisia as before is first mover and can commit to some price at which she will pay for Armin's product depending on its quality: $p^C = \bar{p}q$. She is simply buying whatever quality he provides at the price given.
- Armin responds to her price either by rejecting the contract (if the price is too low) or by delivering a good of a quality level of his choosing.
- Patrisia measures the quality of the good and pays $p^C = \bar{p}q$.
- This ends the game.

Because the contract is complete Patrisia knows that she will get exactly what she pays for, nothing more and (more important to her) nothing less. So she realizes that she did not need to have take or leave it power – dictating the quality as well as the price – in order to transfer the entire rent from the transaction to herself. She could have offered a different contract simply specifying a price: She would purchase any q that Armin provided at a price $p = 4q\underline{u}$. This means that Armin could pick not only point **c**, but also any other point on the isocost line c_1 through point **c**. This line is now Armin's constraint, and he would like to be on the highest indifference curve as possible.

Which point would he pick? Recalling that points to the right and below are better for Armin (higher price and lower quality), point **c** is the point he would



(a) The complete contracting game

(b) Both complete and incomplete contracting

Figure 10.11: **A comparison of complete and incomplete contracts** In Panel a, the complete contracting game in which the principal either has take it or leave it power or simply price setting power, the Nash equilibrium is the allocation where the lowest (steepest) feasible isocost ray is tangent to the agent's participation constraint, which occurs at point **c** with price and quality combination (p^C, q^C) (C for *complete*). The participation constraint in this case is $u = u_0 = 0$ the iso-utility indifference curve rather than the iso-value curve because the game is one-shot. In panel b, the Nash equilibrium (p^N, q^N) is found at the point of tangency of the lowest feasible isocost ray with slope $\frac{q}{p}$, where q and p are the Nash equilibrium price and quantity, and the best-response function with slope $-\frac{2u}{p^2}$. (p^N, q^N) is the *incomplete contract* solution.

c_1 (complete contract): $q = \dfrac{p}{4\underline{u}}$

c_2 (incomplete contract): $q = \dfrac{p}{8\underline{u}}$

Participation constraint: $q = 1 - \dfrac{\underline{u}}{p}$

Best-response function: $q = 1 - \dfrac{2\underline{u}}{p}$

pick (all of the other points in c_1 are worse than **c**, in fact they give him negative utility.

So he would deliver his goods to Patrisia, she would measure their quality, finding **q** to be one-half, and pay him the price $2\underline{u}$ that is $p = 4q\underline{u}$, as promised.

Checkpoint 10.11: Why a complete contract matters

- Explain why the availability of verifiable information about quality and hence the feasibility of a complete contract means that Patrisia is now constrained by Armin's participation constraint, not his incentive compatibility constraint.
- How much per period would Patrisia be willing to pay to rent her "miracle machine" that verifiably measures quality, assuming that she purchases from 100 subcontractors in a period.

M-Note 10.6: The complete contracting outcome

We use Armin's participation constraint to find the complete contracting outcome:

$$\begin{aligned} \text{Participation constraint} \quad u(p,q) &= p - \frac{u}{1-q} = 0 \\ \Rightarrow p &= \frac{u}{1-q} \end{aligned} \tag{10.20}$$

In the case of complete contracts, Armin's participation constraint is the same thing as his willingness to sell: for any given level of quality, Armin has a minimum price at which he is willing to sell to Patrisia. This is given by Equation 10.20

To find the maximum of q/p that is consistent with Equation 10.20 (Armin having at least zero utility) we can substitute in Armin's participation constraint:

$$\text{Maximize} \quad \frac{q}{p} = \frac{q}{\frac{u}{1-q}} = \frac{(1-q)(q)}{u}$$

To find the price selected by this constrained optimization problem we first differentiate this expression with respect to q:

$$\begin{aligned} \text{Product rule} \quad \frac{d\frac{q}{p}}{dq} &= \frac{1}{\underline{u}}((1-q)(1)+(q)(-1)) \\ &= \frac{1-2q}{\underline{u}} \end{aligned}$$

We then set this equal to zero and solve for q^c and by inserting this value into Equation 10.20 we find the value of p^c:

$$\begin{aligned} \frac{1-2q}{\underline{u}} &= 0 \\ 1 &= 2q \\ q^c &= \frac{1}{2} \\ \therefore p^c &= 2\underline{u} \end{aligned}$$

The quality of the product is the same as in the incomplete contracting case, but Patrisia is paying half the price she did in the incomplete contracting case. She is appropriating all the rents from the interaction with Armin, her sub-contractor, and he is no better of than at his fallback option namely $u = 0$.

10.12 Features of equilibria with incomplete contracts: Summing up

Compared to a complete contract, the incomplete contracting contingent renewal Nash equilibrium has six important characteristics, shown in Table 12.1. These characteristics are general features of contingent renewal incomplete contracts, and do not depend on the utility function, the fallback option or any of the other special assumptions we have made in the Benetton model:

1. *Equilibrium rents*: Under the incomplete contract, the agent receives a rent above his next best alternative. In Table 12.1 you see that he has a utility of $2\underline{u}$ for each period and that because he provides $q^N = 0.5$ the transaction is expected to last for the inverse of this, two periods, giving him a rent of $4\underline{u}$.

2. *Price making*: The principal is a price-maker, not a price-taker. The reason why the principal does not treat the price as given is that she can benefit from changing the price, given the contractual incompleteness concerning the quality of the good. Price-making in the incomplete contracting context does not derive from any non-competitive aspect of the assumed market structure, such as monopoly power. The Benetton model can include many principals and many agents and there are no barriers to entry by either type, but contractual incompleteness allows the principals to benefit by setting prices.

REMINDER As we explained in Chapter 9 price taking means taking some price (or wage) as given and not varying it as a way of doing better (increasing profits, reducing cost).

3. *Competitive equilibrium without market clearing*: Because the agent receives a rent, we know that unlike the complete contracting case, his next

Contract over q, p	*Incomplete*	*Complete*
Constraint on principal's optimization	ICC	PC
Nash equilibrium price paid	$4\underline{u}$	$2\underline{u}$
Nash equilibrium quality provided	$\frac{1}{2}$	$\frac{1}{2}$
Principal's payoff per period, $\frac{q}{p}$	$\frac{1}{8\underline{u}}$	$\frac{1}{4\underline{u}}$
Agent's utility per period	$2\underline{u}$	0
Enforcement rent?	$4\underline{u}$	0
Principal a price maker?	yes	no
Market clearing in equilibrium?	no	yes
Principal's short side power?	yes	no
Durable interactions?	yes	no
Pareto-efficient Nash equilibrium?	no	yes

Table 10.3: **Complete and incomplete contracts: A summary of differences** The numerical entries are from the model in which the agent's utility is given by Equation 10.1

best alternative cannot be to just walk across the street and contract with some other principal on identical terms. Instead, he will have to search for a new partner, during which time he will be without a transaction. But for this to be the case there must be other agents also searching for buyers: if he were the only buyer without a contract then given the ordinary turnover from jobs (quits, deaths and other reasons for leaving a job) he would find a similar job without delay. So the fact that agents on the supply side of the market who are transacting with principals are getting a rent is diagnostic about the state of the market: there must be excess supply. It is not the case that there is excess supply *because* agents are receiving rents. We will turn to the question – why does the market not clear when contracts are incomplete – in Chapter 11.

EXAMPLE A fever (elevated body temperature) is diagnostic of some kind of illness, most often an infection. The fever is evidence of the infection; not the cause. Similarly rents are diagnostic of excess demand or excess supply, not the cause.

REMINDER In Chapter 9 we showed that rents exist when there is excess demand or excess supply, so that there are quantity constrained buyers or sellers. Rents are a money measure of the more conventionally quantity-measured excess supply or excess demand. Correspondingly, when rents exist it must be true that the market does not clear.

4. *Durable Transactions*: The principal (buyer) and agent (supplier) will interact over many periods, even though there are many identical buyers and suppliers. Competitive equilibrium with contingent renewal incomplete contracting will be characterized by a series of durable *bilateral* trading islands rather than a sea of anonymous traders engaged in one-shot interactions in spot markets.

5. *Endogenous claim enforcement through the exercise of short-side power*: The buyer (principal) in the contingent renewal contract minimizes costs by threatening to terminate the ongoing relationship with the seller (agent). Because of this threatened sanction, the agent acts in the principal's interests, costs to benefit the buyer by providing a product of higher quality.

6. *Pareto-inefficient Equilibrium*: Because the buyer maximized taking the supplier's best-response function as the (incentive compatibility) constraint, rather than the supplier's participation constraint and because the two constraints differ, the non-contractual equilibrium will not be Pareto-efficient.

Checkpoint 10.12: Why can't they just agree?

If moving from the Nash equilibrium to point **f** in Figure 10.9 is a Pareto improvement (both Patrisia and Armin benefit), why don't they just agree to each change their strategies so that they can make the move?

10.13 Incomplete contracts and the distribution of gains from exchange

In the model discussed so far, the agent could choose any level of quality. By restricting the agent to two feasible quality levels, we get a simpler model that focuses on the most important aspects of the problem and that we will find useful in other contexts, such as Chapter 11.

An incentive compatible price offer

In this stripped-down version of the scenario, Armin the agent and supplier of the variable quality shirt may offer either low or high quality, at a disutility cost $\underline{c}$ and $\bar{c}$ respectively where $\underline{c} < \bar{c}$. Patrisia (the buyer) never mistakenly thinks that a high quality shirt is low quality. But if Armin has provided her with a low quality shirt she will detect this with probability t.

Here is the game.

- Patrisia is first mover and offers a price p for Armin's product (not knowing whether it is of high or low quality).
- Given the price offered, Armin decides to produce high or low quality and delivers the good to Patrisia.
- If he has delivered a low quality good then with probability t Patrisia detects this and refuses to pay, in which case Armin must sell the good to another buyer at a lower price.
- If the good is of high quality, or if it is of low quality but not detected by Patrisia, she pays for the good.

To ensure that Armin will provide high quality, Patrisia must offer a price high enough so that his expected income from offering high quality, $p - \bar{c}$, is not less than his expected income when he offers low quality. Patrisia has to make it a best response for Armin to provide high quality.

The game with just two levels of quality is similar to the games in Chapter 1 where the players had just two choices, for example fishing 10 or 12 hours. So here the incentive compatibility constraint (ICC) that limits Patrisia's optimization problem is not a range of prices p that she must offer if she wants quality

q. It is a single price that will make it in Armin's interest to choose high over low quality.

The interaction along with the agent's payoffs are shown in the game tree in Figure 10.12. The right-hand side branch gives the result for when the agent provides high quality: he bears the cost of producing high quality $\bar{c}$ and is paid the price p.

The left-hand branch shows how the game proceeds if he produces low-quality. There are two outcomes that might occur if he chooses not to produce high quality:

- With probability t Patrisia detects the low quality, refuses to pay Armin, and he gets his fallback price p^z selling the good to some alternative buyer and so has income $p^z - \underline{c}$
- With probability $(1-t)$ Patrisia does not detect the low quality, so she pays him p and his income is $p - \underline{c}$.

His expected income is just the income he receives in these two events, multiplied by the probability of each of them occurring. This is the right-hand side of Equation 10.21.

$$\text{ICC:} \quad p - \bar{c} \geq (1-t)(p - \underline{c}) + t(p^z - \underline{c}) \tag{10.21}$$

To find the lowest price Patrisia can offer that will induce the Armin to provide high quality, we rearrange Equation 10.21 to isolate p (as shown in M-Note 11.7):

$$\text{Nash equilibrium price:} \quad p^N = \frac{\bar{c} - \underline{c}}{t} + p^z \tag{10.22}$$

Patrisia will set the price p^N and Armin will provide high quality. This is a Nash equilibrium because at the price p^N Armin would not do better by providing low quality, and given the incentive compatibility constraint based on what she knows about Armin's behavior, Patrisia cannot do better than to offer p^N. If she offered a higher price than p^N she would be throwing away money. If she paid less than p^N, then he would produce low quality.

M-CHECK So that we could express Equation 10.22 as an equality rather than an inequality, we have assumed that as long as the expected income from providing low quality is not higher than the expected income from providing high quality, Armin will provide high quality.

From the equation for the equilibrium price, you can see that the least price that Patrisia can offer compatible with Armin supplying high quality will be:

- higher, the greater is the difference in cost to provide high rather than low quality
- higher, the greater is the agent's fallback price, p^z
- lower, the better the principal at detecting low quality, that is, the higher is t.

./figures/information/info_gametree.pdf

Figure 10.12: **The quality game, with the agent's payoffs.** The game tree (like the one for the Ultimatum Game in Chapter 2) gives the order of play (from the top down), the actions that each participant can take at each node (branching point) in the tree, and the agent's payoffs that will result from each path through the tree.

Incomplete contracts, enforcement rents and profits

These results replicate some of the economics you have already learned about exchange with incomplete contracts. But in one respect this model goes further: it provides us with a measure of *how incomplete* a contract is, and why this matters. It gives us a measure of how asymmetric the relevant information is. Armin knows if he has produced a low quality good; but Patrisia will discover this only with probability t. So if we denote the quality of Armin's knowledge as 1 (one-hundred percent) then the difference between this and the quality of Patrisia's knowlege (t) or $(1-t)$ is an indicator of the extent of information asymmetries. The accuracy of her information – t itself – is a measure of how complete the contract is.

We now can see the effect of contractual incompleteness on the distribution of income between the principal and the agent. To do this, let's assume that Patrisia can sell a high quality shirt for some given price p^B (for Benneton). Then her maximum willingness to pay Armin for a good shirt is p^B. If she were to sell at that price her profits would be zero. We do not need to consider the case in which she ends up with a low quality good and has to (attempt to) sell that. If she pays Armin p^N he will not deliver low quality goods.

Armin's minimum willingness to sell is the price at which he could have sold a low quality good, plus compensation for the extra cost of producing high quality, or $\bar{c} - \underline{c} + p^z$. The difference between her willingness to pay p^B and his willingness to sell $\bar{c} - \underline{c} + p^z$ is the total rent to be gained from exchange of one unit. This is the quantity that will be divided up between them in the form of her profits and his rent. How it is divided up will depend on how complete the contract is.

Figure 10.13 plots Equation 10.22, showing that the Nash equilibrium price differs for different degrees of contractual completeness. It illustrates how the distribution of the gains from exchange depend on the extent of information asymmetry and contractual incompleteness. The vertical distance between the agent's willingness to sell and principal's willingness to pay is the "pie" that is to be divided up into the two "slices" the principal's profit and the agent's rent.

Equation 10.22 shows that if the contract were complete ($t = 1$ so that low quality could be detected with certainty), Patrisia need do nothing more than to meet Armin's participation constraint paying his minimum willingness to sell price

$$\text{Complete contract price} \qquad p^C = \bar{c} - \underline{c} + p^z \tag{10.23}$$

Even at this low price he would provide high quality because in this case his best response to that price is to produce high quality. To see this, think about Armin's options. He could produce high quality and get $p^N - \bar{c}$ for sure, or produce low quality and get $p^z - \underline{c}$ for sure.

If the contract is incomplete, meaning $t < 1$, however, Figure 10.13 shows that Patricia will offer Armin a price greater than his willingness to sell. You can also see that the size of the resulting rent increases more the more incomplete the contract is.

As expected, the more complete the contract, the larger is the share of the principal's profits. Do not conclude, however, that from the agent's standpoint the more asymmetric the information the better. There is some level of information asymmetry ($\underline{t}$ in the figure) below which the best the principal can do would be to pay the agent more than the price at which she can sell the trademarked good to a consumer. In this case there is no way for the transaction to satisfy the *principal's* participation constraint. So she will stop purchasing shirts from suppliers, trademarking them, and selling them to consumers. With no exchange taking place there are no gains from trade to be shared.

We can use this simple model of the quality problem to better understand a fast-growing kind of work in many countries: the gig economy.

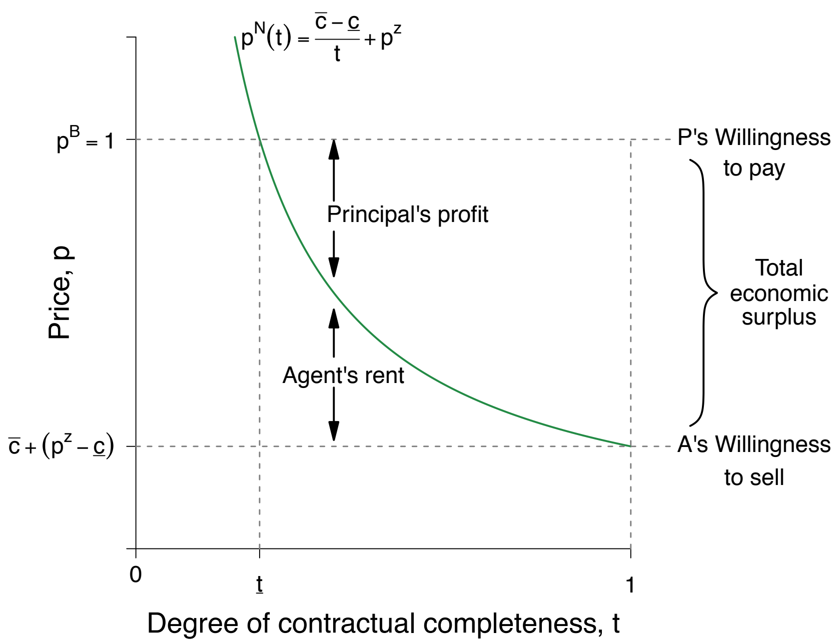


Figure 10.13: **Contractual incompleteness and the distribution of the economic surplus.** The total rent made possible in the transaction difference between the principal's willingness to pay for a high quality good and the agent's willingness to sell a high quality good, or p^B minus $\bar{c} - \underline{c} + p^z$. We set $p^B = 1$ so that we can interpret the quantities shown as percentage shares of the price sold.

Checkpoint 10.13: Quality testing

Explain why having a perfect test of the quality of the goods that Armin produced accomplishes the same result as being able to enforce a complete contract.

M-Note 10.7: From the ICC to the Nash equilibrium price

We find the Nash equilibrium price, by rearranging the incentive compatibility constraint, Equation 10.21.

$$\text{ICC:} \quad p - \bar{c} \geq (1-t)(p - \underline{c}) + t(p^z - \underline{c})$$

Express this as an equality because the principal would never pay the agent more than necessary to secure high quality:

$$p - \bar{c} = t(p^z - p) + p - \underline{c} \tag{10.24}$$

Add $\underline{c} - p$ to both sides:

$$\underline{c} - \bar{c} = t(p^z - p)$$

$$\text{Divide by } t: \quad \frac{\underline{c} - \bar{c}}{t} = p^z - p \tag{10.25}$$

Solve for p, which is:

$$p^N = p^z + \frac{\bar{c} - \underline{c}}{t} \tag{10.26}$$

The agent's fallback is $\bar{c} - \underline{c} + p^z$ so the enforcement rent the agent receives, $p^N - (\bar{c} - \underline{c} + p^z) = \frac{1-t}{t}(\bar{c} - \underline{c})$.

This is equal to the difference between the costs of providing high and low quality $\bar{c} - \underline{c}$ multiplied by $\frac{1-t}{t}$, which is the ratio of the probabilities of escaping termination and being terminated if providing the low quality good.

10.14 Application: Complete contracts in the gig economy

A 'gig' for a musician or comedian is a single appearance for which they will be paid not by the hour, but an agreed sum for the performance. The *gig economy* is not about jokes and tunes, however, it refers to the combined activities of Uber or Lyft drivers, TaskRabbits, UpWorkers, Mechanical Turkers and others who transport people and goods, home-assemble online purchased furniture, and perform other well-defined tasks for which they are paid a fixed rate.

FACT CHECK App-based urban transportation services like Uber use network-based software of the type used by eBay to profit from the massive expansion of smart phones with GPS capabilities. In New York City, four major app-based companies provide point to point service, in addition to Uber: Lyft, Juno, and Via. There are now twice as many Uber cars serving New York than there are licensed taxis. In 2017, these companies provided 160 million rides, and by February 2018, they were making twice as many trips as were conventional taxis. There are 6 times as many app-based vehicles providing service as there are taxis. Most app-based drivers work full time.

In many legal jurisdictions, gig workers are considered private contractors and not employees. They provide their own cars or tools and gain access to their gigs by means of a two-sided platform that connects those who will pay for the gig, and those who perform it. The company – Uber, for example – sets the prices and determines the number of people are allowed to use their app.

The gig economy is a small portion (in the U.S. not more than 2 percent of employment) even of those high income economies where, for example, ride services like Uber and Lyft have made significant inroads against conventional taxi firms. The gig economy is growing because modern information technology makes it easier to match buyers and sellers – drivers and those needing a ride, or those needing a task done in their home – and others with the time and skill do the job.

App-based ride hail, delivery, and other parts of the gig economy provide an illuminating contrast with the model of hidden actions with variable quality studied in this unit. The key difference is that in some cases the tasks performed are sufficiently well-defined and easily measured so that a virtually complete contract is possible: if the person is not delivered from the hotel to the airport, the Lyft driver does not get paid; if the Ikea shelves purchased online are not assembled properly, the TaskRabbit does not make a penny.

This is equivalent to the simple model of the quality problem in which the principal is the person who has engaged the gig worker (the agent) to do a job. But in this case the signal of quality, $t = 1$ so that if the job is not done, the principal does not pay.

An important feature of the gig economy is that the only way that drivers or taskers can get gigs is through the platforms owned by a few firms such as

Uber, Lyft, TaskRabbit, Mechanical Turk and others. This means that those performing the gigs have no real bargaining power. If a Rabbit tasker objects to the terms, there will always be another tasker to take her place, but few if any other ways that the disgruntled tasker could find a gig.

The computer platform that allows those who need a gig performed to connect to those performing the gigs makes possible substantial mutual benefits in putting together gig workers who have free time and the skills, a vehicle, or other equipment required with those willing to pay for a completed gig. But because there are few platforms and many potential gig workers, they typically receive very little pay for often difficult and onerous work.

A result is that gig performers in this economy face extraordinary economic insecurity: they are not guaranteed a fixed schedule of hours and pay, nor do they receive health insurance benefits, maternity leave, holiday pay, or pension contributions through their employer. Working full time, the hourly earnings of the vast majority of app based drivers for ride hail services in New York City would place them below the official poverty level for New York.

What accounts for the low pay. The answer is that gig companies do not need to substantially pay more than the gig worker's next best alternative. Beyond compensating the tasker for her time and trouble, they do not need to motivate the Tasker to do the job as specified: if it is not done, the Tasker will not be paid. A result is that the gig economy can often produce services at a lower cost and price than are available from conventional firms that, as we will see in the next chapter, pay their employees more than their next best alternative.

FACT CHECK In reality, gig workers are often paid considerably *more* than their next best alternative, as is evident from the rapid increase in the number of people ready to become Uber, Lyft and other gig drivers.

We can illustrate what this means with the 'two levels of quality' model just introduced. Think about some task, for example, assembling a bicycle that a person has just purchased in a kit. A Rabbit tasker might take on this job for an agreed upon fee. If the task is not performed – the bicycle does not work properly when assembled, for example – this information will be available to the purchaser and the tasker will not be paid. In terms of the model, the probability that low quality (a non-functioning bike) will be detected, t, is much closer to 1 than in conventional jobs.

HISTORY The gig economy resembles what was called "putting out" in 18th century Great Britain. In this system, women spun on home spinning wheels using cotton fibers provided by a "putting out" company. They were paid by the amount produced, just as gig workers are paid by the task completed. With many households ready and able to take up spinning, and in most cases just a single putting out company, the spinners had little bargaining power, and were paid low prices for their products.

But even in the gig economy contracts are not complete and so exchanges are made possible by a combination of the (incomplete) contract itself along with two aspects of buying and selling less frequently studied by economists – the exercise of power that we discussed above and the effect of social norms.

10.15 Application: Norms in markets with incomplete contracts

EXAMPLE According to Lisa Bernstein in the diamond industry :

> . . . disputes are resolved not through the courts and not by the application of legal rules announced and enforced by the state . . . [but rather by] an elaborate, internal set of rules complete with distinctive institutions and sanctions.

When contracts are complete, you get what you pay for. So for a given price, there is little economic reason to be concerned about one's exchange partner's psychological makeup or moral commitments. If you do *not* get what you paid for you get your money back at no cost to yourself. This is what a complete contract means: its terms are enforced – if necessary – by the courts, not by the parties to the exchange. And this is also why we care about who we interact with a lot more in cases where contracts are incomplete.

To see this put all of the people with whom you have any economic interactions into two groups: those whose names you know and those whose names you do not know. You probably do know the names of your employer, your doctor, the person you consult with for legal advice, and perhaps your car mechanic. These are all exchange in which the contract is substantially incomplete. Do you also know the name of the gas station attendant, the checker at the super market, or the clerk at the store from whom you purchase milk?

When contracts are incomplete, parties to an exchange – whether buyers or sellers – will favor social interactions where exchange is *personal* and *durable*. Exchange is *personal*, rather than *anonymous*, when the parties to exchange have personal knowledge of each other, such as personal histories and knowledge of whether the other party is trustworthy or untrustworthy. Exchange is *durable*, rather than one-shot, when it results in long-term repeated interactions between the parties to exchange, such as when you regularly go to hair stylist you like, or you trust a car mechanic or a babysitter you've known for a long time.

REMINDER Recall Bernstein's explanation that diamond traders address the fact that quality is not easily determined (mentioned in the introduction of this chapter), by relying on an "internal set of rules", rules that have informally emerged from *within* the diamond industry.

HISTORY The economic historian Avner Greif (1994) analyzed the divergent cultural and institutional trajectories of the traders from Genova, Italy and North African Maghrebi traders in the late medieval Mediterranean from this perspective. The Maghrebi traders had what Greif terms a "collectivist" system of contractual enforcement whereby none of them would ever deal with anyone who had ever failed to fulfill their contractual obligations. The individualism of the Genovese traders, on the other hand, precluded the high levels of cooperation and loyalty to one another on which the Maghrebi system depended. But the the limits of Genovese individualism also provided an impetus for their development and perfection of an ultimately more successful system of state and other third-party enforcement of contract terms.

The relationship between contractual incompleteness and market structure can be seen in the contrasting structures of the rice and raw rubber trade in Thailand. Buying and selling in the the wholesale rice market – where the quality of the product is easily determined by the buyer impersonal – buyers and sellers hardly know one another – in contrast with the personalized exchange based on trust in the raw rubber market. In the raw rubber market quality is impossible to determine at the moment of purchase. As a result, buyers purchased rubber repeatedly from the same sellers rather than shopping around, a strategy that gave them the ability to exercise the kind of short-side power Patrisia has over Armin.

Similarly, in villages like Palanpur (in India), wheat and rice as well as seeds and fertilizer are standardized, easily measured, commodities and are subject to relatively complete contracting. These inputs are bought and sold in region-wide markets in which transactions are governed by little more than the going price and the budget constraints of the participants. The markets are impersonal and anonymous.

By contrast, exchanges concerning labor, credit, the use of land, and the services of farm assets such as bullocks take place almost entirely within the village, and often within the same caste. Moneylenders in villages rarely extend loans to those not known to them or not living in the village. The village markets in goods or services with *incomplete contracts* are *personalized*.

Experimental evidence

A number of experiments also show differences in behavior depending on the possibility of complete contracting.

Economists have investigated a variety of experimental markets to understand the decisions people make when buying and selling goods.Experimental researchers can change the "institutions" under which participants interact. Remember, institutions are rules of the game, so in an experimental game, changing institutions just means changing the rules. For example, one experiment could have the institutional structure of *complete contracts*, and another *incomplete contracts.*

FACT CHECK The sociologist Peter Kollock investigated "the structural origins of trust in a system of exchange, rather than treating trust as an individual personality variable" with similar results. Using an experimental design based on the exchange of goods of variable quality, Kollock found that *trust* in and *commitment* to trading partners as well as a concern for one's own and others' *reputations* emerges when product quality is variable and non-contractible but not when quality is subject to complete contracts.

Brown, Falk, and Fehr designed a market experiment to explore the effects of contractual incompleteness on patterns of trading. The good exchanged varied in quality, with higher quality more costly to provide. In the complete contracting condition, the experimenter enforced the level of quality promised by the supplier, while in the incomplete contracting condition the supplier could provide any level of quality (irrespective of any promise or agreement with the buyer).

Buyers and sellers knew the identification numbers of those they were interacting with, so they could use information they had acquired in previous rounds as a guide to whom they would like to have as trading partners, and the prices and quality to offer. Buyers had the opportunity to make a private offer (rather than broadcasting a public offer) to the same seller in the next period, thus attempting to initiate an on-going relationship with the seller.

Table 10.4: **Contractual incompleteness and market social structure: experimental evidence**

Structure of interactions	*Complete contracts*	*Incomplete contracts*
Duration	one shot	contingent renewal
Offers	public	private
Price determination	haggling, offers rejected	price setting by short-sider
Traders	anonymous relationship	trust, retaliation for cheating
Market networks	many thin connections	bilateral trading islands

Very different patterns of trading emerged under the complete and incomplete contracting conditions. In the complete contract condition, 90 percent of the trading relationships lasted less than three periods (and most of them were one-shot). By contrast, under the incomplete contracting condition only 40 percent of the relationships were less than three periods, and most traders formed *trusting* relationships with their partners. Buyers in the incomplete contracting condition offered prices considerably higher than the cost of providing quality (just as in the principal-agent shirt quality model). When buyers were disappointed by the quality supplied, they terminated the relationship, withdrawing the implied enforcement rent from the supplier. Other differences are summarized in Table 10.4. The behavioral differences in complete and incomplete contracting treatments were particularly pronounced in later rounds of the game, suggesting that the subjects updated their behaviors according to experience.

These experimental results suggest that there may be a two-way relationship between trust, reciprocity and other social preferences on the one hand and the degree of contractual completeness on the other.

- Where contracts are incomplete – as in the above experiment – economic interactions may endure over long periods during which people develop trusting and reciprocal relationships; while this is unlikely to be the case where contracts are complete.
- Where people are trusting and reciprocal, making the contract "as complete as possible" may not be worth the legal costs and possible offense to one's trading partners. But if people are entirely self interested, trying to complete the contract may be the only way to do business.

10.16 Conclusion

The most important organizations governing exchanges in modern economies are firms, whose managers combine other peoples' labor and (what Adam Smith called) "other people's money," neither of which are subject to complete contracting, to produce and market goods and services. Labor and credit markets are typical of the many important exchanges in which what is transacted are not well-defined and easily measured objects, like the nuts and apples in Ronald Coase's example in the headquote for this chapter. In these markets the transaction involves something quite different and much more difficult to enforce – the promise to repay the loan – or to measure – the promise to work hard on the job.

Coase put it this way: "what are traded on the market are not, as is often supposed by economists, physical entities, but the rights to perform certain actions . . . the objects of exchange are complex bundles of obligations and claims concerning who should do what under what conditions."

In the next two chapters we use the principal-agent model you have just learned to study how the owners and managers of firms – as employers and borrowers – structure the rights to perform actions concerning other people's labor and other people's money, respectively.

Making connections

Limited information: Asymmetric and/or non-verifiable information about the quality of goods or other aspects of an exchange results in contracts that are incomplete (they do not cover all that matters to one of the parties to an exchange and or they are unenforceable by the court), a common feature of modern economies that will be important in the remaining chapters.

Incomplete contracts, external effects, and coordination failures: Because contracts are incomplete one or more actors will not take appropriate account of the effect of their actions on others; these external effects are similar to the environmental external effects (e.g. over-fishing) in earlier chapters in that they result in coordination failures, that Pareto is inefficient outcomes.

Optimization limited by incentive compatibility and participation constraints: In principal-agent interactions the fact that contracts are incomplete means that the relevant constraint is not the agent's participation constraint but instead her incentive compatibility constraint. This is the reason why the Nash equilibrium cannot be Pareto efficient.

Mutual benefit and conflict over distribution: Like other economic actors, when principals and agents interact they each do so in order to gain something they value, so exchanges are mutually beneficial; but there is also a conflict about the distribution of these mutual benefits.

Price making: Principals are price makers, not price takers (and as we will see, wage makers and interest rate makers).

Non-clearing markets: In the Nash equilibrium of the market for the variable quality good there is excess supply (some agents are quantity constrained and unable to sell their products). Competitive markets that do not clear in equilibrium – including those with excess demand such as the credit market – are a feature of principal agent models.

Power and social norms: When contracts are incomplete, the private exercise of power and social preferences such as trust and reciprocity (along with contracts) provide the basis for mutually beneficial exchange.

Experimental evidence: Behavioral experiments can clarify how the kinds of contracts – complete or incomplete – affects the social structure of exchange.

Important ideas

(in)complete contract	disutility (of effort/quality)	enforcement rent
in-group favoritism	endogenous distribution of gains	verifiability
time & future states	measurability	authority
motivation	repeated games	strategic (a)symmetry
information (a)asymmetry	power	social norms, trust & fairness
principal	agent	conflict of interest
(im)perfect information	hidden actions	hidden attributes
adverse selection	moral hazard	lemons problem
insurance	contingent renewal contract	isocost curve
indifference curve	iso-value curve	Solow condition
Best-response function (BRF)	Marginal rate of substitution	fallback
enforcement rent	price-making	non-market clearing
quantity constraints	durable transaction	endogenous claim enforcement
Pareto (in)efficiency	Nash equilibrium	short-side power
long-side of the market	no-shirking condition	

Mathematical notation

Notation	**Definition**
q	agent's provision of quality
p	price the agent receives from the principal
$u()$	agent's utility function
$\underline{u}$	agent's disutility of providing quality
$v()$	agent's value function of the ongoing commercial relationship
t	principal's termination schedule
T	length of the contract for a given termination schedule
$\bar{c}$	agent's disutility cost of high quality
$\underline{c}$	agent's disutility cost of low quality

Note on super- and subscripts: N: Nash equilibria; C: complete contract; z: fallback position.

Discussion questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

See reference list.

11

Work, Wages & Unemployment

> If a workman moves from department Y to department X, he does not go because of a change in prices but because he is ordered to do so . . . for certain remuneration [the worker] agrees to obey the directions of the entrepreneur . . . the distinguishing mark of the firm is the suppression of the price mechanism. (387, 389)
>
> Ronald Coase, "The Nature of the Firm" (1937)

DOING ECONOMICS

This chapter will enable you to:

- Explain why the employment contract is incomplete and how the labor market differs from markets in which contracts are complete.
- Understand how, when an employer hires a worker, both may be better off as a result, why there will a conflict over the distribution of these mutual benefits, and how a more complete contract favors the employer in this conflict.
- Show how the employer chooses a wage to maximize the profits and explain how along with the threat of termination the employment rents that workers receive motivate them to work hard and well.
- Show that in competitive Nash equilibrium: i) the employer exercises power over the worker, ii) the wage and effort level are Pareto-inefficient, and iii) involuntary unemployment will exist.
- Analyze wages, markups, unemployment and profits using a model of firms' price- and wage-making in the whole economy.
- Understand the conditions under which the imposition of a minimum wage will induce a firm to hire more workers rather than fewer.

11.1 Introduction

On the morning of January 5, 1914, a little known mechanic turned automobile producer named Henry Ford shocked his colleagues and competitors by announcing that he would pay his workers a minimum of five dollars for an 8-hour day, at once shortening the work day and more than doubling the hourly rate of pay for the vast majority of his workers. Ford was not responding to insufficient labor supply: A reporter arriving that morning for the press conference at which the announcement would be made noticed a line of several hundred workers seeking employment.

In the weeks following the announcement, the queue outside the gates swelled to over 12,000, almost as many as were working inside. Remarkably, profits rose, supported by a more than a doubled increase in output per hour of production labor. Ford would become a household name around the world and the combination of high wages and assembly-line work came to be called "Fordism."

For the lucky workers who had been in the right place at the right time, the basic facts of work life inside the plant changed beyond recognition. The previous year Ford's labor force had averaged 13,623. During the course of that year 50,448 had walked out the door, most had quit. 8,490 had been fired. The year following the announcement, employment had grown by a third, but the number quitting had fallen to a tenth of its earlier level, and only

./figures/Employment/rivera_cropped.jpg

Figure 11.1: **Workers at the Ford plant in River Rouge, near Detroit, USA.** The mural is one of many that Henry Ford's son and successor, Edsel paid Diego Rivera, a member of the Central Committee of the Communist Party of Mexico, to create.

27 employees had been fired. Changes of this magnitude clearly cannot be explained by cyclical variations in supply and demand in the local labor market. It seems unlikely that Ford doubled the wage to attract better workers or to retain those workers in whom the company had invested expensive training. A Ford superintendent boasted that "two days is . . . ample time to make a first-class core molder of a man who has never seen a core-molding bench in his life."

Figure 11.2: **Made by happy (or at least well paid) workers: A model T Ford. Salt Lake City, UT. Harry Shipler.**

Exactly why Ford raised wages and shortened the work day remains a mystery. More important, the success of his gamble is a puzzle, for it contradicts the view that profit maximization entails paying employees a wage as low as possible consistent with their showing up, that is to say, satisfying their participation constraint and nothing more. Ford's five-dollar day would not make sense in this model because five dollars a day was much better than most workers' fallback option. The fact that people were lining up for jobs at Ford even before the wage increase tells us that even then they were paid much more than was required by their participation constraint.

11.2 Employment as a principal-agent relationship

The most likely reason why Ford doubled the wage is that he understood that raising the wage can reduce the cost of labor. A principal agent model explains how this seemingly paradoxical statement could be true. The key idea is that labor – the activity that produces cars – is not something you can measure in hours on the job. Instead it consists of tasks done and, like the quality of goods in the Benetton model, these tasks cannot be written into a complete contract. Getting the job done may require paying workers a lot more than their fallback option.

Incomplete employment contracts

To see how a principal-agent model might explain why Ford's radical move worked, recall that a **principal-agent relationship** arises when two conditions hold:

- *Conflict of interest*: the actions or attributes of the agent affect the payoffs of the principal in such a way that there is a conflict of interest between the principal and the agent.

- *Incomplete contract*: the agent's actions are not subject to an enforceable contract either because the they are not known to the principal, or, if known, are not verifiable for some other reason.

The aspect of the exchange between Ford and his workers that fits these two conditions is the workers' effort on the job, completing the tasks required to produce Model T's (the only car Ford produced at the time).

	Benetton Model	*Ford model*
Principal	Buyer (e.g. Benetton)	Employer (e.g. Ford Motor Company)
Agent	Producer & seller	Worker
Non-contracted action by agent	Quality of the good provided, q	Effort and care by the worker, e
Contingent renewal	Termination of sub-contract, $t(q)$	Termination of employment, $t(e)$
Agent's fallback option	Another buyer after searching	Another job after job search
Principal sets	Price and termination, $p, t(q)$	Wage and termination, $w, t(e)$
Short side of the market	Demand (excess supply of goods)	Demand (excess supply of labor)

Table 11.1: **A comparison of two principal agent models: quality and effort.** What we have labeled the Ford model is also called the labor discipline model and sometimes the "efficiency wage model."

- The conflict arises because Ford profits if his employees work harder or faster, while the workers preferred to work at a slower pace both for their safety and in order to go home a bit less exhausted at the end of the day.
- But the workers' effort – manifested in literally hundreds of tasks performed per day, many as part of teams of workers, was not something that Ford could measure and write into an enforceable contract.

To see why the employment contract is necessarily incomplete, think back to three of the reasons why this is the norm rather than the rule in a modern economy given in Chapter 10:

- *Asymmetric or non-verifiable information*: Think about cyber-loafing: texting with her friends and web-surfing at work. The extent of this may be unknown to the employer, and even if it is known, the evidence of it may not be something that could be used to enforce a contract.
- *Time*: The worker takes a job today and the employer would like to renew her employment over a period of months or even years. The employer has no way to determine the tasks he would like her to do under all of the possible conditions that might arise over this period.
- *Measurement*: For most work tasks there are no measures of work done – e.g. quantity and quality of task completion – that are precise enough to be the basis of an enforceable contract. This is the case both because the tasks are difficult to measure, and their completion typically depends on the efforts of more than a single worker.

HISTORY Herbert Simon (1916-2001) who, like Coase, won a Nobel Prize in economics, provided the first model of the firm along these lines. In a 1950 article, he represented the employment contract as an exchange in which the employees transfer authority over their work tasks to the employer in return for a wage. Simon stressed the advantage to the employer of this arrangement given the unavoidable uncertainty about the tasks that would be required over the course of the contract, and therefore the high cost of agreeing to a complete contractual specification of the activities to be performed.

Here is some evidence that the labor contract is generally incomplete. In cases where "work done" *can* be written into a contract, employers benefit; but contracts of this kind – called piece rates – are extremely rare. Some examples where paying piece rates seems to have increased worker effort:

PIECE RATE A piece rate contract compensates a worker not by hours worked, but by "pieces" of output produced.

- When a pay system for auto glass installers in the U.S. shifted from hourly wages to paying piece rates, output per worker rose by one fifth.
- When British Columbia tree planters randomly assigned to piece rate compensation – they were paid by the number of trees planted – they outperformed by 20 per cent other planters randomly assigned to a fixed wage.

But even with the substantial growth of the gig economy – where some of the contracts approximate piece rates – the number of piece rate workers as a fraction of all workers in the U.S. economy is not more that five percent. This is in part due to the fact that industries that once extensively used piece rates — clothing and shoe production, for example — now employ very few people. And sectors in which "work done" is almost impossible to measure – caring for others, personal security, knowledge production and distribution – have grown significantly.

FACT CHECKTunisians who worked for others some of the time for a fixed wage were half as productive as when the same person worked their own farm and owned the output of their own work.

EXAMPLE Another piece of evidence that employment contracts do not cover everything that the employer expects to get by hiring worker: "work to rule" is a trade union strategy in which workers perform those tasks that are specifically required in the contract, and nothing more. Very little production takes place under these conditions.

EXAMPLE What it means for a contract to be incomplete is illustrated by the difference between hiring someone to care for your child for an afternoon, and hiring an Uber driver to take you somewhere. If the Uber driver does not show up, or shows up half an hour late, he will not get paid. If you get home and your child is miserable, they you may wonder if the baby sitter cared well for her, but you will pay her anyway. Complete contract case: You were prepared to purchase a particular service from the Uber driver, and it was not delivered. Incompllete contract case: You hired the baby sitter for a block of time, and hoped she would do a good job.

A model of employment as a principal-agent relationship

Based on the key idea that work effort is not subject to contract, here is the principal agent game between the employer and worker. While the agent's action is now work effort per hour of employment e rather than product quality q and the principal will pay the agent a wage w rather than a price p, the structure of the game is the very similar Benetton model as is clear from Figure 11.1 and Table 11.1.

1. The employer, the principal, is first mover. He announces to the worker a wage w and the offer to renew the contract at the end of each period unless the worker is terminated for insufficient effort, which occurs with a probability that is inversely related to the effort provided by the worker, $t(e)$. We call the wage rate and termination schedule introduced in this step the employer's labor discipline strategy.
2. The worker, one of a group of identical workers who may be employed by the principal, responds to the employer's offer by selecting a level of effort to expend per hour of employment, e.
3. The principal then chooses the total hours of workers' time that he would like to employ, h, the agent is employed, and production takes place.
4. At the end of the period the employee is renewed with probability $1-t(e)$ and terminated with probability $t(e)$. If terminated the employee receives her fallback option, which for now we assume to be zero. If the worker is terminated this ends the game.

5. Conditional on the worker not being terminated, repeat steps 1-4 above with the values of w, e, and the termination schedule unchanged.

6. Repeat the previous step until the worker is terminated, which ends the game.

We call this the labor discipline model (or the Ford model in recognition of the car maker's $5 day). To understand the game we need to ask about each of the above steps: what is the actor attempting to accomplish and what do they know at that particular stage?

1. The employer selects the labor discipline strategy e, and $t(e)$ to *maximize* his profits (which requires minimizing the average cost of a unit of effort $\frac{w}{e}$), *knowing* the worker's best-response function $e(w)$.

2. When selecting a level of effort to perform the worker is *maximizing* the expected value of her employment, *knowing* the employer's labor discipline strategy as well as her fallback option.

3. When selecting the level of employment to hire (h) the employer is *maximizing* his profits and *knows* the cost of a unit of effort (which is the result of the labor discipline strategy he has implemented), how worker effort contributes to output, and the demand for the firm's product.

We take up these three steps in the next section. But because the employer has to know the worker's best-response function before deciding on a labor discipline strategy, we take up step 2 before turning to step 1 and finally step 3.

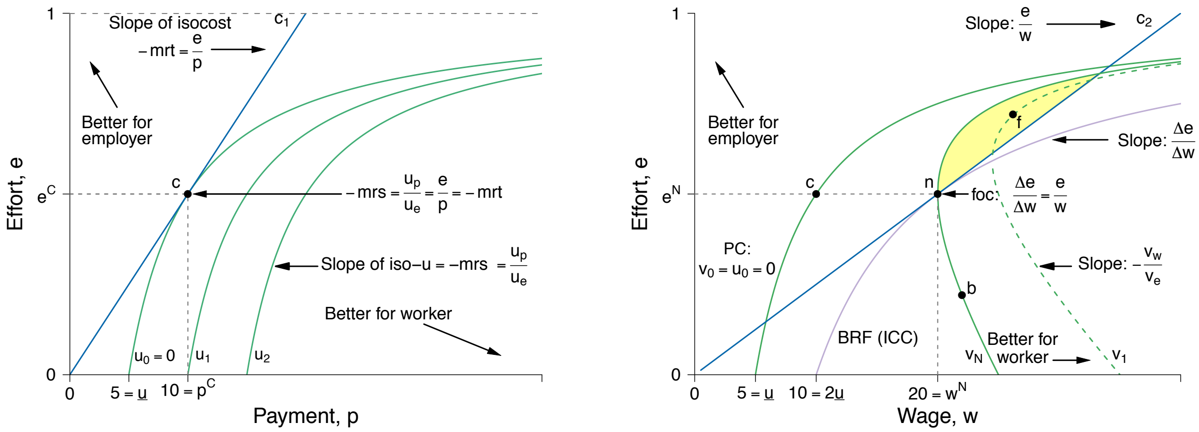
11.3 Nash equilibrium wages, effort, and hiring

Because the Ford model is very similar to the Benetton model and uses the same mathematical functions, you can quickly learn to use the model by studying the comparison in Table 11.1 and the representation of the model in Figure 11.3. The functions describing the workers utility and best-response function are shown in M-Note **??**.

The workers' best response and employer's labor discipline strategy

In the left panel of Figure 11.3, we show indifference curves representing the per period utility function of the worker. The horizontal axis is the payment to the worker whether it be in the form of a wage or a price (for example if piece rate payment is possible). Included is the curve labeled u_0 which gives the combinations of e and p such that the worker's utility is zero, that is equal to his next best alternative or fallback option.

In the right panel we show two of the worker's iso-value curves. Points **b** and **n** of the curve v_N for example show combinations of the wage and effort by



(a) Utility-based indifference curves and the complete contract

(b) Iso-value curves and the incomplete contract

Figure 11.3: **Comparison of complete and incomplete labor contracts** The figure shows how the Benetton model introduced in Chapter 10 can be adapted to represent the relationship between an employer and a worker. Both panels show the employer's and the worker's very different evaluations of points in the figure, the employer preferring more effort and less pay for the worker and the worker preferring the opposite in panel a, and preferring high wages in panel b. The per period utility-based indifference curves in the left panel along with the employer's termination function are the basis for the multi-period iso-value functions in the right panel. The point of maximum value for a given price (like point **n**) gives us the worker's best-response function $e(w)$. The rays from the origin show the best the employer can do under the incomplete contract (point **n** in panel b) and the complete contract (point **c** in panel a). The yellow shaded area is the Pareto-improving lens showing all of the pairs of wages (w) and effort (e) that are preferred by both employer and worker over the Nash equilibrium, w^N, e^N. We explain this lens and the complete contracting case in Section 11.5.

the worker that have the same expected value to the worker, taking account of both the per period utility of those two combinations of effort and pay but also the effect of working harder (at point **b**), namely extending the expected duration of the job (by reducing the probability of termination). The worker expects the same value from point **n** as point **b**, even though she is working harder at point **n** for a lower wage, because she expects that she will be terminated much sooner if she is working less hard, i.e. at point **b**.

The point where the iso-value curve is vertical (for example point **n**) gives the effort level (the vertical axis coordinate of the point) that is the worker's best response to the wage (the horizontal axis coordinate of the point). The best-response function is composed of points like **n** where the indifference curves (not shown) are vertical. The worker's best-response function is the incentive compatibility constraint (ICC) on the employer's profit maximizing strategy.

$$\text{Worker's best-response function (ICC)} \qquad e = e(w) \qquad (11.1)$$

In the right hand panel of Figure 11.3 the iso-value curve labeled v_0 gives the combination of effort levels and wages such that the worker receives a utility of zero, which is the same as his fallback option. This curve is identical to the curve labeled u_0 in the left panel because in order for the job to be worth zero to the worker, it must be that the utility of having the job for a single period (shown by the indifference curves in the left panel) is also zero.

M-Note 11.1: Review of model set-up and solution with specific utility and profit functions

The functions and solutions for this labor discipline model are set up and solved almost identically in Chapter 10. Here is a brief review of the set up and solutions. Make sure you are able to reproduce these results yourself. You can check the M-Notes in Chapter 10 for guidance.

Complete contract model:

- Employer chooses price p and verifiable effort level e to maximize $\frac{e}{p}$ subject to worker participation constraint
- Worker utility function $u(p,e) = p - \frac{\underline{u}}{1-e}$
- Worker participation constraint $u(p,e) \geq 0$
- Nash equilibrium effort level $e^c = \frac{1}{2}$, price $p^c = 2\underline{u}$

Much of solving the complete contract model is outlined in M-Note **??**.

Incomplete contract model:

- Employer chooses wage w and hours h to maximize profits subject to worker best-response function
- Given w, worker chooses effort level e to maximize present value of utility $V = \frac{u(w,e)}{t(e)} = (w - \frac{\underline{u}}{1-e}) \times \frac{1}{1-e}$
- Worker first order condition: $u_e = t_e v$
- Worker best-response function from maximizing v: $e(w) = 1 - \frac{2\underline{u}}{w}$ (see M-Note 10.3)
- Employer maximizes $\frac{e}{w}$, subject to $e(w)$, the worker's best-response function (the employer's ICC)
- Employer first order condition (Solow condition): $\frac{\Delta e}{\Delta w} = \frac{e}{w}$
- Nash equilibrium wage $w^N = 4\underline{u}$, effort level $e^N = \frac{1}{2}$

We solve for equilibrium hours hired h later in the chapter, as we have not done this before.

Turning to what the employer cares about, the cost of effort, that is $\frac{e}{w}$, is the same as a ray from the origin ($\frac{e}{w}$ is its slope). The employer would like to be on a steeper isocost ray, meaning higher $\frac{e}{w}$ or what is the same thing lower cost of a unit of effort $\frac{w}{e}$. But points above the worker's best-response function are not incentive compatible: there is no way that the wage and effort combination given by these points could come about. In other words the employer would like to, minimize the cost of a unit of effort (c_l) as follows:

$$\text{Minimize} \quad c_l = \frac{w}{e} \tag{11.2}$$

$$\text{Subject to the ICC} \quad e = e(w) \tag{11.3}$$

The Solow condition

In the right panel of Figure 11.3 this means finding a point on the employee's best-response function (the employer's incentive compatibility constraint or ICC) that is on the the steepest possible isocost ray. This will be point **n**, where the isocost ray is tangent to the best-response function, that is, where the slopes of these two lines are equal:

$$\text{Slope of ICC} = \frac{\Delta e}{\Delta w} = \frac{e}{w} = \text{Slope of isocost ray} \tag{11.4}$$

Equation 11.4 gives us the solution to the constrained optimization problem shown in Equations 11.2 and 11.3. It is the rule that tells the employer the wage that will minimize his cost of a unit of effort, called the Solow condition after the macroeconomist Robert Solow, who first demonstrated it. The condition can be restated as: choose the wage such that the marginal effect of raising the wage is equal to the average level of quality per dollar of wage spent. In M-Note 11.1 we show how the Solow condition gives us the profit-maximizing wage that the employer will offer.

LABOR DISCIPLINE MODEL The labor discipline model presented here was developed by one of us (Bowles) to try to make sense of the movements of wages and labor productivity – called the great productivity slow down – during the late 1960s and 1970s Its initial purpose was not academic at all but instead was the basis of advice requested by a number of trade unions and public interest bodies seeking to understand the end of "the golden age of capitalism." Other variants of the model – that developed by Shapiro and Stiglitz for example – were motivated by the desire to provide Keynesian ideas about unemployment with a microeconomic foundation without making ad hoc assumptions such as "wage stickiness" (the tendency of wages to maintain their levels despite recessions).

You can determine that the wage and effort level given by the Solow condition are the Nash equilibrium of the labor discipline and effort provision part of the game, so we now give them the N superscripts.

- *Wage:* Given that the worker has adopted the strategy described by her best-response function $e(w)$, the best the employer can do to minimize the cost of effort is to select the wage w^N.
- *Effort:* Given that the employer has offered w^N the best the worker can do to maximize the value of her job is to provide e^N.

This completes the first step of the employer's profit maximizing process: finding the labor discipline strategy that minimizes the cost of effort. Now, knowing the cost of a unit effort $\frac{w^N}{e^N}$ the employer will choose how much effort to employ in producing its goods. To do this it must determine the number of hours of workers' time to hire h^N.

11.4 The employer's profit-maximizing level of hiring

LABOR is the amount of actual work devoted to production, that is the effort provided per hour by a worker, e, times the hours of workers time hired, h, or l = eh. Labor is measured in units of effort (sometimes called "efficiency units") not in hours.

An employer would normally face two questions concerning hours: how many hours a day will each worker work, and what is the total number of hours to be hired (this will determine the number of workers to hire, given the length of the working day). For simplicity we address only the second question, so h is just total hours hired by the employer.

Hiring hours, and employing effort

REMINDER The hours worked in a day makes a big difference to individual workers and their families, and as you know from Chapters 3 and 7, the hours worked during the course of a year differs substantially among countries and changed markedly over the course of the 20th century.

But hours of workers time is not what produces the goods the employer wishes to sell; that is done by workers' effort. So we distinguish between

- the number of hours of workers time hired by the employer h called *hours hired* and
- the total amount of actual work devoted to producing goods, which will be the effort provided by each worker in an hour times the hours hired, or $he^N = l$ called *labor employed*.

To determine the hours hired the employer makes use of two analytical tools about which you already know:

- the *demand function* for the firm's product showing the maximum number of units of the good x that can be sold at price p or $x(p)$, and
- the *production function*, showing the combinations of the amount of capital goods k and labor l that can produce each level of output x, or $x(k,l)$.

As is the case in selecting the level of output that you studied in Chapters 8 and 9, the owners of firm maximize profits by producing a level of output such that marginal revenue equals marginal cost. To see what this implies for the employer's hiring, we introduce a new term, the **marginal revenue product of labor**, which is the change in total revenue associated with a small change in labor hours hired. As is shown in M-Note 11.3 this is the marginal product of labor (derived from the production function) multiplied by the marginal revenue (derived from the demand function).

When deciding on how much labor to employ, therefore, the employer weighs two things:

- *marginal cost of labor* : the effect on total costs of using more labor in production, which we know is $c_l = \frac{w^N}{e^N}$, against
- *marginal revenue product*: the benefits of hiring that worker.

Here the marginal cost of labor does not depend on the amount of labor hired, so the average and marginal cost are the same. In Section 11.12 we introduce what is in some cases the more realistic case — termed monopsony — in which the cost of employing labor increases the more the employer hires. But for now c_l is fixed by the cost minimizing labor discipline strategy of the employer.

If the marginal revenue product of hiring more exceeds the marginal cost of hiring more, the employer will hire more. It will continue hiring more until it reaches the level of employment so that the marginal revenue product equals the *marginal cost of labor* (mcl).

M-Note 11.2: The marginal revenue product of labor (effort)

The information relevant to the problem is:

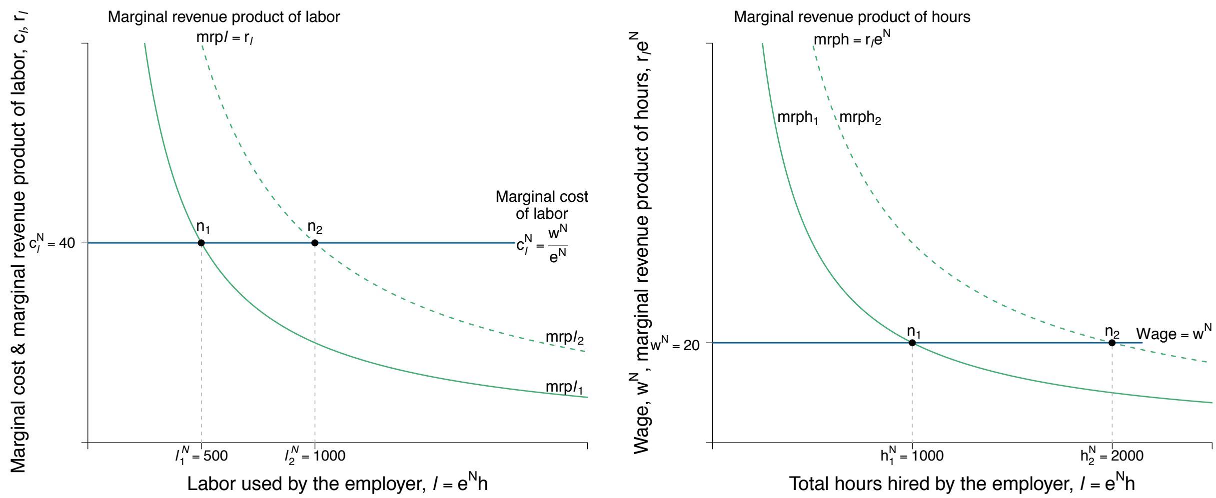
- *The production function* $x = x(k,l)$
- The *inverse demand curve* for the firm's product $p = p(x(k,l))$, and
- the firm's *total revenue* is $r(k,l) = p(x(k,l))x(k,l)$

In doing the differentiation below, keep in mind the two places where the labor variable l appears as an argument in a function: the inverse demand function and the production function.

The marginal revenue product of labor is defined as the change in total revenue associated with a small change in labor devoted to production, that is, the derivative of the revenue function with respect to l:

MARGINAL REVENUE PRODUCT OF LABOR The marginal revenue product of labor is the change in total revenue associated with a small change in labor employed or $\frac{r}{\Delta l}$ where Δl is small, which is the marginal product of labor times marginal revenue.

MARGINAL COST OF LABORThe marginal cost of labor is the change in total costs associated with employing more labor or $\frac{\Delta c}{\Delta l}$ where Δl is small.



(a) Employer's profit-maximizing level of labor employed

(b) Employer's profit-maximizing level of hours hired

Figure 11.4: **Employer's hiring decision**. In panel **a** we show the firm's decision about how much labor (that is total amount of effort) to employ as an input into the production process, given that employer's decision on the wager rate w^N and the resulting amount of effort per worker hour h^N has determined the cost of a unit of effort c^N. In panel **b** we show the same profit maximizing problem in terms of the hours of workers time hired. The marginal revenue product of hours is just the marginal revenue product of effort times the amount of effort per hour e^N. The cost of an hour of a worker's time is just the wage w^N.

$$\text{MRP of Labor} = \frac{\partial r}{\partial l} = \underbrace{\overbrace{\frac{\partial p}{\partial x}\frac{\partial x}{\partial l}x}^{\text{Revenue lost due to lower price}} + \overbrace{p\frac{\partial x}{\partial l}}^{\text{Revenue gained due to increased sales}}}_{\text{Marginal revenue product of labor}}$$

The marginal revenue product can be decomposed into two terms, one negative and the other positive. The first term on the right-hand side is the negative effect of employing more labor (and producing more) on the price of the product at a given level of sales (because the demand curve is downward-sloping). The second term on the right-hand side is the positive effect of employing more labor and producing more revenues (at a given price). Together these two terms constitute the marginal revenue product of labor.

We can rearrange this equation so that it reads

$$\text{MRP of Labor} = \frac{\partial r}{\partial l} = \underbrace{\overbrace{\left(\frac{\partial p}{\partial x}x + p\right)}^{\text{Marginal revenue (increase in x)}} \overbrace{\frac{\partial x}{\partial l}}^{\text{MPL}}}_{\text{Marginal revenue product of labor}}$$

from which we see that the marginal revenue product of labor is the product of two terms:

- the marginal product of labor
- the marginal revenue associated with an increase in x

We illustrate this case in the left panel of Figure 11.4. The horizontal line labeled $c_l^N = \frac{w^N}{e^N}$ is the cost of effort determined by the employer's choice of a cost minimizing labor discipline strategy. The downward-sloping solid curve shows effect of employing more labor (by hiring more hours, given e^N) on the marginal revenue product of labor.

If the employer employs less labor than l_1^N the marginal revenue product of labor is greater than the marginal cost so the employer would make greater profits by employing more. Similarly if the employer were employing more than l_1^N it would see that the marginal cost of labor is greater than the marginal revenue product of labor so it would profit by cutting back on employment. So the profit maximizing level of employment is l_1^N.

In the right panel of Figure 11.4 we translate the decision on how much labor (total effort or eh) the employer will employ into how many hours of workers' time he will hire. The right and left panel provide different presentations of the same information: on the left we see the employer's decision on how much of the input – labor – to use (l_1^N), while on the right we see the way that he implements that decision (given the Nash equilibrium level of effort $e^N = \frac{1}{2}$) by hiring some number of hours of workers' time h_1^N. (M-Note **??** provides a numerical example.)

The level of hiring, h_1^N is a Nash equilibrium because w^N, e^N and hence c_1^N are all Nash equilibria. Hiring h_1^N is the best the employer can do, given those values and the employer's production function and demand function (which determine the marginal revenue product of labor curve).

The marginal product of labor time and the wage

From the right panel of Figure 11.4 we can see that the wage given by the Solow condition determines the level of hiring by the firm. As a result it also determines the marginal revenue product of labor hours. If the cost-minimizing wage rate for the firm had been higher, the marginal revenue product would have been higher at the employer's profit maximizing level of hiring.

The dashed lines in Figure 11.4 show the effect of an increase in the marginal product of labor (which also increases the marginal revenue product of labor in the left panel and the marginal product of worker hours in the right panel). The effect is to increase the level of hiring, so as to depress the (shifted up) marginal revenue product of labor now to the level of the wage.

The wage does not change. The wage was determined by the cost minimizing labor discipline strategy (the Solow condition), and that is not affected by the productivity of labor.

M-Note 11.3: Profit-maximizing employment of labor, given the cost of a unit of effort $\frac{w^N}{e^N}$

The information relevant here is the same as in the previous M-Note except that we now introduce the following:

- the cost of a unit of labor is $c_l^N = \frac{w^N}{e^N}$ given by the Solow condition
- the cost of a capital good is c_k and
- the employer's total cost is $c_l l + c_k k$
- the employer's *profit* is its total revenue - total cost: $\pi = p(x(k,l))x(k,l) - c_k k - c_l l$

To find the level labor to use in production that maximizes profits, we differentiate the profit function with respect to l, and set the result equal to zero. So using the chain rule for differentiation we have:

$$\frac{\partial \pi}{\partial l} = \underbrace{\overbrace{\frac{\partial p}{\partial x}\frac{\partial x}{\partial l}x}^{\text{Revenue lost due to lower price}} + \overbrace{p\frac{\partial x}{\partial l}}^{\text{Revenue gained due to increased sales}}}_{\text{Marginal revenue product of labor}} - \underbrace{c_l}_{\text{Marginal cost of labor}} = 0 \tag{11.5}$$

If this condition (Equation 11.5) is not satisfied it means that either

- the marginal revenue product of labor exceeds the marginal cost of labor in which case the employer should employ more labor or
- the opposite, in which case the employer should employ ess labor.

We can also re-arrange Equation 11.5 to find the following:

$$\frac{\partial \pi}{\partial l} = \underbrace{\overbrace{\left(\frac{\partial p}{\partial x}x + p\right)}^{\text{Marginal revenue for an increase in x}} \overbrace{\frac{\partial x}{\partial l}}^{\text{Marginal product of labor}}}_{\text{Marginal revenue product of labor}} - \underbrace{c_l}_{\text{Marginal cost of labor}} = 0$$

M-Note 11.4: Finding the profit-maximizing hours of labor (effort) hired using specific functions

Here we use a specific revenue function to the profit-maximizing number of hours to hire. To do this we first determine how much total effort the owner would like to devote to production, l^N. Assume that the employer is using some given level of capital stock, $\bar{k}$.Holding constant $\bar{k}$ and varying the amount of labor, the revenue function is:

$$r(l) = 20000\ln(l)$$

As was shown in Equation 11.5, the employer's optimal level of employment is such that the marginal revenue product equals the marginal cost of labor. Thus, the employer's first order condition is:

$$\text{Marginal revenue product of labor} = \frac{\partial r}{\partial l} = \frac{w^N}{e^N} = \text{marginal cost of labor}$$

We – and the employer – already know equilibrium levels of w^N and e^N. Thus, we can solve for l^N:

$$\begin{aligned}\frac{\partial 20000ln(l)}{\partial l} &= \frac{w^N}{e^N}\\ \frac{20000}{l} &= \frac{20}{\frac{1}{2}}\\ l^N &= 500\end{aligned}$$

Knowing that $l = eh$, we can use l^N and e^N to solve for equilibrium hours hired:

$$\begin{aligned}l &= e^N h\\ 500 &= \frac{1}{2}h\\ h^N &= 1000\end{aligned}$$

Checkpoint 11.1: Effects on the level of labor used and hours hired

Now think about what happens if we change some of the parameters in the model:

- How does an increase in the cost of a unit of effort affect the amount of labor that the employer devotes to production? Which shift does this correspond to and how does it affect the hours of workers' time that the employer hires?

M-Note 11.5: Employer's first order conditions: the general case

We've presented the employer's choice of labor discipline strategy and work hours as separate problems. But, in reality, the profit maximizing employer solves both simultaneously. To see this, we will work with a general model.

As before, the employer chooses the wage w and hours h, given the worker's best-response function $e(w)$, and gets profit π. We will assume that the employer is using some given level of capital stock, $\bar{k}$. Holding constant $\bar{k}$ and varying the amount of labor, the profit function of each employer is:

$$\begin{aligned}\pi &= \text{total revenue - total cost} \\ \pi &= r(e(w)\cdot h) - wh\end{aligned}$$

As we have defined $l = he$, we can rewrite the profit function as:

$$\pi = r(l) - wh$$

Given the worker's best-response function $e(w)$, employers will now choose w and h to maximize profits, resulting in two first order conditions:

$$\pi_h = r_l e - w = 0 \quad (11.6)$$

$$\pi_w = r_l h e_w - h = 0 \quad (11.7)$$

Here we are using the notation that x_y is the partial derivative of x with respect to y. So $r_l = \frac{\partial r}{\partial l}$. We can use both equations to derive the Solow condition:

$$\begin{aligned}\pi_h &= r_l e - w = 0 \implies r_l = \frac{w}{e} \\ \pi_w &= r_l h e_w - h = 0 \implies r_l = \frac{h}{he_w} = \frac{1}{e_w} \\ \therefore \pi_h &= 0 \implies \frac{w}{e} = \frac{1}{e_w} \implies \frac{e}{w} = e_w\end{aligned}$$

$$\text{giving the Solow condition:} \quad \frac{e}{w} = e_w \quad (11.8)$$

The Solow condition tells us that the employer wishes to equate the marginal effect of wages on the amount of effort the worker provides (the right-hand side of Equation 11.8) to the ratio of effort to wages (or the average effort per dollar of wages). This can be seen in figure 11.3 to be the point where the slope of the worker's best-response function equals the slope of the employer's isocost ray, which gives us the equilibrium wage and resulting effort level.

Given this equilibrium level of w^N and the resulting e^N that workers will choose from their best-response function, Equation 11.6 can be rearranged to show the profit-maximizing level of hours hired:

$$\pi_h \quad = \quad r_l e^N - w^N = 0 \implies w^N = r_l(h)e^N$$

This tells us that the employer chooses how many hours to hire where the marginal revenue product of hours equals the wage. This can be seen in Figure 11.3.

11.5 Comparing the incomplete and complete contracts cases

To understand better the Ford model of employment under complete contracts, we now provide a contrast with a hypothetical case of complete contracting in which the employer can effectively purchase the worker's work, not just her time.

Complete contracting

A complete contract would require that the worker deliver some specified level of effort. This would be approximated by cases in which piece rate compensation – mentioned earlier –is possible or a few very routine tasks such as data entry in which a person could be paid by the keystroke.

We show the complete contracting case in the left panel of Figure 11.3. As in the case of the incomplete contract, the employer would seek to minimize the cost of a unit of effort. So, just as in the right panel, he would seek to be on the steepest isocost ray. But now he would just offer a price for some quantity of effort, rather than hiring the worker by the hour, and he will offer the lowest price consistent with her willingness to sell.

The employer's cost minimization would be constrained by the worker's minimum price at which she is willing to sell effort. To determine this remember that the worker's next best alternative if not interacting with this employer is to get nothing. So for each level of e provided, the worker's minimum willingness to sell is given by the indifference curve u_0. This is the participation constraint limiting the employer's cost minimizing attempts. The point on that participation constraint that is on the steepest possible isocost ray is indicated by point **c**, that is, where the isocost ray is tangent to the participation constraint.

REMINDER Point **c** in Figure 11.3 is derived in exactly the same way as that we derived the Nash equilibrium in the Benetton model, that is point **c** in Figure 10.11

MATH CHECK The fact that The level of effort provided in the Nash equilibrium of the complete contracting case, e^c, is the same as in the incomplete contracting case is a coincidence due to the functional forms we are using. We could have $e^c > e^N$ or $e^c < e^N$

Table 11.2 summarizes the contrast between the hypothetical case of complete contracting and the more realistic incomplete contracting case. Concerning model set-up in the first 4 rows, in both cases the principal maximizes profits, is free to set any price or wage that it wishes, and is competing with other firms.

The major difference in the model setup is that the constraint on profit maximization differs: as you have just seen, it is the participation constraint if the contract is complete and the incentive compatibility constraint (the worker's

Model characteristics and results	*Incomplete*	*Complete*
Employer maximizes	Profits	Profits
Subject to (employer's constraint)	ICC	PC
Competition among firms and workers	yes	yes
Flexible wages and prices?	yes	yes
Employment rent for worker?	yes	no
Market clearing (no unemployment)?	no	yes
Pareto efficient Nash equilibrium?	no	yes
Employer's short side power?	yes	no

Table 11.2: **Complete and incomplete contracts: A summary of model structure and characteristics of the Nash equilibrium.** The first four lines refer to model structure while the remaining four refer to results. Remember ICC means incentive compatibility constraint and PC means participation constraint. Notice that the worker is better off (receiving a rent) in the Pareto-inefficient incomplete contract case.

best-response function) when the market is incomplete. A major difference between the two models follows from this.

Workers receive employment rents

The employer offers the worker a wage which along with a termination schedule that results in the worker preferring to keep the job rather than lose it, given her fallback option. The reason is that if the principal offered a transaction that was no better than the agent's next best alternative (i.e. a transaction on the participation constraint) the agent would receive no rent and would not care if the transaction ended. The result would be that the quality supplied by the seller and the effort supplied by the worker would be whatever they pleased (maybe zero in both cases) for there would be no fear of job loss.

The existence and the magnitude of the rent can be seen in Figure 11.3: it is the distance **cn**. To see this, we ask: is she better off at the Nash equilibrium than in her reservation position, namely $v = 0$? By comparing the iso-value curve at the Nash equilibrium with her iso-value curve at her participation constraint, we can see that the answer is yes, she does get a rent.

To determine the size of the rent we ask: at the level of effort that she works in the Nash equilibrium e^N how much less could she be paid and still be no worse off than at her participation constraint? The answer is given by the distance **cn** in Figure 11.3.

Checkpoint 11.2: Employment rent in two dimensions

The distance *cn* is a money measure of the employment rent, but it can also be measured in effort: If the worker is being paid w^N how much more effort could he provide and still be no worse off than her fallback option. Use figure 11.3 to show this quantity.

The labor market does not clear: There is excess labor supply

The existence of the employment rent means that the labor market does not *clear*: there are identical workers who are without a job and for whom $v = 0$ who would prefer to be employed receiving v^N, but are unable to make a transaction with an employer. Those workers unable to make a transaction are *quantity-constrained*, that is, they are unable to purchase or sell as much as they want – as many hours of labor – at the going terms of exchange. The quantity of labor that is supplied by workers at the equilibrium wage exceeds the quantity of labor demanded by employers, which means there is excess supply of labor or unemployment.

We explained in the Chapter 10 that the workers' rent is evidence *that* unemployment exists, not the reason *why* it exists. We will see why it exists later in the chapter.

Checkpoint 11.3: Undercutting

To check that point **n** in Figure **??** is really a Nash equilibrium we need to consider the unemployed, and whether they – like the employer and the worker – are doing the best they can. Suppose an unemployed worker went to an employer and promised to work as hard as a member of his current workforce but at a slightly lower wage. How would the employer respond? Why would this unemployed worker *not* get the job?

Pareto inefficiency

The Nash equilibrium exchange (e^N, w^N) is *Pareto-inefficient*. This is true for exactly the same reason that the outcome of the equilibrium exchange in the Benetton model of Chapter 10 is Pareto inefficient: the principal in both cases is maximizing profits subject to a best-response function (incentive compatibility constraint), *not* a participation constraint. In both cases given that the contract is incomplete, the participation constraint is not the relevant limit on the principal's cost minimization.

We can see from figure 11.3 that the Nash equilibrium allocation (w^N, e^N) is not Pareto efficient. The yellow Pareto-improving lens shows possible allocations that like point **f** are both preferred by the employer, who would get more effort for a lower wage, and also preferred by the worker, who would be on a higher iso-value curve. This Pareto-improving lens must exist as long as the isocost ray and the iso-value curve are not tangent: and they cannot be tangent, as the iso-value curve must vertical if it is on the worker's best-response function and the isocost ray cannot be vertical.

The reason why the Nash equilibrium in the complete contracting case – point **c** in the left panel of Figure 11.3 – *is* Pareto efficient is that if contracts are complete the employer is limited by the participation constraint. Remember that Pareto inefficiency results when an actor does not appropriate account of the effect of their actions on others. Being limited by the participation constraint forces the employer to do this.

Here is why. The participation constraint is the workers willingness to pay curve, and its slope – the marginal rate of substitution – is the workers *own* evaluation of the trade off between working harder and being paid more. Because at point **c** the employer's isocost ray is tangent to the workers participation constraint, it follows that the price of effort to the employer $\frac{p}{e}$ is exactly the cost to the worker of providing that effort, the worker's marginal rate of substitution.

Exactly the same reasoning applies to aspects of the job other than the wage, including workplace amenities such as flexible work hours, a respectful and safe work environment. The employer will take *some* account of workers' preferences with respect to these amenities, because labor discipline depends on the worker having a lot to lose if she is dismissed. For example, if there is an inexpensive way to make the job more valuable to the worker – like installing air conditioning, preventing managers for sexually harassing workers, or providing paid parental leave – then the employer will see this as a possible cost-cutting measure. The employer could introduce the inexpensive amenity and as a result be able to reduce the wage without reducing the worker's employment rent. But, the extent of "worker-friendliness" of the job dictated by profit maximization profits is not Pareto-efficient.

To see why, think first about the ideal case. If the employment contract were complete, the employer would be constrained by the worker's participation constraint. Because the participation constraint is the worker's own utility at their fallback, the employer will maximize profits by evaluating the importance of workplace amenities (relative to the wage or other things that the worker cares about) *exactly as the worker does.*

But this result does not hold when effort is not subject to contract because in this case it is not the participation constraint the limits the employer's optimization process. Workplace amenities are no different from wages in this respect. We have already seen that the profit-maximizing employer's offer (e^N, w^N) will be Pareto-inferior to some other combination of e and w characterized by small increases in effort and wages (the points in the yellow shaded Pareto improving lens). The same reasoning applies to working conditions: a small improvement in workplace amenities accompanied by a small increase in effort would be Pareto-improving.

REMINDER: SHORT AND LONG SIDE OF A MARKET THAT DOES NOT CLEAR You learned in Chapter 9 that the *short side* of a market is the side on which at a given price the number of desired transactions is least. The *long side* is the side on which the number of desired transactions is greatest. The short side of a market may be buyers (as in the case of quality) or sellers (as in the case of financial institutions supplying credit, as we will see in Chapter 12.)

M-Note 11.6: The Nash equilibrium is not Pareto efficient

Mathematically, we can show that the Nash equilibrium is Pareto inefficient using what we know from the first-order conditions of the employer and the employee, both of which must be satisfied for the an outcome to be a Nash equilibrium. At the equilibrium, you can see in M-Notes 11.1 and 11.5 that the first-order conditions of the employer and the employee require:

$$\pi_w = 0 \tag{11.9}$$
$$v_e = 0 \tag{11.10}$$

But it is also the case that:

$$v_w > 0 \tag{11.11}$$
$$\pi_e > 0 \tag{11.12}$$

Equations 11.9 and 11.11 mean that the first derivative of the profit function with respect to the wage provided to the worker equals zero, but the marginal value to the worker of the wage being higher is positive. Equations 11.10 and 11.12 mean that the first derivative of the worker's value function with respect to their effort is zero, but the marginal profit to the employer for an increase in effort by the worker is positive.

Stated differently, this means that at the equilibrium values of w and e, the employer is indifferent to small variations in the wage but the worker strictly prefers a wage increase (Equation 11.9) and that the worker is indifferent to small variations in the effort level, but the the employers profits increase if the worker works harder (Equation 11.10).

This being the case, there exists some (sufficiently small) values $(\Delta e, \Delta w)$ where Δ means a "small change" such that:

- $v(e^N + \Delta e, w^N + \Delta w) > v(e^N, w^N)$, and
- $\pi(e^N + \Delta e, w^N + \Delta w, h^N) > \pi(e^N, w^N, h^N)$.

So a small increase in effort accompanied by a small increase in the wage would be a Pareto-improvement over the Nash equilibrium. The employer and employees, however, cannot realize this Pareto-improvement through bargaining because effort levels are not contractually enforceable.

The employer exercises short-side power over the employee

Recall two things from previous chapters.

- From Chapter 9 recall that when a market does not clear the short side of the market — it can be either supply or demand – is the side on which the desired number of transactions is least. So in the labor market where there is excess supply (unemployment) it is the demanders – the employers – who are on the short side.
- From Chapter 10 recall that principals on the short side of a market can exercise power over those on the long side with whom they transact by threatening to terminate the transaction if the agent does not act in ways consistent with the principal's interest.

In the Ford model, the employee works harder than she would in the absence of the threat to take away her enforcement rent. By exerting more effort than she otherwise would, she contributes to the profits of the firm.

POWER For B to have power over A, it is sufficient that, by imposing or threatening to impose sanctions on A, B is capable of affecting A's actions in ways that further B's interests, while A lacks this capacity with respect to B. Short-side power is the power exercised by principals when they are on the short side of a non clearing market.

To see this go back to Figure 11.3 and suppose the employer set a wage of 5, which is equal to the disutility the worker experiences if she shows up at work and does nothing. The employer also threatens to terminate her employment if he is not satisfied with her level of effort. How hard would she work?

She would not work at all because she would be as well of without a job than with one if the wage is 5. When, instead, he pays her w^N she does work and contributes to the employer's profits. We can conclude then that:

- by threatening to impose a sanction on her (deprive her of her rent)
- the employer induced the worker to do something she would not otherwise have done (work) that
- furthered the employer's interests (raised his profits), and that
- she lacked the capacity to advance her interests by threatening the employer.

The last point holds because were she to threaten to terminate her relationship with the employer, he could find an identical worker among the those looking for work, so this would inflict no cost on him. This confirms that the employer's relationship to the worker included the exercise of power, as defined.

11.6 Employment rents and the workers' fallback option

We have so far assumed that the worker's next best alternative to having her current job is to have utility zero. We now turn to what the worker's fallback option really is.

Employment rents

People without work can expect to receive assistance both from friends and family and from the government, we call this assistance the unemployment benefit. Those out of work also search for jobs, and most find employment after some period of time. We are interested in employment rents primarily because they explain the behavior of the employed (not the unemployed). But to do this we have to put ourselves in their shoes and think what would happen if they were terminated for insufficient effort.

Imagine that you have a job paying w^N in which you are working at effort level e^N and you think hypothetically about what you would experience were you to lose your job. How would your life be different? A great many things would change including that you might lose your friendships at work and perhaps be less respected by others; in the U.S. you would most likely lose your health insurance. Here we focus on two important changes:

- *Lost income*: You would no longer have your wage w^N; instead you would have instead some kind of unemployment benefit B, but this would be less than your wage.
- *More free time and reduced disutility of work effort*: You would no longer be working harder than you would like.

The difference between how much you value what you have as an employed worker and what you *would have* if you lost your job is your *rent*. It is the maximum amount you would be willing to pay not to lose your job.

SPELL OF UNEMPLOYMENT A spell of unemployment is a duration of time spent unemployed. Again, don't always trust English here: it is not magic or fantastical. A "spell" is simply a duration of time that we index with the value s.

The situation of the worker is depicted in Figure 11.5 which shows:

- the wage that the worker will receive if she remains employed
- the disutility of effort that she will expend on the job at that wage
- the unemployment benefit per week that she would receive if her job were terminated

We track over time both the employed worker and the hypothetical same worker were she to lose her job, considering the case in which her job is terminated at the end of week 2, and she experiences a spell of unemployment of 26 weeks, after which she is finds another job on terms identical to the job she lost.

To calculate her rent per week on the job we compare two terms:

- the *value of her job* which is w^N if this is paid weekly minus $\underline{u}$ the disutility of effort when working at the level e^N for a week and
- her fallback option, which is weekly the unemployment benefit B; also termed her *reservation wage*, as it is the least wage at which she would agree to take a job (meaning, to show up but not work).

The difference between the value of her job (net of the associated disutility of effort) and the reservation wage is her rent per week.

$$\text{Per week employment rent} \qquad = w^N - \underline{u} - B \tag{11.13}$$

If she anticipates being out of work for a spell of s weeks then her rent is

$$\text{Worker's rent} \qquad = s(w^N - \underline{u} - B) \tag{11.14}$$

To see why the prospect of losing your employment rent can provide strong motivation to act in accordance with your employers wishes we now look at the magnitude of these rents.

FACT CHECK The expected length of a spell of unemployment varies over the business cycle. In the U.S. just prior to the recession associated with the global financial crisis or 2008, for example, about one in 6 of the unemployed and been jobless for 27 weeks or more; two years later almost half of those without work had been unemployed that long.

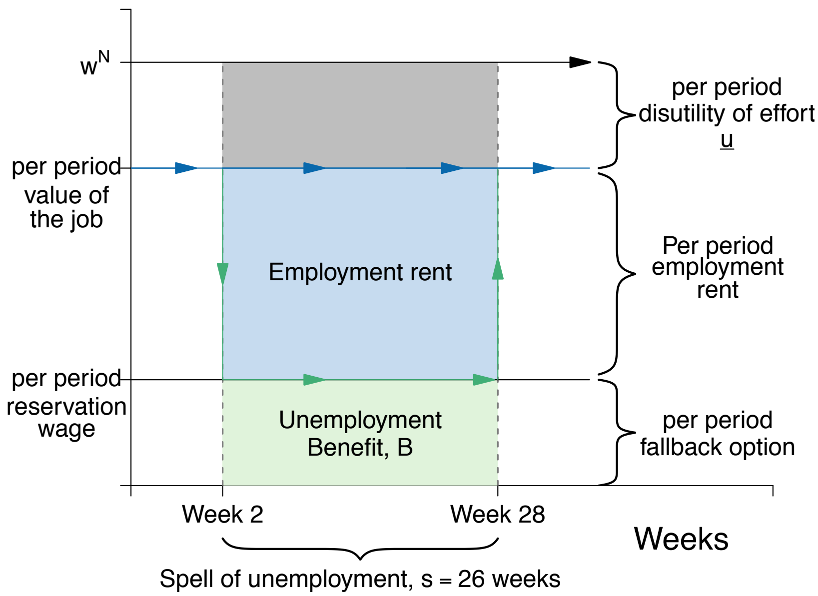


Figure 11.5: **The opportunity cost of working and the employment rent.** The blue arrows show the actual course of events of the person who has her job. The green arrows across are her hypothetical trajectory that she imagines would occur were she to lose her job. These two courses of events are identical before week 2 and after week 28, so the employment rent is the difference between the two over the period from week 2 to week 28.

Employment rents are substantial

Estimating the size of the employment rent is a challenge. To do this we cannot simply compare the economic situation of workers currently employed with the unemployed because the unemployed are different people who, if employed, on average would earn less than those currently with jobs.

An entire firm closing or a mass layoff of workers provides a natural experiment that can help. When a factory closes because the owners or managers have decided to relocate production to some other part of the world, for example, virtually all workers lose their jobs not just those who might be particularly "unemployment prone."

Louis Jacobson and his co-authors exploited a natural experiment to estimate the magnitude of employment rents. They studied experienced, full-time workers hit by mass job losses in the US state of Pennsylvania in 1982. In 2014 dollars, those displaced had been averaging about $55,000 in earnings in the year prior to losing their jobs. Workers who were fortunate enough to find another job less than three months after they lost their job took worse paying jobs, averaging only $35,000: a loss of $20,000 in the first year after the firings. So our hypothetical story about the worker imagining a spell of unemployment followed by a return to an equally good job does not reflect what these workers experienced.

Four years later they were still making $12,000 less than other workers who had been making the same initial wage, but whose firms did not have mass firings. In the five years that followed their layoff they lost the equivalent of an entire year's earnings.

Another challenge in measuring employment rents is that taking account of how people feel about losing a job and being unemployed, rather than just the monetary losses, the rents may be considerably larger. The reason is the social stigma, indignity and unhappiness resulting from being without work. One study using a data set allowing comparisons of the same individual when she is with and also without work, found that the subjective effect of joblessness was much larger than the loss of income. The amount of income that would compensate typical individuals for the social esteem and other costs of being out of work is actually greater than the income loss itself. So the true cost was more than twice the income loss.

REMINDER in Chapter 3 we looked at studies measuring subjective well-being or "happiness." The studies here use the same measures to show loss of happiness, depression and more from losing a job.

It is no surprise, then, that as the expected duration of a spell of unemployment increases reducing workers' fallback option and boosting job rents, workers work harder.

Evidence on job rents, unemployment, and work effort

Edward Lazear, chief economic adviser to former US President George W. Bush, investigated a single firm during the global financial crisis of 2008, to see how the workers reacted to the sharp increase in unemployment The firm specializes in technology-based services, such as insurance-claims processing, computer-based test grading and technical call centres, and operates in 12 US states.

The nature of the work made it possible for the management of the firm to track the productivity of workers, which is a measure of worker effort. It also allowed Lazear and his colleagues to use the firm's data from 2006-2010 to analyze the effect on worker productivity. Figure 11.6 shows the results. Lazear and his co authors found that productivity increased markedly as the duration of spells of unemployment rose during the financial crisis. This was particularly the case for branches of the firm located in places that experienced an especially large increase in unemployment

This raises a question: as the expected spell of a bout of unemployment rose after 2008, employers could have cut wages and still maintained sufficient employment rents to motivate workers to work as hard as before. Productivity would not have risen but labor costs would have fallen due to the decrease in wages. Why did they not cut wages?

Another economist, Truman Bewley wanted to know why employers typically do not cut wages even during recessions. He interviewed more than 300 business people, labour leaders, business consultants and careers advisers in the northeast of the US. He found that employers chose not to cut wages because they thought it would induce an angry push back from workers, who would then work less hard raising the cost of labor (meaning actual work

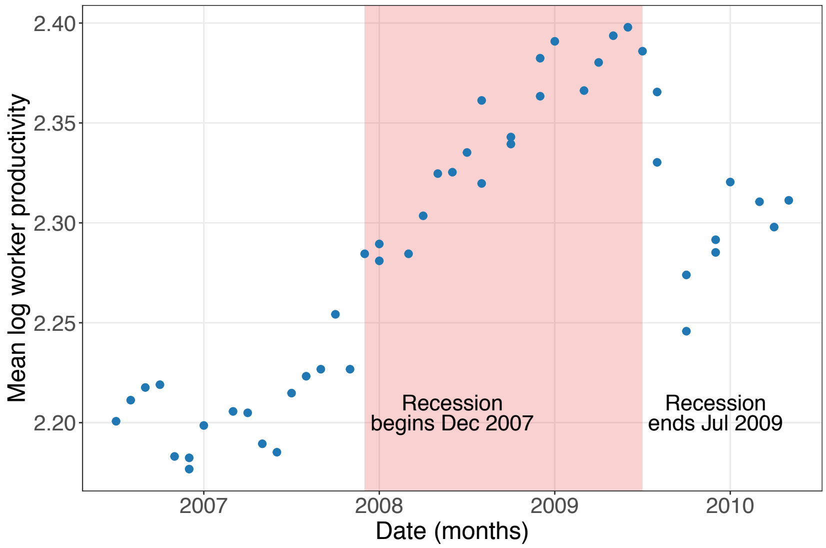


Figure 11.6: **Labor productivity before, during, and after the global financial crisis**. Monthly data are from a single firm in technology-based services during the period 2006-2010 where productivity can be measured with completion of computerized tasks. Productivity is measured as the log of mean worker output per hour. For workers with a job, productivity increases after the start of the recession and starts to decrease again after the recession ends. Source: Lazear, Shaw and Stanton (2016).

done). The employers thought it would ultimately cost the firms more than the money they would save from lower wages.

11.7 Connecting micro to macroeconomics: A no-shirking condition

The work speedup in response to the recession of 2007-9 makes it clear that macroeconomic factors alter the environment in which the firm selects a labor discipline strategy and workers respond. This occurs because the macroeconomy affects the fallback option of workers.

We now introduce a model in which the worker's fallback option depend on the level of employment in the economy as a whole, and the unemployment benefit. The microeconomic problem of labor discipline – from the employer's perspective, getting employees to work hard and well – provides the basis for a model that connects wage setting, work effort (and resulting productivity), and profits to the level of unemployment, unemployment insurance and other aspects of public policy.

Incentive compatibility in the "no-shirking game"

SHIRKING A dictionary definition of the verb "to shirk" is "to avoid a duty or responsibility." When an employee does not work as hard as the employer requires, economists call this "shirking." Workers sometimes required to work at an exhausting or even unsafe pace might call this an abuse of the language.

The set up is as follows. The employer, who as before is the principal and first mover sets a "no-shirking" level of effort $\underline{e}$, and announces that he will terminate without pay the worker if she is detected providing less. The employer then figures out what wage is necessary to motivate the worker to work at the no shirking level.

If there was a complete contract in effort, then the least "price" the employer could offer for $\underline{e}$ would be the unemployment benefit the worker would have received when not working (B) plus just enough more to compensate the worker for the disutility of providing the effort $\underline{u}(\underline{e})$ required by the employer. This is the workers willingness to work.

$$\text{Willingness to work: } = w \geq B + \underline{u}(\underline{e}) = \text{Participation constraint} \quad (11.15)$$

This is the least the employer could pay such that if the worker worked at the specified level of effort, she would be no worse of than if she had not accepted the job (not "participated.")

But because effort cannot be purchased under a complete contract and the employer cannot cost-effectively monitor all of the workers all of the time, the worker may be able to exert no effort at and nonetheless get paid. So the employer has to provide incentives for the worker to work. Here is the game (as described by the game tree in Figure 11.7).

REMINDER Earlier you studied incentive compatibility constraints expressed as a best response *function* which gave for each level of effort the least wage that would motivate the worker sufficiently to provide it. Here there is just one level of effort – $\underline{e}$ – so the ICC is an equation giving us just a single value, the least wage sufficient to motivate the worker to provide the no shirking effort level when it cannot be enforced by a complete contract.

- The employer announces a wage, a “no-shirking" level of effort $\underline{e}$, and a probability t that the worker will be terminated if she does not provide $\underline{e}$.
- The worker decides to provide $\underline{e}$ or not (in which case she does not work at all, that is $e = 0.$).
- If she provides $\underline{e}$ then with certainty she will not be terminated and will receive the wage w and experience the disutility $\underline{u}(\underline{e})$.
- If she provides less than $\underline{e}$, then one of two things may occur: with probability t, she is detected and terminated, or with probability $(1-t)$ she escapes detection and is paid w.
- if she is terminated, one of two things two things may occur: with probability j she does not find another job and receives the unemployment benefit B; or she finds another job with probability $1 - j$.

This gives us the four possible paths through the game tree resulting in the four outcomes at the ends of the branches. In the last case (she is terminated by immediately is re-employed), the job that she gets is identical to the job from which she was terminated (workers and employers are identical). To see what this payoff is, we need to know the wage offered by the employer.

The no-shirking wage

This will be the no-shirking wage, that is, the smallest wage the employer can pay that will motivate the worker to supply effort of $e = \underline{e}$. To determine the no-shirking wage, the employer considers two numbers:

NO-SHIRKING WAGE The no-shirking wage is the wage that is just sufficient for the worker to provide effort at the level specified by the employer rather than shirk.

- *The worker's payoff if she does not shirk*: the wage she gets minus the disutility of effort $w - \underline{u}$

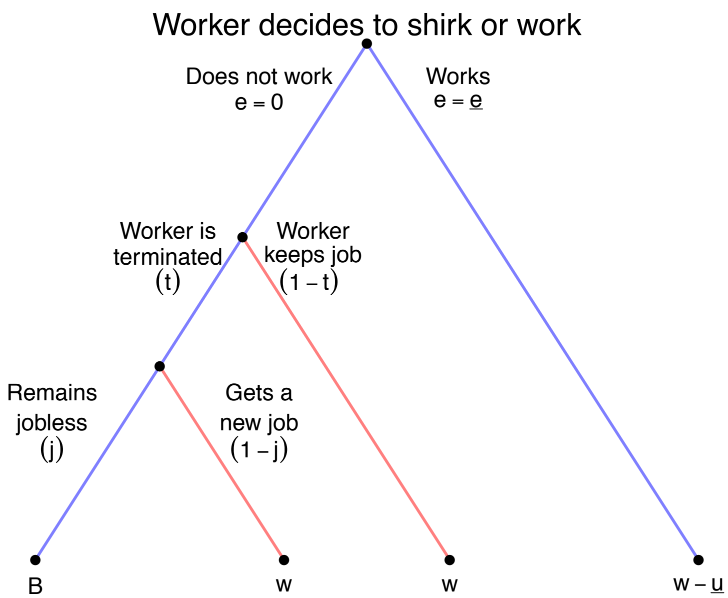


Figure 11.7: **The sequence of moves in a one-shot game determining the payoffs to shirking.** In response to the employer's statement of the no shirking level of effort $\underline{e}$, the termination probability t and the wage w, the decides whether to exert $e = 0$ or $e = \underline{e}$. If she exerts effort $e = \underline{e}$, she obtains the payoff for working and incurring disutility of effort, $w - \underline{u}$. If she exerts no effort, $e = 0$, then there is a probability t she will be terminated and a probability $1 - t$ she will not. If she is not terminated, she gets the wage w (and incurs no disutility of effort). If terminated, she will remain jobless (be unemployed) with the probability j and get unemployment benefit, B, or she will be re-employed with the probability $1 - j$ and receive the same value of the job that she would have experienced had she not been terminated, $w - \underline{u}$.

- *The worker's expected payoff if she does shirk*: this depends on the probability she gets fired (t), if fired, the likelihood she will remain jobless (j), and her unemployment benefit (B) if she and remains without work.

To motivate the worker to work, the employer therefore needs to offer a wage high enough so that the payoff from working is not less than the expected payoff from not working, that is.

$$\text{Payoff from working at } e = \underline{e} \quad \geq \quad \text{Expected payoff from shirking at } e = 0$$
$$\text{ICC: } w - \underline{u} \quad \geq \quad (1-t)w + t(1-j)(w-\underline{u}) + tjB \qquad (11.16)$$

Equation 11.16 is the *incentive compatibility constraint* ICC for the employer. As we have done in other models we assume that the worker will provide the no shirking level of effort as long as it is *as good as* shirking so Equation 11.16 will be satisfied as an equality (not an inequality).

This means that $w - \underline{u}$ is the value of the job to the worker whether she decides to work at the level required by the employer (the left-hand side of the equation) or not (the right hand side). To see this imagine that you are without work, and that you could bribe some employer to give you a job on the terms laid out by the employer in Figure 11.7. What is the maximum bribe you would pay? It is $w - \underline{u}$.

So if a terminated worker finds another job, this is its value, as is shown in Figure 11.7 at the end of the did-not-work, terminated, found-another-job branch of the tree.

The right-hand side of Equation 11.16 is from the left hand ("shirking") branches of the game tree in Figure 11.7 and comprises the following probabilities, all given the fact that she has shirked:

- $(1-t)w$: *Not terminated*: She is paid the wage (w), weighted by the probability that she will not get terminated $(1-t)$ even though she did not work, *plus*
- $t(1-j)(w-\underline{u})$: *Terminated, immediately finds another job*: She is both fired and also gets a new job the value of which is $(w-\underline{u})$; this occurs with probability $t(1-j)$.
- tjB: *Terminated, remains jobless with unemployment benefit*: This is the unemployment benefit she will get (B) multiplied by the probability that she gets terminated *and* remains unemployed

JOINT PROBABILITY The probability that both of these two events occur – she is terminated with probability t and she remains unemployed with probability j – is the product of the two probabilities, tj. This is called the joint probability of the two events.

M-CHECK If the fact that one of two possible events occurs does not influence whether the other occurs (the events are independent) then the probability that both will occur – called their joint probability – is $P-1P_2$ if the two events will occur with probability P_1 and P_2.

We can re-arrange Equation 11.16 and isolate all the w terms (shown in M-Note 11.7) to find the no-shirking wage:

$$\text{No-shirking wage} \qquad w^N = B+\underline{u}+\frac{1-t}{tj}\underline{u} \tag{11.17}$$

Equation 11.17 is a restatement of the incentive compatibility condition called the *no-shirking condition*. We use the N superscript, because given:

- the workers disutility of providing effort $\underline{u}$,
- her chance of being detected and fired if she does not work t,
- her chance of remaining without work if she is terminated, j
- and her unemployment benefit if she remains jobless, B

the employer setting the wage w^N and the worker providing effort $\underline{e}^N$ is a Nash equilibrium because:

- w^N is the least wage the employer can offer consistent with the worker working (it minimizes the cost of effort)
- given the wage offer the worker cannot do better than to provide effort $\underline{e}^N$.

We can rearrange the no shirking condition to isolate the worker's employment rent:

$$w^N-\underline{u}-B = \frac{1-t}{tj}\underline{u} = \frac{\underline{u}}{j}\left(\frac{(1-t)}{t}\right) = \text{Employment rent} \tag{11.18}$$

M-Note 11.7: Re-arranging the ICC to find the no-shirking condition

To find the no-shirking condition, we can rearrange the incentive compatibility constraint for the employer, equation 11.16:

$$\text{ICC: } w-\underline{u} \quad \geq \quad (1-t)w+t(1-j)(w-\underline{u})+tjB$$

The employer will pay the worker just enough for the worker to exert effort rather than shirk, and no more. So we will express the ICC as an equality:

$$w-\underline{u} \quad = \quad (1-t)w+t(1-j)(w-\underline{u})+tjB \tag{11.19}$$

Collect all terms with w on one side:

$$\begin{aligned} w-(1-t)w-t(1-j)w &= t(1-j)(-\underline{u})+tjB+\underline{u} \\ \text{Factor out } w\text{: } w(1-(1-t)-t(1-j)) &= t(1-j)(-\underline{u})+tjB+\underline{u} \\ \text{Simplify: } w(1-1+t-t+tj) &= -t\underline{u}(1-j)+tjB+\underline{u} \\ w(tj) &= -t\underline{u}+t\underline{u}j+tjB+\underline{u} \\ w(tj) &= tj(B+\underline{u})+\underline{u}(1-t) \end{aligned}$$

Solve for w:

$$w^N = \frac{tj(B+\underline{u})+\underline{u}(1-t)}{tj}$$

$$w^N = B+\underline{u}+\frac{1-t}{tj}\underline{u} \tag{11.20}$$

This is the no shirking condition.

Checkpoint 11.4: No-shirking wage

Consider Equation 15.8. What will happen to the *no-shirking wage* if the following changes occur?

- the probability of getting caught and terminated (t) decreases?
- the probability (j) that a worker will remain unemployed increases?
- the unemployment benefit (B) increases?
- the disutility of effort ($\underline{u}$) increases?

11.8 Incomplete contracts & the distribution of gains from exchange

We can now see how the extent of contractual completeness affects the distribution of rents between the employer and the worker.

REMINDER In Chapter 10 we did a similar analysis on the distribution of gains between the agent (the seller) and the principal (the buyer, Benetton in the example).

A measure of contractual completeness is the likelihood that a worker who does not work (sets $e = 0$) will be detected and terminated. To see how the completeness of the contract affects the extent of the worker's rent return to Equation 11.18. Suppose hypothetically that the contract were complete, so $t = 1$: the worker would receive no employment rent. The employer would pay a wage just sufficient to satisfy the worker's participation constraint, as shown in Equation 11.15. This is because if the contract is complete the worker is not renting her time but instead effectively selling her effort, so the employer can pay the worker her minimum willingness to sell her effort. In this case the participation constraint and the incentive compatibility constraint coincide.

To study the distribution of rents between the employer and the worker suppose that the per period output produced by the worker providing the no-shirking level of effort is γ which can be sold for a price of 1 Euro. We assume there is no other input than labor effort, so wages paid is the employer's only cost. Then the wage divides the average revenue from one period of "no shirking effort" into the worker's per period wage and the employer's per period profit.

Figure 11.8 shows how the revenue generated by the worker's effort is divided. We can see that the more incomplete is the contact the larger will be the share of the revenue that goes to the wage (composed of the worker's willingness to work plus the employment rent).

We can also see that there is some low level of contractual completeness $\underline{t}$ below which no mutually beneficial interaction between the worker and the employer is possible. The reason is that for $t < \underline{t}$ the no shirking wage would exceed the revenue produced by the workers' effort,that is $w^N > \gamma p$. As a result, there would be no way that the employer could pay a wage sufficient to get the worker to work and sell the resulting product at a profit. As a result, no workers would be hired.

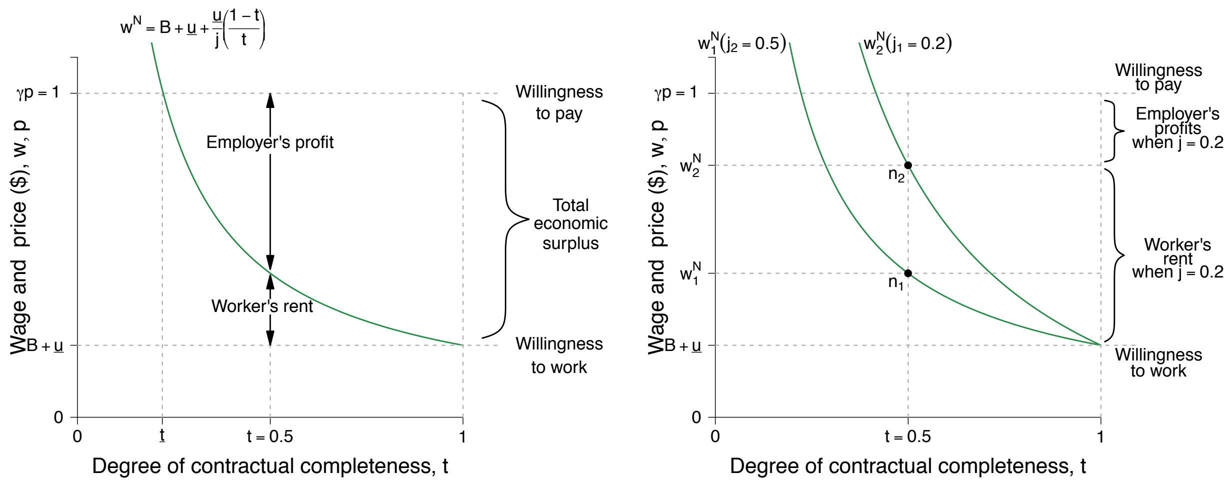
Comparing the two curves in panel b, we can also see the effect of labor market conditions on the distribution of rents to the employer and the worker. When the probability that the terminated worker remains without a job is 0.5 (as in panel a, for the lower curve in the panel b) the employer gets most of the rent.

When that probability falls to $j = 0.2$ the worker is less concerned about finding work if she is terminated, so the employer must pay her more to ensure that she works at the no shirking level. In this case most of the rents go to the worker, not the employer.

11.9 Application: Contract enforcement technologies

Charlie Chaplin's Great Depression film "Modern Times" ridiculed the efforts of “time and motion" men with stop watches and clip boards trying to speed up the pace of factory work. Today, it is getting easier for employers to monitor how much effort a worker is putting in, at least in some jobs. The "items-per-minute" of check-out staff can be checked in real time by supervisors just by looking at their computer monitor (17 items per minute is the minimum to keep your job at one outlet).

Software called a ‘keylogger’ can record the keystrokes of data entry personnel. In Chapter 12 you will read about the device installed in cars financed by



(a) The distribution of rents and profits for a given termination probability (b) An increase in the probability of finding a job increases worker rents

Figure 11.8: **Degree of contractual completeness and employment rents**. The worker's minimum willingness to sell (provide the no-shirking effort level) is $B + \underline{u}$. The employer's maximum willingness to pay is $p\gamma$. Remember from Equation 11.18 that the worker's rent is $\frac{\underline{u}}{j}\left(\frac{(1-t)}{t}\right)$. In panel a. the likelihood that the worker will remain jobless if they are terminated is $j = 0.5$, (an expected unemployment spell of 26 weeks in a 52-week year, so $1 - j = 0.5$). The figure shows for a given termination probability ($t = 0.5$) what the distribution of the surplus between the worker and employer will be given those values. Panel b. shows what happens if the probability of remaining jobless decreases to $j = 0.2$ (an expected unemployment spell of 11.2 weeks).

an auto loan: they can remotely disable the starter of the car if loan repayments fall behind. These devices can make a contract more complete either by covering more of that the employer wants (key strokes, items checked out) or by making the contract more enforceable (the remote ignition disabler). You know from Figures 10.13 and 11.8 that improving the contract in these ways will increase the share of the gains from exchange that will go to the principal – the buyer in the Benetton model or the employer in the Ford model. Correspondingly, as a result, a lesser share of the rents go to the supplier or the worker.

A similar device has been in use now for decades – trip recorders on trucks. Trip recorders vastly improved employers ability to monitor the actions of the truck drivers.The devices provided the company with verifiable information on the speed, idle time, and other details of the operation of the truck about which there was a conflict of interest between the driver and the company. For example, the cost of operating the trucks (paid by the company) was is higher the faster the trucks are driven. Drivers preferred to drive faster than the cost minimizing speed, and to take longer breaks. Drivers who owned their trucks — called owner-operators — were residual claimants on their revenues minus costs and so they internalized the costs of fuel and depreciation. driving at the cost minimizing speeds and realizing significant savings as a result. Based on their lower costs prior to the innovation of trip recorders, owner-operators successfully competed with company fleets whose drivers were employees not owners.

RESIDUAL CLAIMANT The residual claimant is whoever gets what is left over (the residual) from the revenue (or other benefit) of a project when all of the costs that have been contracted for are paid. If a land owner charges a tenant a fixed rent for the use of his land, the tenant is the residual claimant; if he hires a worker at a fixed wage to grow crops for him, then the land owner is the residual claimant.

When the trip recorders came in, companies were able to write contracts based on the speed at which the truck was driven, and to provide drivers other incentives to act in the companies' interests. By improving the companies' contractual opportunities the use of trip recorders induced drivers in trucks with recorders to drive slower and reduced the costs of operation of the company owned trucks. The result was a substantial decline in the market share of owner-operators.

The innovation of trip recorders therefore resulted in:

- **Less competition**: There was greater market concentration in the trucking industry, fewer firms and fewer owner-operators, and so less competition

- **Inequality** Many owner-operators could previously make a good living, constituting a kind of 'middle class' of the trucking industry with incomes greater than the drivers employed by the large companies, but much less than the incomes of the owners and mangers of those companies. By making possible a more complete contract between the companies and the drivers, the trip recorders contributed to inquality in the industry.

Unlike other on-board computers, (the electronic vehicle managements systems, or EVMSs), the trip recorders provided no improvement in coordination between truckers and dispatchers, as the information was available to the company only on the completion of the trip. The sole purpose of the trip recorders was to make the contract more complete with respect to drivers' behaviors in which there was a conflict of interest between the drivers and the companies.

Checkpoint 11.5: Non-compete clauses and work effort

Suppose the owner of a firm is considering imposing a non-compete contract on his workers and he asks you to explain how it might affect the wage he will have to offer to motivate his employees to work. Use the equation for the no shirking condition to provide a reply. Which variable (or variables) in the equation might the imposition of the non-compete clause affect?

Monitoring knowledge work

Employers now use a vast array of software to monitor the work activities of employees: ActivTrak, InterGuard, Veriato 360, Teramind, WorkSmart, Prodoscore, and Work Examiner are all software packages use to track the activities of workers on-site in company buildings or at home when they log in to company servers.

The software may monitor your use of email, messaging, browsing, and access to social media. Or, it may take a screenshot of your workstation and use your webcam to ensure you're at your workstation. The software will continue to take photographs every ten minutes to ensure you're on task, keep track of your keystrokes as you type, and monitor the movements of your mouse. With many white collar workers isolated a home during the COVID 19 pandemic

Veriato advertised that its devices could provide their employer with information on "What hours employees are working, how much they're working, what they're spending their time on"

Though none of these technology can provide verifiable information on all of the dimensions of effort, they provide data that employers can use to assess what kinds of worker activities correspond to worker productivity. In so doing, they make the employment contract more complete, re-allocating the rents from the worker to the employer.

Monitoring and trust: when monitoring backfires

Another challenging facet of this problem is that workers respond differently when they are monitored relative to when they are allowed to remain autonomous and make their own decisions about their productive tasks. For example, a survey of German citizens during the COVID pandemic in 2020 found that more were willing to comply with government advice about social distancing, getting a vaccination and installing a tracing phone app if compliance was voluntary than if it was enforced by law. The "enforcement-averse" Germans tended to be those who did not trust the government or believe its scientific reports.

Economists have researched the choice of monitoring methods, something we have not explored here (as we just assume that termination is lower if you work hard or provide high quality goods, without exploring how the employer acquires the relevant information). The research shows that when principals monitor agents closely and impose minimum quotas of work the principals tend to activate workers' negatively reciprocal feelings and sow distrust among their workers. As a result, close or intrusive monitoring may backfire. Workers value their autonomy in non-monetary ways, and so taking away that value, reduces the value of the job to the worker and as a result, the diminishes the force of the employers labor discipline strategy.

FACT CHECK Evidence reviewed in Chapter 16 suggests, furthermore, that workers especially value knowing that their employer *could have* monitored them or penalized them in some way, but then refrained from doing so. In this way, workers reciprocate trust with effort, and distrust with shirking.

11.10 Equilibrium unemployment and the wage curve

The no-shirking variant of the Ford model is connected to the macroeconomy because the likelihood of remaining unemployed if fired (j) depends on the level of unemployment in the economy-wide labor market. The more unemployed people there are looking for work, the less is the chance that any one worker finds a job (in any week or other given time period).

The economy-wide wage curve

To represent this in our model we can write the probability of remaining jobless, j, as a decreasing function of the fraction of the labor force that is employed or H. If labor supply is 1 and $H = 1$ this means that everyone seeking

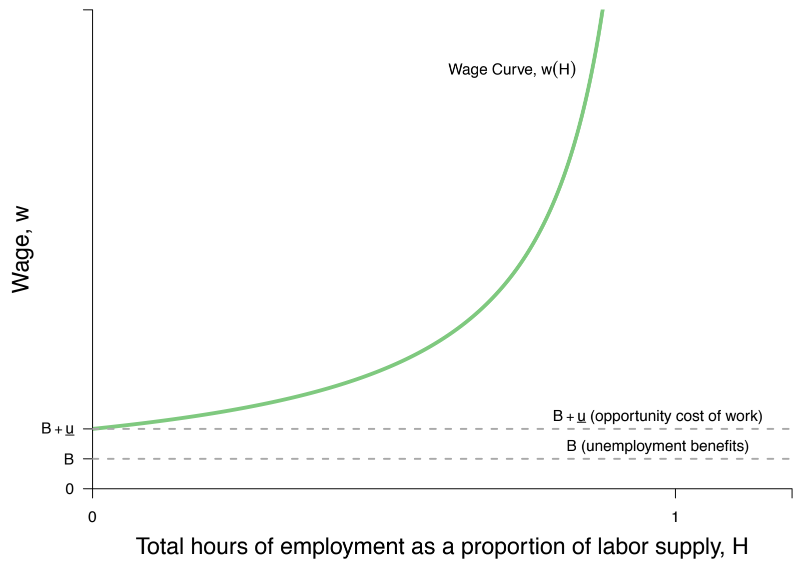


Figure 11.9: **The economy-wide wage curve**. The no-shirking wage $w^N(H)$ is the worker's willingness to sell $B+\underline{u}$ plus the workers employment rent. Notice that as unemployment falls (H rises) the rent required to motivate workers to provide effort increases. For a given fraction of employment, H_a, the corresponding wage will be $w_a^N(H^a)$ as shown by point **a**. Remember from Equation 11.18 that the worker's rent is: $\frac{\underline{u}}{j}\left(\frac{(1-t)}{t}\right)$ from which you can see that when $H=0$ (so $j=1$) we have the Nash equilibrium wage when $H=0$: $w^N(0)=B+\underline{u}+\frac{\underline{u}(1-t)}{t}$.

work has a job, and $H=0$ means that nobody is employed. Therefore, we write the probability of remaining jobless as $j=j(H)$ and we simplify by letting $j=(1-H)$ so if everyone looking for a job is employed $H=1$ then the terminated worker will immediately get a job with certainty ($j=0$). (In M-Note 11.11 we show how a relationship between j and H (j decreasing as H increases) can be derived using information of how frequently people quit their jobs and how employers find new employees.)

We developed the no shirking version of the Ford model by looking at the relationship between a single employer and a single worker. But that single worker is a member of a team of identical employees working for the employer. And we can extend the model further to represent the economy wide labor market.

To do this we write $j=j(H)=(1-H)$ and use the no-shirking wage equation to get:

$$\begin{aligned} w^N &= B+\underline{u}+\frac{1-t}{t\,j(H)}\underline{u} \\ w^N &= B+\underline{u}+\frac{1-t}{t(1-H)}\underline{u} \end{aligned} \tag{11.21}$$

We show the resulting economy-wide no shirking wage in Figure 11.9. This is called the *wage curve*.

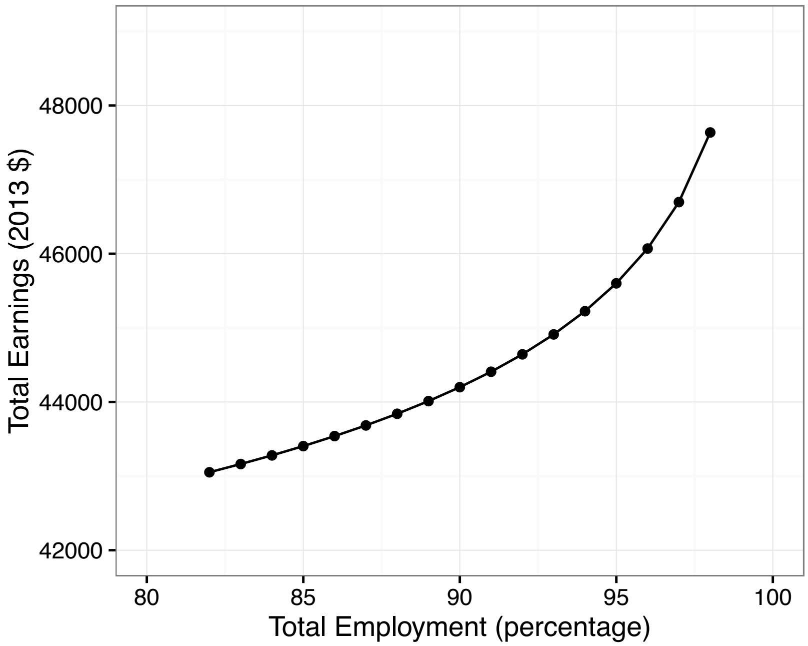


Figure 11.10: **An economy-wide wage curve estimated for the United States economy: 1979-2013**. The sample of 2.34 million workers are males aged 26-64 over the years 1979-2013. Earnings are in 2013 dollars.

The wage curve exists

By using data on unemployment rates and wages in local areas and over different periods of time, economists can estimate and plot the wage curve for an economy. Real wages tend to vary with the level of employment as the labor discipline model predicts. An example is shown in Figure 11.10, a wage curve estimated from data for the US.

As the labor discipline model predicts, Figure 11.10 shows that workers do better when more workers are employed – the wage curve is upward-sloping — and the effect of limited unemployment in pushing up wages is stronger, the less unemployment there is — the curve is steeper closer to $H = 1$.

Workers do better when unemployment is low not only because fewer workers experience unemployment, but also because those who are employed (the vast majority) are receiving higher wages. The level of employment in the economy as a whole therefore affects the distribution of income between workers and employers.

Equilibrium unemployment: There is no market-clearing wage

Equation 11.21 and Figure 11.9 give us the information we need to ask: can there be an equilibrium in which the labor market clears?

The answer is no. The reason is not any of the following:

- *"Sticky" or inflexible wages*, so that that workers' pay cannot adjust to excess supply of labor, that is, unemployment. Wages do adjust, which is what the wage curve shows.
- *Lack of competition* so that the unemployed cannot compete for jobs with the employed. They can, but as we saw their offers to work as hard as the current employers for less pay will not be accepted by employers.
- *Trade unions* that sustain wages making it unprofitable for firms to hire the unemployed. There are no trade unions in this model.

The reason why there will be unemployment in the equilibrium of this model is that a market clearing wage does not exist.

It appears from the figure that as H approaches 1 – market clearing, that is no unemployment – the wage curve becomes nearly vertical. We can confirm that this is the case by evaluating the no shirking wage when unemployment $(1-H)$ is zero, that is, $j=0$:

$$w^N = B + \underline{u} + \frac{1-t}{t \times 0}\underline{u} \tag{11.22}$$

Equation 11.22 shows us that if H = 1 we would have to evaluate the wage, $w^N(1)$, with division by zero. So the wage would have to be infinite to motivate workers to supply effort if everyone who want a job had one. This is impossible and so full employment is impossible in our model.

Equation 11.9 also allows us to understand why in the labor discipline model labor markets cannot clear in equilibrium, that is, supply of labor cannot equal demand for labor in a market with incomplete contracts.

To see why this is the case lets imagine that in an economy composed of many identical workers and employers, labor supply equaled labor demand, so the markets cleared and there is no unemployment The reasoning progresses as follows:

- *No unemployment*:*If* the labor market *clears*, supply of labor equals demand for labor and there would be no involuntary unemployment.
- *Immediate re-hiring*: Without unemployment, anyone who is terminated for insufficient work could immediately find a new job $(1-j)=1$ and $j=0$
- *Impossible infinite wage*: But, from Equation 11.17 or Equation 11.22 with no joblessness, $j=0$, and the lowest wage that would deter shirking would be $w^N = \infty$, that is, an infinite wage.
- *Firms shut down and create unemployment*: Firms cannot afford an infinite wage, so firms would lose money, shut down, put people out of work, and there would be involuntary unemployment (people who want a job at the market wage can't get a job).

- *Contradiction*: The existence of unemployment contradicts the initial premise that labor markets cleared.

The contradiction at the end of the chain of reasoning shows that labor markets cannot clear in the labor discipline model.

11.11 The whole-economy model: Profits, wages, and employment

If labor market clearing is not what determines the wage w^N and employment level H, what does? To answer the question we need to clarify some terms. In this chapter we use w or "the wage" to mean the real wage, or $\frac{w^n}{p}$, the nominal wage divided by the price level. When we say that employers "set real wage" we mean that they set the nominal wage for some given level of prices.

But to determine the level of the real wage and employment of the entire economy – not just a single firm – we use a model of product markets and goods prices as well as labor markets. The reason is that the real wage depends on the output and pricing decisions of firms owners, not just their decisions setting the money wage. The Nash equilibrium real wage and employment level must be the the result of best responses not only for workers and their owners as employers, but also for the owners as sellers of the products produced.

M-CHECK We want to determine two outcomes, the real wage w and the equilibrium level of employment H^N. But so far we have just a single equation, the wage curve. To determine the values of two variables we need two equations. The competition condition provides the second equation.

The wage curve gives us all of the possible Nash equilibria in the labor market. But the wage curve – by itself – cannot determine which of these combinations of w and H will occur.

To do this we one more piece of information. Think of this as another line in the wage curve figure, as shown in the margin as Figure 11.11 whose intersection with the wage curve will tell us which of the possible combinations of wages and hours of employment will be a Nash equilibrium.

To find this addition to the figure, we need an answer to the following question: given the level of labor productivity and the extent of competition in product markets, what is the wage that will result in the number of firms neither increasing nor decreasing? This is the wage such that total employment in the economy is *constant*.

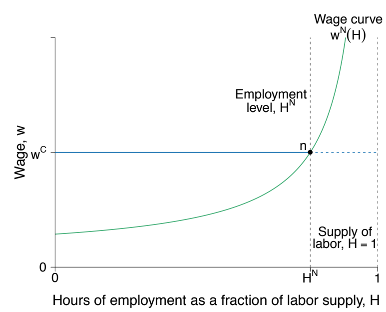


Figure 11.11: **Looking for a second equation in the wage curve figure**.

We will discover that there is only one such wage.

Barriers to entry and the competition condition

We use the model of competition in Chapter 9 to study a long-run equilibrium in which the number of firms is neither increasing nor decreasing. Because some firms re-locate or cease to exist for reasons unrelated to the model, new firms must enter for the numbers of firms to remain constant. Attempting to enter a market is a gamble: it requires investment that may not pay off if the entry attempt fails.

Whether the owners of a firm will attempt to enter depends on two things:

- the economic profit ($\hat{\pi}^E$ that the owners can expect to receive if they are successful, and
- the probability b that they will fail, not receiving that profit, but instead losing their investment.

If there are few firms in the market there will be little competition among firms and prices will be high relative to costs (which in this case are simply wages as we assume no other costs). So expected profits will be sufficient to motivate owners to attempt entry. If there are more firms and competition is greater prices will be lower so firms will not enter.

The competition condition tells us the unique relationship of prices to wages – that is the real wage rate w^c – that ensures the number of firms will neither increase nor increase. This wage provides us with the extra information that, along with the wage curve determines the total level of employment in the economy.

The competition condition requires that the expected (accounting) profit rate of an entering firm are equal to the opportunity cost of capital or $\hat{\pi}^A = \rho$, or what is the same thing, that the expected rate of economic profit $\hat{\pi}^E \equiv \hat{\pi}^A - \rho = 0$. Recall from Chapter 9 that p is the price at which the good sells, w^n is the money wage, b (for *b*arriers to entry) is the probability that a firm attempting to enter the market will not be able to sell its product and fail, and a_L is the amount of labor time (of workers working at the no-shirking level) required to produce a unit of output.

The owner pays the employees at the beginning of the period and sells the resulting product at the end of the period. Because we assume that there are no inputs other than labor, the wage bill is the level of investment the owner must devote to the firm. This is called "working capital" to distinguish it from the value of the capital goods used in production which here we abstract from. In M-Note 15.2 we show the real wage consistent with the competition is:

$$\frac{w^n}{p} = \frac{(1-b)\gamma}{(1+\rho)} = w^c = \text{the real wage that meets the competition condition} \tag{11.23}$$

The real wage will be higher:

- the more competitive the economy is, that is, the lesser are the barriers to entry, b
- the greater is the productivity of labor (γ) and

- the lower is the opportunity cost of capital (ρ).

Figure 11.12 shows the competition condition. If the wage is higher than w^c, then the expected (accounting) profit rate of a firm considering entering the industry will fall short of the opportunity cost of capital and so no firms will enter. Even if incumbent firms' profit rates allow them some economic profits on average, as in Chapter 8, some incumbent firms will fail for chance reasons unrelated to our model: a bad managerial decision, becoming a target of some other firm's predatory pricing, the death or departure of some critical personnel. With some firms leaving and none entering, the number of firms will fall, and output and employment will decline. If the wage is lower than w^c the opposite process occurs. Firms will enter, producing more and hiring more labor.

M-Note 11.8: Barriers to entry and the competition condition

The competition condition tells us how the the level of the wage consistent with the number of firms being constant depends on the degree of barriers to entry. Using the notation introduced in the text, we can write the expected accounting profit rate as the ratio of expected net revenues to the capital invested both expressed in per unit terms. So:

$$\text{The accounting rate of profit} = \frac{p(1-b) - w^n a_L}{w^n a_L}$$

And then equating this expression to the opportunity cost of capital ρ and rearranging the equation we get an expression for the real wage:

$$\frac{w^n}{p} = \frac{(1-b)}{(1+\rho)a_L}$$

Taking account of the fact that the inverse of the labor input requirement for a unit of output $a_L^{-1} = \gamma$, the productivity of labor, we have the following:

$$\frac{w^n}{p} = \frac{(1-b)\gamma}{(1+\rho)} \equiv w^c = \text{the real wage that meets the competition condition}$$

M-Note 11.9: A special case: "Unlimited competition"

If the economy were to be characterized by what Cournot called "unlimited competition," namely the complete absence of barriers to entry so that $b = 0$, then (continuing from M-Note 15.2) we have the following relationship between wages and the opportunity cost of capital, ρ:

$$\begin{aligned} w^c &= \frac{\gamma}{(1+\rho)} \\ \text{Raise to the power } -1 \quad \frac{\gamma}{w^c} &= 1+\rho \\ \text{Isolate } \rho: \frac{(\gamma - w^c)}{w^c} &= \rho \end{aligned} \tag{11.24}$$

The left-hand side Equation 11.24 is the accounting rate of profit, expressed in terms of one hour of labor, or hourly productivity minus the wage (that is profits per hour) divided by the capital invested per hour of labor (that is, the wage itself). So this is equivalent to:

$$\pi^E = \pi^A - \rho = 0$$

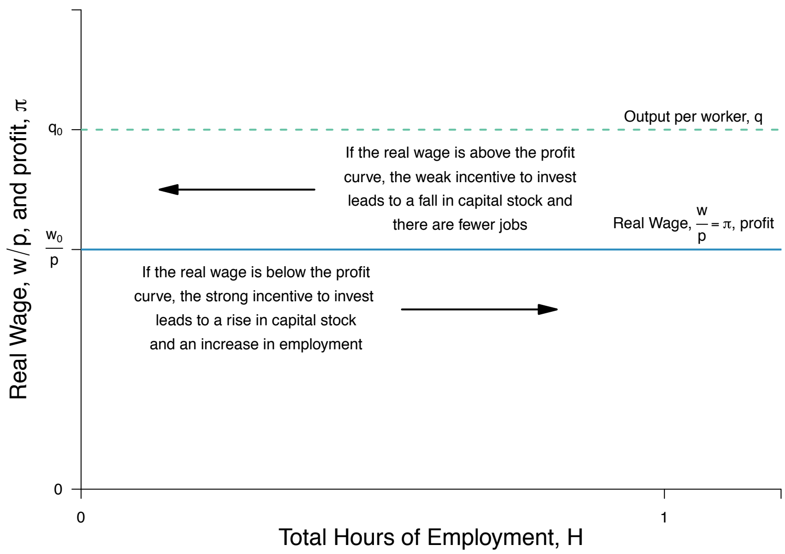


Figure 11.12: **The competition condition and the real wage.** Given the extent of barriers to entry, output per worker hour, and the opportunity cost of capital, the real wage indicated by the competition condition divides the output per worker between wages and profits in such a way that the number of firms does not change. It is called the competition condition because its level depends on the extent of competition which is greater the lower are the barriers to entry so that the number of firms competing is greater. Greater competition (lower barriers to entry, more firms) allows a higher wage, shifting up the blue line.

> This expression is the **competition condition** for an economy with unlimited competition so that economic profits are eliminated. It is also called the **zero profit condition**.

Profits, wages and employment

Putting the competition condition together with the wage curve in Figure 11.13 we have a model of the economy-wide labor market. But this is not a conventional model of demand and supply. It is true that:

- at wages above the competition condition employment contracts, below the competition condition demand employments expands and
- above the wage curve workers supply the effort required by their employer, below the wage curve they do not.

But be careful not to confuse the wage curve with a supply curve. The wage curve is not the “supply of labor." The wage curve is *not* the answer to the question: for each possible real wage, what is the amount of labor supplied? Instead, the wage curve is the answer to the question: at each level of employment, what is the lowest wage consistent with employees providing effort on the job, that is, working? The wage curve divides the space in the figure into two regions.

- *No production*: At wages below the wage curve no production can take place because workers are providing no effort.
- *Feasible production*: At wages above the wage curve workers are working,

so production is feasible (though firms may be leaving the industry if the wage is too high).

Similarly, the competition condition is *not* a demand curve. The competition condition does *not* provide the answer to the question: if the wage is w how many hours of labor will be hired? Instead, the competition condition divides the figure into two regions:

- At wages below the competition condition employment is *expanding* because firms are *entering* as long as the wage is on or above the wage curve.

- At wages above the competition condition employment is contracting because firms are *exiting*.

We now bring together the competition condition and the wage curve. These two curves depict outcomes that satisfy two of the biggest challenges for a modern capitalist economy:

- *Investment*: In the real economy production requires more than the working capital to hire workers. The wage must be such that owners of firms have the incentive to invest – constructing machinery and buildings necessary to employ people (represented by the competition condition): this is ensured by being on or below the competition condition.

- *Work*: The wage and level of employment must be such that workers have the incentive to work hard and well. This is ensured by being on or above the wage curve.

These requirements cannot be ensured by government order: firms cannot be ordered to invest, and workers cannot be ordered to work. Because thousands – even millions – of people – each independently pursuing their own objectives – must act in ways consistent with these objectives, the incentives have to be right.

In Figure 11.13 the wage curve and the competition condition divide the space shown into four regions. In only one of them – on or below the competition condition and on or above the wage curve – are the conditions for both investment and work met. We call this the feasible production region, outside the feasible production region there is *no production* (as shown in the figure). In the other three regions it is the case that either:

- Firms are dis-investing, that is leaving so in the long run none will be hiring or

- Workers are not working, so firms will not be employing anyone, or

- both.

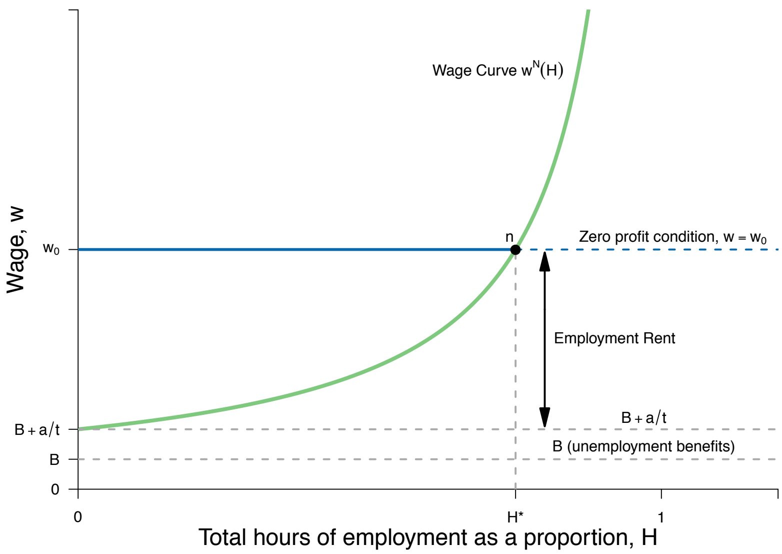


Figure 11.13: **Equilibrium in the product and labor markets of the whole economy.** Point **n** indicates the Nash equilibrium wage and employment levels, from which the level of profits and unemployment (that is $1 - H^N$) can be calculated.

Narrowing down our attention to the feasible production region, are any allocations within it Nash equilibria? Recall that a Nash equilibrium is a situation in which, given the strategies adopted by others, one cannot do better by changing one's own strategy. In the economy as a whole, this means that at the equilibrium level of employment:

- *owners of firms could not make more profit by changing the real wage*: they are paying the least wage consistent with workers working
- *workers could not do better by working harder or less hard*: they are working at the level required by the employer, and this is the best they can do
- *firms leaving the economy are exactly offset by firms entering; and no firm can benefit by expanding or shrinking its capital stock and hence their demand for labor.*

REMINDER Equilibrium sounds like a good thing, better than disequilibrium. But, this model can explain conditions under which in equilibrium, wages would be extraordinarily low, unemployment substantial, and inequality unacceptably high. Here and in general, equilibrium means nothing more than that there are no forces that will change the situation other than exogenous changes in the underlying structure of the model.

The first two points mean that the Nash equilibrium must be a point on the wage curve. The last point means that the Nash equilibrium must be on the competition condition. Therefore the Nash Equilibrium is the intersection of the two curves.

Figure 11.13 shows that the labor market does not clear: the total supply of labor does not equal the demand for labor. When the proportion of the labor market H^N is employed there remains a proportion of workers represented by the complementary proportion $1 - H^N$ who remain involuntarily unemployed. Those workers would like a job, but cannot get one.

We term this a long-run model because the process underlying the competition condition – the expansion or contraction in the size or number of firms – requires building or scrapping equipment and buildings: a process that takes months or even years to complete. Starting from the point **n** in the figure, shifts in aggregate demand in the economy associated with an export boom, or a collapse in investment, for example, will not shift the wage curve or the competition condition; but they can displace the economy away from the Nash equilibrium to higher or lower levels (respectively) of total employment than H^N. The unemployment that exists at the Nash equilibrium is termed structural unemployment – meaning unemployment that results from the fundamental structure of the economy including its technology and institutions as represented by the two equations of our model. On any given day or year the level of unemployment may exceed or fall short of structural unemployment as a result of the business cycle or anything else that results in the actual unemployment rate being different from the structural rate. The difference between structural and realized unemployment is called cyclical unemployment.

REMINDER You might recall from an introductory macroeconomics course that **aggregate demand** is thought of as the economy-wide demand for goods and services rather than just demand in one industry or market.

STRUCTURAL UNEMPLOYMENT Structural unemployment is the unemployment that results from the fundamental structure of the economy including its technology and institutions as represented by the two equations of our model. To learn more about cyclical and structural unemployment read Carlin, Wendy, and David Soskice. 2016. Macroeconomics: Institutions, instability and the financial system. Oxford: Oxford University Press.

Checkpoint 11.6: Total employment

Explain the effect in the long run on total employment of the following

a. A change in technology making it more difficult to detect a shirking worker
b. An increase in the unemployment benefit
c. An increase in labor productivity
d. The introduction of a new risk-free government bond with guaranteed payments of a substantial percentage of the value of the bond.

11.12 *Monopsony, the cost of inputs and the level of hiring*

In a monopsonistic labor market one or just a few firms hire a large fraction of the labor of some particular kind and the more they hire, the more they have to pay to retain or motivate workers to work. The firm is interested not simply in getting workers to show up (meeting their participation constraint): they want to provide incentives for the worker to work (meeting their incentive compatibility constraint). And because there are costs of finding and training new workers, they also want to set wages so that few workers will quit their jobs.

MONOPSONYA firm is termed a monopsony if it is the only buyer (or just one of a small number of buyers) in a particular market for some good or service. "Monopsony" comes from the Greek root "mono" ("single"), and "psony" ("buyer"). A firm is said to have monopsony power if it can influence the price at which it buys by purchasing more or less of the good or service.

Employers are **wage makers**. Firms set wages so as to minimize the cost of acquiring the labor effort that is an input into their production. In the case presented in panel b of Figure 11.3 this labor-cost minimizing wage is independent of how many hours of labor the firm hires, as can be see in Figure 11.4. But for monopsonistic employers there is another dimension of wage making: They know that their hiring decisions make the wage be what it is.

A monopsonistic labor market: wage making

The distinguishing characteristic of the monopsonistic employer is that, taking the employers eye view of the labor market, he can pay less if he hires less. This makes it similar to a monopoly, duopoly or other firm selling on a market with a limited number of competitors: selling less is a profitable strategy, in the case of sellers, because it allows sales at higher prices.

The difference is that in the case of monopsony it is in the market for *inputs* that competition is limited and the objective of restricting *output* is to limit buying, and hence *reduce* the prices of inputs. Of course a firm can be both a monopoly and a monopsony, with limited competition in both the market for its output and the market for its inputs.

Why does restricting the amount of hiring lower the labor costs of a firm? Here are two major reasons.

- *Company town or neighborhood labor market*: A large fast-food chain may find that it employs a significant proportion of the available low-wage workers in the neighborhoods where it operates, particularly if the lack of low cost public transport makes it difficult for workers who live in one neighborhood to work for firms located some distance away. In this case if a firm employs more in a neighborhood this may put noticeable upward pressure on local wages at the low-wage end of the labor market.

- *Balancing quits and new hires*: Owners and managers face the following problem: they would like to maintain a level of employment that is just enough to produce the profit-maximizing level of output. But in any given period, say, a month, some fraction of workers will quit or retire. These workers will have to be replaced by new hires from among those applying for work. The firm will set the wage to balance the number of workers leaving and new hires to sustain the desired level of employment. But because a larger firm will lose a greater number of employees in a month, in order to attract a sufficient number of replacements, a firm that seeks to maintain a higher level of employment will have to pay a higher wage.

For these two reasons, the lowest wage at which a monopsonistic employer can pay to motivate and retain workers is greater if he wishes to hire more, so the wage depends on the amount of hours hired or $w = w(h)$. The average cost of labor is the wage: the workers are identical and they all are paid the same wage. This means that the firms do not practice price discrimination, that in this case would be termed wage discrimination (studied in Chapter 9). So the wage that all workers receive is the average wage.

But for a monopsonist, the wage is higher if he hires more. This means that the the effect on total costs of hiring additional hours of workers time — marginal cost of labor = $\frac{\Delta w(h)h}{\Delta h}$ — is greater than the average cost of labor

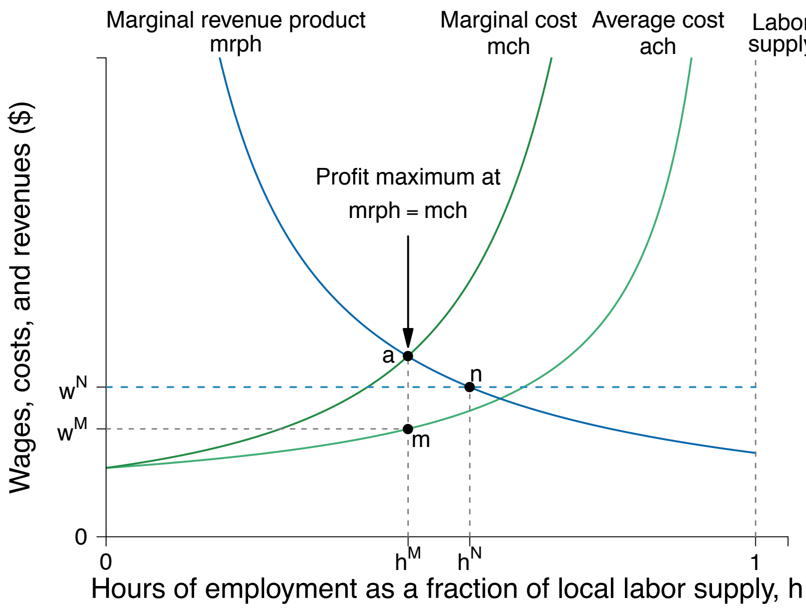


Figure 11.14: **The monopsonist's profit-maximizing level of hiring.** Employment as a fraction of labor supply, h, is plotted on the horizontal axis, and marginal revenue product of hours hired, the wage per hour, and marginal cost of hours of labor on the vertical axis. Recall that for the firm without monopsony power the wage is determined by the Solow condition and is w^N and is independent of the number of hours of workers time hired. So w^N is the average and marginal cost of workers time, and in that case the profit-maximizing level of employment (h^N) is given by the intersection of marginal revenue product (the downward-sloping blue line) and the horizontal green average and marginal cost line at line w^N. The rule for choosing the profit-maximizing level of hiring does not change: Profit is still maximized at the employment level where the marginal revenue product of labor equals the marginal cost. But because in the monopsony case the marginal cost exceeds the average cost of labor, this is now point **a**, not point *n*. And the profit maximizing amount of hours hired is now $h^M < h^N$ which means that the employer has to pay then pays the wage on their average cost of hours of work for that level of hiring, namely at point **m**, namely $w(h^M)$.

(the wage). This is because hiring an additional hour adds to the total labor costs in two ways:

- the wage paid to the particular worker for the additional hour *plus*
- the effect of this additional hiring on the wages of all of the *other* workers currently employed by the firm.

These two bullets appear in M-Note 11.10 as the two terms in Equation 11.25 labeled marginal cost.

We illustrate this monopsony case in Figure 11.14. To facilitate the contrast with the no-monopsony case we also show the outcome of where monopsony is absent studied in Figure 11.4. In this case the employer had used the Solow condition to set a wage w^N, and was hiring under conditions in which the least cost wage did not vary with its employment level. Equating w^N to the marginal revenue product of hours of labor at point **n** the employer hired h^N.

When monopsony is present, as in Figure 11.14, the employer will maximize profits by employing the number of hours such that the marginal revenue product of worker hours is equal to the marginal cost. Figure 11.14 differs from Figure 11.4 in that the marginal cost of labor is no longer the wage, set by the Solow condition: it is now higher than the wage because the average cost of labor function $w(h)$ is rising with more hiring. Figure 11.14 shows that the monopsonistic firm hires up to the point that marginal revenue product equals the marginal cost of labor, that is, up to point **a**, where the firm hires

h^M. From the figure you can see that the effects of monopsony are:

- *The firm hires fewer hours* : Relative to a labor market in which the firm is not a monopsonist, a monopsonistic employer will restrict employment, hiring $h^M < h^N$ as can be seen comparing points **a** and **n** in Figure 11.14.
- *Lower wages paid to workers*: Because the monopsonistic employer hires fewer hours, workers who *are* hired are paid less than they would be if the cost of motivating and retaining labor were independent of the level of hours hired and if, as a result, the employer had hired more. Remember, the employer's hiring decision was given by point **a** – where the marginal revenue product of hours equal the marginal cost of hours. But the wage paid is determined by the average cost of hours, that is point **m**. So $w^M(h^M)$, is less than w^N.

The model of the monopsonistic employer's hiring provides a lens with which to study a controversial topic: the effects of minimum wages.

M-Note 11.10: Labor hours hired by a monopsonistic employer

Here we study the number of hours of workers' time the employer will hire, assuming that the amount of effort per hour is given (equal to e^N as determined by the Solow condition). So the marginal product of an hour is just the marginal product of effort divided by a constant, namely e^N.

All of the information is the same as in M-Note 11.3, except that the total cost of hiring labor is no longer just wh because the wage rate, w, depends on how many hours of labor is hired. Labor costs therefore increase the more labor is hired. So, the profit function is now:

$$\pi = p(x(h))x(h) - w(h)h$$

To find the profit-maximizing level of hiring, as in 11.3, we differentiate the profit function with respect to the amount of hours hired, h, and set the result equal to zero.

$$\frac{\partial \pi}{\partial h} = \underbrace{\frac{\partial p}{\partial x}\frac{\partial x}{\partial h}x(h) + p(x(h))\frac{\partial x}{\partial h}}_{\text{Marginal revenue product of hours}} - \underbrace{\left(\frac{\partial w(h)}{\partial h}h + w(h)\right)}_{\text{Marginal cost of labor}} = 0 \tag{11.25}$$

The marginal revenue product of hours is unchanged, but the marginal cost of hours is now the average cost of hours – the wage itself – plus a term that captures the effect of hiring more on the wage. Because the marginal cost of hiring is higher than the average cost in the monopsonistic firm, but not in the case of the fixed wage, the profit-maximizing monopsonistic firm will hire fewer hours of labor.

Checkpoint 11.7: Monopsonistic hiring

- What is monopsony? How is it different from, similar to monopoly?
- Give two reasons why a firm may be monopsonistic in a labor market
- Give examples of forms of monopsony other than in labor markets (think of Walmart or Alibaba).

- Explain why the monopsonistic firm will hire fewer workers, or purchase fewer of other inputs, than an otherwise identical firm that is not a monopsonist.

11.13 Monopsony and the cost of hiring (non-shirking) labor

figures/Employment/nytimes_minwage.png

Figure 11.15: **A New York Times editorial, 1987**

In January 1987, a headline in the NY Times read, "The right minimum wage: $0.00." It reflected a consensus among economists at the time: they generally opposed minimum wages. The demonstration of how minimum wages were a misguided policy appeared in introductory textbooks using a standard intersecting supply and demand graph. The figure showed that in the absence of any government intervention the market would *clear*, with the quantity of labor supplied equal to the quantity of labor demanded and anyone not working was doing so "voluntarily" as they would rather have leisure than work at the going market wage.

With a minimum wage imposed at a level above the "market-clearing wage", the reasoning progressed, the supply of labor would exceed demand for labor and unemployment would result. We used a figure of exactly this type to teach what happens when prices are greater than market-clearing prices in Chapter 9. But, as we have shown in this chapter and in Chapter 10, labor markets and markets with incomplete contracts generally do not clear in equilibrium.

Since the early 1990s economists specializing in the labor market have studied the effect of increases in the minimum wage on the level of employment, primarily in the United States. One of the pioneering studies by David Card and Alan Krueger examined the increase in the minimum wage from $4.25 to $5.05 in the state of New Jersey by comparing firms on either side of the border of New Jersey and Pennsylvania (which did not have an increase in the minimum wage). They found that the increase in the wage by almost 20% for service sector workers did *not* increase the rate of unemployment nor did it decrease the number of hours hired.

FACT CHECK We have modeled monopsony – a market for inputs with a single dominant buyer – but the results are similar for an *oligopsony*, that is, a few wage making firms. Alan Krueger and Orley Ashenfelter found that many service-oriented U.S. firms, such as H & R Block, McDonalds, Burger King, Jiffy Lube, had "no-poaching" agreements that meant that the firms would not try to hire workers from other the firms. By limiting employees' alternative employment opportunities, this results in downward pressure on wages.

Their method of looking at hiring before and after the wage increase in establishments in neighboring states, cities, or counties where a minimum wage increase was either present or absent became the standard way of evaluating these policies (with underlying sophisticated statistical methods). To the surprise of many economists, the estimated job losses associated with a minimum wage increase have been either small or even nonexistent. In some cases, the data suggest even a small *positive* effect of the minimum wage on hiring.

As the evidence accumulated, economists began to reconsider their conclusions about the adverse employment effects of the minimum wage. In 2015 an elite panel of distinguished economists brought together by the Booth School

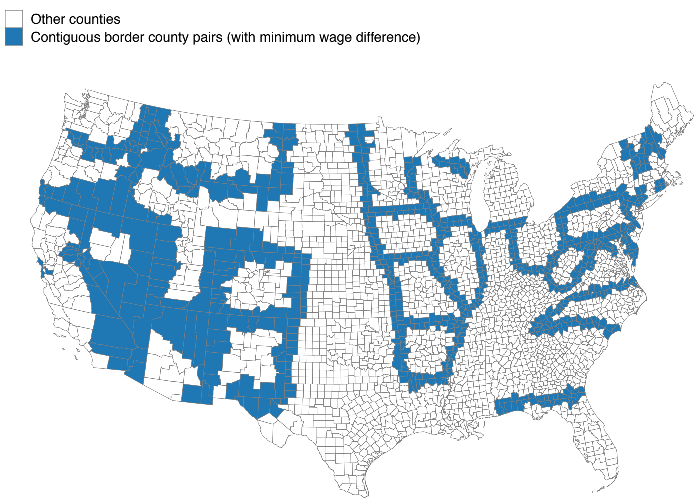


Figure 11.16: **A research design to determine the employment effects of the minimum wage.** This map of the United States shows (in blue) the paired counties that were compared in one study, in most cases those located on the boundary of two states with different levels and changes in the levels of the minimum wage

of Business at the University of Chicago were asked whether they thought that raising the minimum wage to $\$15$ (more than double the national minimum wage then in force) "would make it noticeably harder for low skilled workers to find employment." Only 26 percent thought that it would.

HISTORY Paul Samuelson, the first American to win the Nobel Prize in economics, wrote a path-breaking introductory economics textbook in the aftermath of the Great Depression and the Second World War. In his chapter introducing the model of supply and demand among competitive price taking buyers and sellers, he warns: "the demand for labor in the United States cannot be analyzed by the methods of this chapter."

While the size (and even the existence) of the adverse employment effects of the minimum wage continue to be debated, many economists are now skeptical about the adequacy of the simple supply and demand model when applied to the labor market. We would like to see if the models you have learned so far provide any clues about why in response to the minimum wage the substantial hiring cut backs expected by many economists have failed to materialize in the data. There are two.

- The *labor discipline model* shows that a wage increase will motivate employees to work harder, and this could partially offset the additional costs associated with a minimum wage increase.
- The *model of the monopsonistic firm's hiring* shows that the firm will restrict its hiring if employing more workers raises the wage. But, if the minimum wage is higher than the monopsonistic wage w^M over the range of hiring it might want to do then the minimum wage becomes the new marginal and average cost of hiring. This removes the monopsonistic motive to hire fewer workers, because hiring fewer workers would not allow the firm to pay a lower wage.

Monopsony, labor discipline, and the no-shirking wage

The worker's best response function (for example, in Panel b of Figure 11.3 shows that paying higher wages will increase worker the effort they provide;

FACT CHECK Decio Coviello and co-authors studied the effect of minimum wages on the earnings and individual productivity of employees in a large U.S. retailer that employs a substantial fraction of all department store employees. The (unnamed) firm operates in all 50 states and the researchers could study more than 10,000 workers in over 300 stores. The data allowed them to contrast wages and productivity in nearby stores which differed along two dimensions:

- one or more of the locations before and after a minimum wage increase; and
- locations in which the minimum wage increased or stayed the same.

The researchers found that an increase

and this could partially offset the costs of a wage increase imposed by law. The same model also provides a reason why many employers are monopsonists, that is to say why the cost of hiring labor would be less if the employer hires fewer hours.

The reason provided by the Ford model is similar to the "company town' logic above. The firm's own hiring

- reduces the pool of unemployed workers in the local labor market and this
- decreases the probability j that a fired worker will remain unemployed; which
- raises the fallback option of employed works (what they get if they are fired) which therefore
- raises the least wage consistent with the worker providing the no shirking effort level.

M-Note 11.11: A monopsonist's cost of (non-shirking) labor

To illustrate the idea of a monopsony with the labor discipline model mathematically, we consider the case in which there is just a single firm, which hires some fraction h of the relevant labor supply for this particular kind of employment in the location in question.

We write the probability of the terminated worker remaining unemployed as $j(h)$. So, modifying the no-shirking condition (Equation 11.17) to take account of the fact that the probability of remaining unemployed depends on total employment, we have the following no-shirking wage:

$$w^N = \frac{u}{t\,j(h)} + B \tag{11.26}$$

The no-shirking model didn't allow workers to quit their current jobs, so we now need to introduce the idea of worker quits.

Worker quits and the probability of remaining unemployed

Let the rate of current workers leaving the firm (quits) in any period be q, so the total number of quits will be qh. The firm needs to hire workers to fill the positions it has lost due to quits. The firm will therefore hire qh in every period to replace those who have left. Let us consider the case where a terminated worker may be rehired by his former firm perhaps in a different establishment: for example, getting fired at one restaurant branch may not stop you from getting a job at another branch of the same restaurant chain. We can then calculate the probability that a terminated worker finds a job. Take the firm's hiring qh each period and divide it by the number of unemployed $(1-h)$. This is the probability a worker finds a new job: $\frac{qh}{1-h}$. We also restrict $qh < 1-h$, which means that the firm's hiring (qh) must be less than the unemployment rate $(1-h)$.

The probability a terminated worker remains unemployed, is therefore:

$$j = 1 - \frac{qh}{1-h} \tag{11.27}$$

M-Note 11.12: Why restricting hiring reduces labor costs

Non-compete clauses and the probability of remaining unemployed

A non-compete clause means that the terminated worker can apply to only a fraction $\mu < 1$ of the job vacancies. As a result, the probability of finding work following termination falls to

$$\text{Probability of finding work} = \frac{\mu qh}{1-h}$$

which is less than without the clause, because $\mu < 1$. This is how non compete clauses increase j the probability that a terminated worker will not find work and remain unemployed.

To find the effect of the firm's hiring (a change in h) on the probability that an unemployed worker will remain unemployed (j) we differentiate Equation 11.27 with respect to the number of workers (h), giving:

$$\begin{aligned}
\frac{\partial j}{\partial h} &= \frac{(1-h)(-q)-(-qh)(-1)}{(1-h)^2} \\
&= \frac{-q+qh-qh}{(1-h)^2} \\
&= -\frac{q}{(1-h)^2}
\end{aligned} \tag{11.28}$$

Equation 11.28 is negative, which means that if the firm hires more (increases h), j will be less, meaning the likelihood of the worker remaining unemployed will decrease. If the firm hires fewer hours the likelihood that the terminated workers will be unable to find a job increases. This illustrates the firm's monopsonistic status in the labor market: the fewer hours it hires, the worse will be the employee's fallback position, and hence the lower will be the no-shirking wage. (The case where the terminated worker will not be re-hired in her former firm can be modeled in similar fashion. What matters is that Equation 11.28 is negative, which it will be in this alternative treatment.)

This is why the average cost of labor, which in this model is just the no-shirking wage, increases with the level of hiring. If the average cost of labor is rising, then the marginal cost of hiring a non-shirking worker must exceed the average cost. M-Note 11.11 shows how the economy-wide wage curve can be re-purposed to become the firm's average cost of labor curve in a local labor market. The result is that the individual monopsonistic employer faces an average cost of hiring labor curve that is a city-wide or other local version of the economy-wide wage curve (for example Figure 11.10, based on U.S. data) introduced earlier.

11.14 The effects of a minimum wage on hiring and labor earnings

EARNINGS This term — sometimes called "labor earnings" — refers to income from employment by a firm, government or some other employer, whether in the form of wages or salaries.

It remains only to show that a monopsonistic firm – that is, one with a rising cost of hiring (non-shirking) labor – may find it more profitable to *increase* hiring when a minimum wage is imposed, rather than *decreasing* hiring.

Checkpoint 11.8: The monopsonist's rising cost of (non-shirking) labor

Explain why the monopolists average cost of labor increases the more labor that he hires and suggest one or more facts about the relevant labor market that would make it increase a) more steeply and b) less steeply.

There are two possible effects of interest:

- Effects on the *total wages* paid to workers and
- Effects on the *level of hiring*.

Effects of the minimum wage on hours hired

With respect to the level of hiring, there are three cases to consider.

- *No effect*: the firm is already paying a wage higher than the minimum wage, in which case we say that the mandated wage is not binding.
- *Decreased hiring*: the minimum wage results in the firm hiring less labor
- *Increased hiring*: the minimum wage results in the firm hiring more labor.

The first case will apply to many firms but it is not really of interest because the minimum wage law has no effect. Figure 11.17 shows the firm's hiring decisions where the minimum wage is binding and it results in a reduction in hiring. The dark green curve – both dashed and solid portions – is the average cost of labor defined by the no-shirking condition. The marginal cost of an hour of labor is the light green dashed curve. Why are the average cost of hours and marginal cost of hours curves at least partially dashed? The answer is the purple horizontal line – the minimum wage. As long as the average cost of hours – the no-shirking wage – is less than the minimum wage, then the average cost of hours curve is no longer the least cost the employer can pay to employ (non-shirking labor). Paying $w^M(h)$ – the average cost of hours – is sufficient to motivate the worker to work, but it is not legal. The employer is required by law to pay more than it needs to to avert shirking.

This means that the average cost of hours curve is no longer relevant. This is why it is dashed. The solid minimum wage is now the average cost of hours curve from for levels of hiring less than h_g – the amount of hiring at which the average cost curve is equal to the minimum wage.

Because the average cost (meaning the minimum wage) is constant over this range $(0 - h_g)$ the marginal cost is equal to the average cost: hiring an additional hour increase total costs by the amount paid for the additional hour, w_d^{min}, and there is no effect on the wages paid to the other workers. This is the reason why the marginal cost curve is also dashed.

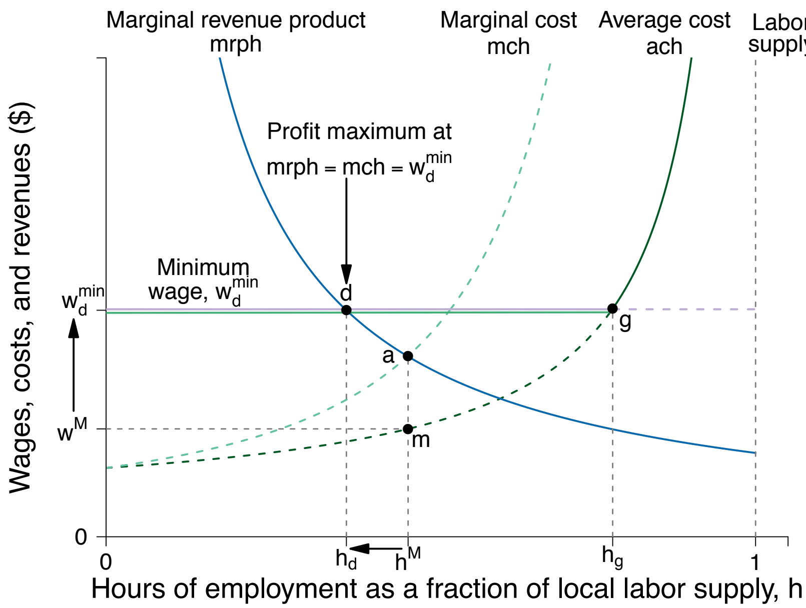


Figure 11.17: **Monopsony hiring level with a minimum wage, Case I: A large wage increase reduces hiring**. In this case, a minimum wage like w_d^{min} that is higher than the intersection at **a.** results in a *decrease* in employment for the monopsonist from h^M to h_d. For any employment level less than h_g the minimum wage is the lowest wage the firm can offer by law and it exceeds the least wage that it would otherwise pay given by the average cost of hours of hiring. So the average and marginal cost curves for hiring levels less than h_g are irrelevant: the average and marginal cost of an hour of a worker's time is the minimum wage.

For levels of hiring greater than h_g the wage required to deter shirking (the average cost of labor curve) exceeds the minimum wage. So to the right of point **g**, it is the minimum wage that is irrelevant, which is why that part of the line is dashed. Instead, we would have a solid increasing marginal cost curve for employment levels above h_g (not shown). See figure 11.18 for a case where the solid portion of the marginal cost curve is visible.

The effect of the minimum wage is, over some range of hiring, to break the link between firm hiring and the cost of sustaining labor discipline and hence the average cost of an hour. Over this range the employer's incentive to restricted hiring so as to maintain lower hourly labor costs disappears.

But the increased cost that the minimum wage imposes on the employer may lead to a reduction in employment as in Figure 11.17 or an increase as in Figure 11.18.

In the first case (Figure 11.17) the minimum wage results in a decreased level of hiring, h_d relative to the monopsony's preferred hiring at h^M.

In the second case, in Figure 11.18 the minimum wage results in an *increase* in employment relative to the monopsony's hiring. The difference between the two cases is that the increase in pay over the wage that the monopsonist would have paid in the absence of the minimum wage in Figure 11.17 is much larger than the increase s shown in Figure 11.18.

FACT CHECK The model in which monopsony in the labor market explains why imposing (or increasing) a minimum wage may have a positive (rather than negative) effect on firm hiring predicts that where hiring is dominated by just a few firms we would see positive employment effects, and where monopsony is more limited, we would find a negative effect. This is exactly what a study of the retail sector of the U.S. economy (think Walmart) found: where hiring was very concentrated, the minimum wage effect on employment was positive, and where concentration was limited the opposite was observed.

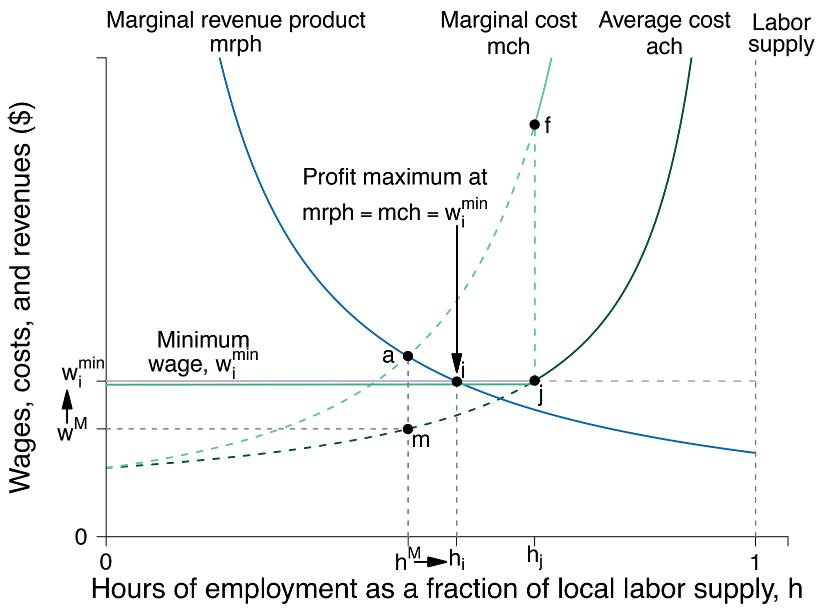


Figure 11.18: **Monopsony hiring level with a minimum wage, Case II: A smaller wage increase increases hiring**. The firm chooses the level of employment for a given level of effort per hour. The fraction of labor supply hired by the firm is h. The local labor supply is assumed constant and equal to 1. The fixed wage given by the minimum wage w_i^{min} results in marginal costs becoming vertical at point **j** where the minimum wage intersects the original average cost of labor curve. Marginal cost therefore intersects marginal revenue product at point **i** and determines the amount of labor hired at h_i. In this case, the minimum wage w_i^{min} is higher than the monopsony wage w_M and the amount of labor hired increases from h^M to h_i.

Effects of the minimum wage on labor earnings

We have shown that the imposition of a minimum wage (or an increase in the minimum wage) can result in either a decrease or an increase in the amount of hiring done by a monopsonistic firm. The other question about the effects of the minimum wage that we asked at the outset is whether taking account of both the change in hours and the change in the wage increases or decreases the total earnings of affected workers. There are two cases:

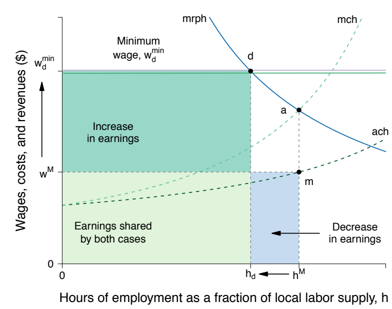


Figure 11.19: **Change in worker earnings from a minimum wage**. The minimum wage reduces hours hired, but those hours for which workers are hired earn more.

- If the effect is to increase the hours hired, then the total earnings of workers must rise (because in this case more hours are hired at a higher wage. This is shown in Figure 11.18
- If the effect is to decrease the hours hired then the effect on total earnings (hours times wages) could be either positive or negative. Figure 11.17 illustrates as case in which the gain in pay more than offsets the loss in hours, so total earnings increase.

The empirical evidence from the U.S. and the U.K, finds that increasing the minimum wage raises the incomes of low wage workers, and thereby reduces inequality. The reduction in inequality is one of the main reasons why many economists have come to have a more positive view of minimum wages as a labor market policy.

Effects of the minimum wage on labor effort and turnover

Firms' responses to an increase in the minimum wage go beyond a change (including a possible increase) in the level of hiring.

The imposition of (or increase in) a binding minimum wage will raise worker effort level for the following reason. Prior to the imposition of the minimum wage, the firm pays the no-shirking wage and workers are working at the firm's designated no-shirking level of effort. Now the firm is required to pay a higher wage, so it will pay more than is sufficient to motivate workers to work at the initial no-shirking level. We can illustrate this response by returning to the no-shirking condition:

$$\text{Minimum wage > no-shirking wage} = w^N = \frac{\underline{u}(\underline{e})}{tj} + B \qquad (11.29)$$

Here the term $\underline{u}(\underline{e})$ makes it explicit that the disutility of effort depends on the level of required no shirking effort.

Faced with paying a wage that is higher than the no shirking wage, the owners of the firm do not have the option to lower the wage (which would violate the law). So instead they can adjust upwards the minimum ("no-shirking") level of effort the worker must do to avoid being terminated if observed by the employer. Given that they must pay the minimum wage in any case, this gains the employer greater effort per hour hired, at no cost.

The minimum wage thus inverts the logic of the no shirking condition. Initially Equation 11.29 told the employer the least wage (w^N) it could pay and still motivate workers to provide effort at the level $\underline{e}$. With a minimum wage greater than w^N, the no-shirking condition, instead, tells the employer what is the greatest amount of effort and associated disutility $\underline{u}(\underline{e})$ that the firm could incentive compatibly require as a condition for not being terminated.

At the higher imposed minimum wage, workers will have the motivation to conform to this new higher no-shirking level of effort. So, the cost of paying the minimum wage will be partly offset by the effort increase by workers.

A second effect – also favorable from the standpoint of the firm's owners – is that the higher wage paid by the firm will reduce the number of employees leaving the firm. This will save the firm on the costs of recruiting and training replacements.

But while the positive effect on effort and the reduction in turnover are both favorable to the firms profit making objective, they cannot fully offset the cost increase imposed by the minimum wage. If that were the case, it would mean that the firm had been paying a less than profit-maximizing wage prior to the imposition of the minimum wage, and it would have already raised the wage just to gain higher profits.

FACT CHECK When minimum wage laws mandated increased pay for nursing assistants and other poorly paid nursing home staff in the in the U.S. the health of those in their care dramatically improved. This is the finding of a study of 7,700 nursing homes in 1,136 counties, comparing the health of residents in nursing homes that had experienced wage increases that had not been not experienced by nursing homes in neighboring counties. The author calculates that a ten percent increase in the minimum wage nationally would lead to approximately 15,000 fewer nursing home deaths annually.

FACT CHECK Arindrajit Dube and his co-authors found that for groups on which the minimum wage was expected to have a major impact (teens and restaurants) an increase in the minimum wage

- eliminated a substantial number of below minimum wage jobs;
- increased the number of somewhat above minimum wage jobs by about the same number, and
- very slightly increased jobs paid more than the minimum wage.

They also found that the minimum wage increase resulted in substantial reductions in quits as fewer people left their jobs, and a resulting reduction in new hires.

A broader picture: The minimum wage in context

To understand the effect of the minimum wage in an entire local labor market or in the whole economy we would need to take account of additional likely effects.

- *Firm profits*: The unavoidable negative effects on the firm's profitability would in the long run be expected to lead to the exit of firms which would both reduce hiring and could increase the degree of monopsony in local labor markets. You can see from Figure 11.13 that starting at point **n** raising the wage above w^N would be expected to reduce employment in the long run.
- *Changing demand*: The redistribution of income from relatively high income profit recipients, who save a substantial portion of their income to low wage workers who spend most of what they earn, could increase aggregate demand, leading firms to hire more workers, at least in the short run.

EXAMPLE Considering a substantial increase in the minimum wage in the U.K., Sajid Javid, the Conservative Party Chancellor of the Exchequer (Minister of Finance or Secretary of the Treasury in other countries) commented: "Monopsony in the labour market is much more common than was previously considered the case. As such a minimum wage is - within reason - not a distortion to the market but a correction to it. Minimum wages . . . promote productivity, and a more efficient allocation of labour."

While the positive effects on aggregate demand might be substantial for a minimum wage increase in an entire economy they are unlikely to result in any significant increase in jobs resulting from a minimum wage increase enacted at the level of a city, state or province. The reason is that the increased earnings of those benefiting from the minimum wage would be for the most part be spent on goods and services produced by workers in the rest of the world, not locally.

We return to a consideration of policies that can increase both wages and employment, and thus reduce inequality, in Chapter 15.

Eliminating policies that restrict labor market competition

Minimum wages are not the only policy that can effectively raise the earnings of low wage workers. An alternative to imposing a minimum wage is to make labor markets more competitive so as to limit the monopsonists' ability to hold down wages by restricting output. Figure 11.14 shows the comparison of the competitive labor market in which employers cannot depress wages by hiring less, with the wage w^N and the case of monopsony with the wage $w^M < w^N$.

FACT CHECK A non compete clause in the contract of a \$ 13 per hour packer at Amazon in 2015: "During employment and for 18 months after the Separation Date, Employee will not, directly or indirectly, ... engage in ... manufacture, marketing, or sale of any product or service that competes or is intended to compete with any product or service sold, offered, or otherwise provided by Amazon ... ". A "no-poach" clause in the contracts of McDoalds franchise holders commits them not to hire anyone who has worked for another McDonalds outlet within the past 6 months.

In Chapter 9 we described some of the rent seeking strategies that owners of firms adopt to reduce product market competition and as a result increase profits. Similar competition-limiting strategies are adopted in labor markets, seeking to enhance profits by allowing employers restricting employment so as to reduce wages.

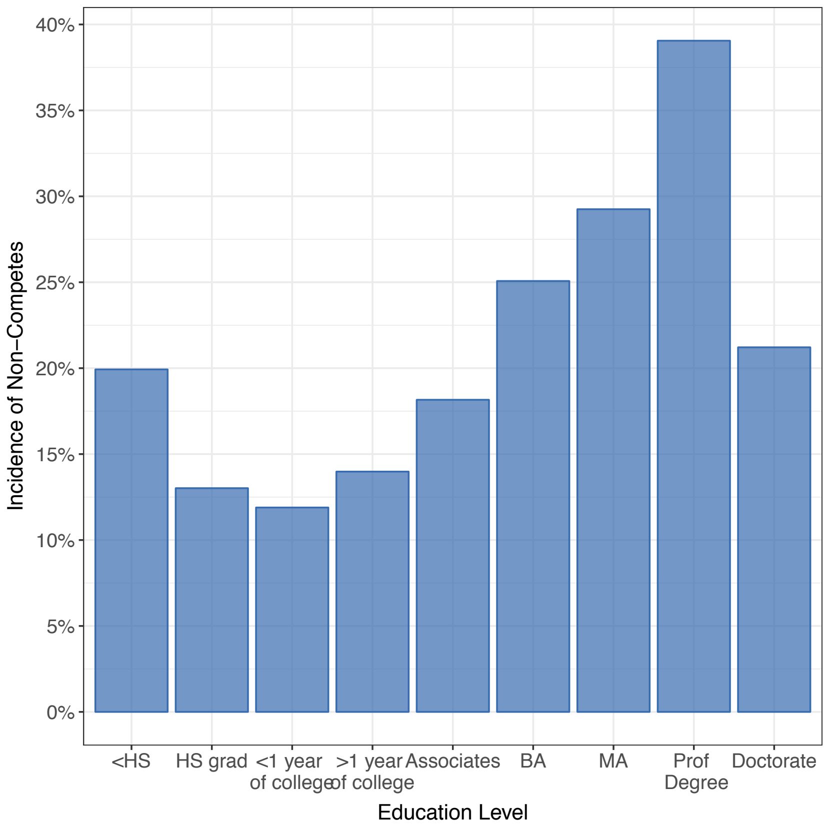


Figure 11.20: **Fraction of workers whose contacts have non-compete clauses.** About 1 in 5 workers in the U.S. (in 2014) are subject to non-compete clauses. Initially justified as a means of protecting trade secrets and technical information, the clauses are now widely applied for warehouse work, fast food establishments and lines of work in which the intent appears to be limiting the alternative employment opportunities of the company's current workforce.

M-Note 11.13: How a non-compete clause reduces the no shirking wage

A non-compete clause means that the terminated worker can apply to only a fraction $\mu < 1$ of the job vacancies. As a result, the probability of finding work following termination falls to:

$$\text{Probability of finding work} \quad = \quad \frac{\mu q h}{1-h}$$

which is less than without the clause, because $\mu \quad < \quad 1$. This is how non compete clauses increase j the probability that a terminated worker will not find work and remain unemployed.

Chief among these are *non-compete clauses* that prevent a either a current worker or a terminated worker from going to work for a employer competing with her current employer. These clauses prohibit, for example, a worker who was fired from or quit a job with McDonalds from taking any job in the fast food business, substantially limiting her chances of finding work. We show the effect in M-Note **??**. Prohibiting non-complete clauses in contracts would make labor markets more competitive and by improving workers' fallback options put upward pressure on wages.

11.15 Conclusion

Developments in both theory and empirical studies have given economists reason to think about not just the economics of labor markets, but also the sociology and politics.

The stimulus for much recent theoretical research on labor markets came from dissatisfaction with the micro-economic aspects of macroeconomic models of aggregate employment and unemployment. Macro-economists were prominent among the early innovators. Models based on incomplete contracting for effort or other aspects of the labor exchange explained how a competitive equilibrium could exhibit involuntary unemployment, thereby narrowing the gap between theory and empirical observation.

We do not know what Henry Ford had in mind when he announced the $5 day. The fact that output per worker hour more than doubled following the increase suggests that workers' effort rose substantially. (Ford increased the level of supervision along with the wage, so the likelihood that slack work would be tolerated undoubtedly fell). Whether the workers' increased effort was a response to the carrot of Ford's seeming generosity or to the stick of closer supervision and increased employment rents, we cannot say.

Making connections

Contracts between workers and employers, whether negotiated individually or collectively (by a labor union) cover important aspects of the interaction – hours, wages and working conditions – but not others – how hard the person works, and many aspects of the quality of the work. Complete and incomplete contracts are contrasting institutions or rules of the game.

Price making and wage making. We have studied two categories of price making: wage setting by monopolists facing limited competition in the labor market, and the choice of a labor cost minimizing wage as part of the employer's labor discipline strategy. In Chapter 15 we will introduce wage setting through trade union bargaining.

Institutions, fairness and Pareto efficiency The Nash equilibrium under incomplete contracts is Pareto-inefficient, but the worker is better off than she would be with complete contracts.

Incentive compatibility, participation constraints, and constrained optimization Because the labor contract is incomplete, the employer is constrained by the incentive compatibility constraint governing how hard the worker works, and in order to maximize profits pays the worker more than the minimum to just satisfy her participation constraint.

Conflicts over mutual gains made possible by exchange The labor market is characterized by conflict over the mutual gains from employment. This conflict is not resolved in a contract, but is determined "on the ground" in the strategic interactions between the employer and the worker. A more complete labor contract favors the employer in this conflict.

Limited competition and inequality. As in Chapter 9 we found that limited competition among firms – in this case in the market for workers time, not for the output of the firm – raises the profits of owners.

Non-clearing markets, rents and power Employers are on the short side of a non clearing market; employees and the unemployed are on the long side. The employer's threat to terminate the rents that employed workers receive are the basis the employer's power and labor discipline strategy.

The whole economy Integrating the labor discipline model with the Cournot model of competition among firms in product markets provides the basis for analysing wages, the markup, profits and unemployment in the whole economy and to understand why the labor market does not clear and why limited competition on product and labor markets will result in lower wages and greater unemployment.

Evidence Data on job rents, how wages and work effort vary with the level of employment, and employers monitoring of employees work effort are consistent with the predictions of the model.

Important ideas

(in)complete contract	disutility (of effort)	employment rent
worker/employee	employer	labor discipline model
self-employment	Solow Condition	contingent renewal contract
unemployment benefit	reservation wage	fallback
Pareto (in)efficiency	Nash equilibrium	short-side power
long-side of the market	no-shirking condition	excess labor supply
non-market clearing	workplace amenities	wage curve
profit curve	aggregate labor market	opportunity cost of capital
zero profit condition	labor supply	labor demand

Mathematical notation

Notation	Interpretation
u	utility function of the worker
$\underline{u}$	base disutility of work
y	income received by the worker
e	worker's provided effort
e^N	Nash equilibrium effort
$\underline{e}$	no-shirking level of effort
p	price that the worker receives from a good she produces
w	wage that the worker receives
$\bar{w}$	worker's reservation wage
w^N	Nash equilibrium wage
w_M	monopsony wage
B	unemployment benefit
c	employer's cost per unit of effort
c^N	Nash equilibrium cost per unit of effort
t	termination probability for the worker
t^N	Nash equilibrium termination probability
z	worker's fallback position
s	length of unemployment spell
v	value of job
v^N	Nash equilibrium value of job
H	employment rate
H_M	monopsony employment rate
h	hours of labor hired
h^N	Nash equilibrium hours of labor hired
l	total labor $(h \cdot e)$ hired by an employer
j	probability of remaining unemployed/jobless
U	unemployment rate
q	quit rate
π	employer's profit function

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

See reference list.

12

Interest, Credit & Wealth Constraints

... managers ... of other people's money than their own ... cannot well be expected that they should watch over it with ... vigilance...

Adam Smith, *Wealth of Nations*, 1776

12.1 Introduction

Her daughter's asthma was acting up and the ten year old had a fever of 103.5 degrees Fahrenheit (39.7 Celsius). Mary Bolender of Las Vegas (USA) knew she had to get her to the emergency room fast.

But that didn't happen.

Her car would not start. There was nothing wrong with the 2005 Chrysler van. Nothing except that the starter had been remotely disabled by the lender from whom she had borrowed the funds to buy the car. Her monthly check to the lender was three days late. The lender, C.A.G. Acceptance of Mesa Arizona had agreed to lend her the funds only after installing a device that can be remotely activated to make it impossible to start the car.

In 2014, about 2 million vehicles in the U.S. had these devices installed. They do more than make the car unusable if you get behind in your loan repayment: some devices send out loud beeps that become more aggressive as the due date for the check to the lender nears. The devices also can track the location of the vehicle: some include a "geo-fence" that informs the lender if the borrow is no longer driving to and from their job.

The tracking capabilities allow the lender to quickly repossess the car if the borrower does not pay up. For the lender this is a big improvement over the old days in which they hired 'repo men' (polite version: repossession agents) who would cruise neighborhoods looking for the vehicles of borrowers in default, and then tow them away.

DOING ECONOMICS

This chapter will enable you to:

- Explain how legal institutions – the rules of the game affecting bankruptcy and limited liability – and the nature of the information available to lenders result in incomplete credit contracts.
- Understand why the Nash equilibrium of a borrower-lender interaction is Pareto-inefficient – even with unlimited competition.
- Show that people without wealth may be excluded from the credit market, while wealthier borrowers can finance larger projects of lesser quality.
- Provide empirical examples of the credit market disadvantages of those with limited wealth.
- Explain how barriers to entry limit competition in the credit market, reduce the total rents available to lenders and borrowers and redistribute the rents in favor of the lenders.
- Indicate the similarities and differences between the lender-borrower relationship and other principal agent relationships you have studied, namely Benetton and the subcontractor, and the Ford Motor Company and the worker.
- Show how the credit market provides a way of understanding both the Keynesian multiplier and the way that that monetary policy can affect aggregate demand, – particularly expenditure on housing and consumer durables.

Figure 12.1: Interest rates run in the many hundreds of percent annually.

To get her van back on the road, Bolender had to pay 389 dollars, money that she didn't have that day when her daughter needed treatment. (As far as we know, the child is fine.) "I felt absolutely helpless" she said. John Pena the general manager of the company that installed the device in her van sees it differently. Without the new technology, he said "we would be unable to extend loans. . ."

Of course borrowers do not lack ingenuity, and ways of disabling the disabler have circulated on the internet. Oscar Fabela was ready when he and his date left a film only to find that his car would not start and the device was screeching like a burglar alarm. The two of them got the car going in no time but the date was not impressed. "It didn't end well," he said. In response to borrower push-back, one device maker has produced a fake disabler (named the Decoy) that is installed visibly along with the hidden real device so that people will think they have turned the device off when it is actually still installed.

CREDIT MARKET A *credit market* is a market in which what is exchanged is a contract for a loan or credit. The seller of the contract for the loan is the *lender* and the buyer of the contract for the loan is the *borrower*.

Bolender's experience indicates the possible misuse of the devices: on the day she needed to go to the hospital, the lender could not legally have repossessed her van, but they were able to disable it. But Pena, the lender, is right in stressing that without the devices many more people would be rejected when they need a loan, or would be charged extraordinary interest rates.

Borrowing to buy a car is a lot easier than borrowing to buy a meal or pay the rent. The reason? When you buy a car on borrowed money, the loan contract generally gives the lender the right to take the car if you do not repay on the contracted schedule. The car in this case is what is called **collateral**, that is a transfer of ownership of some thing of value to the lender should the loan contract be violated by the borrower. Lending to someone to pay for food or to pay the rent is much riskier for the lender because there is typically no collateral. These are what are termed unsecured loans.

Where lenders find it difficult to ensure repayment or to seize the collateral, they typically do not lend money. But if they do, then the probability of not getting the money back is built into the interest charged. In New York City people in need of cash take out short-term loans to be repaid when their next pay cheque comes in. These payday loans bear interest rates between 350 per cent and 650 per cent per year. The legal maximum interest rate in New York is 25 percent.

COLLATERAL When a borrower signs a contract to borrow funds, they often agree to put up collateral that can be claimed by the lender in case the borrower is unable to re-pay the loan. Typically, the collateral is capital that the borrower already owns, such as retirement funds or home equity.

We know a lot about these practices in New York because in 2014 the "payday syndicate" offering these loans was charged with criminal **usury** in the first degree and the prosecution described their illegal dealings in some detail. In Illinois, the typical short-term borrower is a low-income woman in her mid thirties ($24,104 annual income), living in rental housing, borrowing between $100 and $200 and paying an average annual rate of interest of 486 percent.

USURY When a lender charges unreasonably, unethically, or illegally high interest rates, then they are said to be committing usury. Usury, or even charging any interest at all on a loan, is condemned in the scriptures of both Islam and Christianity. . In the United States, individual states have their own usury laws which put a cap on how much interest lenders can charge borrowers in order to protect consumers from usurious lending practices.

Seemingly extraordinary aspects of modern economies such as "payday loans" at usurious and illegal interest rates and "starter interrupt" devices installed in cars are explained by the incomplete or unenforceable nature of contracts. The promise to repay the loan is sometimes no more enforceable than an employee's promise to work hard and well. These two examples – payday loans and starter interrupt devices – are just a small window into the workings of a credit market, the subject of this chapter.

We will see that a sizable fraction of populations for which we have data are either unable to borrow at all or unable to borrow as much as they would like at the interest rates being charged by lenders. At the going price (the rate of interest) they are unable to borrow the quantity they would like (or even at all): They are called quantity constrained borrowers.

Think how odd this is. You would definitely find it strange if you were told that half of all families were unable to purchase some commodity, say, milk, or the amount of milk that they wished to buy at the going price of milk. In the next section we will see that this is exactly what occurs in credit markets: large numbers of individuals are unable to "buy" any credit at all or the amount they would like at the interest rates being charged.

12.2 Evidence on credit and wealth constraints

REMINDER One of the main conclusions of Chapters 10 and 11 was that markets with incomplete contracts do not clear in equilibrium: in the Benetton model some sub-contractors don't get contracts and in the Ford model, some workers don't find jobs. Like these sub-contractors and workers, some would-be borrowers in the credit market are quantity constrained.

In the credit market, many who would like to borrow:

- are excluded from borrowing entirely (unable to borrow), or
- face limits on how much they can borrow (can't borrow as much as they wish) or
- pay extraordinarily high rates of when they do succeed in getting a loan.

Those who cannot borrow at all are termed **credit market excluded**. The excluded along with those who face limits on how much they can borrow and or face very high rates of interest are termed **credit constrained**. Where borrowers face these limitations due to their lack of wealth or low income, they are termed **wealth constrained.**

WEALTH CONSTRAINTS Where borrowers face credit constraints due to their lack of wealth or low income, they are said to face **wealth constraints**.

How do we know if a family or business is credit constrained?

Quasi-experimental measures of credit constraints

CREDIT CONSTRAINTS A family or business is said to be credit constrained if they either

- are excluded from borrowing entirely or
- face limits on how much they can borrow or
- pay extraordinarily high rates of when they do succeed in getting a loan.

Economists look for situations that are like experiments in which we study the actions taken by people who are similar except that some unexpectedly get access to funds (for example an inheritance) and some do not. If a person is not credit constrained then having extra funds should not lead to any change in business practices or behavior: if a change would have raised profits, then funds could have been borrowed for the purpose.

So if the surprise arrival of additional funds changes behavior of it means that prior to their good luck they were credit constrained. Here is some evidence from the U.K..

- In one study the inheritance of $10,000 doubled a typical British youth's likelihood of setting up in business.
- Another study found that holding other things constant people were much more likely to go into self-employment shortly after inheriting wealth, and that inheritance leads the already self-employed to increase the scale of their operations considerably.
- Research on home owners found that a 10% rise in value of housing assets that could be used as collateral in the United Kingdom increases the number of start-up businesses by 5%.

Here is another example.

In 2011 the government of Nigeria invited aspiring entrepreneurs under the age of 40 to submit business plans for either start-up firms or improvements on existing firms offering the most promising proposals training in business and networking opportunities. Among those plans judged to be superior, a random selection would be offered an unconditional grant equivalent to almost $50,000. Twenty-four thousand plans were submitted and 720 plans were selected by a lottery from a group of these proposals that were judged to be well conceived.

By comparing those randomly selected to receive the grant with those who were not selected, this program – called YouWiN!! – provides an ideal experiment for assessing the ways that credit market exclusion and wealth constraints limit businesses. If the young entrepreneurs had been able to borrow all they wished at the going rate of interest, then the grant would have had no effect on how they ran their business. It just would have made them richer, allowing them to consume more, or to save and use their $50,000 to purchase stocks in other companies or lend it to others.

But, tracking what the companies did through 2015, we know that is not what happened. They had good projects that they previously could not finance or they had projects which they had succeeded in financing, but on a scale less they would have been profitable. The lucky winners used their grants to expand their businesses. They purchased more equipment and other capital goods and as a result hired more labor. New firms were created, and they were profitable and survived.

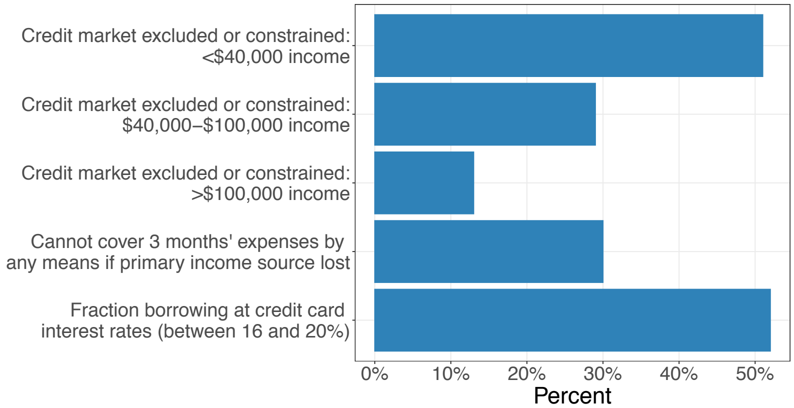


Figure 12.2: **Credit market exclusion in the United States in 2019.** Data Source: These data are from U.S. households surveyed in the 2019 and 2020 Survey of Household Economics and Decision making (SHED). This survey, conducted by the United States Federal Reserve Board, measures the economic well-being of U.S. households and identifies potential risks to their finances. `https://www.federalreserve.gov/consumerscommunities/shed.htm`

Even more dramatic evidence of credit market exclusion comes from an experiment in Sri Lanka. A sample of 408 very small businesses were randomly divided into those that received a grant worth about 100 US dollars and the control group that received nothing. The researchers then collected information on the subsequent investments, sales and profits of the two groups. The firms that received the grants earned extraordinary profits equal to about 60 percent of their capital stock annually. If additional funds made profits of this magnitude available, these firms surely would have borrowed if they could have (the rate of interest that banks charge in Sri Lanka between 12 and 20 percent). But they could not.

Who is constrained and how many are they?

Another way to measure the extent of credit constraints is simply to ask people if they have been denied credit, or if they *believe* they would be were they attempt to borrow. This survey-based evidence along with the quasi-experimental data provide an estimate of the extent of credit constraints.

The U.S. Federal Reserve Board (the nations central bank) surveyed U.S. families in 2019 and 2020, with results for 2019 shown in Figure 12.2. (The results for 2020 may be atypical due to economic impact of the COVID-19 pandemic.) They found, for example, that among loan applicants with incomes less than $40,000 over half were credit constrained, that is unable to borrow at all, or not able to borrow the amount they wished. More than half of all families were paying rates close to 20 per cent a year on their credit card debt (four or five times the rate of interest that home owners were paying on their mortgages.) Thirty percent reported they would be unable to cover their expenses from any source (including savings and borrowing) if they lost their primary source of income for three months (as many did during the pandemic).

Further evidence that credit constraints are common, even in high income countries, comes from another study using a quasi-experimental strategy. It exploited the fact that credit card borrowing limits are often increased automatically (and from the card-holder's viewpoint, unexpectedly). If borrowing increases in response to these exogenous changes in the borrowing limit, we can conclude that the person was credit constrained.

Based on the borrowing behavior of U.S. families, the authors found "that increases in credit limits generate an immediate and significant rise in debt." Gross and Souleles estimate the extent of credit limits as follows:

> It is plausible that many of the one-third of households without bankcards are liquidity-constrained ... Of the two-thirds with bankcards, the over 56 percent who are borrowing and are paying high interest rates (averaging around 16 percent) might also be considered liquidity-constrained, lacking access to cheaper credit. Combined with the households lacking bankcards, they bring the overall fraction of potentially constrained households to over 2/3.

FACT CHECK In studies of household surveys, one person in the household is typically designated as the main decision-maker or *household head*. Differences in the characteristics of the household head have been shown to be relevant for a variety of economic outcomes, as the study in Italy shows.

A study of Italian households found that those who did not borrow either because they were denied credit or believed they would be refused credit, were more likely to be larger poorer families, with an unemployed, less well educated, female, and younger head of household.Moreover, by comparison to families unlikely to face credit constraints, poorer, younger, families with more uncertain sources of income (self-employment rather than pensions, for example) tended to avoid holding risky assets, consistent with the view that credit-constrained people enjoy lower expected income from the investments they do make.

In sum, credit constraints are common in both high and lower income economies. A principal-agent model of lending and borrowing with an incomplete credit contract will explain why this is the case, and illuminate its consequences for how the economy works. To build up the model we start with two cases that to not involve borrowing with an incomplete credit contract:

- A person wealthy enough to own and operate a risky investment project without borrowing and
- A person who must borrow the funds to invest in the project but (unrealistically) secures a loan with a complete contract.

The project we consider is risky in the sense that it may fail or succeed and while the operator of the project (the borrower) can influence which of these is more likely to happen there is no way to ensure that the project will with certainty yield a positive income.

We continue the analysis of risk in the next chapter, where we introduce what is termed "risk aversion" namely a preference for the "sure thing" over a risky bet with the same expected payoff. (A risk averse person would prefer $100 with certainty than a coin flip to see if she receives either nothing or $200.)

The lack of risk aversion is termed risk neutrality, which is what we assume in this chapter.

12.3 The wealthy owner-operator case

A risky project

Antonio invests in a project that requires an amount \$ $\bar{k}$ to carry out. We will call this amount \$1 but it could represent \$1 thousand or \$1 million dollars. Imagine that the "project" is a machine that has a dial on it by which Antonio can regulate its speed of operation, ranging from 0 to 1. The machine produces goods in proportion to the "speed" at which it is run. But the faster that Antonio runs the machine, the greater is the likelihood that the machine will break and destroy itself and all of its output. Going forth, we will use the words "project" and "machine" interchangeably.

We assume that f, the probability that the machine will break (i.e. fail), is simply the speed at which it is run (so f represents both fail and fast). It surely will *not* break if it is not operated ($f = 0$) and it surely *will* break if it is run at top speed ($f = 1$). If the machine does not fail while in operation, it becomes worthless at the end of the period. The goods produced are available at the end of the period under the condition that the machine has not failed.

So increases in the speed of the machine f represents *both*:

- *Higher potential income:* Greater income from the investment (higher output for higher f, therefore more goods sold which is the basis of Antonio's income)
- *Greater chance of failure:* higher probability of failure when he runs the machine faster.

How Antonio evaluates this trade off depends on how he feels about taking risks, and this may depend on his ability to borrow funds or access some kind of insurance if the project fails, or other risks to which he is exposed (ill health for example). We take up attitudes towards risk – called risk aversion – in the next chapter. Here we assume that Antonio wants to maximize his expected income (so he does not care about the risk, he is said to be "risk neutral").

Machines differ in how good they are. We let q, a positive constant, represent the *quality* of the project: higher quality projects result in more income for any given speed at which the machine is run, lower quality projects in less income. The project results in either:

- *Success*: income of qf if it succeeds.
- *Failure*: 0 income if it fails.

M-CHECK The probability of failure depends only on the speed at which the machine is run, *not* its quality (q). We could represent quality as a reduction in the probability of failure at any given speed, so that, for example, q varies from 0 to 1 and the probability of failure is $(1-q)f$. In this case the top quality machine (that is $q = 1$) even when run at top speed would be indestructible! But, ruling out that case, this alternative model would produce similar results to what we have assumed here.

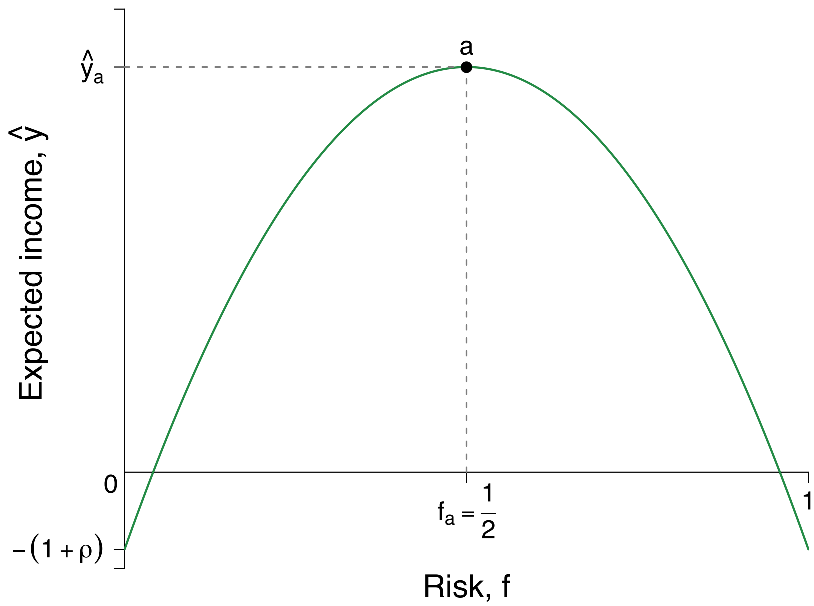


Figure 12.3: **Risk and expected income.**
If the machine is not operated at all ($f = 0$) then the expected income is negative, namely the opportunity cost of the $1 project. Starting from a low level of risk, increasing the risk taken initially increases the expected income ($\hat{y}$). Eventually adopting an even more risky strategy leads to lower expected income. The maximum expected income, $\hat{y}_a(f) = \frac{q}{4} - (1+\rho)$ is achieved when the operator chooses $f = f_a = \frac{1}{2}$. Read M-Note 12.1 to make sure you understand the calculations.

The owner-operator's choice of a risk level

Antonio will sell the output of the machine, so we have Antonio's expected revenues from operating it:

$$\begin{aligned} \text{Expected revenues} &= 0 \times \text{Pr. failure} + qf \times \text{Pr. success} \\ \hat{r}(f) &= 0 \times f + qf(1-f) \\ \hat{r}(f) &= qf(1-f) \end{aligned} \tag{12.1}$$

His expected income $\hat{r} -\hat{}$ is an indication of expected value – from the project must include the opportunity cost of the $1 that he could have invested in some other project $1+\rho$ where ρ is the opportunity cost of capital. If the owner-operator had not bought the machine and instead had invested the one dollar it cost at the risk-free interest rate ρ, he would have had (with certainty) $\$1+\rho$ at the end of the period.

His expected income from investing in the project is therefore

$$\begin{aligned} \text{Expected income} &= \text{Expected revenues - Opportunity cost of the investment} \\ \hat{y}(f) &= qf(1-f) - (1+\rho) \end{aligned} \tag{12.2}$$

In Figure 12.3 we show Antonio's expected income from the project and how it depends on how fast he runs the machine.

Antonio is the owner-operator – he owns the machine and he owns any output that it produces – so he would vary the speed at which he runs the machine (f) to maximize his expected income on the project and set $f = f_a = \frac{1}{2}$ and (inserting this value in Equation12.2) have expected income of $\frac{q}{4} - (1+\rho)$, as illustrated by point **a** in Figure 12.3, and shown in M-Note 12.1.

Having determined how fast he will run the machine if he invests in the project, Antonio now has to decide whether to do it. Because we have assumed that his next best alternative is the risk free return $1+\rho$ we can see from the figure that he makes an economic profit on the investment (income greater than opportunity cost of capital). So he should undertake the project.

You will will see from Figure 12.3, and as is shown in M-Note 12.1 that the opportunity cost of capital ρ affects whether Antonio will undertake the project, but not the speed at which he runs the machine if he does.

M-Note 12.1: Maximum expected income and project quality

Starting with Equation 12.1, Antonio will maximize $\hat{y}(f)$ with respect to f. To do this, he will impose the first order condition for a maximum $\frac{d\hat{y}}{df} = 0$:

$$\begin{array}{rrcl} \text{Expected income} & \hat{y} & = & qf(1-f)-(1+\rho) = qf - qf^2 - (1+\rho) \\ \text{First Order Condition} & \frac{d\hat{y}}{df} & = & q - 2qf = 0 \\ \text{Isolating } f & f & = & \frac{q}{2q} = \frac{1}{2} \end{array}$$

So Antonio finds the maximum expected income by running the machine at $f = \frac{1}{2}$, or one half its maximum possible speed (at $f = 1$ it would be guaranteed to break).

We can now substitute $f = \frac{1}{2}$ into the expected income, $\hat{y}$:

$$\begin{aligned} \hat{y} &= q\frac{1}{2}\left(1-\frac{1}{2}\right)-(1+\rho) \\ &= \frac{q}{4}-(1+\rho) \end{aligned}$$

Therefore, Antonio's expected income is a positive function of the quality of the project. Notice that the opportunity cost of capital ρ does not enter into the speed that maximizes the income of the project.

Checkpoint 12.1: When would Antonio decide not to invest in the project?

- Redraw Figure 12.3 showing a case in which Antonio would decide not to invest in the project.
- If $\rho = 0.05$ what is the smallest value of q the quality of the machine such that Antonio would decide to invest in it?

FACT CHECK Though a probability of failure of 50% may seem high for Antonio's project, in the United States data from the Bureau of Labor Statistics says that roughly 50% of new businesses fail within their first five years of operation. About 20% fail within the first year. Businesses seeking to produce an innovation of some kind - a new app or a new pharmaceutical – are especially likely to fail. But our model does not attempt a realistic picture of risks, it is instead a way to think through the logic of incomplete contracting in risky situations.

12.4 Complete credit contracts: A limiting case

To introduce credit, we now consider the case in which Antonio has no wealth and thus cannot self-finance his project (pay for the initial investment himself), but needs to borrow the funds from a lender, Parama. To allow us later to clarify the difference that the incompleteness of the credit contract makes, we begin with a hypothetical case in which the lender can include in the loan contract the degree of risk that the borrower takes. So the the "speed dial" on the machine is not only visible to the lender, but also the information it shows is verifiable and can be used to enforce a contract.

Interest, repayment, bankruptcy and limited liability

Parama is in the business of making money, so she will want to make a profit from her loan. At the end of the period for which the loan is granted, in addition to requiring Antonio to repay the principal – the amount of the loan, \$1 in this case – she will require an additional amount, called the interest on the principal. The interest rate i is a percentage of the principal that is added to the amount the the borrower is required to repay. At the end of the period, Antonio is required to repay Parama the amount δ, called the **interest factor** which is:

$$\begin{aligned} \text{Interest factor} &= \text{Principal + interest} \\ \delta &= \text{Principal}(1 + \text{interest rate}) \end{aligned}$$

But just as operating the machine is risky for Antonio, lending money to him is also risky for Parama. The reason is that in most legal systems laws concerning bankruptcy and limited liability mean that if the project fails the lender may not be able to recover the loan. If the project fails, then the lender may not take the borrower's house or other assets except those specifically pledged as collateral for the loan. We simplify by assuming that if the project fails, Antonio pays back nothing: an aspect of the model that is crucial to what follows.

INTEREST FACTOR The interest factor is one plus the rate of interest.

REMINDER Unfortunately, there are two meanings of the word *principal* that we have to use in this chapter and we don't want to confuse them. First, *principal* is used to describe a decision-maker involved in a principal-agent relationship, such as the lender-borrower relationship we describe. Second, the word principal also has a particular meaning in *finance*: the principal is the initial sum or investment amount lent out by a lender to a borrower, such as the proportion of \$1 that Parama lends to Antonio in our example.

Antonio therefore has two possible incomes:

- *Project does not fail:* With probability $1 - f$ the project succeeds, giving Antonio revenue of qf. He must also re-pay the principal his loan plus interest – that is, the interest factor (δ) – so his income is: $qf - \delta$
- *Project fails*: With probability f the project fails, so he receives no revenue and, because of limited liability, he does not pay back the loan. His income is zero.

Antonio's expected income is the sum of these two incomes – that is, $qf - \delta$ and zero – weighted by the probability that each occurs:

$$\hat{y}(\delta, f) = \underbrace{(1-f)(qf-\delta)}_{success} + \underbrace{f \cdot 0}_{failure} \tag{12.3}$$

The lender's expected economic profits

Whether the project fails or succeeds, the opportunity cost of the funds Parama lends to Antonio is $(1+\rho)$. We assume that she, the lender, wishes

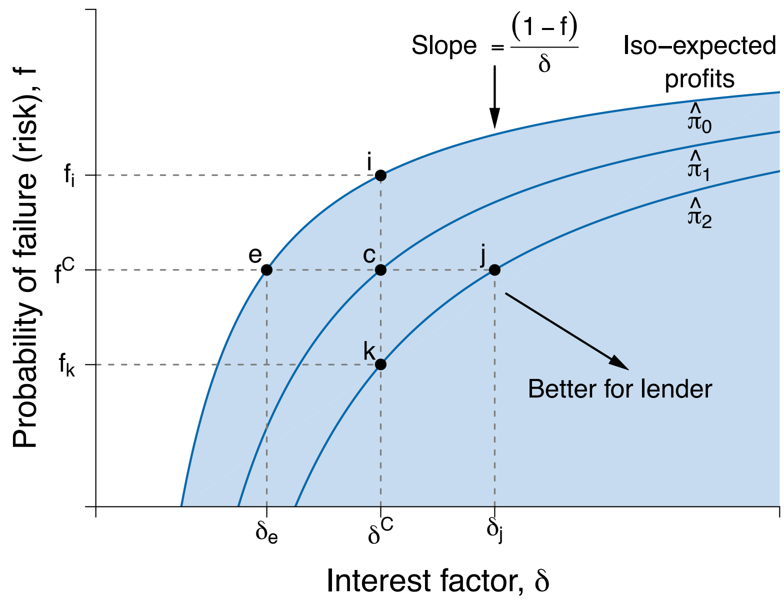


Figure 12.4: **The lender's isoexpected profit curves.** The points making up a given isoexpected profit curve indicate combinations of risk (f) and the interest factor (δ) that according to Equation 12.4 result in the same profit.

The slope of the isoexpected profit curve is $\frac{\Delta f}{\Delta \delta} = \frac{1-f}{\delta}$, as shown in M-Note 12.2.

to maximize expected economic profits, meaning the expected repayment of the loan minus the opportunity cost of the funds lent. Like Antonio there are two levels of income she may receive:

REMINDER:ECONOMIC PROFITS Economic profit of a project is its income minus its costs including the opportunity cost of the funds devoted to the project. We use the term 'profits' here to mean economic profits, not accounting profits (which does not take account of the opportunity cost of the funds involved.)

- *Antonio's project does not fail.* This occurs with probability $1-f$, and in this case Parama receives δ.
- *Antonio's project fails*: This occurs with probability f and Parama then receives zero.

So we have Parama's:

$$\text{Expected profit} \quad \hat{\pi}(\delta, f) \quad = \quad \overbrace{(1-f)\delta}^{success} + \overbrace{f \cdot 0}^{failure} - \quad \overbrace{(1+\rho)}^{opportunity\quad cost} \quad (12.4)$$

To summarize so far, here is the complete credit contract game (we use the C superscript to refer to the complete contracts case):

- Parama, the principal, announces both an interest factor δ^C and the speed at which the machine is to be run, f^C.
- Antonio, the agent, either accepts or rejects the contract
- If he accepts he operates the machine at the specified speed, and it fails with probability f^C
- If the machine fails both Parama and Antonio receive nothing.
- If it does not fail they receive respectively δ^C and $qf^C - \delta^C$

This ends the game.

Figure 12.4 shows the lender's isoexpected profit curves (indicated by $\hat{\pi}$) with higher expected profits represented by the curves that are to the right and lower. Why is this the case?

Remember, isoexpected profit curves give Parama's evaluation of every point in the space, irrespective of whether that point could occur. So to answer the question think hypothetically. If the borrower were to take some given level of risk (f), then the lender can expect to make more profit the higher is the interest factor. Comparing points **e**, **c**, and **j** we can see that if the borrower were to run the machine at a speed f^C the lender would make greater profits if the interest factor, δ is higher, with $\hat{\pi}_0 < \hat{\pi}_1 < \hat{\pi}_2$ on the corresponding iso-expected profit curves. Similarly, comparing points **i**, **c**, and **k** for a given interest factor δ^C, the lender would make more profit if the borrower were to take less risk.

The iso-expected profit curve labeled $\hat{\pi}_0$ is special because it represents the combinations of f and δ such that the expected profit of the lender is zero (meaning accounting profits are just sufficient to offset the opportunity cost of the funds used for the project). The curve labeled $\hat{\pi}_0$ divides the space in the figure into two regions.

The blue shaded area in the figure are the outcomes satisfying the lender's participation constraint (because for these combinations of f and δ her expected economic profit is positive). She would not be interested in engaging in any lending on terms that lie outside the blue shaded area. We will see later that this explains why some prospective borrowers are unable to secure a loan at any rate.

M-Note 12.2: The lender's $mrs(\delta, f)$ and the iso-expected profit curve's slope

Using Equation 12.4, we can find the slope of the lender's isoexpected profit curves:

$$\text{P's expected profit} \qquad \hat{\pi}(\delta, f) = (1-f)\delta - (1+\rho)$$

To find the slope of the iso-expected profit $\frac{df}{d\delta}$, we want to find the changes in f and δ that are consistent with no changes in the lender's expected profits, or staying on the same iso-expected profit curve. To do this, we need to find the total derivative of the lender's expected profit function and set it equal to zero:

$$d\hat{\pi} = \frac{\partial \hat{\pi}}{\partial \delta} d\delta + \frac{\partial \hat{\pi}}{\partial f} df = 0$$

This means that, for any two points on an iso-expected profit curve, the difference in expected profits associated with a difference in the interest factor ($\frac{\partial \hat{\pi}}{\partial \delta} d\delta$) is exactly compensated by the opposite signed difference in expected profits associated with the difference in the risk level ($\frac{\partial \hat{\pi}}{\partial f} df$), so that the total difference in expected profits is zero.

We can use our specific expected profit function and rearrange to find $\frac{df}{d\delta}$, the slope of the indifference curve:

$$\begin{aligned} d\hat{\pi} &= (1-f)d\delta - \delta df = 0 \\ \text{Re-arranging} \quad \delta df &= (1-f)d\delta \\ \frac{df}{d\delta} &= \frac{(1-f)}{\delta} \end{aligned} \tag{12.5}$$

Equation 12.5 says that for $f < 1$, the slope of the lender's iso-expected profit curve is positive because $\delta > 0$.

Recall that the lender's marginal rate of substitution is the negative of the slope of her iso-expected profit curve, which we can therefore say is the following:

$$mrs(\delta, f) = -\frac{(1-f)}{\delta}$$

The borrower's participation constraint and the lender's profit maximization

The lender will propose a contract specifying both the speed at which the machine will be run f and the interest factor to be paid at the end of the loan period, δ. The lender's offer must be sufficiently attractive to the borrower so that he will accept. This depends on what the borrower's other money making opportunities are.

For simplicity we assume that Antonio has no other opportunities so his fall-back option z is zero. So using Equation 12.3, the participation constraint limiting the lenders contract is

$$\begin{aligned} \text{Borrower's expected income} &\geq \text{Borrower's fallback option} \\ (1-f)(qf-\delta) &\geq z = 0 \end{aligned} \tag{12.6}$$

The participation constraint divides the space of contractual terms in Figure 12.5 into hypothetical contracts that Antonio would accept (the green shaded area) and that he would reject. Contracts whose terms (f and δ) satisfy the participation constraint as an equality are on the green ray from the origin. These contracts provide Antonio with the minimum expected income (that is zero) sufficient for him to accept the loan.

REMINDER As before, for simplicity we assume that a contract is accepted if it is not worse than the alternatives.

As shown in M-Note 12.3 we can rearrange Equation 12.6 as:

$$\text{Participation constraint:} \quad f \geq \frac{\delta}{q} \tag{12.7}$$

This is also Antonio's willingness to sell.

In what follows we express Equation 12.7 as an equality, or

$$f = \frac{\delta}{q} \tag{12.8}$$

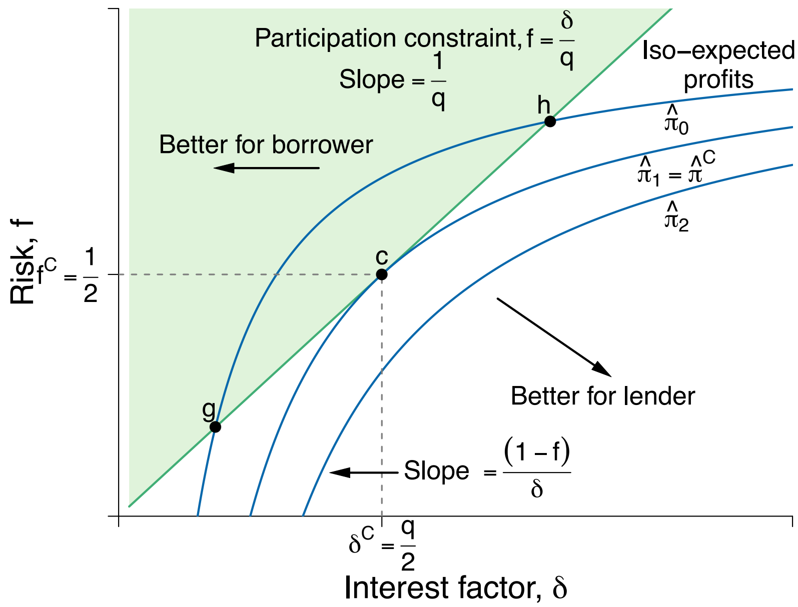


Figure 12.5: **The lender's expected profit maximizing choice of a contract (f^C, δ^C) when the contract is assumed complete.** The lender receives higher expected profits at points that are lower and to the right. The figure shows that the best that the lender can do is to offer the borrower a contract indicated by point **c**, where her highest feasible iso-expected-profit curve is tangent to the borrower's participation constraint. At point **c**, the lender requires the borrower to run the machine at half speed ($f^C = \frac{1}{2}$), and the lender charges the borrower an interest factor of $\frac{q}{2}$. Points **g** and **h** both lie on the borrower's participation constraint, so the borrower is indifferent among them and point **c**. But, the single lender would not choose points **g** or **h** as they are not profit-maximizing (lying on $\hat{pi}_0$) as the lender can make a take-it-or-leave-it offer to the borrower and increase their expected profits to $\hat{pi}_1$.

We call this his willingness to sell because for any limit on Antonio's risk taking (that is, f) that Parama would like enforce in the contract, the participation constraint tells us the greatest interest factor that Parama can charge, given some quality level (q) of the project. Equation 12.8 tells Parama that, for a given quality, if she wants Antonio to agree to a contract that requires him to run the machine slower, she will have to charge him a lower interest factor.

The lender can now find the expected profit maximizing contract by finding the point on the participation constraint (PC) on the highest iso-expected profit curve. This is point **c** where

$$\begin{aligned}
\text{Marginal rate of substitution} &= \text{Marginal rate of transformation} \\
\text{- Slope of iso-expected profit curve} &= \text{- Slope of PC } = -\frac{\Delta f}{\Delta \delta} \\
-\frac{(1-f)}{\delta} &= -\frac{1}{q}
\end{aligned} \tag{12.9}$$

M-CHECK From Equation 12.9

$$\frac{(1-f)}{\delta} = \frac{1}{q}$$

using Equation 12.8

$$\begin{aligned}
\frac{(1-f)}{fq} &= \frac{1}{q} \\
\frac{(1-f)}{f} &= 1 \\
(1-f) &= f \\
\text{we get } f^C &= \frac{1}{2}
\end{aligned}$$

As shown in the M-Check, we can use Equation 12.8 to replace the δ by fq in Equation 12.9, giving us the result that the lender will set $f^C = \frac{1}{2}$, as above using the C superscript to indicate the complete contract case.

Notice that the level of risk implemented when risk can be determined by the lender in an enforceable contract is identical to the risk chosen by the owner-operator. In M-Note 12.3 we show that this is the case because the optimizing problem solved by the owner operator is mathematically identical to that solved by the lender who is able to enforce a level of risk by contract. The owner operator's chosen level of risk is by definition Pareto efficient because he has maximized his expected income, so he cannot be made better off. And there is nobody else involved.

In the interaction between the lender and the borrower, the outcome is also Pareto efficient. This is because the lender maximized profit subject to a constraint given by the borrower's utility level (his participation constraint) and therefore the lender implemented a Pareto-efficient outcome.

REMINDER As we have seen in Chapters 4, 5 and 10, when one actor maximizes utility or profit subject to a constraint that the other's utility level not be less than a given minimum value, the result is by definition Pareto-efficient.

The risk levels that result in the two scenarios – owner operator and lender borrower – are the same. The reason is that in both cases the decision maker – Antonio in the owner operator case and Parama the lender in the lender-borrower case – was in a position to capture the entire gains – that is the total expected rents – from the transaction. In both cases the maximum possible total expected rent was $\frac{q}{4} - (1+\rho)$. This being the case they wanted it to be as large as possible, which Antonio achieved by setting $f = \frac{1}{2}$ and which Parama achieved by charging him an interest factor of $\frac{q}{2}$ and requiring Antonio to set $f = \frac{1}{2}$.

The difference between the two cases is that in the owner operator case Antonio had enough wealth to carry out the project himself so he got all of the rents. In the borrower lender case he was without wealth and had to borrow funds from Parama; and she got all of the rents. Antonio's wealth (or lack of it) is what explains the difference in the distribution of the gains from the exchange.

Except for this important difference in who got the rent, complete contracting allows the lender to implement an outcome that mirrors what the borrower would do if he were both the owner and operator of the machine. In this sense, complete contracting erases the distinction between lender and borrower, and reinstates the world of the owner-operator. The results change when we turn to real-world, incomplete credit contracts.

M-Note 12.3: A complete credit contract mimics the owner operator

To study the case of complete contracting shown in Figure 12.5,we start with the borrower's participation constraint Equation 12.6 satisfied as an equality, so:

$$(qf-\delta)(1-f)=0 \tag{12.10}$$

For Equation 12.10 to be true, either $(qf-\delta)=0$ or $(1-f)=0$.

- $(1-f)=0$: If this is true, then it must be that $f=1$, in which case the machine will fail with certainty. So this cannot be true.
- Therefore, the participation constraint can be rewritten $(qf-\delta)=0$.
- Or, equivalently $\delta = qf$
- And rearranging this we have $f=\frac{\delta}{q}$.

With respect to the lender, we need to consider her expected profit function:

$$\text{Lender's profit} \quad \hat{\pi}(\delta, f) = (1-f)\delta - (1+\rho) \tag{12.11}$$

We can now substitute $\delta = qf$ into Equation 12.11 such that:

$$\hat{\pi}(f) = (1-f)qf - (1+\rho) \tag{12.12}$$

The function the lender maximizes here is the same as in the owner-operator case, namely Equation 12.2 and in M-Note 12.1. As in that case, the lender will choose f to maximize Equation 12.12 and set $f = \frac{1}{2}$. Then inserting this value into Equation 12.12 we see that the lender has an expected profit of $\hat{\pi} = \frac{q}{4} - (1+\rho)$, the same as in the case of the owner operator.

12.5 The general case: incomplete credit contracts

A principal-agent model based on an incomplete credit contract can explain why lenders use devices like car disablers for borrowers in default or why so many families are unable to borrow the amounts they would like. We begin with borrower – Antonio – who has no wealth; later we will show why borrowers like Antonio are likely to be excluded from the credit market entirely.

The credit contract will be incomplete if either or both of two critical pieces of information are not known by (or unverifiable for) the lender.

- *Hidden attributes:* The quality (q) of the borrower's project – in our example how good the machine is.
- *Hidden actions:* The level of risk that the borrower takes (f) – in our example, how fast he runs the machine.

Either or both could be the case, but to focus on a concrete example (the speed of the machine) we assume that the quality of the project is known to the lender (no hidden attributes), but that the level of risk that the borrower takes is a hidden action. Information about the speed at which the machine is run is either asymmetric (only Antonio knows it) or non-verifiable (Parama may know it as well, but cannot use it in court to enforce a contract). So the lender can no longer contractually set the degree of risk taken by the borrower, and can set only the interest factor, δ.

Here is the game:

- *The lender* is the principal, who knows the borrower's best response choice of a risk level for each level of the interest factor ($f(\delta)$), and as first mover she determines and proposes the interest factor (δ) that will maximize her expected profits.
- *The borrower* (the agent) then selects the level of risk to take (the speed of the machine, $f(\delta)$).

REMINDER Recall that a **principal-agent relationship** arises when two conditions hold:

- *Conflict of interest*: the actions or attributes of the agent affect the principal's objective in such a way that there is a conflict of interest between the two.
- *Incomplete contract*: the agent's actions or attributes that are of interest to the principal are not subject to an enforceable contract.

REMINDER As in the Benneton model, the principal's name begins with a P (here, Parama, the lender) and the agent's name begins with an A (here, Antonio, the borrower). We will use "principal," "lender," and "Parama" interchangeably (and likewise "agent," "borrower," and "Antonio").

REMINDER In Chapters 10 and 11 we presented similar games, first explaining the agent's best response function (the quality of goods to provide, how hard to work) and then using as the incentive compatibility constraint on the principal's choice of a price or a wage. We follow the same logic here.

- *Chance* then intervenes: the project either fails with a probability f or does not fail with probability $1-f$.
- *The borrower* then repays the interest factor to the lender if the project has not failed (the machine has not destroyed itself) and repays nothing if the project failed.

This ends the game.

Because the first mover, the principal, uses the borrower's best response function as the incentive compatibility constraint for her expected profit maximizing problem, we begin with that.

The borrower's best response

Given the interest factor set by the lender, δ, the borrower will choose f to maximize his expected income (repeating a slightly rearranged Equation (12.3)):

$$\underbrace{\hat{y}(\delta,f)}_{Expected\quad income} = \underbrace{qf(1-f)}_{Expected\quad revenues} - \underbrace{\delta(1-f)}_{Expected\quad repayment} \tag{12.13}$$

The final term in Equation 12.13 makes it clear that the lender, by extending credit to the borrower is unavoidably providing a kind of insurance. This is because:

- the level of risk taken is not enforceable by contract and
- bankruptcy and limited liability laws make it unlikely that loans invested in failed projects will be repaid.

The lender bears the costs of the risk decisions made by the borrower: if the machine fails, the borrower does not repay the loan. So in choosing a risk level the borrower does not 'own' all of the consequences of his choice.

Recall that one-half is the risk level chosen when the decision is made by someone who bears all of the risk, that is, by the owner-operator. Because the borrower does not bear all the consequences of the risk level that he chooses he will run the machine at a value of f greater than one half. How much faster depends on the interest factor that the lender will choose. The tradeoff faced by the borrower is that running the machine faster runs the risk of failure, and getting nothing; but it also increases the probability that the loan will not be repaid (due to the failure).

The borrower will choose a value of f by comparing the effect of running the machine faster on both expected revenues and expected repayments, the two terms on the right hand side of Equation 12.13.

REMINDER: EXTERNAL EFFECTS The fact that the lender bears some of the risk resulting from decisions made by the borrower is another example of a negative external effect, similar to the positive external effects that the worker's hard work or the quality of the sub-contractor's good confers on the principals (the employer and the buyer) with whom they interact. Other negative external effects you have studied include the downward pressure one firm's sales on the demand for another firm's products in the Cournot model, and the reduction in the fish caught by one fisherman resulting from the fishing of another.

- *Effect on expected repayments*: You can see from the last term on the right of Equation 12.13 that by running the machine faster the borrower reduces the likelihood that the borrower will have to repay the interest factor δ, which is the *benefit* of running the machine faster. The higher the interest factor, the greater is this incentive – the gain to the agent of not having to pay back the loan – so the faster the agent will run the machine.

- *Effect on expected revenues:* Remember, running the machine at one-half is the speed that maximizes that expected revenues of the project. Because the borrower will set f at a level greater than the risk level that would maximize revenues, we know that increasing f will reduce revenues. We show in M-Note 12.4 that the marginal cost of running the machine faster is $-q(1-2f)$.

These two effects of running the machine faster are challenging to understand because both involve reducing something: in one case reducing a "good" and in the other reducing a 'bad'. Running the machine faster reduces something that the borrower considers to be a good, namely expected revenues (that's bad). It also reduces something the borrower considers to be a "bad" the expected repayment (in the eyes of the borrower, that's good).

The borrower's expected income is the expected revenue minus the expected repayment. To determine the expected income maximizing level of risk to take, the borrower will adopt the following rule, illustrated in Figure 12.6:

- If the benefit of increasing f (reducing the expected repayment) exceeds the cost (reducing expected revenues), increase f. For example, if the interest factor is δ_b and the machine is run at speed f_h, then some faster speed would increase the borrower's expected income.

- If the benefit of increasing f falls short of the cost, decrease f, so if the machine is being run at speed f_e and the interest factor is δ_n, then the machine is being run too fast.

- If the benefit of increasing f is the same as the cost of increasing f (as occurs at points **b**, **n**, and **e** in the figure), then implement this level of f.

M-Note 12.4: The borrower's first order condition

The borrower's best-response function is derived by finding for each interest rate that the lender may offer, the level of the borrower's chosen level of risk that will maximize his expected income $\hat{y}(\delta, f)$. To find this, we differentiate Equation 12.13 with respect to f and set the result equal to zero, or

$$\begin{aligned}
\underbrace{\hat{y}(\delta, f)}_{Expected \quad income} &= \underbrace{qf(1-f)}_{Expected \quad revenues} - \underbrace{\delta(1-f)}_{Expected \quad repayment} \\
&= qf - qf^2 - \delta + \delta f \\
\frac{d\hat{y}}{df} &= q(1-2f) + \delta = 0
\end{aligned}$$

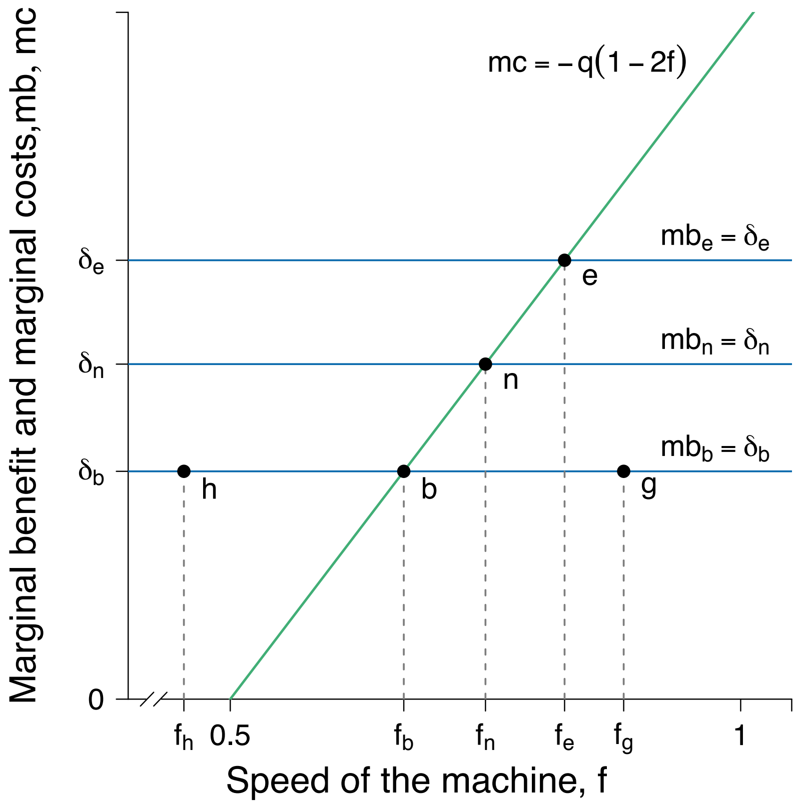


Figure 12.6: **Marginal benefits and marginal costs of choosing greater risk** (f) **given the interest factor** (δ) **determined by the lender**. Marginal costs are higher for greater levels of risk (f). We show the marginal costs only for the case where f is 0.5 or greater, so that $(1-2f)$ is negative. For $f<0.5$ the marginal cost is negative, that is, running the machine faster increases the expected revenues from the operation of the machine. No operator – either as an owner or an agent using borrowed funds – would operate the machine slower than $f=0.5$. Note also that if δ is zero, then the borrower will run the machine at half speed (just like the owner-operator case). The figure shows that, for a low value of δ, the borrower will run the machine slower than for a higher value of δ. The lettered points here have counterparts in Figure 12.7, where we look at the same problem, but from a different vantage point.

The value of risk that maximizes the borrower's expected income is determined by:

$$-q(1-2f) = \delta$$

This means that the expected income maximizing speed of the machine is the value of f satisfying the following rule, expressed in terms of marginal benefits (MB) and marginal costs (MC) of increasing f:

$$\text{MB of increasing f} = \delta = -q(1-2f) = \text{MC of increasing f} \tag{12.14}$$

which, rearranged gives us

$$\text{A's best response} \quad f(\delta) = \frac{1}{2}+\frac{\delta}{2q} \tag{12.15}$$

You can see from Equation 12.15 that:

- The borrower's best-response to the interest factors offered by the lender shows that at an interest factor of zero, the borrower will run the machine at a speed of one half, $f=\frac{1}{2}$, just as did the owner-operator and the borrower with complete contract.

M-CHECK

$$\begin{aligned}
\text{Borrower's FOC:} \quad \delta &= -q(1-2f)\\
\text{Multiply out} \quad \delta &= -q+2qf\\
\text{Isolate f term} \quad 2qf &= \delta+q\\
\text{Divide through by } 2q \quad f &= \frac{\delta+q}{2q}\\
\text{Borrower's BRF} \quad f(\delta) &= \frac{1}{2}+\frac{\delta}{2q}
\end{aligned}$$

- The slope of the borrower's best-response function is $\frac{1}{2q} > 0$, so the borrower takes on more risk (higher f) if the lender chooses a higher interest factor.
- The level of risk chosen by the borrower f is less the higher is the quality q of the project (given the same δ).
- As the slope of the best response function (how steep it is) is $\frac{1}{2q}$, the larger the quality q, the flatter the best response function will be. In other words, the higher the quality, the lesser the effect of the interest factor in raising the risk level.

Checkpoint 12.2: Risk and quality

Why does the borrower choose a higher level of risk for lower quality projects? Think about what changing q would do to the slope of the marginal cost curve.

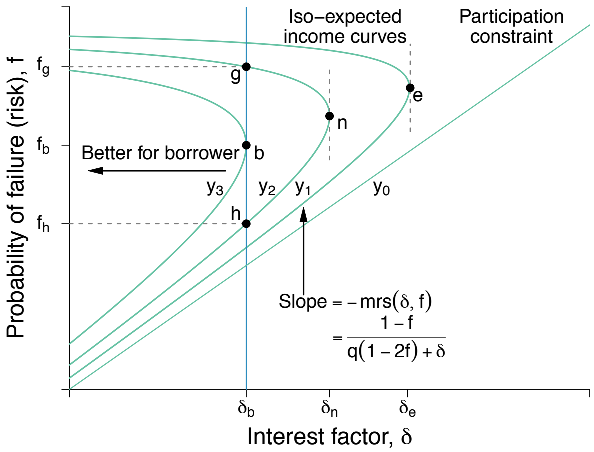
12.6 The Nash equilibrium level of risk and interest

We now illustrate the borrower's best response with the help of his iso-expected-income curves shown in Figure 12.7. Each point in the figure represents some hypothetical outcome of the interaction between the principal and the agent– a combination of a degree of risk f and an interest factor δ (no matter how unlikely). An iso-expected income curve (as in Chapter 10) gives us all such combinations that result in the agent having some given level of expected income.

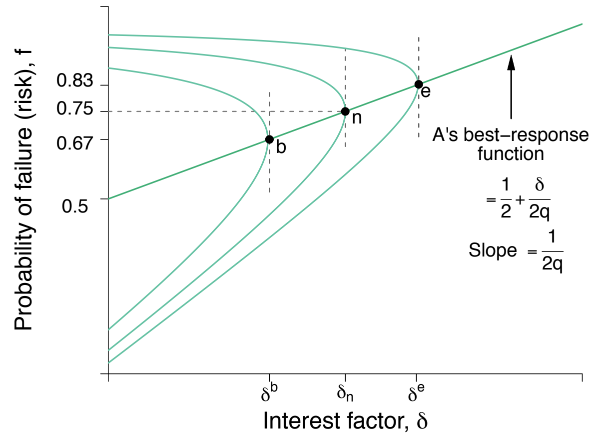
REMINDER: BEST-RESPONSE FUNCTIONS
In developing the Benetton model we first derived the the sub-contractor's (the agent's) best-response function by finding the quality level that equated the marginal benefits and costs of their action (quality of good supplied) in Equation **??**. We then showed how the same best response function could be derived from the agent's iso-expected-income curves in Figure 10.6. We follow the same procedure here. Having shown how equating the marginal costs and benefits of running the machine faster provides us with an expression for the borrower's best-response function (Equation 12.15), we now represent the same ideas using the borrowers iso-expected-income curves.

In the left panel, starting with the participation constraint (which you have already seen in Figure 12.5) where the expected income is zero, the expected income is greater for curves that are closer to the vertical axis, as indicated by the horizontal arrow, that is $0 = \hat{y}_0 < \hat{y}_1 < \hat{y}_2 < \hat{y}_3$. This is because for a given probability of failure, that is, comparing points on the curves horizontally, a lower interest factor raises the borrowers expected income. As we did in Chapters 1 and 4, we can think of the iso-expected-income curves as contours in a map of a hill that is sloping up to the left (the horizontal arrow is pointing up).

For a given interest factor, that is, for points along a vertical line e.g. **g**, **b**, and **h** the comparison is more complicated. These points have counterparts in Figure 12.6 in which we compared the marginal benefits and costs of running the machine faster. At point **g** in both of these figures, for example, at the given interest factor δ_b, Antonio is running the machine too *fast*. In Figure 12.6 it was clear that the marginal benefits of running it faster fell short of the marginal costs. Here you can see that by slowing down the machine to f_b Antonio reaches a higher "contour" moving from iso-expected income of $\hat{y}_2$ at point **g** to $\hat{y}_3$ at **b**.



(a) The borrower's iso-expected income curve



(b) The borrower's best-response function

Figure 12.7: **The borrower's iso-expected income curves and best response to the lender's choice of** δ. Points **n**, **b**, and **e** in this figure are the same outcomes as **n**, **b**, and **e** in the Figure 12.6, that is successively higher rates of risk (speed of the machine) adopted by the borrower as the interest factor goes from low, to moderate, to high.

With the same interest factor δ_b, at point **h**, on the other hand, Antonio is running the machine too *slowly*. He can increase his expected income (moving from $\hat{y}_2$ to iso-expected income curve $\hat{y}_3$) by increasing the speed at which he runs the machine from f_h to f_b and moving to point **b**. Again, return to Figure 12.6 and notice that at **h**, marginal costs are lower than marginal benefits and if he moves to **b** they will be equal and, as shown in Figure 12.7 Panel a, he will increase his income from $\hat{y}_2$ to $\hat{y}_3$.

The borrower will not choose points such as **g** and **h** and will instead choose a point like **b**, **n**, or **e** at which his marginal benefits equal his marginal costs. For each of points **b**, **n**, and **e**, for a given level of interest factor δ_e, δ_n, or δ^b, the choice of corresponding risk level that maximizes that borrower's expected income is that point at which the iso-expected income curve is tangent to the vertical line corresponding to that interest factor.

If you think of the line **gbh** as a "trail" across the hill shown by the contours, it rises from **h** to **b**, is flat at **b** and then descends to **g**. So **b** is the highest point on the trail. That is why **b** is a best response. In the right panel of Figure 12.7, we see that the best-response function is made up of points like **b**, **n**, and **e** at which the iso-expected-income curve is vertical.

Checkpoint 12.3: The iso-expected profit

Consider the following questions:

1. Explain in your own words why the iso-expected income curves upward sloping for low values of f and downward sloping for higher levels of f?
2. Explain why the iso-expected profit curves are bowed out (away from the vertical axis) while the iso-expected value for the employees in the previous chapter are bowed in (towards the vertical axis).
3. Use the borrower's best-response function (Equation **??**) and what you know

about the slope of the borrowers iso-expected income curves to explain why the points on the best-response function are where the iso-expected income curves are *vertical*.

M-Note 12.5: The Borrower's $mrs(\delta, f)$

As in M-Note 12.2, we can take the total derivative of the borrower's expected profit function, equation 12.3, and set it equal to zero to find the slope of the iso-expected income curve and its negative, the borrower's *marginal rate of substitution* ($mrs(\delta, f)$):

$$\begin{aligned} \hat{y}(\delta, f) &= (qf - \delta)(1-f) \\ d\hat{y} &= -(1-f)d\delta + (q - 2qf + \delta)df = 0 \\ \text{Re-arranging: } (q - 2qf + \delta)df &= (1-f)d\delta \\ \frac{df}{d\delta} &= \frac{(1-f)}{q - 2qf + \delta} \end{aligned}$$

This is the slope of the borrower's iso-expected income curve. Given that $f < 1$, the numerator of the slope must always be less than zero, so the sign of the denominator will determine the sign of the slope of the iso-expected income curve. That is:

$$\begin{aligned} q(1-2f) + \delta &\lesseqgtr 0 \\ \delta &\lesseqgtr -q(1-2f) \end{aligned} \qquad (12.16)$$

There are three relevant cases:

- $\delta > -q(1-2f)$: The denominator is positive, so the slope of the iso-expected income curve is positive.
- $\delta < -q(1-2f)$: The denominator is negative and so the slope of the iso-expected income curve is negative.
- $\delta = -q(1-2f)$: The denominator is equal to zero, so the iso-expected income curve is vertical where it intercepts the borrower's best-response function.

We already know that this last case is a point on the borrower's best-response function because it is the same as Equation 12.14, the rule for the borrower selecting the speed of the machine:

$$\delta = -q(1-2f)$$

As the borrower's marginal rate of substitution is the negative of the slope of her iso-expected income curve, we can therefore say it is the following:

$$mrs(\delta, f) = \frac{-(1-f)}{q - 2qf + \delta}$$

The lender's expected profit maximization

The lender seeks to maximize her expected profits. But now that we have dropped the unrealistic assumption that she could enforce her chosen level of risk taking on the borrower, she bound by a tighter constraint than the participation constraint of the borrower (requiring merely that his expected income be at least zero). Because the contract is now incomplete, the lender will have to provide him with incentives to operate the machine more prudently than he otherwise would.

We show the interaction of lender and borrower in Figure 12.8. The lender is restricted to points on the borrower's best response function. Profit increases as f is less and δ is greater (that is, down and to the right). She will select the interest factor for which the borrower's best response function is tangent to her highest possible iso-profit curve, point **n** in the figure. This is the value of δ such that the marginal rate of substitution (from the principal's) iso-expected profit curve is equal to the marginal rate of transformation (from the borrower's best-response function):

$$mrs(\delta, f) = mrt(\delta, f)$$

This rule for selecting the interest factor is equivalent to finding the δ such that the marginal benefit to the lender of raising the rate of interest (more repayment if the machine does not fail) is equal to the marginal cost to the lender (increased probability of no repayment due to the faster pace of the machine). To show this equivalence, we express the mrs and mrt as the slope of the lender's iso-expected profit curve and the borrower's best response function, respectively.

REMINDER We have already (in M-Note 12.2) derived the $mrs(\delta, f)$, which is the negative of the slope of the lender's iso-expected profit curves, and the slope of the borrower's best-response function which is the negative of the marginal rate of transformation or $-mrt = \frac{\Delta f}{\Delta \delta} = -f_\delta = -\frac{1}{2q}$

$$\begin{aligned} mrs(\delta, f) = -\frac{(1-f)}{\delta} &= -\frac{\Delta f}{\Delta \delta} = mrt(\delta, f) && (12.17)\\ (1-f) &= \delta \frac{\Delta f}{\Delta \delta} \\ \text{And because} \quad \frac{\Delta f}{\Delta \delta} &= \frac{1}{2q} \\ \text{we have } (1-f) &= \frac{\delta}{2q} && (12.18)\\ \text{which is to say: Marginal benefit} &= \text{Marginal cost} \end{aligned}$$

The marginal benefit of raising the interest factor (δ) for a given level of risk is $1-f$ because this is the probability that the interest factor will be paid. The marginal cost of raising the interest factor is $\delta \frac{\Delta f}{\Delta \delta}$ because raising δ increases on the risk taken by the borrower by $\frac{\Delta f}{\Delta \delta}$, which results in a loss of δ to the lender if the machine fails.

As shown in M-Note 12.6 we can use the borrower's best-response function to eliminate f from Equation 12.18 to find the Nash equilibrium interest factor that the principal will choose:

$$\delta^N = \frac{q}{2} \tag{12.19}$$

To find out how the borrower responds to this interest factor we substitute the lender's chosen interest factor (δ^N) into (12.15). Solving for f gives us the Nash equilibrium risk level:

$$f^N = \frac{3}{4} \tag{12.20}$$

We know that the outcome indicated by point **n**, namely (f^N, δ^N) is a Nash equilibrium (hence the N superscripts) because:

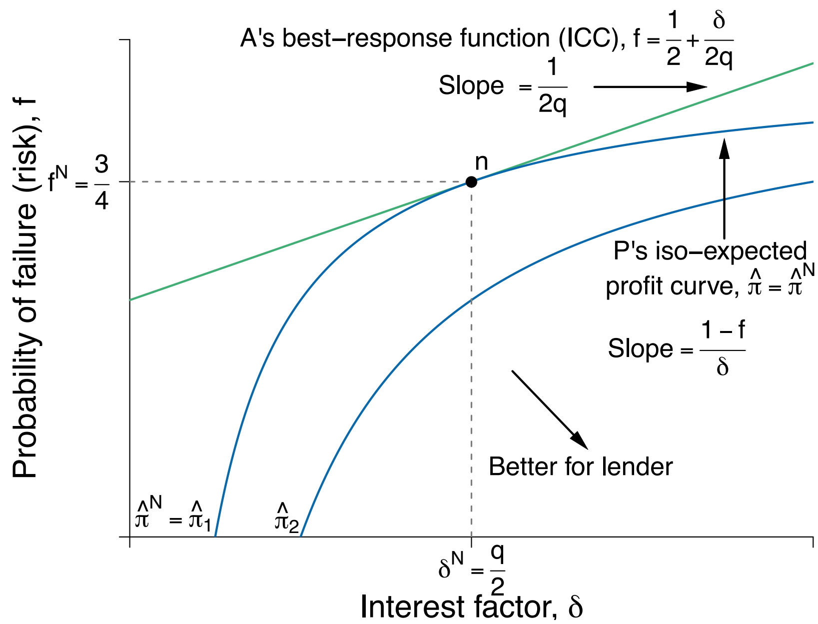


Figure 12.8: **The Nash equilibrium of the lender-borrower interaction.** The agent (borrower) best responds to the principal (lender) by choosing a level of risk (f) on his best-response function based on the interest factor (δ) he is offered by the lender. The Nash equilibrium outcome is (δ^N, f^N) at the tangency of the borrower's best response and the lender's isoexpected profit curve where the lender's $mrs(\delta, f) = mrt(\delta, f)$. This figure is similar to those in the previous two chapters showing the best response function of the agent (the sub contractor, and the worker) and the iso-expected profit curves of the principal (Benetton and Ford Motor Company). But there is a difference: here the straight line is the agent's best response function (the incentive compatibility constraint), and the curves (the iso-expected profits) represent the principal's objectives, not the the incentive compatibility constraint that limits their profit making.

- Given the interest factor that the principal has offered, namely, (δ^N), then $f^N(\delta^N)$, is the best the borrower can do; and
- Given strategy adopted by the borrower as described by his best response function $f(\delta)$, then δ^N is the best the lender can do.

Notice in the second bullet, that the lender best responds to the *strategy* of the borrower — the best response function – not to the *action* of the borrower, that is, level of risk that he takes in the Nash equilibrium f^N. If it were the case that the borrower would choose f^N whatever level of the interest factor the lender chose, then, as we will see, she would charge a much higher interest factor. But this is not the case: the risk chosen by the borrower depends on the interest factor chosen by the lender.

Checkpoint 12.4: The best-response and $mrs = mrt$

Confirm that the lender's choice of interest factor, $\delta^N = \frac{q}{2}$ by substituting the agent's best-response function: $f = \frac{1}{2} + \frac{\delta}{2q}$ into Equation 12.18.

M-Note 12.6: Nash equilibrium risk and interest factor

We used the condition for $mrs = mrt$ in equations 12.9 and 12.18 to find the Nash equilibrium value of δ. Here we do the same derivation starting from the profit maximization of the lender by differentiating the profit function of the lender with respect to δ (the variable over which the lender has control) and setting it equal to zero. The lender will maximize her expected profit, $\hat{\pi}(\delta, f)$, subject to the borrower's best-response function. To do this, we write the lender's expected profits are as in (Equation 12.4), but f now depends on δ, giving the new expected profit function (where $f_\delta = \frac{\partial f}{\partial \delta}$):

$$\begin{aligned}
& & \hat{\pi}(\delta, f) &= \delta(1 - f(\delta)) - (1+\rho) \\
&\text{First Order Condition} & \frac{d\hat{\pi}}{d\delta} &= (1 - f(\delta)) + \delta(-f_\delta) = 0 \\
&\text{Re-arranging} & \frac{1 - f(\delta)}{\delta} &= f_\delta
\end{aligned}$$

This tells us that, at the Nash equilibrium, the slope of the iso-expected profit curve is equal to the slope of the best-response function.

We can now substitute in the best response function $f(\delta) = \frac{1}{2} + \frac{\delta}{2q}$ to find the interest factor that maximizes the lender's expected profit:

$$\begin{aligned}
\frac{1 - (\frac{1}{2} + \frac{\delta}{2q})}{\delta} &= f_\delta \\
\frac{\frac{1}{2} - \frac{\delta}{2q}}{\delta} &= \frac{1}{2q} \\
\frac{1}{2} - \frac{\delta}{2q} &= \frac{\delta}{2q} \\
\frac{1}{2} &= \frac{\delta}{q} \\
\delta^N &= \frac{q}{2}
\end{aligned}$$

This is the Nash equilibrium level of interest that the lender chooses. To complete the analysis we need to find the borrower's Nash equilibrium best-response level of risk ($f(\delta^N)$). So we substitute δ^N into Equation **??**:

$$\begin{aligned}
&\text{Borrower BRF} & f(\delta) &= \frac{1}{2} + \frac{\delta}{2q} \\
&\text{Substitute in } \delta^N = \frac{q}{2} & f &= \frac{1}{2} + \frac{\frac{q}{2}}{2q} \\
& & f^N &= \frac{1}{2} + \frac{1}{4} = \frac{3}{4}
\end{aligned}$$

12.7 Characteristics of the incomplete contract Nash equilibrium

In the credit market, just as we have seen in the labor market and the market for goods of variable quality (Chapters 10 and **??**) the incompleteness of the contract between the principal and the agent leads to results quite different from what is the case in exchanges covered by complete contracts, like the bread market. These differences start with the constraint that limits the principal's expected profit maximization process. We show this in the first row of Table 12.1 and illustrate the two cases in Figure 12.9. By way of review and application of what you have learned so far, we now explain the remaining entries in the table.

Contract over interest and risk	*Incomplete*	*Complete*
Constraint for Principal (lender)	ICC	PC
Rent for the borrower?	yes	no
Pareto-efficient Nash equilibrium?	no	yes

Table 12.1: **Complete and incomplete contracts: A summary of differences in the Nash equilibrium of the credit market**. The ICC and the PC are the incentive compatibility constraint and the participation constraint, respectively. The case of the complete contract is shown in Figure 12.5 ; that of the incomplete contract is shown in Figure 12.8 and the figures in the remainder of the chapter. The two cases are contrasted in Figure 12.9.

The borrower receives a rent

To see that the borrower receives a rent we can use Figure 12.9 and observe that the Nash equilibrium (point **n**) is to the left of point **g** on his participation constraint. We know that expected income is higher to the left of the participation constraint (because the interest factor is lower) so he must expect an income greater than zero, his fallback option. So he receives a rent.

To determine how large is his rent we proceed as we did in Chapter 11, by asking: hypothetically holding constant the level of risk he is taking in the Nash equilibrium (f^N) how much higher could the interest factor be without violating his participation constraint. The answer is given by the horizontal distance between points **g** and **n**, that is $\delta_g - \delta^N$.

The principal's profits decrease in the incomplete contracting case relative to complete contracting: substituting f^N and δ^N into Equation 12.4 gives $\hat{\pi}^N = \frac{q}{8}$ rather than the expected profits of $\hat{\pi}^C = \frac{q}{4}$ in the complete contracting case.

The total rent to be divided between lender and borrower is also smaller in the incomplete contract case than it is in the complete contracting case because the borrower runs the machine at a speed faster than that which would maximize the expected total income from the project. But, even though the total rent is lower with incomplete contracting, unlike the complete contracting case, the borrower *shares* some of the rent when the contract is incomplete: the rent is not entirely taken by the lender as it was when contracts were complete.

The Nash equilibrium is Pareto-inefficient

In the complete contracting case as we have already seen, the result must be Pareto-efficient because the constraint on the lender's profit maximization is the borrower's participation constraint.

When the contract is incomplete, the borrower's best-response function not his participation constraint limits the lender's options. To see why the resulting Nash equilibrium (point **n** in Figure 12.10) must be Pareto-inefficient notice that the lender is better off at points below and to the right of her isoprofit curve $\hat{\pi}^N$. The borrower is better off at points to the left of his iso-expected income curve $\hat{y}^N$.

There is a set of points, such as point **b**, that are Pareto improvements over

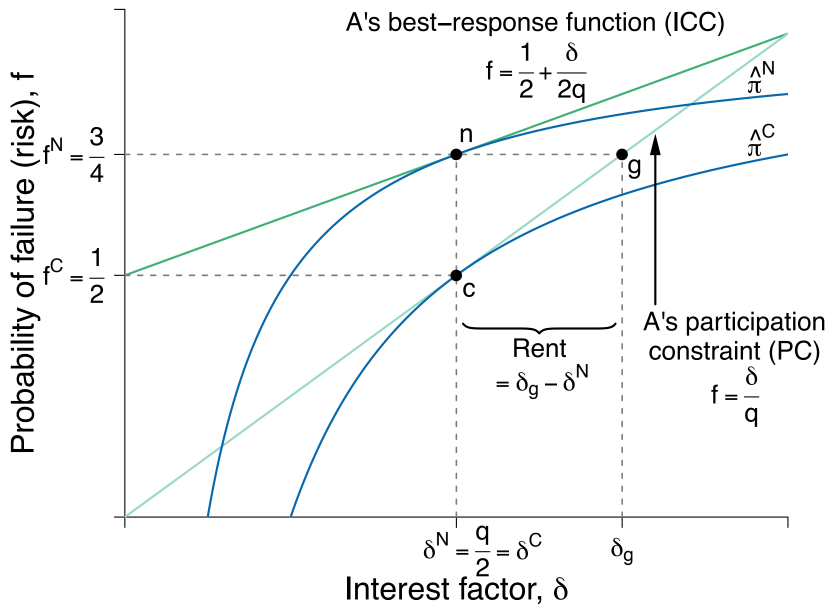


Figure 12.9: **A comparison of the complete and incomplete credit market contract outcomes.** Point **n** corresponds to the incomplete credit contract Nash equilibrium with risk $f^N = \frac{3}{4}$ and interest factor δ^N. Point **c** corresponds to the complete contract outcome with risk $f^C = \frac{1}{2}$ and interest factor δ^C.

point **n**. At point **b** in the lens, the lender would earn higher profits on $\hat{\pi}_b$ and the borrower would get higher expected income on $\hat{y}^B$ showing that point **b** is Pareto-superior to the Nash equilibrium at point **n**. The Pareto-improving lens must exist for the same reason that contractual incompleteness led to inefficiency in the equilibrium of the employee-employer interaction in Chapter 11 or the interaction of buyer and seller of the variable quality good in Chapter 10. At the Nash equilibrium, point **n** in Figure 12.10 the lender's iso-expected profit curve is tangent to the borrower's upward sloping best-response function. This means that it cannot also be tangent to the borrowers iso-expected income curve – as a Pareto efficient outcome would require – because at point **n** this is vertical (as we showed in M-Note 12.5 and in Figure 12.7).

REMINDER: PROVING PARETO INEFFICIENCY In M-Note 11.6 in Chapter 11 we used the mathematical properties of the first order conditions for the principal's and agent's constrained optimization problem to demonstrate why the Nash equilibrium in a similar case (employers and workers) is not Pareto efficient.

Checkpoint 12.5: Pareto efficiency and Pareto improvements

a. Make sure that you can explain why it is the case that one person maximizing his profits or utility subject to the participation constraint of another must result in a Pareto-efficient outcome.

b. Can you explain why the set of Pareto-improving points in the Pareto-improving lens must exist when the constraint is the best-response function? *Hint*: think about the construction of the best-response function and what you know about the slope of the borrower's iso-expected income curve at the Nash equilibrium, point **n**.

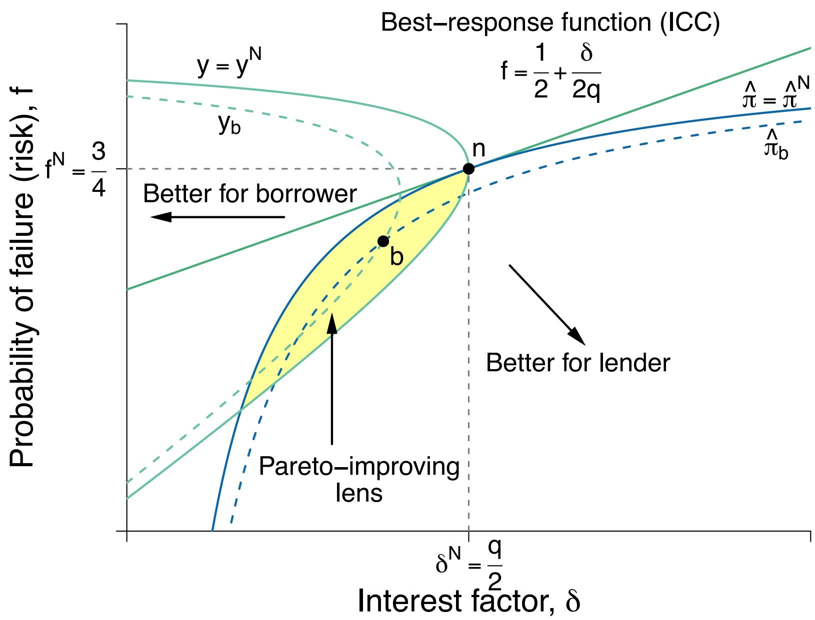


Figure 12.10: **Credit Equilibrium with non-contractual risk level and Pareto-improving lens** At the Nash equilibrium, the borrower's iso-expected income curve is $\hat{y} = \hat{y}^N$ and the lender's isoexpected profit is $\hat{\pi}^N$. The yellow-shaded area is the Pareto-improving lens. At point **b** in the lens, the lender would earn higher profits on $\hat{\pi}_b$ and the borrower would get higher expected income on $\hat{y}_b$ showing that point **b** is Pareto-superior to the Nash equilibrium at **n**.

Case: Differing rules of the game	*Borrower's PC or ICC* $f(\delta,q)$	*Risk,* f	*Interest factor,* δ	*Expected rents per period*		
				Income (borrower)	*Profit (lender)*	*Total*
Owner-operator (no loan)	–	$\frac{1}{2}$	–	–	–	$\frac{q}{4} - (1+\rho)$
Complete contract	PC: $f = \frac{\delta}{q}$	$\frac{1}{2}$	$\frac{q}{2}$	0	$\frac{q}{4} - (1+\rho)$	$\frac{q}{4} - (1+\rho)$
Incomplete contract	ICC: $f = \frac{1}{2} + \frac{\delta}{2q}$	$\frac{3}{4}$	$\frac{q}{2}$	$\frac{q}{16}$	$\frac{q}{8} - (1+\rho)$	$\frac{3q}{16} - (1+\rho)$

Table 12.2: **Nash equilibrium outcomes in the borrower lender model depend on the rules of the game: ownership and the nature of the contract.** Remember, for the lender's profit maximization process, the PC is the participation constraint for the borrower and ICC is the borrower's incentive compatibility constraint. The fact that the interest factor is the same under a complete and an incomplete contract is an artifact of the particular model we are using: the interest factor under the complete contract could be either higher or lower than under the incomplete contract.

c. Explain each of the three differences between the complete and incomplete contracting case in Table 12.1 and why the difference in the contract produces this difference in outcome.

The rules of the game matter: Total rents and their distribution

Table 12.2 summarizes the cases addressed so far, and demonstrates how the three different rules of the game – owner-operator, borrower-lender with a complete contract, and borrower-lender with an incomplete contract – support different results. Two points are important:

- *Distribution of the rents*: Both the owner who operates the machine without any other player involved, and the lender having a complete contract with a borrower get all of the rents (last column, first two rows). The borrower gets his fallback option (which is zero). But in the incomplete contract case (the third row), the borrower receives a third of the total rents.

- *Total gains from exchange*: The owner operator case and the complete contract maximize the total rents available to the participants, meaning the output of the machine minus the opportunity cost of the funds used. This is because the person making the decision about the risk taken – the owner operator or the lender – is the residual claimant on the income of the project. Because they get all of the rents (the previous bullet) they 'own' the consequences of their decisions. Comparing in the final column the third row with the first two, the total gains from the exchange are twenty five percent less when the contact is incomplete. This is because the borrower does not own all of the consequences of his risk taking, and as a result takes 50 per cent more risk.

12.8 Many lenders: Competition and barriers to entry

While an urban neighborhood may have a single pay-day lender and many small towns just a single bank or money-lender, most prospective borrowers – whether individuals or firms – can shop around. Many people, as we saw, use their credit cards as a source of loans.

We have so far represented Parama as the sole lender from whom Antonio can borrow. We now introduce additional lenders competing with Parama in the credit market. To do this we embed the principal-agent model of the lender and borrower in a model of the entire market.

How many competitors there will be depends on the barriers to entry in the credit market and the opportunity cost of capital. As in Chapter 12, we take the perspective of a wealthy individual or firm considering entering the credit market. The entrant will consider the profits that they can expect to make if they successfully enter the market as well as the probability b that they will fail in their attempt, due to the barriers to entry. There are two categories of risk that the entrant faces: it may *fail* to transact any loans, or it may *succeed* in entering the credit market and then make a loan that is not repaid.

FACT CHECK Barriers to entry in the credit market are anything that makes it likely that an attempted entry into the market – by a new credit card company, for example – will fail. Getting a license to operate as a bank, for example, requires meeting a set of demanding standards, often including a substantial level of assets. Entering the credit card lending market faces a challenging chicken and egg problem: merchants will not honor cards that few customers have, but customers will not have a card that few merchants recognize.

Suppose the entrant considers devoting some amount to their entry attempt, say, a million Euros. Then their expected profits, in millions of Euros, are:

- *if they fail to enter*, which occurs with probability b, they lose $1 + \rho$, which is the opportunity cost of the assets they devoted to their project
- *if they succeed in entering which occurs with probability* $(1 - b)$ their expected revenues minus opportunity cost of capital is the same as the

lenders already in the market, called the *incumbent firms*:

$$\text{Incumbents' expected profits} \quad \hat{\pi} = \delta(1-f)-(1+\rho) \quad (12.21)$$

where $(1-f)$ is the probability their loan is repaid.

They pay the opportunity costs of their entry attempt $1+\rho$ with certainty (whether they fail or succeed), while they gain receipts only if they do not fail. So the costs are *known* and the revenues are *uncertain*.

They will decide to attempt to enter as long as their expected revenues do not fall short of costs – or so that their expected profits are greater than zero:

$$\begin{aligned} \text{Entrant's expected profits} \quad \hat{\pi}^b &= \overbrace{(1-b)[\delta(1-f)-(1+\rho)]}^{success} - \overbrace{b(1+\rho)}^{failure} \geq 0 \\ \hat{\pi}^b &= (1-b)\delta(1-f)-(1+\rho) \geq 0 \end{aligned}$$

The b superscript is a reminder that this is a possible entrant in the market facing barriers to entry. If this condition is not satisfied, firms will not enter the credit market. And because there is always some exit of firms for reasons outside our model, if:

$$\hat{\pi}^b = (1-b)\delta(1-f)-(1+\rho) \leq 0$$

then, no new firms will enter and the number of firms in the credit market will fall. The competition condition ensuring that the number of firms in the credit market to be constant is therefore:

$$\text{Competition condition} \quad \hat{\pi}^b = (1-b)\delta(1-f)-(1+\rho) = 0 \quad (12.22)$$

Equation 12.22 is called the *credit market competition condition*. It is analogous to the competition condition for the whole economy based on the entry and exit of firms in the market for goods and services (in Chapter 11). In this chapter where there is no danger of confusing the two we will call it the competition condition.

As you can see in the M-Check, we an rearrange Equation 12.22 to show that when there are barriers to entry ($b > 0$) the incumbent firms will make positive economic profits in the credit market equilibrium satisfying the competition condition, or:

$$\begin{aligned} \hat{\pi} &= \frac{(1+\rho)}{1-b} - (1+\rho) \\ &= (1+\rho)\left(\frac{1}{1-b} - 1\right) \end{aligned} \quad (12.23)$$

Equation 12.23 is greater than zero if $b > 0$.

M-CHECK To get an expression for incumbent firms' profits when the competition condition is satisfied we substitute Equation 12.22 (that is, $\delta(1-f) = \frac{(1+\rho)}{1-b}$) into Equation 12.21:

$$\begin{aligned} \hat{\pi} &= \frac{(1+\rho)}{1-b} - (1+\rho) \\ &= (1+\rho)(\frac{1}{1-b} - 1) \end{aligned}$$

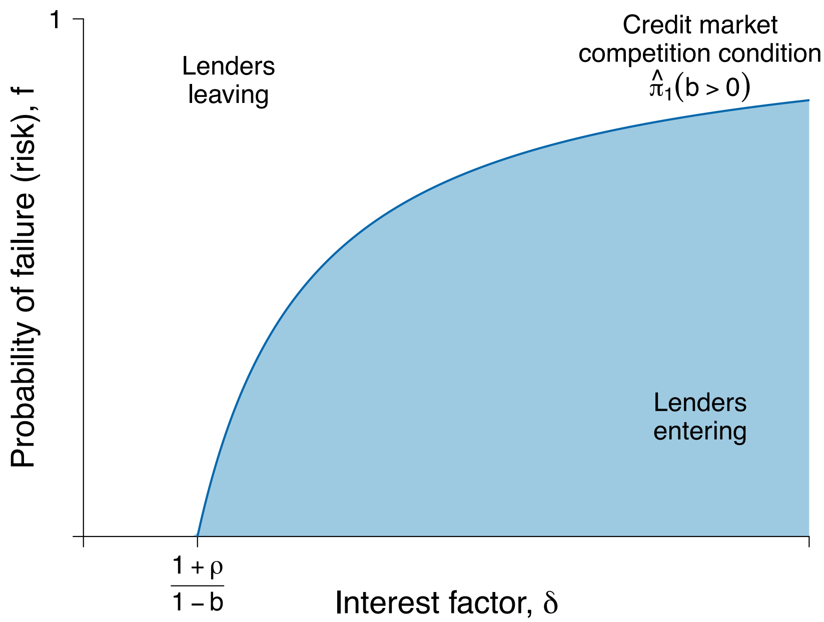


Figure 12.11: **The credit market competition condition for some given level of barriers to entry, b.** The curve is composed of values of the interest factor δ and the level of risk f that satisfy the equation: $\delta(1-f) = \frac{(1+\rho)}{1-b}$. This is the level of expected revenues of incumbent firms such that the expected profit of a prospective new entrant to the market is zero. As you can see from the horizontal axis intercept, lesser barriers to entry and greater competition among lenders would be represented by a competition condition above and to the right of the one shown, as illustrated later in this chapter in Figure 12.17. We have assumed a level of barriers to entry of $b = 0.3$.

If there are no barriers to entry so $b = 0$ – the case we call unlimited competition – the competition condition becomes $\delta(1-f) - (1+\rho) = 0$ which is called the zero profit condition because in this case the expected profits of the entering and incumbent firms are equal and both equal to $\hat{\pi}^b = \hat{\pi} = 0$. You have already seen this in Figure 12.4 where the iso-expected-profits curve labeled $\hat{\pi}_0$ is the competition condition for the case with unlimited competition (so that expected economic profits are zero).

In Figure 12.11 we show the competition condition. This is the iso-expected-profits curve for incumbent firms yielding sufficient profits so that the number of firms in the market does not change. The competition condition divides the space of credit market outcomes into:

- outcomes with positive expected profits for an entering firm so that firms would enter and
- outcomes with negative expected profits firms would not enter and the total number of firms in the market would fall.

Credit market equilibrium with barriers to entry

An equilibrium in the credit market requires that the number of lenders (or the total amount of lending, if lenders are differing in size) be unchanging. So the outcome must be somewhere along the competition condition curve in Figure 12.11. But where? To answer this question we need another condition – some additional information about the relationship between δ and f – that

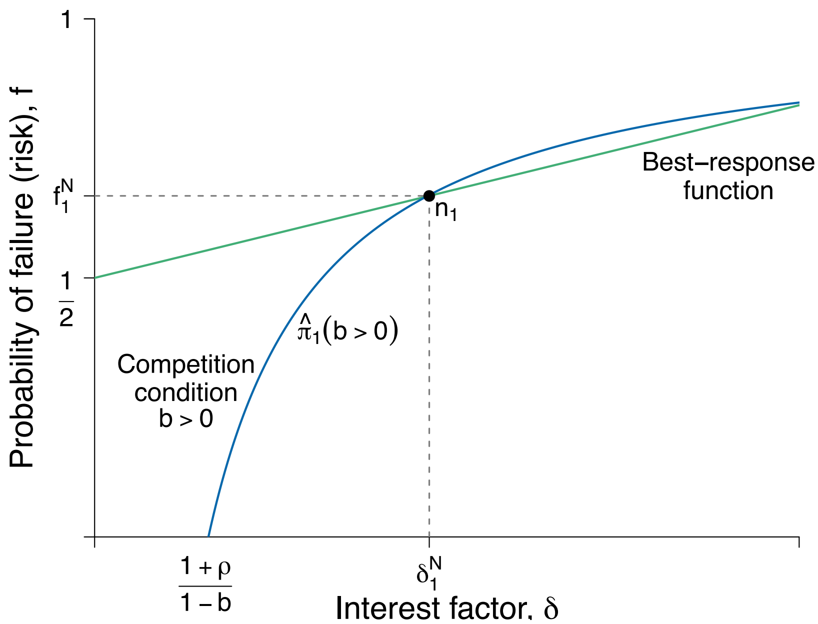


Figure 12.12: **Equilibrium in the credit market give by the borrower's best response function and a competition condition.** For this illustration we used $b = 0.3, q = 10$ and $\rho = 0.05$)

we know must be true. This is provided by the borrower's best-response function.

The interest factor and risk level (δ^N, f^N) in the Nash equilibrium of the credit market as a whole – not just a single lender interacting with a single borrower, like Parama and Antonio – requires two things:

- *Incentive compatibility.* As before, the level of risk taken by borrowers must be incentive compatible given the interest factor charged by lenders, that is, it must be on their best response function, and
- *Competition condition.* And now, we further require that the number of firms in the credit market must be constant, so Equation 12.22 must be satisfied.

In Figure 12.12 we show these two equations and their intersection. You can see that this outcome satisfies both the competition condition (in blue) and the incentive compatibility constraint (in green). But are the lenders maximizing their expected profits?

To explore this, imagine that you are Parama or another lender, charging an interest factor of δ_1^N, with Antonio the borrower taking risk f_1^N. This outcome is at a point where your iso-expected profit curve (the competition condition itself) is not tangent to the borrower's best response function. So you might think that by charging a higher interest factor you could move to a higher iso-expected profit curve (to the right of n_1 on the best response function, not shown in the figure).

But competition in the credit market (even with barriers to entry) has changed the game. Antonio's fallback option is no longer to get zero as it was in his one-on-one interaction with Parama, but instead to borrow funds from another lender. And if Parama did raise the interest factor there definitely would be an alternative lender ready to lend to Antonio, because we know that there is a higher interest factor consistent with Antonio's best response function that would raise her profits above the level dictated by the competition condition. So she could not charge more than δ_0^N and maintain her borrowers. This is another example of the analysis in Chapter 9: competition works by changing the players' fallback options.

M-CHECK You may wonder about the second intersection of the two curves in Figure 12.12. If the outcome were at that point (at the far right of the figure, with no letter marking it) then any lender could increase their expected profits by lowering their interest factor moving along the borrower's best response function to the left, which you can see would be on a higher iso-profit curve (not shown). This would draw in additional lenders, eventually driving the interest factor down to δ_1^N.

WEALTH The stock of things (or the value of that stock) owned by a person or household that contribute to a flow of income or other benefits, is their wealth.

12.9 Wealth matters: Borrowing with equity

You know from the introduction to this chapter that a common practice of lenders is to require that the borrower provide some asset called collateral, the ownership of which will be transferred to the lender if the borrower does not repay. Collateral requirements are common when loans are provided for purchasing a home or a car (the home or car itself is the collateral asset). Those with limited wealth often cannot provide collateral for anything but a house or car loan, and for this reason have difficulty securing credit for other purposes such as starting up a business, or retraining to gain new skills.

Equity and collateral

An alternative arrangement, more common when the loan is to start or expand a business, is for the borrower to share in the risk of the project by investing some of his own **wealth** in the project. One's own wealth invested in a project is called equity. Lending to a borrower who has invested his own wealth in a project protects the lender to some degree from the borrower recklessly risking the funds.

Borrowers generally have some wealth, and if the expected income of the project is greater than the income from the individual's next best alternative, it may be in the borrower's interest to invest equity in the project or to provide collateral. This will alter the borrowing game between the lender and the borrower.

First, if, contrary to our assumption, the lender does not know q, investment of the borrower's own wealth is a credible signal of the borrower's assessment of quality of the project, a *hidden attribute*. As we will see presently, in competitive equilibrium those with less wealth will need superior projects to obtain financing, so the borrower has an interest in overstating a project's quality in order to secure a loan.

Knowing this, the lender would rather lend to a borrower who has risked his own assets by providing equity or collateral, which is a promising sign to the

EXAMPLE: WEALTH Included in wealth are the value of a home, car, any land, buildings, machinery or other capital goods, and any financial assets such as shares or bonds. Debts are subtracted – for example, the mortgage owed to the bank. Debts owed to the person are added. A broader definition of wealth includes what is termed human capital, meaning the individual's skills, connections and other capacities that contribute to a flow of income. We will use the term *wealth* to refer to assets that may be used as collateral or equity.

INCOME is the maximum a household or person can consume over a given period of time without reducing the value of the household's wealth (their stock of assets such as home or car, minus any outstanding debt). The borrower's income from the project is a part of their income.

EQUITY is one's own wealth (rather than borrowed funds) invested in a project. There is a second entirely different use of the term, meaning the character of being fair, as in "an equitable division of the pie."

lender about the borrower's assessment of the quality of the project.

The second reason, and the one modeled here, is that the discrepancy between the objectives of the lender and borrower concerning the choice of the level of risk (this is the *hidden action*) would be reduced if the borrower invested in the project and thus shared some of the risk of failure with the lender. Investing equity or posting collateral does not make the loan contract complete, but we shall see that equity reduces the degree of conflict of interest between the lender and the borrower. It does this because the borrower now "owns" some of the consequences of his choice of risk.

To model this case, we begin with the dyadic (meaning one-on-one) lender borrower interaction in which the borrower has invested some equity in the project before studying the credit market as a whole. As shown in M-Note 12.9, the analysis of a borrower posting collateral is qualitatively similar to the case when the borrower has equity (the main difference being that collateral shifts the best-response of a potential borrower down without changing its slope), and we will follow only the equity case from this point on.

Best response level of risk by a borrower with equity

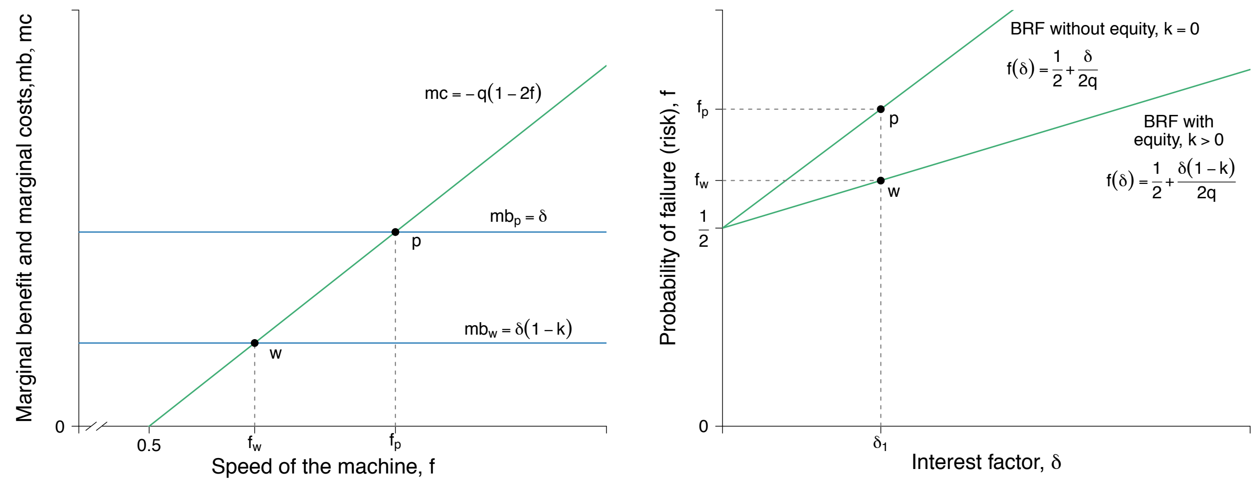
The borrower, Antonio, now has wealth k. Other than this, the structure of the game between him and Parama is unchanged. His assets, worth k, could be invested in some alternative project (possibly a government bond) that after one period (year or other time unit) will with certainty yield him $(1+\rho)k$ in income. If Antonio invested these funds as equity in acquiring the "machine", he would then borrow the amount remaining to fund the \$1 project, that is, $1-k$. His expected returns (including the opportunity cost of the foregone returns on the alternative asset) would be:

$$\underbrace{\hat{y}(\delta,f)}_{\text{Expected income}} = \underbrace{qf(1-f)}_{\text{Expected revenues}} - \underbrace{\delta(1-f)(1-k)}_{\text{Expected repayment}} - \underbrace{(1+\rho)k}_{\text{Opportunity cost of capital}} \tag{12.24}$$

On the right hand side, the expected revenues are the same as before the borrower invested in the project. But Equation 12.24 differs from Equation 12.13 — the case of the borrower with no wealth invested in the project — in two ways, both due to the fact that Antonio now has some of his own capital invested as equity:

- There is an opportunity cost of investing in this project (the last term on the right) and
- He has borrowed only $(1-k)$ so he now will repay Parama the lender $\delta(1-k)$ if the machine does not fail, rather than δ which was the case previously.

In what follows we use the terms wealth and level of equity committed to the



(a) Marginal benefits and costs of increased risk with and without equity. (b) The borrower's best-response functions with and without equity (k).

Figure 12.13: **The choice of a risk level when the borrower has equity.** Marginal costs rise as the borrower takes on more risk (f). For a given level of borrower's equity k and an interest factor (δ_1) **determined by the lender**, the borrower's expected income maximizing level of risk, f, will be lower when the borrower has equity. Comparing points **p** and **w**, Antonio's chosen level of risk is lower when he is wealthy and has equity $k > 0$ with ($f_w = f(\delta(1-k))$ than when he is poor and does *not* have equity ($f_p = f(\delta)$). In panel **b**, the borrower's best-response function with equity is similar to the best response without equity except that its slope is now $\frac{1-k}{2q}$ and so is lower (flatter) than when $k = 0$. Points **p** and **w** correspond to the choices of a poor and wealthy borrower respectively for a given interest factor (δ_1).

project interchangeably: agents devote all their wealth to the project, if they devote any.

M-Note 12.7: A borrower with equity selects the best-response level of risk

We start with Equation 12.24.

$$\begin{aligned} \hat{y}(f,\delta,k) &= qf(1-f) - \delta(1-k)(1-f) - (1+\rho)k \\ &= qf - qf^2 - \delta(1-k)(1-f) - (1+\rho)k \end{aligned}$$

The borrower differentiates his expected income with respect to f to maximize his income.

$$\begin{aligned} \frac{\partial \hat{y}}{\partial f} &= q - 2qf + \delta(1-k) = 0 \\ \delta(1-k) &= -q(1-2f) \end{aligned} \tag{12.25}$$

Equation 12.25 is the borrower's first order condition to maximize his expected income.

The fact that Antonio has an alternative use of his funds does not affect the speed at which he will run the machine. But the fact that he now has less to pay back if the machine does not fail changes his incentives. Here is why. As before, Antonio, the borrower will select f to maximize $\hat{y}$ by equating the marginal benefits and marginal costs of operating the machine faster. But these are now (as shown in M-Note 12.7):

$$\begin{aligned} \text{Marginal benefit} &= \text{Marginal cost} \\ \delta(1-k) &= -q(1-2f) \end{aligned} \tag{12.26}$$

Comparing this rule for the borrower selecting f with the same rule where equity was absent (that is $\delta = -q(1-2f)$) we can see that the right-hand side of Equation 12.26 is identical to when the borrower had no equity.

But the marginal benefit of increasing the speed of the machine is now less. Remember, the reason why running the machine faster is a benefit is not that the expected number of goods that it will produce will be greater. To see this, remember that $f = \frac{1}{2}$ is the speed which maximizes the income of the project, as was chosen by the owner-operator. If $f > \frac{1}{2}$ running it faster reduces the expected revenues, this is the marginal *cost*. The benefit is that running it faster may lead it to fail, in which case the loan will not be repaid. Because the borrower has his own equity in the project the loan to be repaid is less, so the benefit of running the machine faster is also less.

This is why the borrower's equity in the project reduces the conflict of interest between the borrower and the lender.

The marginal benefit and marginal cost of taking on greater risk are depicted in Figure 12.13 in which we compare two cases. In one Antonio (the borrower) is poor and does not have any wealth to use as equity and the second in which he has wealth ($k < 1$) which he invests in the project. The figure shows that, when he has wealth, Antonio's marginal benefits of running the machine faster are lower. Because of this, for a given interest factor δ_k, he will run the machine more slowly at f_w (he selects point **w**). When he is poor, if the lender selects the same interest factor, he will select point **p** and a higher risk level at f_p.

M-CHECK: DERIVING THE BRF WITH EQUITY

$$\begin{aligned} \underbrace{\delta(1-k)}_{mb} &= \underbrace{-q(1-2f)}_{mc} \\ \text{Divide through by } q \quad \frac{\delta(1-k)}{q} &= 2f - 1 \\ \text{Isolate } f \quad 2f &= 1 + \frac{\delta(1-k)}{q} \\ \text{Divide by 2} \quad f(\delta) &= \frac{1}{2} + \frac{\delta(1-k)}{2q} \end{aligned}$$

Rearranging Equation 12.26 – Antonio's rule for selecting f – we have his best-response function, shown in Figure 12.13:

$$\text{Best response with equity} \qquad f(\delta, k) = \frac{1}{2} + \frac{\delta(1-k)}{2q} \tag{12.27}$$

Equation 12.27 is identical to the best response function where there is no equity except for the $(1-k)$ term (as you can see by setting $k = 0$.). Four things to notice about this best response function for the borrower with equity:

- A higher interest factor is still associated with greater risk taking unless the "agent" has financed the project entirely with his own funds ($k = 1$): the best response function is upward sloping. In this case (called "complete equity financing") there is no principal. He borrows nothing.
- as $k \to 1, f \to \frac{1}{2}$: for levels of equity close to complete equity financing ($k = 1$) he will approximate the prudent expected income maximizing risk choice of owner-operator ($f = 0.5$).
- For higher levels of equity of the borrower the best-response risk level for any given interest factor is less (comparing points **p** and **w** in the two panels of Figure 12.13). This is because the marginal benefit of increasing risk is smaller.
- The borrower's best-response risk becomes less sensitive to the interest factor as his level of equity increases: its slope is now $\frac{1-k}{2q}$ rather than just $\frac{1}{2q}$. So, the best-response function is flatter.

12.10 Excluded and credit-constrained borrowers

Using this new best response function so as to take account of the equity that the borrower has invested in the project, we now consider the Nash equilibrium of the entire credit market. Here prospective borrowers differ in the amount of wealth they are able to invest as equity in the project.

Remember for an outcome to be a Nash equilibrium, it must be a point on both:

- the borrower's *best response function* which will vary with the amount of equity he has invested k, and
- the *competition condition*, which will vary with the degree of barriers to prospective lender seeking to enter the credit market, b.

Credit-constrained and excluded would-be borrowers

MARGINAL BORROWER AND EXCLUDED BORROWER A *marginal* borrower is a borrower with just enough wealth to be included in a credit contract with a lender. An *excluded* borrower is a borrower with insufficient wealth to obtain a credit contract with a lender and is therefore unable to obtain credit to fund a project.

We illustrate the market in Figure 12.14. The competition condition with barriers to entry $(b > 0)$is the iso-expected-profit curve labeled $\hat{\pi}_1$. Also shown are three best response functions for borrowers with differing amounts of wealth to invest in their project. The Nash equilibrium for the market is the following:

- *Limited wealth, excluded borrower.* The best response function of those with limited equity (k_1) lie wholly above the competition condition. If lenders extended loans to these borrowers, the expected profits would be negative. As a result, borrowers with wealth k_1 are unable to borrow. They are the *credit market-excluded* or *excluded borrowers*.
- *Modest wealth, credit-constrained marginal borrower*. The borrower whose wealth, k_2, is just enough so that his best-response function is tangent to the competition condition at point **e** is called the *marginal borrower*. This is because he has exactly enough wealth to secure a loan. The interest factor and risk level for the marginal borrower will be $(\delta_2, f(\delta_2))$.
- *Wealthy, credit-constrained borrower paying a lower interest factor*. For the borrower with wealth $k_3 > k_2$ Nash equilibrium is at point **a**, where you can see that he pays a lower rate of interest than the marginal borrower.

Wealth constraints on quality and size of projects

In addition to paying lower rates of interest, borrowers with more wealth will be able to finance larger projects and projects of lower quality. To see this take the lender's eye perspective on borrowers: which would you rather interact with? The simple answer is you prefer borrowers who will not increase their risk taking by very much if you charge a higher interest factor. You would like to find borrowers with flatter best response functions. That is why lenders

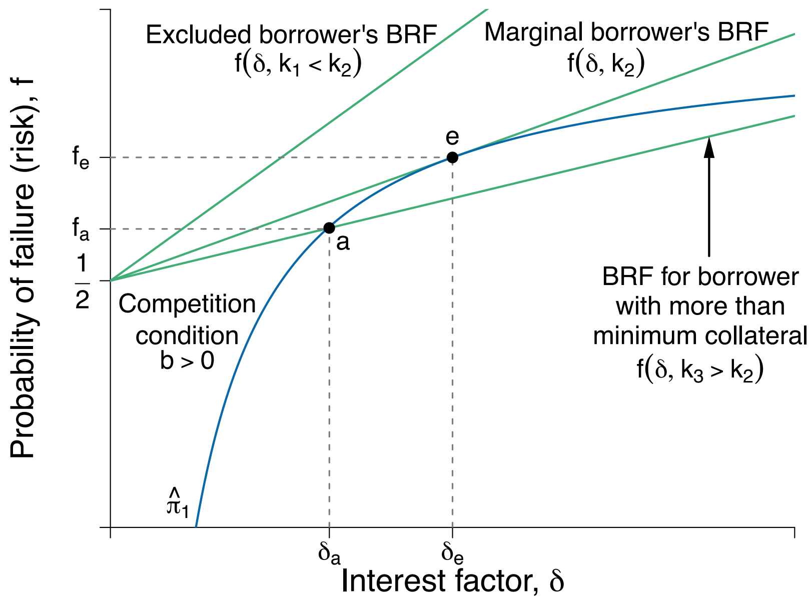


Figure 12.14: **Credit market exclusion and credit constraints.** The competition condition Equation 12.22 is: $\delta(1-f) = \frac{(1+\rho)}{1-b}$ Potential borrowers with $k < k_2$ are unable to find a lender willing to lend to them. Those with more wealth $(k_3 > k_2)$ pay a lower interest factor.

will not transact in any way with the excluded borrowers (their best response functions are too steep).

Now recall that for the borrower who has invested k in the project whose quality is q we have

$$\text{Best response with equity} \qquad f(\delta, k) = \frac{1}{2} + \frac{\delta(1-k)}{2q}$$

$$\text{and, slope of the BRF} = \frac{(1-k)}{2q} = \frac{\text{amount of the loan}}{\text{twice the quality of the machine}}$$

Two conclusions follow:

- *Wealthier borrowers can finance inferior project*s: a wealthier person (higher k) can have a flatter best response function and therefore receive a loan, even if the quality of the project q is less than the quality of a borrower excluded because he has too little wealth.

- *Wealthier borrowers can finance larger projects*: the slope of the best response function depends on the size of the loan, that is, the difference between the project size and the amount the borrower invests. So putting a given amount of equity into the project allows the wealthy borrower to increase the size of the project by the same amount, without affecting the slope of the best response function.

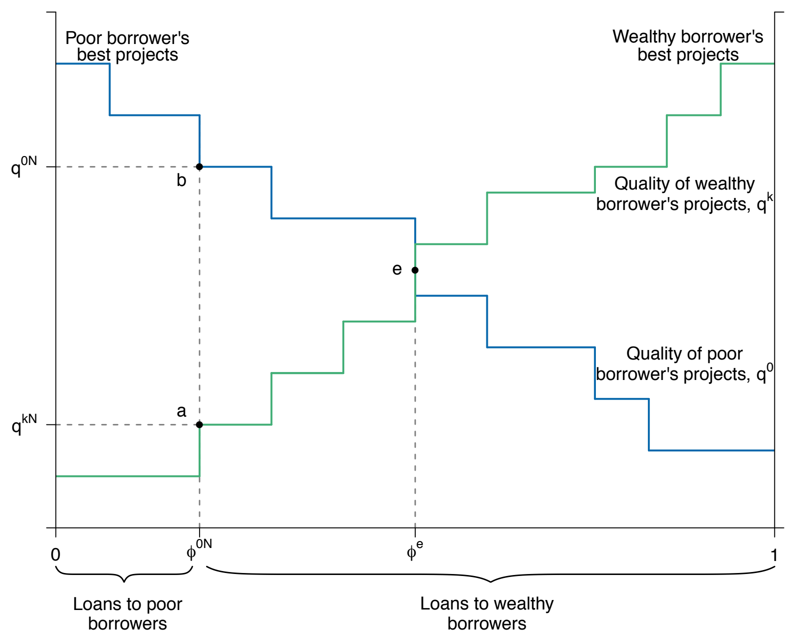


Figure 12.15: **Efficiency losses due to borrower wealth differences.** A total amount of credit (the length of the horizontal axis) is allocated between a borrower with no wealth (superscript 0) and a borrower with substantial wealth (superscript k)invested as equity in various projects wealth. An fraction of the total funds loaned ϕ^{0N} is allocated to finance the projects of the poor. The remainder $1-\phi^{0N}$ finances the projects of the wealthy borrower. The vertical axis measures the quality of the projects. The two step functions show that other than the size of the projects, the both rich and poor have a similar distribution of quality of projects. This need not be the case, and it is unimportant in what follows. The average quality of the projects funded will be maximized if the best excluded project of the poor borrower is worse than the worst included project of the wealthy borrower, and the best excluded project of the wealthy borrower is worse than the worst included project of the poor borrower. This is not true at the distribution given by ϕ^{0N} because the quality of the worst included project of the wealthy borrower ^{kN} is less than the quality of the best excluded project of the poor borrower q^{0N}. The only distribution of funds for which the average quality of project maximizing condition is true is ϕ^e.

12.11 *Why redistributing wealth may enhance efficiency*

Discriminating among potential borrowers in this fashion of course cannot be efficient, as there will be some poor borrowers with good projects that will not be carried out because they are excluded from the credit market, while some wealthier borrowers will obtain financing to carry out inferior projects.

To see the consequences for the efficient use of resources at the societal level, suppose that some given total amount of finance is available, normalized to one, to be divided among projects operated by a wealthy and a not wealthy borrower (these could represent many borrowers of each type). For simplicity we assume that the poor borrower has no wealth, and that the wealthy borrower will invest equity equal to one half the size of the project in any project for which she is granted a loan. We order the projects of each from the best (highest q) to the worst (lowest q).

Figure 12.15 shows the array of projects of the two, the poor borrower's best project on the left with projects of lesser quality arrayed in the step function descending to the right. The wealthy borrower's best project on the right, and his projects of successively lower quality are shown by the step function descending to the left. The height of each step function is the quality of the particular project in question. The horizontal width of the step is the size of the project for the poor borrower and half the size of the project for the rich borrower (because he borrows only half of the project size, providing the rest with his own investment of equity.)

In the figure ϕ^{0N} is the fraction of total loans received by the poor borrower (with $(1-\phi^{0N})$ the fraction going to the wealthy borrower). We use the N superscript because we know from the previous demonstration that, in the Nash equilibrium, the wealthy borrower will succeed in obtaining financing (a loan) for projects that are of lower quality than the best excluded project of the poor borrower.

To see why this is inefficient, think about the consequences of hypothetically shifting some of the loan funds from the rich to the poor, so that the rich would not be able carry out his worst included project, and the poor would be able to carry out his best excluded project. That would replace an inferior rich person's project with a superior poor person's project, as you can see from the figure. This would increase the average quality of the projects funded because $q^{0N} > q^{kN}$.

You can also see that the same would be true of further redistribution of finance towards the poor person's project, until the poor persons projects received a fraction ϕ^e of the total. We called the redistribution of of loans hypothetical, but it could be accomplished by a redistribution of the wealth of the two borrowers, so that they had equal assets. Their projects would then be treated equally in the credit market and the poor borrower would receive ϕ^e of the funds. The average quality of the projects funded would increase.

But there is a second reason why a redistribution of wealth could enhance efficiency, even in the absence of any change in the quality of the projects. If a poor former borrower were to become wealthy enough to finance the project herself, then she would "own" all of the consequences of her own risk taking. As a result she would implement the expected income-maximizing risk level.

To see this, think of a particular borrower whose project has a quality q seeking a loan from a wealthy individual to finance the project. Suppose $q = 8(1+\rho)$, then in a credit contract like the ones we have studied, the two would transact as is shown in the top "before" line of Table 12.3. The lender's expected profit ($\frac{q}{8}-1$) is just equal to the risk-free rate of return ρ, while the borrower's expected income is $\frac{q}{16} = \frac{1+\rho}{2}$.

REMINDER The residual claimant is the person who owns what is left over after all of the contractually obligated costs are paid. In a firm, the residual claimants are the owners who own the revenues of the firm minus the costs of labor, materials, management taxes and other inputs have been paid. If a farmer pays a fixed rent and she farms and owns the crop that she harvests, then she is the *residual claimant* on the value of the harvest minus the rent and other costs incurred. She also owns the loss, if the harvest is worth less than the costs.

Now imagine that instead of the poorer person borrowing from the lender, instead , the government confiscates the assets required by the project from its wealthy lender and gives this amount ($1) to the poor former borrower, who then operates the project as *residual claimant* at the owner-operator level of risk, that is, $f = \frac{1}{2}$. As a result the former borrower now has an income equal to $\frac{q}{4}$ or $2(1+\rho)$.

But at the same time, the government imposes a tax obligation on the beneficiary of this redistribution (the previously poor borrower), requiring him to pay $1+\rho$ at the end of the period. The tax must be paid irrespective of whether the machine fails. (The borrower might have to sell his car or mortgage his

	Total rents	*Owner's Income*	*Operator's Income*
Before	$\frac{3q}{16} = \frac{3(1+\rho)}{2}$	$\frac{q}{8} = \frac{2q}{16} = 1+\rho$	$\frac{q}{16} = \frac{(1+\rho)}{2}$
After	$\frac{q}{4} = \frac{4q}{16} = 2(1+\rho)$	$\frac{q}{8} = \frac{2q}{16} = 1+\rho$	$\frac{q}{8} = \frac{2q}{16} = 1+\rho$

Table 12.3: **Comparison of levels of income before and after re-distribution for an owner of a machine and an operator of a machine.** The "Before" line is the row in Table 12.2 labeled Incomplete contract.' The borrower has no wealth. The "After" row shows the result of redistributing wealth equal $1 to the erstwhile borrower, taxing the borrower an amount $1+\rho$ and transferring these tax revenues to the former owner of the asset, as compensation for the confiscation if her asset. We illustrate this case assuming that $q = 8(1+\rho)$.

home in order to do this.)

The government then transfers these tax revenues to compensate the former owner for her loss, so she now has income $1+\rho$, the same as before. The beneficiary of the redistribution after paying the tax has an expected amount of $1+\rho$ for himself, and is therefore better off. (Recall he made only half this amount as a borrower, *without* the government intervention.) So the former lender is as well off as before, and the former borrower is better off.

There is nothing special about the numbers: all that is required to make a Pareto-improvement possible is that total rents are larger when the former borrow becomes rich enough to run the project as the owner-operator.

By extracting from the beneficiary the tax sufficient to pay compensation to the former owner irrespective of the fate of the project, the government was able to offer the equivalent of an enforceable loan contract to the beneficiary at the risk-free interest rate. What the asset transfer plus the tax accomplishes is to make the owner-operator of the project the residual claimant on all of the risk entailed by his choices (rather than being shielded from risk by the unenforceability of the promise to repay the loan).

The key to the success of the redistribution is that private transactions are governed by limited liability and bankruptcy laws that protect the borrower from risk by placing his other assets (car or home in the above example) beyond the reach of the lender seeking to enforce repayment. The obligation to pay the government's tax is not limited by these provisions. This accounts for the superiority of the owner-operator case, and allows for the Pareto-improving redistribution.

REMINDER In Chapter 13 we shall introduce the idea of risk aversion. When a person is risk averse, she dislikes risk even if it might mean higher expected income.

What this policy application shows is that redistributing wealth from richer to poorer people may allow a Pareto improvement, so that the poor benefit and the rich do not lose. But so far we have left out an important reason why redistribution of this type sometimes fail: the poor are probably more *risk averse* than the rich. They place a higher value than the rich on reducing the risk to which they are exposed. This being the case they may prefer lower expected returns on less risky projects. We return to the question of risk aversion and wealth or income redistribution to the less well off in Chapters 13 and 15.

Checkpoint 12.6

Imagine explaining the credit model to a friend studying economics who has not

Case: Differing rules of the game	*Borrower's PC or ICC* $f(\delta,q)$	*Risk,* f	*Interest factor,* δ	*Expected rents per period*		
				Income (borrower)	*Profit (lender)*	*Total*
Incomplete contract (no competition)	ICC: $f=\frac{1}{2}+\frac{\delta}{2q}$	$\frac{3}{4}$	$\frac{q}{2}$	$\frac{q}{16}$	$\frac{q}{8}-(1+\rho)$	$\frac{3q}{16}-(1+\rho)$
Incomplete contract (no competition): *Example*	ICC: $f=\frac{1}{2}+\frac{\delta}{2q}$	0.75	5	0.625	0.2	0.825
Incomplete contract (unlimited competition): *Example*	ICC: $f=\frac{1}{2}+\frac{\delta}{2q}$	0.65	3	1.225	0	1.225

Table 12.4: **The difference that credit market competition makes: A numerical example** This example is based on Table 12.2. For the numerical examples we used $q=10, \rho=0.05$ and for the last line $b=0$.

yet read this chapter. Think about the following questions:

a. If there was an alternative outcome that would have benefited both the owner of the machine and the operator, why did they not simply make deal to implement the mutually preferred outcome?

b. Why was the alternative outcome accomplished by a government when it did not happen by means of private exchange?

12.12 Competition, barriers to entry and the distribution of rents

We turn now from the effect of differing levels of wealth among borrowers and the effects of redistribution of wealth to consider another aspect of the credit market that is a subject of public policy. This is the degree of competition in the credit market. We begin with a borrower without wealth, like the one depicted in Table 12.2, now shown in Table 12.4 contrasting the case of the one-on-one interaction between Parama and Antonio with a situation in which borrowers and lenders interact in a credit market with no barriers to entry or what we have termed unlimited competition.

The difference between the final two rows of Table 12.4 underscores the effect that limits to competition has on the distribution of income. Looking just at the relationship between the lender and a single borrower, two differences are noticeable:

- *The total rents available to the borrower and lender*: Competition among lenders (the bottom row) results in a substantial increase in the size of the total "pie" — from 0.825 to 1.225. The reason is that competition forces lenders to charge lower interest factors – δ drops from 5 to 3 – and in response borrowers adopt lower (more nearly total-rent-maximizing) levels of risk – f falls from 0.75 to 0.65.

- *The distribution of rents between borrower and lender:* Unlimited competition eliminates the rents of the lenders, they receive expected profits equal to the opportunity cost of capital, all of the rents go to the borrower. As a result, their accounting profits, that is, their income derived from their loan

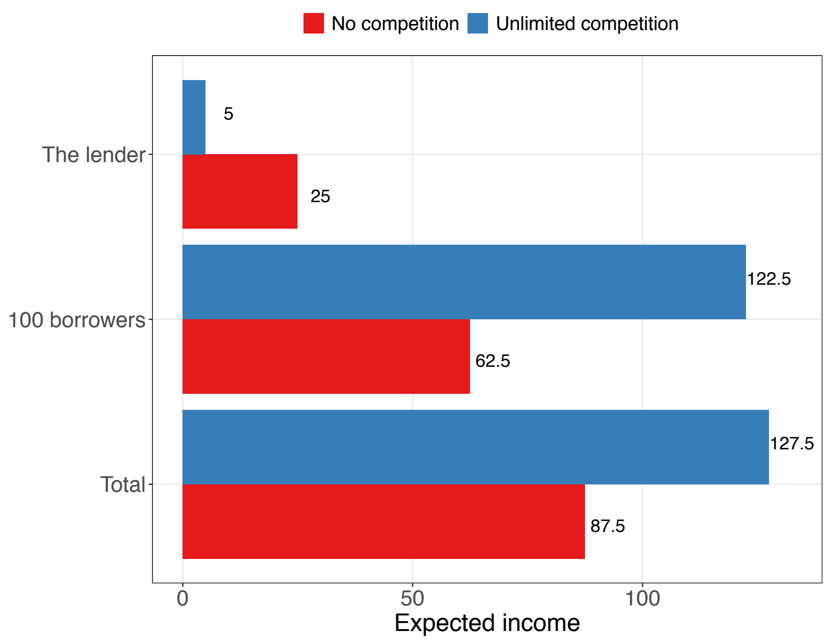


Figure 12.16: **Effects of competition on total income of a single lender and her set of borrowers, and its distribution.** The data for the figure come from the numerical examples in the last two rows of Table 12.4 with two additional pieces of information. We show the lender's income based on transactions with 100 borrowers identical to the single borrower shown in the table. And the lender's income is correctly measured by his accounting profits, namely $\pi^A = \delta(1-f) - 1$, not by his economic profits ($\pi^E = \delta(1-f) - (1+\rho)$. So, for example, the top blue bar, the lender's income of 5 under unlimited competition (the case in which economic profits are zero) is 100 transactions times the size of the loan (1) times 0.05.

of $1 is, is $\pi^A = \rho = 0.05$;

The entries in Table 12.4 give the results for the lender and just one of her borrowers. To see how this affects the distribution of income in the economy as a whole, we need to take account of the fact that banks and other lending institutions interact with very large numbers of borrowers. So think about the lender, Parama, and the, lets say, 99 *other* borrowers to whom Parama has extended loans, like she did to Antonio.

The results in Figure 12.16 show that the elimination of competition increases the income of the lender by five-fold while cutting borrowers' incomes to approximately half of their level under competition. Total income is reduced by one-third.

The contrast between these extreme cases – no competition versus unlimited competition – provide another illustration of the effect of limited competition: in the markets for goods and services

- increasing economic profits and reducing consumer surplus in Chapter 9 and
- reducing the real wage in the model of the whole economy in Chapter 11

As well as:

- the effect of limited competition in the labor market – monopsony – in reducing employment and lowering wages.

The comparison of the extreme cases – unlimited or no competition – suggests that public policies to reduce barriers to entry in credit markets would raise total income and reduce the inequality between lenders and borrowers. To see how this would work, in Figure 12.17 we study the credit market under two levels of competition:

- *barriers to entry:* the status quo with substantial barriers to entry ($b > 0$), and
- *unlimited competition*: a possible result of a government's competition policy in which barriers to entry are eliminated.

The left panel shows the competition conditions under under these two assumptions. In the right panel, point $\mathbf{n_1}$ is the status quo Nash equilibrium while $\mathbf{n_0}$ is the outcome under unlimited competition, that is $b = 0$. If policies could be implemented to shift the competition condition to the left (as shown in panel a) the outcome would be

- A reduction in the interest factor charged by the lender, from δ_1^N to δ_0^N,
- A reduction in the risk taken by the borrower, from f_1^N to f_0^N

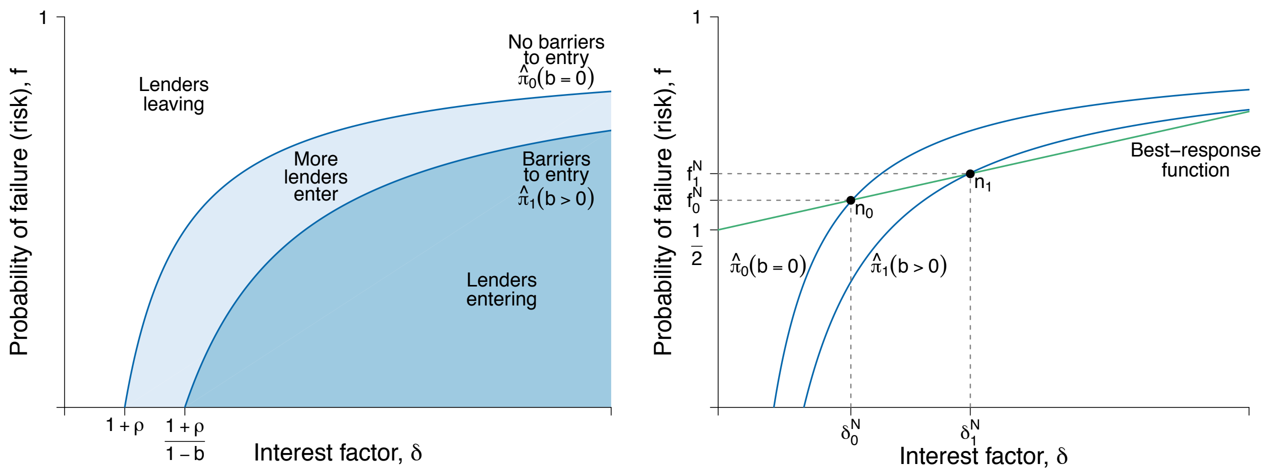
resulting in

- An increase in total income (you know the total income is maximized at $f = \frac{1}{2}$)
- An increase in the expected income of the borrower (you know from Figure 12.7 that expected income is higher at points to the left on the borrower's best response function) and
- A decrease in the expected income of the lender (because the competition condition has shifted to the left, n_0 is on a lower iso-expected profits curve than is n_1.

In Figure 12.18 we introduce borrower's equity at the level of $k = 0.5$ and we show how variation in the level of barriers to entry affect the total rents and the expected income and profits of the borrower and lender respectively.

12.13 Application: From micro to macro: The multiplier and monetary policy

We saw in Chapter 10 that when account is taken of the incomplete nature of contracts, competitive markets need not clear in equilibrium; and in Chapter 11 that this is true specifically of the labor market, which in long term equilibrium is characterized by structural unemployment. Here we show how principal agent models of the credit market provide additional foundations for a unification of microeconomics and macroeconomics.



(a) A decrease in barriers to entry shifts the competition condition

(b) Changes in barriers to entry change the Nash equilibrium

Figure 12.17: **Nash equilibria with unlimited competition and barriers to entry.** The green line is a borrower's best response function. Two alternative competition conditions are shown, one for the case of unlimited competition (b = 0) and the other for positive barriers to entry. You can see that with greater barriers to entry the Nash equilibrium level of the interest factor and the level of risk taken are both higher.

The microeconomic foundations of the Keynesian multiplier

An essential concept for macroeconomic policy is the Keynesian multiplier. The Keynesian multiplier indicates the total effect on aggregate demand generated by a single unit of exogenous change in expenditure (for example in the form of a government transfer, investment or net export demand).

Think about a fall in export demand, which results in fewer of the exported goods being produced and a reduction in employment, meaning some workers lose their jobs. The Keynesian model shows how this income shock to the employee then sets off ripple effects throughout the economy amplifying the initial effect. This occurs because with reduced income, there is a second-round effect: the worker spends less money on goods and services, so the income of the local market or grocer now also falls, along with the incomes of other people from which the she would have purchased goods and services had she not lost her job. A third round follows: the grocer purchases less from his suppliers, and so on. The multiplier is a measure of the extent to which an initial shock is amplified by successive rounds of reduced expenditure.

But if the unemployed worker could have readily borrowed sufficient funds to sustain her previous level of consumption until she found another job, the lost income of the unemployed worker would be where the process ends. There would have been no amplification of the shock. National income would fall by the amount of the reduction in export demand. There would be no second- and third-round ripple effects.

To explain the multiple rounds of additional expenditure that is essential to the Keynesian multiplier, it is common to assume that there are many "hand to mouth" households, whose consumption expenditure rises and falls with its income, which seems to be the case. By one estimate: "half of households

M-CHECK: In figure 12.18, for $b = 0.58$, the profit rate of the incumbent firms would have to be so high that the borrower with $k = 0.5$ would be the marginal borrower. For $b > 0.58$ there is no profit rate of incumbent firms that would be both sufficient to attract entering firms and consistent with incumbent firms transacting with borrowers.

AGGREGATE DEMAND is the sum of expenditures on goods and services produced in a country, including demand from the rest of the world.

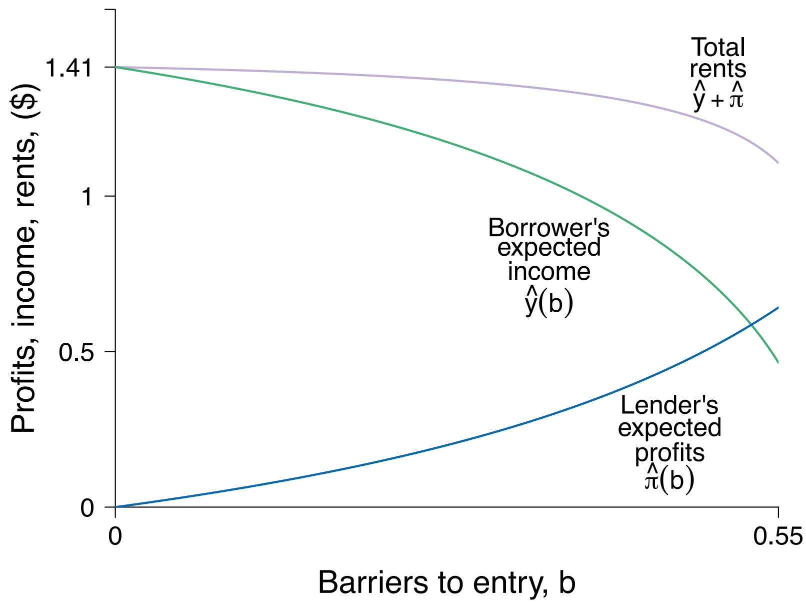


Figure 12.18: **Barriers to entry in the credit market and the distribution of rents** To show the effect of barriers to entry, we set specific values for the rest of the model and vary barriers to entry (b). In the figure the quality of the machine, $q = 10$, the level of equity is $k = 0.5$ and the opportunity cost of capital is $1 + \rho = 0.05$.

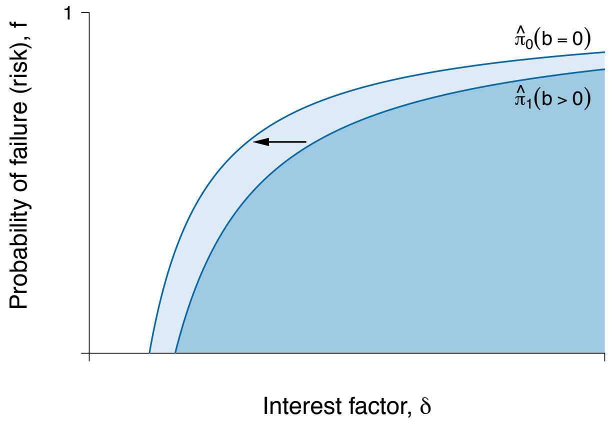
follow the 'rule-of-thumb' of consuming their current income."

The model of the credit market with incomplete contracts that you have learned provides one explanation of where this rule of thumb might come from and why these hand-to-mouth households are so common. A great many families are either excluded from the credit market entirely, or are unable to borrow except in small amounts or at prohibitively high rates of interest.

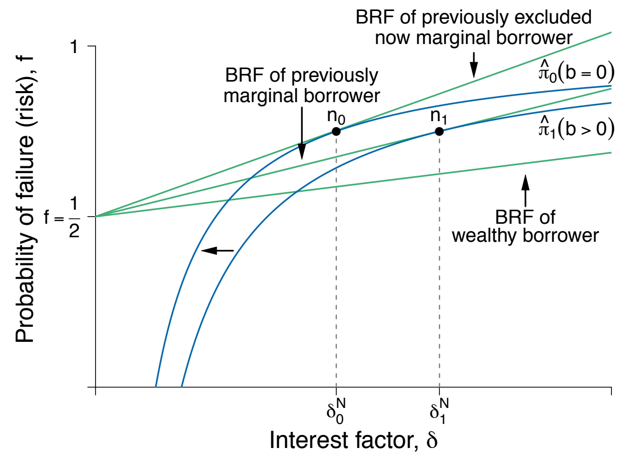
The microeconomics of monetary policy

The credit market model along with the labor market model in the previous chapter explains how **monetary policy** can stimulate investment and support higher wages and employment in the Nash equilibrium of the economy. Recall that if the credit market is competitive, then the zero profit condition requires that the expected profits of the lender, $\hat{\pi}$, be equal to the opportunity cost of capital, $1 + \rho$. In the U.S. the Federal Reserve System sets a target federal funds rate of interest which effectively determines the interest rates charged by commercial banks throughout the economy. The same is done in other countries by the central bank, for example, the Bank of England or the Reserve Bank of India. The rate of interest at which banks lend is one of the main determinants of the *opportunity cost of capital*.

To see how this affects the macro-economy, we use Figure 12.19 where the solid blue curve is the initial zero profit condition, $\hat{\pi}_1$. Now suppose the Federal Reserve System decides to lower the federal funds rate. This will shift the competition condition up and to the left as is shown by the the second



(a) Shift in competition condition when ρ decreases



(b) Change in marginal borrower

Figure 12.19: **The competition condition and the central bank interest rate.** When the central bank (reserve bank) decreases the interest rate, the competition condition shifts from $\hat{\pi}_0$ (solid blue curve) to $\hat{\pi}_1$ (dashed blue curve). As a result, a borrower who was previously excluded prior to the policy change is now the marginal borrower in the credit market (at the tangency of $\hat{\pi}_1$ to the new marginal borrower's BRF). The interest factor that a wealthy borrower pays decreases from δ_0^W to δ_1^W.

competition condition, $\hat{\pi}_0$.

There are two effects of the change in policy:

- *Include previously excluded borrowers:* The prospective borrower without wealth, who was previously excluded from the credit market, can now borrow (we have assumed that the reduction in the opportunity cost of capital was just sufficient to include her in the credit market as a marginal borrower) and
- *Lower interest for previously included borrowers*: The two borrowers shown who were able to borrow previously are now able to borrow at lower interest rates. The wealthy borrower for example was previously borrowing at the interest factor of δ_0^W and after the reduction in the opportunity cost of capital can now borrow at δ_1^W.

The effect of these two changes in the credit market will be to increase spending: the previously excluded borrower will now spend the funds borrowed, and the other two borrowers will now enjoy higher profits on the projects that they previously financed with their loans. They can also finance larger projects or projects of lesser quality.

MONETARY POLICY Monetary policy implemented by a country's central bank (in the U.S. the Federal Reserve System) affects the rate of interest at which businesses and others can borrow and the amount of borrowing, thereby regulating aggregate demand to moderate the business cycle.

We have modeled the effect of monetary policy on investment in projects similar to our "machine." The primary effect of lower interest rates on spending, however, is not to expand investments such as building new plants, office buildings, and equipment but instead on the construction of new homes and the purchase of cars and other consumer durables (typically bought on credit).

A slight modification of the credit market model is required to study housing and consumer durable credit, to take account of the fact that lenders typically require collateral rather than equity. (The modifications are shown in M-Notes

Figure 12.20: **Mortgage interest rates and new private housing construction, U.S. 1977-2006.** The vertical axis measures the change in the number of new home units on which construction began between one quarter (e.g. Q1, meaning Jan to March) of year and the same quarter a year earlier. The horizontal axis measures the change in the real mortgage interest rate over the year before the current year. The real mortgage interest rate takes account of the effect of inflation on the cost of repaying the loan. The dots in the upper left, for example show that in quarters that saw a big drop in the real mortgage interest rate over the previous year there was big increase in housing starts. We also show best fit estimated (red) line based on these data. The correlation between the two variables shown is -0.397, meaning the larger the change in the mortgage rate, the greater is the drop in the number of housing starts. (Ninety-five percent confidence intervals for this estimate are $(-0.301, -0.493)$). From the slope of the red line we can calculate that a 10 percent increase in the real mortgage interest rate is associated with a 0.307 percent drop in housing starts.

12.8 and 12.9). But the underlying mechanism is the same: a reduction in the opportunity cost of capital (when the central bank lowers its lending rate) lowers the Nash equilibrium interest rate on mortgages and car loans, and allows some previously excluded borrowers to secure a loan.

Figure 12.20 illustrates this channel for monetary policy affecting aggregate demand. It shows that lower mortgage interest rates are associated with higher levels of new home construction. As we shall see at the start of Chapter 13, the global financial crisis resulted in many people defaulting on home loans and going bankrupt, the central banks of the world stepped in by engaging in monetary policy to stimulate demand for loans by changing the central bank lending rate.

M-Note 12.8: The case of collateral : Expected lender's profits and borrowers income

Similar to the case of equity, the borrower has collateral of value k, and the structure of the game between him and the lender otherwise is unchanged. The borrower will still borrow \$1 for the project, but now additionally will post k as collateral.The collateral may be the borrower's home if the loan is a mortgage for the purchase of the home. It could also be a car being purchased on credit. Therefore we assume that the borrower can still use the collateral until it is claimed, so there is no opportunity cost of devoting k to collateral unless the project fails. Thus, his expected income will consist of:

- As before, in the case the project succeeds (with probability $(1 - f)$), the borrower receives expected returns minus expected repayment of the loan
- Now, additionally, if the project fails (with probability f), the borrower will have to give the lender the collateral, k.

Putting this together, we have:

$$\hat{y}(f,\delta,k) \quad = \quad qf(1-f)-\delta(1-f)-fk$$

The lender's expected profits will be as in equation 12.4, but now, instead of getting nothing in the case that the project fails, the lender will receive the value of the collateral, k:

$$\hat{\pi}(\delta,f,k) \quad = \quad \delta(1-f)+fk-(1+\rho) \tag{12.28}$$

M-Note 12.9: The case of collateral : Interest factor and risk decision

Following the procedure in M-Note 12.4 and the corresponding M-Check, the borrower differentiates his expected income with respect to f and sets it equal to zero to maximize his income, which gives us his first order condition and best response function:

$$\begin{aligned}
\frac{\partial \hat{y}}{\partial f} &= q-2qf+\delta-k=0 \\
\text{FOC: } \underbrace{\delta}_{MB} &= \underbrace{-q(1-2f)+k}_{MC} \\
\text{BRF: } \quad f(\delta,k) &= \frac{1}{2}+\frac{\delta-k}{2q}
\end{aligned}$$

We can see that, instead of changing the slope of the best-response function of the borrower as we saw in the case of equity, collateral shifts down the best-response of the borrower.

Now, following the procedure in M-Notes 12.2 and 12.6, we can solve for the Nash equilibrium level of risk and interest factor using the condition that, at the Nash, the slope of the iso-expected profit curve is equal to the slope of the best response function, which is now:

$$\frac{1-f(\delta)}{\delta-k} \quad = \quad f_\delta$$

We can now substitute in the new best response function $f(\delta,k)=\frac{1}{2}+\frac{\delta-k}{2q}$ and solve to find the interest factor that maximizes the lender's expected profit (where $f_\delta=\frac{\partial f}{\partial \delta}$):

$$\begin{aligned}
\frac{1-(\frac{1}{2}+\frac{\delta-k}{2q})}{\delta-k} &= \frac{1}{2q} \\
\delta^N &= \frac{q}{2}+k
\end{aligned}$$

This is the Nash equilibrium level of interest that the lender chooses. To complete the analysis we need to find the borrower's Nash equilibrium best-response level of risk ($f(\delta^N)$). So we substitute δ^N into his best response function:

$$
\begin{aligned}
f(\delta^N) &= \frac{1}{2} + \frac{\delta^N - k}{2q} \\
&= \frac{1}{2} + \frac{\frac{q}{2} + k - k}{2q} \\
f^N &= \frac{3}{4}
\end{aligned}
$$

Thus, in the case of a single lender, f is unchanged from the case without collateral, derived in M-Note 12.6. This is because, as we can see from the borrower's first order condition, the marginal cost of increasing f has increased by the amount k, or the cost of losing his collateral. But, the lender has also raised the interest factor δ by the amount k, which means the borrower's marginal benefit of increasing f has also increased by exactly this amount. In other words, the borrower's incentive to increase f in order to avoid paying a higher δ is canceled out by the prudence induced by the increased stake the borrower has in the project.

12.14 Application. Why cotton became king in the U.S. South following the end of slavery

In the U.S. South, prior to the Thirteenth Amendment to the U.S. Constitution that abolished slavery (1865), it was said that "cotton was king." But it was not until after the Civil War that cotton truly ascended to the throne among crops: In the quarter of a century following the demise of slavery, the production of cotton relative to corn (the main food crop) increased by 50 percent.

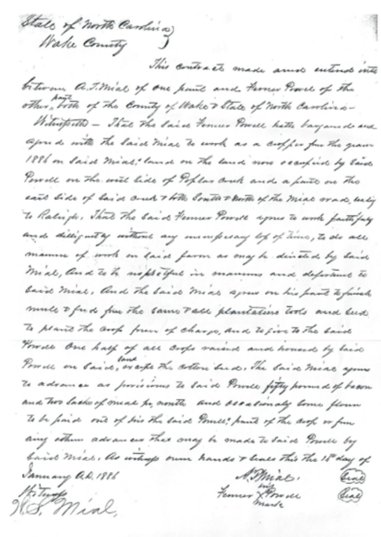
The intensification of the cotton monoculture puzzled observers at the time and since, as it coincided with a slight downward trend in the price of cotton relative to corn. Moreover, there were no changes in the technical conditions of production that would have offset the adverse price movement. In fact, the growth of corn yields appears to have outpaced cotton yields during this period.

Nor can the shift from corn to cotton be explained by changes in factor supplies: The Cotton South experienced a serious labor shortage following the war, which should have led some farmers to abandon cotton in favor of corn, as the latter was a much less labor-intensive crop.

What then explains the growing dominance of cotton?

To answer this we need to investigate the structure of local *credit markets*. To finance the crop cycle, most farmers – mostly poor share-croppers and rental tenants, many of whom were former slaves – purchased food (including corn) and other necessities on *credit* during the growing season. Because there typically was a single merchant in each area, the prices at which the farmers accumulated their debt were inflated by the *monopoly power* of the merchant-lender.

The loans were repaid when the crop was sold at the end of the season.


State of North Carolina)
Wake County)
This contract made and entered into between A.T. Mial of one part and Fenner Powel of the other, both of the County of Wake & State of North Carolina – Witnesseth – That the said Fenner Powell hath bargained and agreed with the said Mial to work as a cropper for the year 1886 on said Mial's land on the land now occupied by said Powell on the west side of Poplar Creek and a part on the east side of said creek & both South & North of the Mial road, leading to Raleigh. That the said Fenner Powell agrees to work faithfully and diligently without any unnecessary loss of time, to do all manner of work on said farm as may be directed by said Mial, And to be respectful in manner and deportment to said Mial, And the said Mial agrees on his part to furnish mule & feed for the same & all plantation tools and seed to plant the crop free of charge, and to give to the said Powell One half of all crops raised and housed by said Powell on said land, except the cotton seed. The said Mial agrees to advance as provisions to said Powell fifty pounds of bacon and two sacks of meal per month and occasionally some flour to be paid out of his the said Powell's part of the crop or from any other advances that may be made to said Powell by said Mial. As witness our hands & seals this the 18th day of January A.D. 1886

A.T. Mial (Seal)

Fenner his X mark Powell (Seal)

Witness

W.S. Mial

Figure 12.21: **A share cropper's contract from North Carolina in 1886**, signed with an X by the formerly enslaved African American Fenner Powel (at the bottom, right) in which he agrees to give the landowner M.S. Mial half of the crop he grows and "to be respectful in manner and deportment to said Mial."

SHARE CROPPER A *share cropper* is a farmer who cultivates land owned by another person with whom he or she contracts to give a share (usually one half) of the crop produced.

LIEN A *lien* is a property right that entitles the holder to the good. In the case of *crop liens*, the owner of the lien (the lender) was entitled to, or had a property right over, a share of the farmer's crops at harvest time.

Most farmers were too poor to provide collateral, so the merchant-lenders secured their loans by means of a claim (called a *lien*) on the farmers' future crop in case of default. The system was therefore called a *crop lien system*. According to its most prominent researchers, Roger Ransom and Richard Sutch, the crop lien system favored cotton:

REMINDER In Chapters 8 through 9, we saw the use of different factor inputs – capital and labor effort .

> In the view of the merchant, cotton afforded greater security for such loans than food crops. Cotton was a cash crop that could readily be sold in a well-organized market; it was not perishable; it was easily stored .... For these reasons the merchant frequently stipulated that a certain quantity of cotton be planted .... It was the universal complaint of the farmers that the rural merchants predicated his willingness to negotiate credit on the condition that sufficient cotton to serve as collateral had been planted.

The crop lien system that came to prominence in the US south after the civil war was an ingenious solution to the problem of providing credit to asset-poor borrowers. It substituted the farmer's unenforceable promise to repay the loan in the future by an action observable by the lender prior to the granting of credit, namely having already planted cotton on which the merchant had first claim.

Taking account of the relative resource costs and prices of the two crops, Ransom and Sutch estimate that the cotton farmer purchasing corn on credit could have increased his income by 29 percent by shifting resources from cotton to corn. But this was precluded by fact that because the farmer had little wealth, he needed credit, and for the same reason, credit was conditioned on planting cotton. The result, according to Ransom and Sutch was that:

> The southern tenant was neither owner of his land nor manager of his business . . . his independent decision making was limited to the mundane and menial aspects of farming. The larger decisions concerning land use, investments in the farm's productivity, the choice of technology, and the scale of production were all made for him.

The story of how cotton became king in aftermath of the end of slavery could not be told if the contract to repay a loan was complete and enforceable. Cotton was the lender's solution to a problem of an incomplete or unenforceable contract.

12.15 Why and How Wealth Matters

A widely circulated legend has it that the F.Scott Fitzgerald once said to Ernest Hemingway "The rich are different from you and me." To which Hemingway is said to have replied "They have more money." The first thing that is wrong with this charming conversation between the two great American authors is that it did not happen. But that is just the beginning.

Having greater wealth conveys quantitative advantages – it determines the location of one's budget constraint and gives you a larger feasible set of goods you can buy. But if contracts are complete that's all it does. The rich "have more money" than others.

In an ideal world of complete contracts, all participants in the economy face the same contractual opportunities (and hence the same prices) irrespective of their holdings. The poor are constrained to buy less than the rich, but they transact on the same terms.

By contrast, where contracts are incomplete, wealth confers qualitative advantages including greater personal autonomy and less being subjected to the will of others – being an employer rather than an employee, for example. Substantial wealth also makes it more likely that an individual will be a lender rather than a borrower in the credit market, or Benetton rather than a subcontractor. Substantial wealth, in other words, opens up opportunities to be principal rather than an agent.

Even among borrowers in the credit market have seen that that in the Nash equilibrium of the credit market wealthier people:

- pay lower interest rates
- can borrow more and so finance larger projects, and
- can finance projects of lesser quality.

People who lack wealth either:

- are not able to engage in contracts or projects that are available to the wealthy, or
- enter contracts on less favorable terms (higher interest factors) than wealthier borrowers.

Stepping back from our simple illustration of the "machine" as a project, the terms on which one can borrow spell the difference between being a principal rather than an agent in other economic interactions. And this often means experiencing one's economic life as one who makes decisions and gives directives as opposed to one who carries them out and may feel fortunate to have a job at all. Therefore when credit contracts are incomplete, wealth differences have qualitative effects, excluding some and empowering others.

The most obvious reason why wealth influences the contracts one can engage in is that only those with sufficient wealth can undertake projects on their own as owner-operators, as illustrated by the first case we studied. Those with enough wealth to start their own businesses have the rights to their profits (they are residual claimants) and can control what they do. They own the results of their decisions and hence there are no external effects of their actions so their projects yield the maximum possible rents.

A second reason why wealth influences the type of contracts that one is offered is that wealth ownership reduces the conflicts of interest and misaligned incentives arising from contractual incompleteness in principal-agent relationships. Wealthier people have access to superior contracts because their wealth allows contracts that more closely align the objectives of principal and agent. An example is when the borrower has sufficient wealth to provide collateral or put her own equity in a project. The borrower who provides collateral or equity to his project experiences enhanced incentives for the following:

- to supply effort for the project to succeed,
- to adopt risk levels preferred by the principal,
- to reveal information to the principal and to act in other ways that advance the principal's interests but that cannot be secured in a contract.

People lacking wealth may acquire education and other forms of human capital on less favorable terms than the rich. In the U.S. and many other countries, university students from families with limited wealth accumulate massive debts, for example, the repayment of which often constrains them to major in fields with high expected incomes, even if their interests lie elsewhere. We will consider the problem of financing university studies – and the policy of free tuition – in the next chapter.

Similarly, in residential housing markets, those with sufficient wealth are more often owners rather than renters and therefore benefit directly in increased values of their own property from the actions they take to improve their property and their neighborhood. The asset-poor are more likely to be renters; unable to borrow the funds necessary to buy their homes.

12.16 Conclusion

While our model was about an imaginary "machine" that was an income making opportunity that the borrower could choose to take (or not), the reality that the model was built to illuminate includes less benign circumstances. Examples are desperately poor people paying usurious interest rates on payday loans as as to be able to purchase necessary medicines, former slaves taking on a new form of bondage in their subjugation to lenders, and people's educational choices and later life chances being limited by the prospect of paying off student loans.

In the next chapter, we continue the analysis of risk and differences in wealth, introducing the important fact that people prefer certainty over risk if all else is equal, that is, people are risk averse. But as we have seen taking risks is essential to making the best of one's economic opportunities. We also see that limited wealth is associated with risk aversion resulting a vicious circle of enduring poverty. Not only do wealthy people *have* more money, but they *treat* their money differently when it comes time make investments with it and bear risk when doing so.

Making connections

Incomplete contracts: The rules of the game and limited information Bankruptcy, limited liability and the limited information available the lender make the level of risk taken by the borrower and the promise to repay a loan difficult to enforce.

Tradeoffs The tradeoff faced by the lender is that a higher interest factor makes her more profits if the machine does not fail, but it will induce the borrower to operate the machine at greater speed incurring more risk of failure. The tradeoff faced by the borrower is that running the machine faster runs the risk of failure, and getting nothing; but it also increases the probability that the loan will not be repaid (due to the failure).

Mutual gains (rents) from exchange and conflicts over their distribution Borrowing and lending – like buying and selling, and hiring and working – allow both parties to the exchange to do better than were the exchange not to occur. The outcome of the conflict over distribution of these rents depends on the rules of the game.

Competition, barriers to entry, and the distribution of income As in the market for goods, barriers to entry limit the number of competing lenders, reducing the borrowers' expected income, raising the lenders' expected profits and reducing the sum of these two types of rents.

Participation and incentive compatibility constraints and optimization As in other principal agent models, these two constraints limit both the borrower's and the lender's optimization process

Principal agent models, Pareto-inefficiency and quantity constraints. Like the market for labor and for goods of variable (and difficult to monitor) quality, the Nash equilibrium of the borrower-lender interaction is not Pareto efficient. Those excluded from borrowing are, like the unemployed, quantity constrained.

Nash equilibria: Dyadic and market-wide. The equilibrium of the credit market requires lenders' interest factors to be a best respond to borrowers best (risk taking) response function and the number of lenders in the market must be such that firms decisions to enter and exit the market result in a constant number of incumbent firms.

Evidence: empirical relevance and history The market-excluded and credit-constrained borrowers predicted by the model are evident in empirical studies. The model helps understand the rules of the game imposed economic hardships on former slaves in the U.S. after the Civil War.

Public policy: redistributing wealth, enhancing competition The model allows an assessment of the effects of policies such as wealth redistribution and reducing barriers to entry both on the size of rents to be shared between borrowers and lenders and on their distribution.

Important ideas

(in)complete contract	credit	rent
risk	interest factor	interest rate
borrower/agent	lender/principal	fallback
owner-operator	unlimited competition	contingent renewal contract
Pareto-(in)efficiency	Nash equilibrium	credit-market excluded
entrant (lender)	competitive market	
incumbent lender	monetary policy	monetary policy
project quality	inequality	redistribution
residual claimant	marginal borrower	credit-market constrained
Keynesian multiplier		

Mathematical notation

Notation	Definition
$\hat{y}$	expected income of a borrower
$\hat{r}$	expected return of the owner-operator
f	risk of the project (failure probability and speed of the machine)
q	project quality
δ	interest factor (one plus interest rate)
z	the borrower's fallback position
$\hat{\pi}$	lender's expected profit (incumbent firms)
$\hat{\pi}^b$	lender's expected profit (prospective entrant firm)
b	probability that a lender attempting entry will not succeed
k	wealth of the borrower invested as equity in the project
ρ	opportunity cost of capital
K	size of the project if different from 1.
ϕ	fraction of total funds loaned to poorer borrowers

Note on superscripts: C: Complete contract; N: Nash equilibrium (incomplete contract); other superscripts and subscripts refer to the wealth level of the borrower or to particular points in figures.

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

Part IV

Economic systems and policy

> Which of these systems [central planning or market competition] is likely to be more efficient depends on the question under which of them can we expect that fuller use will be made of the existing knowledge. And this, in turn, depends on whether we are more likely to succeed in putting at the disposal of a single central authority all the knowledge which ought to be used but which is initially dispersed among many different individuals, or in conveying to the individuals such additional information as they need in order to enable them to fit their plans in with those of others.

Friedrich Hayek, 1945, "The Use of Knowledge in Society," *The American Economic Review*

In this last part of the book you will have the opportunity to apply some of the analytical tools and models you have learned to understand how the institutions making up an economic system work as a whole, and how they might be made to better.

But there remains a set of tools that you need to do this: the analysis of risk, and the related study of inequality. Risk will be important because understanding the success of capitalism in raising material living standards requires an analysis of a risky process at which capitalism as a system has excelled: innovation.

Inequality is related to risk because the analytical tools to evaluate risky choices that a single individual may make can be re-purposed to study how we evaluate an unequal situation. For example how a person evaluates a gamble where both a 'good outcome' and a 'bad outcome' are possible can be also be used to evaluate a situation where 'wealthy' and 'poor' poor are the outcomes for different people. Because along with innovation capitalism also generates significant level of inequality, having the tools to analyse both is essential to understanding toady's world.

In Chapter 14 we consider ideal models of a decentralized market economy in which contracts are complete. Two quite different variants of this approach are called "perfect competition" and "efficient bargaining". They are similar in that under some conditions they implement Pareto efficient outcomes. We call these models utopian because the they rely on assumptions – unlimited competition, complete contracts and efficient bargaining – that are remote from actual economies as we know them.

We also revisit one of the greatest debates in the history of economics between advocates of this ideal market economy and proponents of an equally utopian conception of central economic planning of the type practiced by the Soviet Union from the time of the First Five Year Plan in 1928 until the end of Communist Party rule in 1991. The epigraph above is from Friedrich Hayek's contribution in which he introduced an important bit of realism to that debate — that information is scarce and incomplete. On this basis he advocated markets over planning because of the superior use that markets make of the necessarily limited information available to economic actors.

Then in Chapter 15 we provide a more empirically based model of how the capitalist economy works, one that takes account of elements missing from the utopian versions of perfect competition and efficient bargaining. We model a "second best" world in which the idealized the assumptions of the utopian models of central planning and perfect competition are replaced by more realistic starting points.

Capitalism, like central planning and other economic systems, is a way of coordinating how we produce and distribute the goods and services on which we live, and how we interact with each other and with nature in the process of doing this. What distinguishes capitalism from other economic systems is that it is based on privately owned profit making firms hiring employees to produce goods and services for sale on a market in competition with other firms. Managers of firms direct the uses of what Adam Smith called "other people's money" (the investors stake in the firm) and other people's labor (that is the employee's work), neither of which is typically subject to a complete contract.

CAPITALISM is a set of institutions making up an economic system that is based on privately owned profit making firms hiring employees to produce goods and services for sale on a market in competition with other firms

We ask what it is about capitalism that accounts for "the great hockey stick of history," namely the sharp upturn of output per capital and rising living standards that occurred in many countries with the emergence of the capitalist economy. We show that in many economies and epochs capitalism has been an "innovation machine" in part because it is the wealthy owners of firms who are in a position to make key decisions about risky decisions concerning new technologies and products. The benefits of innovation and sometimes unfair inequalities are thus inextricably linked in the capitalist economy.

Putting together our models of markets for goods and services, credit, and labor, we have a picture of how the capitalist economic system as a whole works. The results – the division of output between profits and wages and the level of employment, unemployment and output – determine the level of income inequality.

Public policies based on the microeconomics of capitalism can contribute to human well-being by addressing market failure and unfair inequality. In the final chapter we return to our starting point in Chapter 1 – the classical institutional challenge. We show how the tools you have learned can contribute to the design of public policies in pursuit of these objectives, and why these policies sometimes fail.

13

A Risky & Unequal World

> "It is not certain that nothing is certain."
>
> Blaise Pascal, 1670 *Pensées* (Thoughts)

DOING ECONOMICS

This chapter will enable you to:

- See how experimental methods allow us to study attitudes towards risk empirically, including gender differences in risk taking.
- Use utility functions and indifference curves to understand why people buy insurance and how this can facilitate people taking risks in ways that raise their expected incomes.
- Explain why people with less income or wealth (and as a result with limited opportunities to borrow) may be especially reluctant to take risks, and how this will reduce their income on average, perpetuating inequalities.
- See that both insurers and the insured benefit when insurance is purchased and that conflicts will exist over the distribution of these mutual benefits.
- Apply the model of risky decision making and insurance to questions of public policy, such as tuition for higher education and tax and transfer policies to reduce both risk exposure and inequalities.
- Identify limitations of the model of risky decisions including cases where probabilities of the occurrence of uncertain events re unknown and the alternative insights based on loss aversion.
- Pose ethical questions about economic injustice in terms of feasible sets, indifference curves, and Adam Smith's Impartial Spectator.

13.1 Introduction

"My daughter purchased a home for $120,000 just before the housing crash [of 2007-09]," Virginia Mayou wrote to an advice column at *USA Today*. "She now owes [the bank] about $89,000. She needs to sell, but the home is valued at less than she owes. She is a single mom with teenage children and doesn't have funds to pay off the mortgage. What are her options?"

Mayou's daughter did not know she was taking a risk in purchasing her home. She had every reason to expect that the value of her home would continue increasing. House prices in the U.S. had doubled over the decade before she borrowed from the bank to purchase the home. The home loan, called a mortgage, did not seem risky to the bank either. The reason is that Mayou's daughter had given the home itself as the collateral, meaning, that if for some reason she failed to pay back the loan the bank could take ownership of her home.

But that collateral would no longer be enough to offset the unpaid portion of the mortgage. For many people, the best they could do after the mortgage crisis began was to give the bank the keys to the house and walk away, leaving the bank with the loss.

House prices on average would fall by more than a quarter between the peak of the housing market in the summer of 2006 and the low in early 2012. Mayou's daughter's loss was even greater.

She was an unfortunate and common casualty of the global financial crisis in which she was a small and unwilling player. The median wealth – the wealth

figures/risk/housing.png

of the family for which half of households are wealthier and half are poorer – fell from $107,000 to $57,800, a drop of almost a half. Median wealth losses of Hispanic and African American households were far greater.

Households like the Mayou's were betting and perhaps even expecting – on the grounds of recent experience – the value their home to increase. For most families the value of their home was most of their wealth. Their home served as a potential source of funds – through sale or additional borrowing with the house value as collateral – in case of some health, job loss, or other emergency. For any family, to find out that their home was worth half as much as they had thought was an unimaginable disaster.

The ripple effects from the downturn in housing prices in the summer of 2006 quickly turned into a financial tsunami. Households sought to restore their declining wealth, cutting back on purchases. The car industry in Detroit was the first major sector to be devastated as people decided they could not afford a new car. Sales of cars and small trucks collapsed, from 16 million in 2007 to 9 million in 2009. Experienced auto workers who were counting on their jobs until retirement were on the street looking for work. General Motors and Chrysler were headed for bankruptcy.

REMINDER Remember from Chapter 12 that wealth is the stock of things (or the value of that stock) owned by a person or household that contribute to a flow of income or other benefits. Income is the maximum a household or person can consume over a given period without borrowing or reducing the value of the household's wealth (their stock of assets such as home or car). The two differ because income is a *flow* – that is it is some amount *per period of time* while wealth is a *stock* – that is a set of things or their value with no time dimension.

Few were spared, and the effects went beyond losses in the value of a person's home, or GM stock, or another asset, or losing a job. When the leading French bank PNB Paribas on 9 August, 2007 announced that it could no longer pay back its loan holders their report pointed to an entirely new dimension of uncertainty: it had become "impossible to value certain assets fairly regardless of their quality or credit rating." As the global financial crisis unfolded it dawned on people that they had no idea what their assets were worth.

What happened to the value of Virginia Mayou's house, or the value of GM stock, or the GM workers jobs, or the Paribas creditors is called a shock, an unexpected difference between what might have been expected and what actually occurred. Shocks can also be welcome, as when the price of one's house rises at a rate more rapidly than expected, and occurred year after year during the housing price bubble that led to the crash of 2007-2009. Or the extraordinary increase in the value of GM stocks between 1995 and 2000.

The fact that expected outcomes are often not realized (do not actually occur) is the foundation of the study of risk. We have already seen its importance in the credit market model of the previous chapter. Here we deepen the analysis by studying people's preferences about risk and how both individual decisions and public policy can mitigate risk.

REMINDER In Chapter 12 two very different realized outcomes occurred depending on whether the machine exploded or not.

13.2 Choosing Risk: Gender differences

REMINDER As we saw in Chapter 2, an *expected value* or *expected payoff* is not some particular outcome that you expect to occur. It is the weighted average of the payoffs that could occur multiplied by the probability that payoff (or outcome) will occur.

People are *not* natural born gamblers: if offered the choice between receiving 100 dollars for sure, or flipping a coin to determine whether you get 1200 dollars or have to pay 1000 dollars, most people would take the sure 100 dollars. But the expected payoff of the bet – the 'good' and 'bad' outcome multiplied by the probability that each will occur – has the same expected value as the sure thing: $\$1200(0.5) - \$1000(0.5) = \$100$.

Choosing a level of risk

RISK AND UNCERTAINTY The term *risk* is conventionally used to describe situations where the probabilities of the possible outcomes are known. The term *uncertainty* describes situations where the decision-maker does not know and *cannot learn* these probabilities.

The coin flip just mentioned is an example of many choices we make in which for one or more of the options before us there is a good and a bad outcome. Think about he Assurance Game (from Chapter 1) representing the problem of planting early or planting late in Palanpur shown in Figure 13.2. For each strategy that Aram or Bina could choose (Plant Early or Plant Late) there is a good outcome (resulting in a higher payoff) and a bad outcome (resulting in a lower payoff). If Aram plants early, the good outcome occurs when Bina also plants early and he receives a payoff of $y^G = 4$ and a bad outcome (when Bina plants Late) of $y^B = 0$. Therefore, for planting early, Aram's risk exposure is $\Delta_E = y^G - y^B = 4 - 0 = 4$. Planting late on the other hand has a good outcome $y^G = 3$ and a bad outcome $y^B = 2$. Therefore, for planting late, Aram's risk exposure is $\Delta_L = y^G - y^B = 3 - 2 = 1$.

If we assume that both Aram and Bina think that the other is equally likely to plant early or late, then planting late is less risky because the difference (the Δ) between the good and bad outcomes is smaller. If the good and bad outcomes associated with a given strategy are equally likely whichever strategy you choose then Δ is a measure of the extent of the risk of the two strategies. His degree of risk is four times as great when Aram plants early as when Aram plants late. The same is true of Bina.

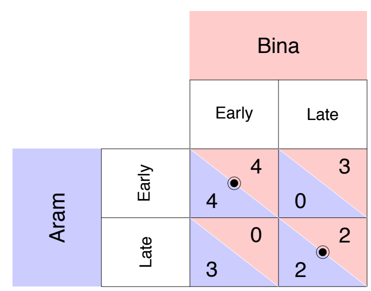


Figure 13.2: **Planting in Palanpur: An Assurance Game.** The *degree of risk* is the difference between the good and bad payoffs. So in Palanpur, Plant Late has a degree of risk of 1 ($3 - 2 = 1$), whereas Plant Early has a degree of risk of 4 ($4 - 0 = 4$)).

We treat the level of risk Δ as something that people choose ranging from everyday actions like whether to carry an umbrella or not, to life setting decisions such as whether to emigrate to another country.

The choice of Δ may refer to any of the following:

- A student choosing to specialize in nuclear engineering (where salaries are high but opportunities few outside the nuclear power industry) rather than in liberal arts where expected salaries are modest, but your training may equip you for a large variety of jobs, exposing you to less risk than the nuclear engineer.
- Relocating to a booming region of your country in search of work (substantial returns if you land a job, negative returns if you do not) rather than

taking one of the available low wage jobs in in your home town (low risk with lower potential returns).

- Going into business as a self employed person – electrician, software engineer – rather than taking a salaried job with predictable course of future raises.
- Producing a single product (high risk, with potential high returns if it is very good) rather than a range of products (low risk, spread across many products).

A lottery

We can describe the choices open to a person by listing what are termed the lotteries in which they may engage. A lottery is a set of outcomes and the probabilities that each will occur. In ordinary language a lottery is a gambling game played in casinos or online, such as Powerball in the U.S., EuroMillions in Europe, or the National Lottery in the U.K. In game theory it is just a way of modeling risky choices involving two (or more) outcomes with given probabilities of occurring.

Aram, Bina and the farmers of Palanpur faced two lotteries: plant early and plant late. In each case there was a risk – the difference between the good and the bad outcome of the lottery. And we assumed that in order to make a decision each of the players assigned some probability that the good and bad outcomes would occur, depending on what the other player did, if they chose plant early or plant late.

Some examples of lotteries are shown in the left hand panel of Figure 13.3. Lottery (L6) in the upper left is the riskiest: It describes a 50-50 chance of winning \$54 and losing \$2. The expected payoff is just the average of these two numbers, or \$26, and the degree of risk is $\Delta = \$54 - (-\$2) = \$56$. By contrast, L1 is not risky at all: If you play that lottery you get \$18 with certainty.

Comparing lotteries 1 through 6, the size of the difference between the good and bad outcomes (that is, the degree of risk or Δ) increases and the expected payoff increases (up to L5).

Gender differences in risky choices

Sheryll Ball and her co-authors studied gender differences among U.S. university students in choices among the six lotteries. We present the results of their experiments in Figure 13.4.

All subjects confronted the same set of expected payoffs and risks given by the six lotteries above. The most common choices differed by gender. You can see from the figure that both men and women are among those choosing the

(a) Experimental Choice of Lotteries

(b) Expected payoffs and degree of risk

Figure 13.3: **Six lotteries: Expected payoff and risk.** In panel a, we show the set of lotteries a person could choose in an experiment. The two numbers in the circles indicate for that particular lottery, the bad and good outcome (in that order). Panel shows the expected payoff of each of the lotteries along with the degree of risk.

riskiest lottery (L6) and also choosing the no risk lottery (L1). But the most common choice (the modal choice) for men was lottery 5. The modal choice for women was lottery 3. On average men were risk takers and women were risk avoiders.

13.3 Risk preferences over lotteries

Why do some people choose riskier lotteries than others. For example, what accounts for the differences in Figure 13.4 between those (both men and women) who choose the the risky L5 and the sure thing, L1. The risk takers must have evaluated the benefits of a greater expected income in L5 highly and not been too concerned about the fact that they might end up with just $2 (when they could have had $18 for sure, if they had chosen L1).

Placing a negative value on being exposed to risks or uncertainty is termed risk aversion. It is a common reaction of people and it could be due to a combination of

- Anxiety about not knowing what will occur and a related personality trait that psychologists term harm avoidance.
- A desire to avoid regret about having made a bad decision
- Diminishing marginal utility of wealth, income or whatever the currency of the lottery is.

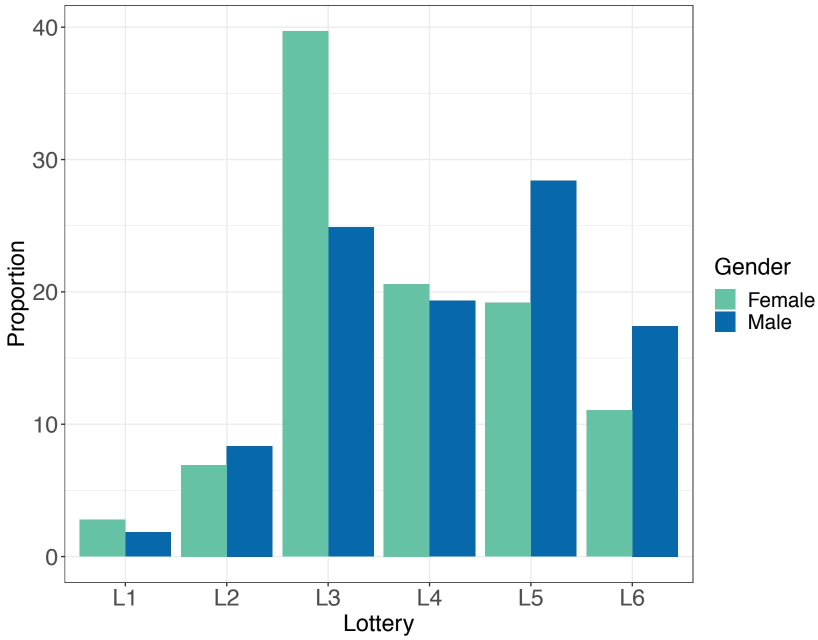


Figure 13.4: **Choices of the level of risk by male and female university students**. Lottery 1 (L1) is risk free, while Lottery 6 (L6) is the riskiest lottery.The height of each bar is the percentage of all women (blue) or men(red) who chose the lottery indicated. The means of the women's and men's choices, respectively were 3.79 and 4.16.

To see why the last bullet is true think about a person with just enough money to purchase one adequate meal in a day. He is then offered a lottery that instead of his one meal for sure, he will with equal probability have no meal or two meals. Think about what you would choose.

You would likely choose the sure thing because one meal is much better than no meal, and while two meals would be nice, the "good" outcome of the risky lottery (two meals) is not good enough to run the risk of having nothing to eat. Where the difference in one's wealth, health, or income associated with risk is substantial – having a job or not, developing extraordinary vision capacities or losing one's sight – then, like no meal at all, the "bad outcome" may be sufficiently catastrophic to motivate avoiding it at almost any cost.

M-CHECK: RISK AND DIMINISHING MARGINAL UTILITY. We show in the online appendix how risk can be analysed using a utility function $u(y)$ in which marginal utility is diminishing in the level of realized income, y. In this framework people do not have feelings of anxiety or excitement about risk itself, instead attitudes toward risk explained entirely by the marginal utility of income. We do not use this approach here because we prefer a more general model that can include all of the reasons – including those studied by psychologists – that people might want to avoid taking risks or might enjoy risk talking. Another advantage of the approach presented here is that it makes use of analytical tools already taught: indifference curves and feasible sets.

Risk-averse indifference curves

To understand the how people make risky choices we describe their evaluations of different outcomes using a utility function in which

- *Expected income is a good:* something the decision maker prefers and wants more of
- *Risk is not a good:* something the decision-maker would like to avoid or is possibly indifferent to, but does not prefer.

Here is the function:

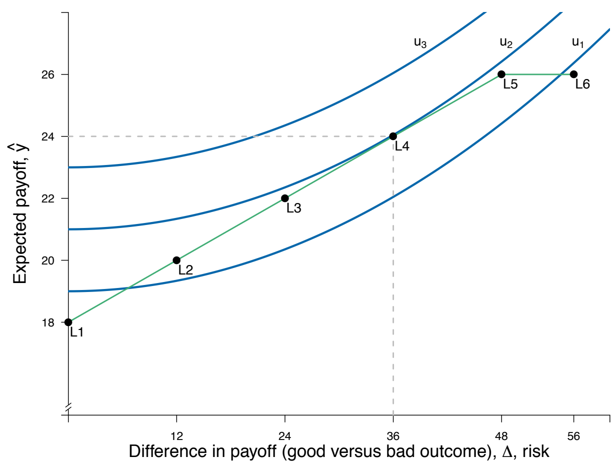
$$\begin{array}{rrcll} \text{person's utility function} & u & = & u(\Delta, \hat{y}) & (13.1) \\ \text{with marginal utility of risk} & u_\Delta & \leq & 0 & \\ \text{marginal utility of expected income} & u_{\hat{y}} & > & 0 & \end{array}$$

We often refer to the marginal *dis*utility of risk and this is just the negative of the marginal utility of risk or $-u_\Delta \geq 0$.

We allow for their being people in situations $(\Delta, \hat{y})$ where the marginal utility of risk $u_\Delta = 0$, so that the person is termed **risk neutral**, that is, indifferent to the level of risk. If the level of risk in a lottery were very small (e.g. a payoff of \$10.01 versus \$9.99) a person might not place any negative value on the risk involved. But for risks of any significant magnitude we assume that $u_\Delta < 0$ so the marginal utility of risk is negative, the person is **risk averse**.

figures/risk/risk_experiments_indifference.pdf

(a) Indifference curves of risk and expected income



(b) Feasible outcomes and indifference curves

Figure 13.5: **Feasible lotteries and indifference curves for a risk-averse preferences**. Because expected payoff is a "good" and risk is a "bad" preferrd outcomes are above and to the left and as a result the indifference curves slope upward: if two outcomes are associated with the same level of utility, then one must have both higher risk and higher expected payoff than the other.

To see if a utility function like Equation 13.1 could explain gender differences in the modal choice of a particular lottery, look at the indifference curves in the left panel of Figure 13.5. The indifference curves are upward sloping because risk is a bad. How steep they are at particular point is a measure of how risk averse the person is under the conditions given by the particular level of Δ and $\hat{y}$ at that point. The slope of the indifference curve (as we explain in the M-Check) is:

$$\text{Indifference curve slope} \quad = -mrs(\Delta, \hat{y}) \quad = \quad \frac{\overbrace{-u_\Delta}^{>0}}{\underbrace{u_{\hat{y}}}_{>0}} > 0 \qquad (13.2)$$

This expression gives the answer to the question: suppose you were required to take on a little more risk (an increase in Δ, the bad, so a move to the right in the figure); how much additional expected income (a move up) would you need to be no worse off than before (that is, to get you back to the same indifference curve)? The answer depends on how bad the risk is (u_Δ) compared to how much you value expected income ($u_{\hat{y}}$).

M-CHECK As before (e.g. in M-Notes 10.1 and 12.2) to find the slope of an indifference curve we totally differentiate the function in question and set the result equal to zero . In this case we set $d\hat{y}(u_{\hat{y}}) + d\Delta(u_\Delta) = 0$ and then solve for the slope of an indifference curve: $\frac{d\hat{y}}{d\Delta} = \frac{-u_\Delta}{u_{\hat{y}}}$

REMINDER The points making up a particular indifference curve are bundles of goods consumed or other results of actions taken associated with identical levels of utility. On a given indifference curve in Figure 13.5, therefore, utility is the same for differing levels of risk and expected income.

Doing the best you can in a risky situation

To take account of the constraints that limit what the decision-maker can do, we define a feasible set to include all of the combinations of the good (expected income, $\hat{y}$) and the bad (the level of risk, Δ) that the decision-maker can implement by their actions. For the experiment just described these are the points in the right panel of Figure 13.5 representing the six lotteries. You can see that for the person whose indifference map is shown, the best they can do is to select Lottery 4, which will expose them to a substantial amount of risk ($36) but also with a substantial expected payoff $24.

How might this way of understanding risky choices explain the gender differences found in the choice of lotteries?

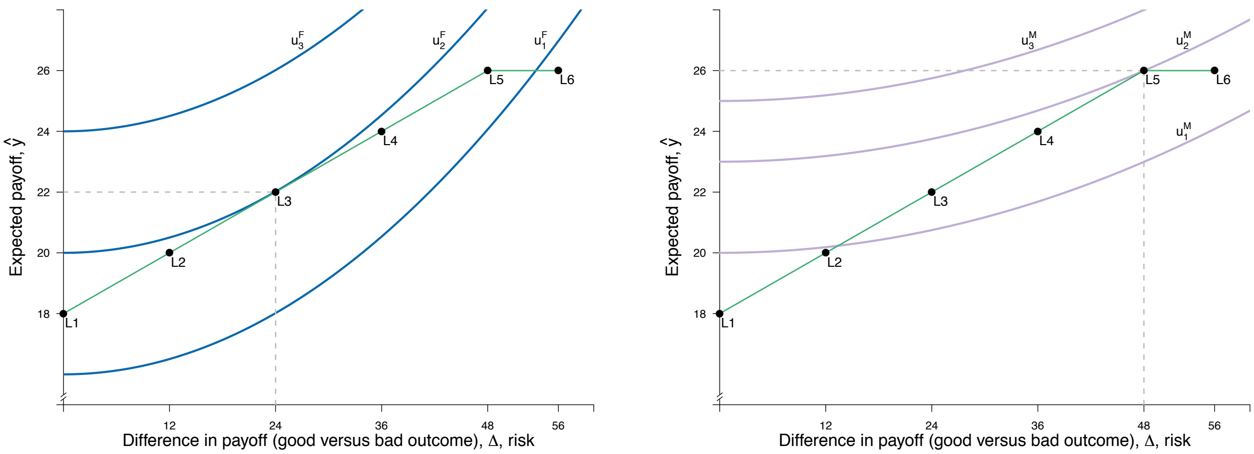
Gender differences in risk aversion?

In Figure 13.6 we provide a possible answer: the two panels show possible indifference maps that would lead to a different choice of lottery. Those in the left panel are steeper, and thus illustrate higher levels of risk aversion that than those on the right, which are flatter, meaning less risk averse. The person depicted on the left would choose L3 the modal choice of women in the experiment, while the person on the right would choose L5, the modal choice of men. Thus a gender difference is risk preferences – women more like the person on the left, men more like the person on the right – could explain the experimental results.

We do not know, of course that the differences between men and women in the experiment are explained by the kind of risk averse indifference curves that we have introduced. Look carefully at their choices over all of the lotteries in Figure 13.4. The expected income of L6 is no greater than L5, and it is riskier. But a substantial number (both men and women) chose L6. They may have placed a positive value on the stakes of the game being large, even if the expected income from taking the chance was no higher. Explaining *their* actions would require "risk loving" indifference curves in which risk is a "good" rather than a bad, as would also describe the preferences of those who take of sky diving and other "thrill seeking" activities.

FACT CHECK: NOVELTY SEEKING Psychologists have identified a personality trait called *novelty seeking* that is associated with risky behaviors and inversely correlated with another trait mentioed above, *harm avoidance*.

Differences in risk preferences between men and women or among other



(a) Indifference curves and choice of a risk averse person

(b) Indifference curves and choice of a less risk averse person

Figure 13.6: **More and less risk averse indifference maps motivating the modal choices of women and men**. In the left panel we show a relatively steep — that is, risk averse — indifference map that could have resulted in the less risky modal choice for women. The right panel shows a flatter set of indifference curves that could have motivated the riskier modal choice of men.

groups can be important because they may explain real-world decisions about risky choices. For example, other things equal, the evidence suggests that men are more likely to choose riskier assets with higher expected returns like stocks and mutual funds for their retirement portfolios than are women. This difference will result in greater incomes on average when men reach retirement.

Risk preferences can affect academic performance (risk takers are more willing to guess answers) and subsequent choices about careers (risk-takers choose STEM fields). Also affected are job choice (going into business on your own versus becoming a teacher), and structure of one's compensation package (a share in a firm's profit vs a fixed salary). These can all contribute to a person's lifetime earnings. If women are on average more risk averse than men, then this will contribute to women on average having lower incomes than men.

FACT CHECK Allison Booth and Patrick Nolen found that in experiments similar to that reported in Figure 13.4 when women are competing with other women in groups of only women, the risk levels they choose do not differ from men. So the context in which the choice takes place makes a difference. Taking account of the effect of context would require letting women have a different utility function when competing in an all-women group, with fatter, less risk averse indifference curves than when competing with both men and women.

Differences in risk aversion may help explain the persistence of inequalities not only between men and women, but more generally between wealthy and wealth-poor people.

13.4 Wealth differences and decreasing risk aversion

People differ in their degree of risk aversion, as the results of the experiment shown in Figure 13.4 suggest. These differences may be the result of:

- *Their type of person*: Some people feel anxiety or distress about uncertainty; others value "surprises." When we distinguish between different types of people we mean people with different utility functions and as a result different entire indifference maps, like the two people contrasted in Figure 13.6.

figures/risk/risk_averse_dara_c.pdf

Figure 13.7: **Decreasing risk aversion as shown in three indifference curves.** Three indifference curves for a person. The slope of the indifference curve at some point $\Delta, \hat{y}$ is a measure of risk aversion.

- *Their situation*: People already facing significant risks and unable to afford any serious loss will be very averse to additional risks. If the same person were wealthier and less exposed to risk, they might be less averse to risk. Differences in the situation faced by a single person could result in differing levels of risk aversion, indicated by the slopes of their indifference curves at different points in the space of expected income and risk. The same person will have differing levels of risk aversion depending on their situation.

In Figure 13.7 we illustrate two important influences of the situation on one particular person's level of risk aversion:

- *Risk exposure*: the degree of risk exposure (how far to the right the person's situation is in the figure). People exposed to substantial risks experience a large marginal disutility of additional risk. This can be seen by comparing the slope of the indifference curve at points **h** and **d**. For a given level of expected income, where risk exposure is greater (at **d**), risk aversion (the slope of the indifference curve) is also greater.

- *Income or wealth*: the person's expected income (how far up in the figure the person is). Those with little income also experience a large marginal disutility of additional risk. This is illustrated by comparing points **e** and **f**. For a given level of risk exposure, the person's expected income is greater (**f**) the person is less risk averse (the indifference curve is flatter) than where expected income is less (**e**).

This is why the indifference curves of any given person are steeper as you move horizontally to the right, and flatter as you move vertically upwards.

The tendency of a person to be less risk averse if she has more income than if she has less income is called **decreasing risk aversion** (it is less if income or wealth is greater). The most important reason for decreasing risk aversion is that a negative shock of a given size is likely to be a much greater loss in well being for a poor person than it would for the same person were they to have greater wealth. For example a 20,000 Euro loss in income – say, due to an injury or illness – will inflict a much greater catastrophe on a person who is poor than it would if they were well off. What this means is that for that for the poor person u_Δ is much greater (in absolute value) so (by Equation 13.2) the indifference curve will be steeper.

DECREASING RISK AVERSION is the tendency of a person to be less risk averse if she has more income (or wealth) than if she has less.

Two additional features of this indifference map are important.

- *The certainty equivalent (y_c)*: The intercept of the indifference curve u_1 and the vertical axis, y_c, has the same utility as points **e, d**, and all of the other combinations of risk and expected income that make up the indifference curve. What is unique about y_c is that there is *no risk* (it is like Lottery 1 in the experiment, a sure thing). So what y_c tells is the level of *certain* income that would be equally valued by the person to each of the other combinations (involving more risk and more expected income).

- *Risk neutrality*: We have said that the indifference curves are upward sloping, but for a person sufficiently wealthy they might be flat (not shown in the figure). Because risk aversion is measured by the slope of the indifference curve and a flat indifference curve has a slope of zero, in this (very wealthy) situation the person would be risk neutral: she would care only about expected income, not about risk. You can also see that at the vertical intercept – that is very small levels of risk, the indifference curves are approximately flat, the person exposed to virtually no risk at all, would not be risk averse.

Recall from Chapter 12 that people of limited wealth may be unable to borrow at all unless at usurious payday loan interest rates. We also saw that access to credit provides a kind of insurance, because the lender bears some of the loss in the case of project failure.

These facts along with risk averse indifference curves can help us explain why poor people may choose not to make risky investments – including investments in their own ability to earn higher incomes such as further training or moving to a distant part of the country. As a result they may end up poorer on average than they would have been had they taken the risk. The result can be a vicious circle or self-reinforcing poverty, contributing to economic inequality and its tenacious persistence. In the next section we provide an example.

Checkpoint 13.1: Risk and Income

Consider the following questions about the discussion so far.

a. Why, if someone is from a poor background, might they not choose to pursue a major in college that they enjoy, but which might be riskier in terms of potential job prospects (some chance of high income, but also a chance of low income)?
b. Why, if someone is from a high income background, might they *not* pursue a more standard major in college and instead pursue what might be otherwise perceived as a riskier major?
c. How would you depict these choices in figures like those in Figure 13.7?

13.5 Application: Risk, wealth and the choice of technology

Farming is one of the riskiest occupations. This is because the farmer's income depends on three things that vary substantially and are out of the farmer's control:

- weather and other environmental conditions affecting crop growth
- susceptibility of crops and livestock to disease
- the prices at which inputs are purchased and (especially) outputs are sold.

Partly for these reasons, researchers have studied farmers to better understand behavior in risky situations. Research on farming families in Indian villages recorded both the types of risk to which the farmers are exposed – uncertainty about the date that the dry season would end and the rains start – and the differing ways that farmers coped with the resulting uncertainty about their incomes. The researchers also recorded the total wealth of the farmers and the forms that the wealth took: land, irrigation equipment, tools, stocks of grain, and draft animals (neutered bulls, that is, bullocks).

EXAMPLE Farmers in the United States can purchase insurance to cover losses due to adverse weather. Insurance against damage by hail has been available in France and Germany for almost two centuries. But for most farmers in poor countries today insurance is not available. Think about the farmers in India. If they could have purchased crop insurance their resulting lesser degree of risk exposure would have allowed them to invest in irrigation pumps rather than bullocks, greatly increasing their income, and perhaps lifting them out of poverty.

There were substantial differences in wealth among the farmers studied: the richest one-fifth of the farmers owned 54 per cent of the total wealth. Rich and poor villagers also differed in *how* they farmed: those with a substantial amount of total wealth favored investments in pumps and other irrigation equipment, while the less wealthy rarely purchased pumps and invested their limited wealth in bullocks. They also found that few of the farmers had access

./figures/risk/bullock_pumps_a.pdf

(a) Anil's choice when he is poor

./figures/risk/bullock_pumps_b.pdf

(b) Anil's choice when he has non-farm income that increases his wealth

to credit in times of need. Instead, to meet their needs they resorted to selling one or more of their bullocks, in which there was a market to sell them.

Figure 13.8: **Bullocks tilling the ground in Karnataka, India**

π_4 rather than at point **b'** on indifference curve u_3 ($u_4 > u_3$).

By not investing in irrigation equipment the less well off farmers missed an opportunity to make substantially more profits: taking account of the costs, an installed pump would have raised profits by 72 percent on average. The reason the farmers avoided this profitable investment is simple: the poor farmers owned bullocks in order to have something they could sell to get them through times of need. The bullocks were a kind of combination savings account and insurance policy! In the villages under study there was no second hand market in irrigation pumps and other equipment, so, unlike a bullock, owning a pump did not provide a buffer against risk.

The fact that the wealthier farmers invested in riskier and more profitable assets (pumps rather than bullocks) had the effect of perpetuating or widening the income differences between them and the other farmers whose low income and lack of access to borrowing made them more risk averse.

Figure 13.9 illustrates this process and how it works for a poor farmer, Anil. On the horizontal axis is the risk undertaken by a choice of investments, with two levels shown: a less risky one with substantial investments in bullocks, the other riskier option with a greater investment in pumps. Thus, Anil must choose between investing in irrigation pumps with risk Δ_i and investing in bullocks with risk Δ_b.

An indifference curve through Anil's choice (point **b**) shows all combinations of risk and expected income that are equally preferred by him (as before points higher and to the left of this curve are preferred). In panel a, we can see that Anil is better off with the bullocks (point **b**) than with the pumps: point **i** on indifference curve u_1 is below and to the right of the indifference curve through point **b** on indifference curve u_2 (remember $u_2 > u_1$).

POVERTY TRAP A *poverty trap* is a self reinforcing set of processes (a vicious circle) that perpetuates low income for a person or group that were the trap broken might enjoy self-reinforcing prosperity.

Now suppose that Anil wins the lottery or somehow obtains substantial wealth, from which he will receive an amount of *non-farm income*. His expected income would be his farm income *plus* his non-farm income (from the newly acquired asset).

In the second state of the world (shown in panel b), Anil is rich and has non-farm income as indicated by the equal difference between points **b** and **b'** or between **i** and **i'** (his difference in income is $\hat{y}_b - \hat{y}_{b'} = \hat{y}_i - \hat{y}_{i'}$). In the state of the world where Anil has non-farm income, he chooses point **i'** on indifference curve u_4 rather than at point **b'** on indifference curve u_3. When he is wealthy he invests in the riskier asset of irrigation rather than bullocks and has higher expected income as a result with expected income $\hat{y}_{i'}$ rather than $\hat{y}_{b'}$ if he had invested in bullocks.

What explains the difference in his choice of assets when he is wealthy? As in Figure 13.7, the indifference curve through point **i'** is flatter (less risk averse) because people with more income are less risk averse. Because of this, when **b** and **i** both increased by the same amount of non-farm income, **i'** ended up on a higher indifference curve than **b'**. In this case, Anil, the same farmer who was once poor, became less risk averse because he became rich.

We can summarize what happened. In the first state of the world, Anil was caught in what is called a **poverty trap** in which his low income and lack of access to borrowing (as protection against risk) led him to choose the less risky but lower profit option: bullocks. The choice made for good reasons of prudence made him safer, but it also kept him poor. If he had had higher income to start with, he would have maintained a high income and invested in the higher return investment: irrigation.

A key part of the story was the level of risk that Anil chose. He had just two options. But in general we have a range of choices about the level of risk we undertake. We can analyze these choices using the same tools of constrained optimization that you have used since Chapter 3.

Checkpoint 13.2: You Win!

Recall the You Win! competition in Nigeria introduced in Chapter 12 where businesses were randomly selected to get the equivalent of $50,000 to invest in their businesses. People invested in minibus taxis, a factory to manufacture paint, and many other opportunities. Using the tools in Figure 13.9 explain why this policy might have worked.

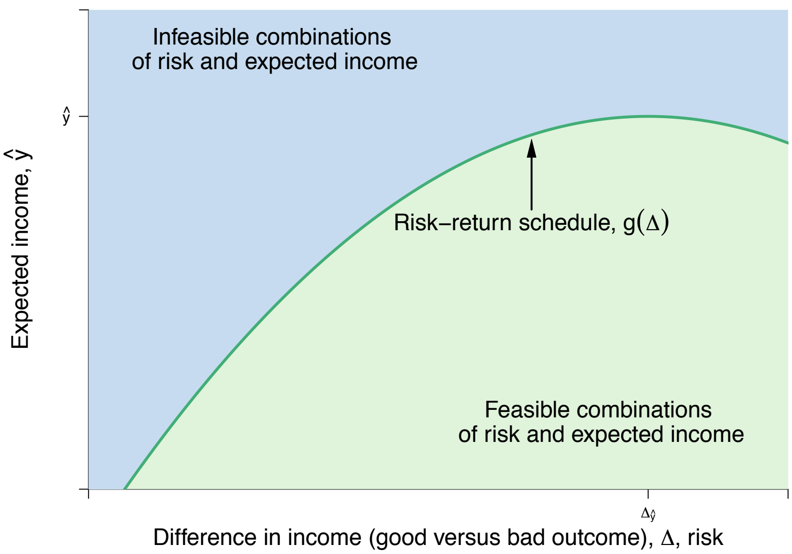


Figure 13.10: **Feasible combinations of risk and expected income.** Feasible combinations of risk (Δ) and expected income ($\hat{y}$) are shown using the risk return schedule, $\hat{y} = \hat{y}(\Delta)$. A combination within the set, for example point **a**, is feasible, but would not be selected by a decision maker because if it is not on the feasible frontier then there must be some other point **a'** with the same expected income and less risk and another point **a''** with the same risk and greater expected income. Both of these points dominate point **a**. Points outside the set are infeasible. Point **m** shows the choice of risk (Δ_m) that maximizes expected income at $\hat{y}_m$ – you can see that the slope is zero as shown by the tangent line.

Look up Episode 702 of NPR's Planet Money Podcast if you find this idea interesting (they report on the You Win! program in the episode).

13.6 Doing the best you can in a risky world

In the previous chapter we modeled the risk-taking choices of the operator of a "machine," choosing the speed at which it is run f. In that model the speed of the machine determined both

- *Difference between the good and bad outcome*: namely, the revenues possible by selling the goods produced by the machine if it does not explode, minus the revenues possible if it does, that is zero; and

- *Likelihood of failure*: The probability that the bad outcome would occur, which increased the faster the machine was run.

In this chapter, we study the choice of the extent of the difference between the good and bad outcome, Δ, but not the probability that each will occur. And for simplicity we will assume that the good and bad outcomes are equally likely: they each occur with probability one-half, independently of any actions that the decision maker takes. We call the difference between the good and the bad outcome the level of risk, or just risk.

./figures/risk/return_maximization_margin.pdf

Figure 13.11: **Risk and expected income.** Recall that in Chapter 12 we modeled risky choices using this relationship between the level of risk taken f and the expected profit of a project.

Feasible choices of risk and return

The terms "returns to risk" or just "returns" refer to the realized income or expected income resulting from having made an investment or some other risky choice. It is what you "get back." Here we let the expected income resulting from a risky choice represent the returns.

RETURNS The terms "returns to risk" or just returns refer to the realized income or expected income resulting from having made an investment or some other risky choice.

Of course most people would like to choose a course of action with high returns and low risk. But not all combinations of risks and returns to risk in terms of expected income are *feasible*. The feasible combinations of expected income ($\hat{y}$) and risk (Δ) are bounded in Figure 13.10 by the risk-return schedule, $\hat{y} = \hat{y}(\Delta)$. (Remember "schedule" is just another word for "function".) Similar to other feasible frontiers you can see that it divides the space of risk and expected returns into combinations that could possibly occur (in light green) and those which will not under any circumstances occur (in blue).

REMINDER The *marginal rate of transformation* is the negative of the slope of the feasible frontier, here given by the risk-return schedule. It is a measure of the *opportunity cost* of the one in terms of the other, that is, the opportunity cost of less risk is lower expected income. Refer to Chapter 3.

Like the risk return schedule in Chapter 12 expected income first rises with risk taking – posing a trade off to the decision maker – then reaches a peak and thereafter falls. The slope of the risk-return schedule is the (negative of the) **marginal rate of transformation** of risk into expected income.

M-CHECK Remember, we use the symbol $\hat{y}_\Delta$ to mean $\frac{d\hat{y}}{d\Delta}$, the derivative of expected income with respect to the choice of risk. This is the slope of the risk-return schedule or feasible frontier. or the (negative of) the marginal rate of transformation.

$$\text{Slope of the risk-return schedule} = -mrt(\Delta, \hat{y}) = \frac{d\hat{y}}{d\Delta} = \hat{y}_\Delta \tag{13.3}$$

The choice of a risk level by a risk-averse person

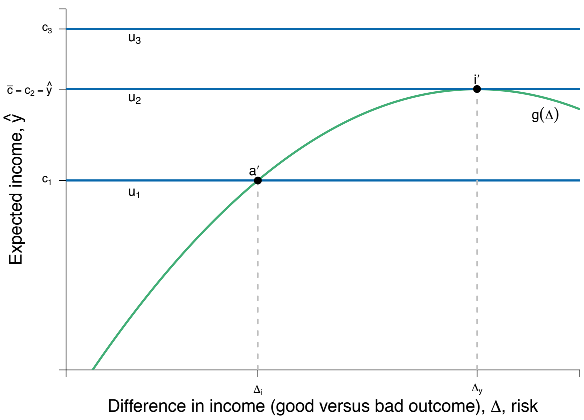
We can use the risk-return schedule along with indifference curves that capture the decision maker's risk aversion to understand the choice of a risk level. The decision-maker will vary Δ to maximize $u(\Delta, \hat{y})$ subject to the risk-return schedule $\hat{y} = \hat{y}(\Delta)$. We show in M-Note 13.1 that this requires choosing the Δ that equates the:

$$\begin{aligned} \text{Slope of the indifference curve} &= \text{Slope of the risk return schedule} \\ -mrs(\Delta, \hat{y}) = -\frac{u_\Delta}{u_{\hat{y}}} &= \hat{y}_\Delta = -mrt(\Delta, \hat{y}) \end{aligned} \tag{13.4}$$

Restricted to feasible combinations of risk and expected income so, what level of risk, then, will the decision-maker choose? He will have horizontal indifference curves like those shown in the left panel of Figure 13.12, and so will select point **m** implementing the level of risk that maximizes his expected income.

REMINDER The slope of an indifference curve is $-\frac{u_\Delta}{u_{\hat{y}}}$ and a risk neutral person does not care about risk so $u_\Delta = 0$. This is why the indifference curves of a risk neutral person are horizontal (zero slope).

The risk-averse person (with $u_\Delta < 0$) shown in the right panel could also pick any point on the risk return schedule (including **a, c, d,** or m). But finding the Δ at which there is a tangency between her risk return schedule and her highest indifference curve, she will select a lower level of risk, with a lower expected return, at point **a**.



(a) Risk-neutral person

./figures/risk/risk_return_dara.pdf

(b) Risk-averse person

Figure 13.12: **Indifference curves and risk choices of a risk-averse (right) and risk neutral (left) person.** A risk-averse person chooses a point like **a** where their indifference curve (u_2) is tangent to their risk-return schedule; as a result they obtain a bundle of expected income and risk $(\Delta_a, \hat{y}_a)$. The risk choice of a risk-neutral person is Δ_m is with expected income $\hat{y}_m$.

Recall that in Chapter 12 when the lender extended a loan to a borrower under an incomplete contract, the borrower choose a level of risk greater than the expected income maximizing level (this is $\Delta = \Delta_m$ here and was $f = \frac{1}{2}$ in Chapter 12). Here, unless she is risk neutral, the decision-maker chooses a level of risk less than Δ_m. Two differences in the models of the two chapters explain the difference in the choice of risk:

- In Chapter 12 we had not yet introduced risk aversion, so borrowers and lenders alike (and the owner operator too) were risk neutral. This is why the owner operator (with horizontal indifference curves) chose the expected income maximizing level of risk, $f = \frac{1}{2}$.
- In Chapter 12 we studied loan contracts in a legal setting (bankruptcy law and limited liability) such that the lender bore the entire risk of non-repayment of the loan if the project failed. We explained that lending under these circumstances is equivalent to also providing insurance to the borrower. The fact that the lender shared the risk with the borrower is the second reason why borrowers took more risk than would have maximized expected income.

In the setting for this chapter – risk averse actors who are not engaged in loan contracts – for risk levels above Δ_m, additional risk-taking is a lose-lose proposition: it incurs more of the "bad" while reducing the "good".

M-Note 13.1: Choosing a level of risk to maximize utility

The decision maker wants to select a level of risk so as to to maximize her expected utility, subject to the feasible set of risk and return, captured by the risk-return schedule:

$$\text{Vary } \Delta \text{ and } \hat{y} \text{ to maximize} \quad u = u(\Delta, \hat{y}) \tag{13.5}$$

$$\text{subject to} \quad \hat{y} = \hat{y}(\Delta) \tag{13.6}$$

Substituting Equation 13.6 into Equation 13.5, we have a maximization problem in one variable, Δ:

$$\text{Vary } \Delta \text{ to maximize} \quad u = u(\Delta, \hat{y}(\Delta))$$

To find the first order condition for a maximum we differentiate this equation and set the result equal to zero:

$$\frac{\delta u}{\delta \Delta} = u_\Delta + u_{\hat{y}} \cdot \hat{y}_\Delta = 0$$

Which rearranged is:

$$\begin{aligned} \text{Slope of indifference curve} &= \text{Slope of risk return schedule} \\ -mrs(\Delta, \hat{y}) = \frac{-u_\Delta}{u_{\hat{y}}} &= \hat{y}_\Delta = -mrt(\Delta, \hat{y}) \end{aligned} \tag{13.7}$$

This is the condition stated in Equation 13.4.

M-Note 13.2: Choosing risk with explicit utility and risk-return schedule

In M-Note 13.1 we analyzed the general case of Justine's decision between risk and expected income, when she is risk averse. Now, we will give explicit functional forms to her expected utility and risk-return schedule.

We assume that her expected utility function is $u(\hat{y}, \Delta) = \hat{y} - 0.5\Delta^2$. Let us assume that her risk-return schedule can be characterized as $\hat{y}(\Delta) = a\Delta - b\Delta^2$. Therefore, Justine's maximization problem is:

$$\text{Vary } \hat{y} \text{ and } \Delta \text{ to maximize} \quad u = \hat{y} - 0.5\Delta^2 \tag{13.8}$$

$$\text{subject to} \quad \hat{y} = a\Delta - b\Delta^2 \tag{13.9}$$

Plugging Equation 13.9 into Equation 13.8, the problem becomes:

$$\text{Vary } \Delta \text{ to maximize} \quad u = a\Delta - b\Delta^2 - 0.5\Delta^2$$

As before, we differentiate this equation and set the result equal to zero to find Justine's first order condition:

$$\text{FOC} \quad \frac{\delta u}{\delta \Delta} = a - 2b\Delta - \Delta = 0 \tag{13.10}$$

Which rearranged is:

$$\begin{aligned} \text{Slope of indifference curve} &= \text{Slope of risk return schedule} \\ \underbrace{\Delta}_{\text{- mrs}} &= \underbrace{(a - 2b\Delta)}_{\text{- mrt}} \\ -\Delta(1 + 2b) &= -a \\ \Delta^{eq} &= \frac{a}{1 + 2b} \end{aligned}$$

Using 13.9, the expected income is:

$$\hat{y} = a\left(\frac{a}{1+2b}\right) - b\left(\frac{a}{1+2b}\right)^2$$

If we set $a = 200$ and $b = 2$ the risk is $\Delta^{eq} = 40$ and the expected income $\hat{y}^{eq} = 4800$.

13.7 How risk aversion can perpetuate economic inequality

Investments are risky because they involve a fundamental transformation. Before making an investment the decision-maker has money or financial assets that can be readily sold. After the investment the decision-maker owns specific assets – buildings and machinery or other assets such as patents or trade marks – dedicated to the production of particular goods and services. Specific assets are harder to sell than general assets. An investment is therefore a gamble that these specific assets – and the goods or services that they can produce – will be a source of profit for their owner.

EXAMPLE Among the farmers in India an irrigation pump is a specific asset (because it is difficult to sell, once initially installed) while bullocks are more general assets (because there is a ready market in "second hand" these animals.)

The future profitability and hence the value of these goods and services may change dramatically due to unforeseen future events and so the choice to invest is a type of decision-making under uncertainty. As in the case of the choice of a college major or the other risky choices we've studied so far, a person with income to invest typically has a choice to invest in more or less risky assets.

Less risky financial assets include cash or U.S. Treasury Bills (called "T-bills" for short) or U.K government "gilts." These are promises to pay a given amount at a future date, an IOU ("I owe you") from the government. A person could also choose to hold moderately risky stocks in well established firms to highly risky venture capital investments in unknown start-up firms or government issued bonds (like T-Bills) issued by unstable governments that might not honor the promise to pay the IOU.

FACT CHECK The name gilt comes from "gilt edge bond" a reference to the low or zero risk associated with them; the British government has never defaulted on a gilt.

None of these is truly without risk. Cash or the government's obligated payment to the owner of a T-bill or a gilt may change in its value – what it can buy – due to an increase or (less likely, decrease) in prices. For simplicity, let's consider a T-bill to be entirely risk-free, having a certain value of $\hat{y}_1 = c_1$ at some future date (c is for certainty equivalent).

This is shown by the vertical intercept of the indifference curve labeled u_1 in the right panel of Figure 13.12 as the "risk-free return." The quantity $\hat{y}_1 = c_1$ is the certainty equivalent of every point – that is every combination of risk and return – making up the indifference curve u_1.

We now introduce Justine, who is considering what to do with her U.S. Treasury Bill. Her fallback option is to just keep her T-bill. The combinations of an expected return and risk making up the indifference curve u_1 are of equal value to her as her T-bill.This is her fallback indifference curve.

(a) Vicious circle

(b) Virtuous circle

Figure 13.13: **How risk aversion can perpetuate economic inequality: vicious and virtuous circles.** The figure on the left-hand side depicts a *vicious* circle (or cycle) where limited income leads to heightened risk aversion, which results in someone choosing to avoid risky investments, which means they get lower expected returns on average and remain with lower levels of income. The figure on the right-hand side, depicts a *virtuous* circle (or cycle) where substantial income leads to lower risk aversion or risk neutrality, which means they can invest in riskier assets, and obtain higher expected income on average.

Justine's investment options are indicated by the risk-return schedule $\hat{y}(\Delta)$. We now that Justine will find point **a** and choose Δ^a with return $\hat{y}_a$ (a for risk **a**verse).

We know from the left panel in Figure 13.12 if instead of risk-averse Justine we had looked instead at risk-neutral Arjun he would have selected point **m** and implemented a level of risk $\Delta = \Delta_m$ where m, recall stands for the maximum feasible level of expected income $\hat{y}_m$, which is the amount Arjun will receive.

We see two things from this model:

- Justine did very well by investing, increasing her certainty equivalent income from the guaranteed income on the T-bill, y_1 to the certainty equivalent income of her risky investment, namely c_2.
- But wealthy and risk neutral Arjun did much better, with an expected income of $\hat{y}_m$. Because he is risk neutral and his indifference curves are horizontal, this is also his certainty equivalent. Because he is risk neutral he would is indifferent between taking the risk Δ_m or holding a T-bill with no risk at all, as long as the expected income of his risky investment was the same as the certain income of the T-bill. Lacking wealth Justine had a totally different view of the options: had she made Arjun's very risky investment (Δ_m) this would have made her even worse off than in her fallback option (the point **m** lies below Justine's fallback indifference curve u_1).

What this means is that over many investment decisions made by the two, or for a population made up of many Arjuns and many Justines, the average return to investment by Justine (or the class of Justines) will fall short of the average return to Arjun (or his class of rich Arjuns). This is one of the ways that income differences and the societal inequalities associated with them are self-perpetuating:

- *Vicious circle*: A *vicious* circle of low income, risk aversion, avoiding risky investments and low expected income, or
- *Virtuous circle*: A *virtuous* circle of substantial income, risk neutrality (or close to risk-neutrality) investing in risky assets, and substantial expected returns.

Figure 13.13 illustrates these two circles.

The different tales of these two circles need not have anything to do with Justine's or Arjun's basic psychology: they could have had the same utility function, how they differed could have been only their initial income. Had Justine been the one with high income and Arjun the one with low income, Justine's circle would have been virtuous and Arjun's vicious.

13.8 How insurance can mitigate risk and reduce inequality

What can be done to break or mitigate these cycles, and to allow more people to benefit from undertaking risky investments? One answer is: insurance.

Insurance can take many forms other than the familiar car and house insurance. Learning how to code in a widely used programming language – like Python or R – means that you will have job opportunities in many sectors of the economy should your current job end. Acquiring this or some other general skill is a form of *insurance* against risky outcomes. In many countries learning English is also a form of insurance as it expands the range of jobs to which one can apply and even countries in which one could seek employment.

INSURANCE is any costly action one can take that reduces the level of risk to which one is exposed.

Insurance reduces risk exposure

We now see that if insurance is available, risk-averse people will purchase it and as a result be willing to take more risks and to benefit from the higher expected returns associated with riskier investments. The reason is that for any given investment project or other decision, insurance reduces the difference between good and the bad outcomes that the person will experience. Insurance makes the bad outcome not as bad because the person who purchases insurance is compensated for the realization of the bad outcome. Insurance makes the good outcome less good, also: whichever outcome is realized, the payment of the cost of insurance means that there will be less income left over for other expenditures by the the insured.

To see why this is so, consider a person, Juliana, who is making a decision involving risk. In the absence of insurance, as in Justine did in the previous example, Juliana maximized her utility by choosing risk level Δ_a (the **a** is the standard utility-maximizing point for risk-averse Juliana) with corresponding expected income $\hat{y}_a$. This is shown in the left panel of Figure 13.14.

Now introduce an insurance contract, which allows Juliana to "buy" less risk, by paying an amount – called the insurance premium – in reduced expected income in return for a reduction in her degree of risk, Δ. (What is reduced is her income left over for other purchases after paying the insurance premium; for simplicity we call this a reduction in expected income). If she has chosen point **a**, then the opportunity to purchase insurance is shown by what is termed the **insurance line** through that point, the orange upward sloping line in the figure.

From point **a** she can move to any point on that line. She is interested in reducing risk, so she will consider moving to the left on the insurance contract line (which we will call the insurance line, for short). To see how this could work let **a'** be some point on the insurance line other than point **a**, and to the left of **a**. Then instead of her current risk exposure and expected income, by purchasing an amount of insurance s at a price per unit of risk reduction p_s she can have:

INSURANCE LINE For some given price of insurance, p_s the insurance line represents the opportunities for risk reduction and it opportunity cost in terms of foregone expected income (due to the payment of the insurance premium $p_s s$.

$$\begin{aligned} \text{Less risk} \quad \Delta_{a'} &= \Delta_a - s \\ \text{and less expected income} \quad \hat{y}_{a'} &= \hat{y}_a - p_s s \end{aligned}$$

We can re-arrange these equations to find the slope of the insurance line:

$$\begin{aligned} \text{Extent of reduction in } \Delta \quad \Delta_a - \Delta_{a'} &= s \\ \text{Extent of reduction in } \hat{y} \quad \hat{y}_a - \hat{y}_{a'} &= p_s s \end{aligned}$$

From which we see that the slope of the insurance line is

REMINDER Recall the price lines in Chapter 3, this line is similar except only risk has a "price" on the market (p_s). Alternatively, we can think of the price of insurance being the amount of the other good (expected income) that the decision-maker has to give up in order to reduce risk exposure, as we compare points along the price line – the same understanding we already have of a marginal rate of transformation.

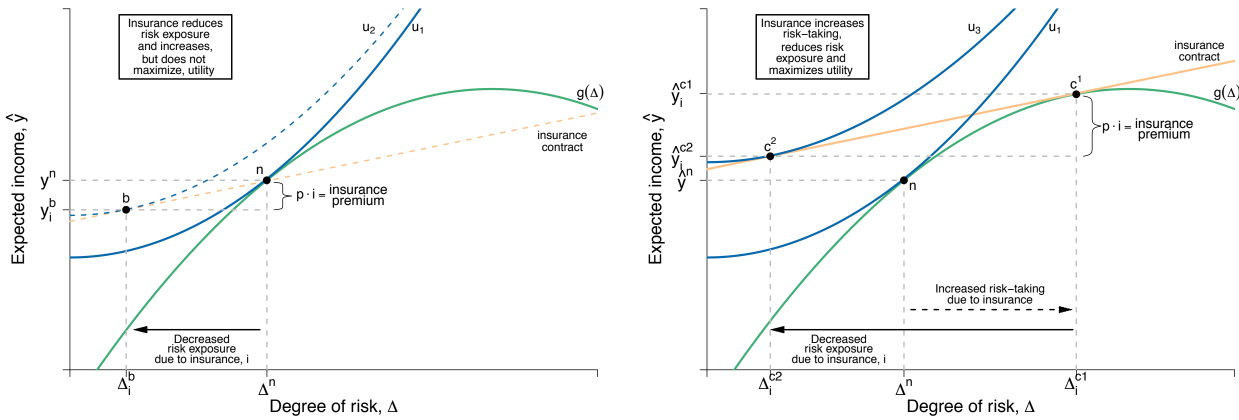
$$\text{Slope of insurance line} \quad \frac{\overbrace{\hat{y}_a - \hat{y}_{a'}}^{\text{Extent of reduction in } \hat{y}}}{\underbrace{\Delta_a - \Delta_{a'}}_{\text{Extent of reduction in } \hat{\Delta}}} = \frac{p_s s}{s} = p_s$$

The slope of the insurance line is the marginal rate of transformation of reduced expected income into reduced risk.

How will insurance affect her choices? You can see from panel a in Figure 13.14 that at her previous choice of a risk level Δ_a Juliana would do better off buying insurance. This is because

$$\text{Slope of indifference curve} = \frac{-u_\Delta}{u_{\hat{y}}} > p_s = \text{Slope of insurance line}$$

Remembering that $-u_\Delta$ is the marginal disutility of risk which is the same thning as how much she would benefit from risk reduction, we can rearrange



(a) Insurance and risk taking that does not maximize utility

(b) Utility-maximizing risk taking and insurance

Figure 13.14: **Effect of insurance on risk-taking and utility.** When insurance is unavailable (which we indicate by the letter **n** for no insurance) the person takes a limited amount of risk and as a result can expect a limited amount of income. The availability of insurance is indicated by the orange "insurance contract" line: an amount of insurance *s* (meaning a reduction in the risk) can be purchased by paying the amount $p_s \cdot s$. The slope of the insurance contract line is p_s so a steeper line means *more costly* insurance. The availability of insurance has two effects: it allows Juliana to reduce her total risk exposure, and because of that it also encourages Juliana to invest in a riskier option, raising her expected income from $\hat{y}^a$ to $\hat{y}_d$.

the above to be:

$$
\begin{array}{rcl}
-u_\Delta & > & u_{\hat{y}} p_s \\
\text{Marginal benefits of risk reduction} & > & \text{Marginal costs of risk reduction}
\end{array}
$$

As before she could choose an investment with risk and expected income of $\Delta_a, \hat{y}_a$, but also buy insurance, and as a result, after taking account of both the insurance premium and the risk mitigation afforded by the insurance, reducing both her expected income and the risk exposure that she experiences.

M-Note 13.3: Choice of insurance starting from a given level of risk and expected return

Suppose $(\Delta_1, \hat{y}_1(\Delta_1))$ is a given level of risk and resulting expected income that Juliana has chosen (such as point **a** in Figure 13.14), with utility $u_1 = u(\Delta_1, \hat{y}_1)$.

When she is given the opportunity to buy insurance, Juliana will choose the level of insurance s that will maximize her utility. Choosing this level s will reduce both her risk (by the amount of insurance she bought, s) and her expected income (by the cost per unit of risk reduction p_s multiplied by the amount of risk reduction s). We index her situation after purchasing the insurance by the subscript 2:

$$
\begin{array}{rrcl}
\text{Experienced risk with insurance} & \Delta_2 & = & \Delta_1 - s \\
\text{Expected income after paying for insurance} & \hat{y}_2 & = & \hat{y}_1(\Delta) - p_s s
\end{array}
$$

Thus, Juliana's maximization problem is as follows:

$$
\text{Vary } s \text{ to maximize} \quad u = u(\Delta_1 - s, \quad \hat{y}_1 - p_s s)
$$

She will find her first order condition for a maximum by differentiating this expression her with respect to s and setting the result equal to zero:

$$\frac{du}{ds} = -u_\Delta - u_{\hat{y}} p_s = 0$$

which, rearranged, means that

$$\text{slope of indifference curve} = -mrs = \frac{-u_\Delta}{u_{\hat{y}}} = p_s = -mrt = \text{slope of insurance line}$$

As a result, Juliana moves down and to the left along the insurance line, mitigating the risk of the investment she chose, and reducing her expected income by an amount equal to her insurance premium. How much insurance will she purchase? She should buy an amount such that the marginal benefit of further risk reduction is neither greater than nor less than the marginal cost. In other words she should buy the amount of insurance such that

$$\text{Slope of indifference curve} = \frac{-u_\Delta}{u_{\hat{y}}} = p_s = \text{Slope of insurance line}$$

From this rule, she will find that the bundle $(\Delta_b, \hat{y}_b)$ is the best she can do if she chooses the investment at point **a**, bringing her to indifference curve u_2.

Checkpoint 13.3: Buying risk

In panel a of Figure 13.14 to explain why Juliana would not prefer some outcome to the right of point **a** along the dashed portion of the insurance line, rather than to the left on the solid portion of the line.

Insurance encourages risk taking

But she could do even better if she reconsidered her initial choice of Δ_a jointly buying insurance and choosing more risk than Δ_a.

How much additional risk should she take, and how much insurance should she buy? To answer this question it is important that the insurance line is not unique to point **a**: from any point that Juliana chooses on the risk-return schedule, this line shows her opportunities to move to less risky states by buying insurance. Her decision to make an even riskier choice and buy more insurance is shown in panel b. of Figure 13.14. The answer is given in two steps:

- point c is her choice of a risk level and expected income (before paying for insurance) that results from that choice, and
- point d is where she will be as the result of her decision to take the risk Δ_c and to purchase an amount s of insurance at a price p_s for a total insurance cost of $p_s \cdot s$ and move to Δ_d after purchasing insurance (and therefore increasing her utility from u_1 at the no-insurance outcome to u_3 taking additional risk and buying insurance).

In this two-step process, "doing the best she can" requires finding two tangencies, not just one as in our usual case so far.

- *First tangency*: To determine her *choice of risk* (point c) she equates the marginal rate of transformation of risk taken into expected income (the slope of the risk-return schedule) to the marginal rate of transformation of premium paid (reduced expected income after paying for insurance) into reduced risk exposure. That is, she finds where the insurance contract line is tangent to the risk-return schedule.
- *Second tangency*: To determine *how much insurance to buy* (point d) she equates the marginal rate of substitution between risk and expected income (the slope of the indifference curve) to the marginal rate of transformation of insurance premium paid into reduced risk exposure (Δ_d). That is, she finds where her indifference curve is tangent to the insurance contract line.

If insurance were more expensive – higher p_s – the insurance contract line would be steeper and she would choose a lower risk level and purchase less insurance. If the price of insurance were so high that the insurance contract line was as steep as the risk-return schedule at point that she chose, then she would buy no insurance at all.

This case illustrates a broader point about the economy as a whole, which we return to later in this chapter: when insurance is available, people are able to take more risks and, on average, enjoy greater expected income. This win-win outcome is possible because Juliana made a risky and high expected income choice, but she could also transfer some part s of her resulting risk exposure to the insurer.

FACT CHECK Using a seat belt while driving in a car is another form of insurance: it reduces the difference between consequences of the good and bad outcome (no crash, crash). Consistent with the main lesson of this section, there is some evidence (see Chapter 16) that drivers wearing seat belts drive faster

In order for Juliana to be buying insurance, someone had to be selling. Could that be Juliana herself or someone like her? Would she be willing to accept greater risk exposure for herself in return for an increase in her expected income? You can see from Figure 13.14 that at the price p_s shown as the slope of the insurance line, Juliana would have no interest in selling insurance. The reason is that selling insurance would mean moving to the right up the insurance line and reaching ever *lower* indifference curves.

In not Juliana, then who?

M-Note 13.4: Utility-maximizing risk taking and insurance

As we saw graphically, Juliana can increase her utility by reconsidering her initial choice of risk, instead jointly buying insurance and choosing a higher level of risk. Using notation similar to M-Note 13.3, subscripts 1 and 2 refer to before and after buying insurance.

Given price p_s, risk schedule $\hat{y}(\Delta)$, and utility function $u(\Delta, \hat{y})$, Juliana will vary her initial risk level Δ_1 and her final risk level after buying insurance Δ_2 to maximize her utility.

The expected income after paying from insurance will be:

$$\hat{y}_2 = \hat{y}(\Delta_1) + p_s(\Delta_2 - \Delta_1)$$

To maximize utility, then, Juliana's maximization problem becomes:

$$\text{Vary } \Delta_1 \text{ and } \Delta_2 \text{ to maximize} \quad u(\Delta_2, \hat{y}_2) = u(\Delta_2, \hat{y}(\Delta_1) + p_s(\Delta_2 - \Delta_1))$$

To derive the optimum, we find her first order conditions of utility maximization:

$$\begin{aligned}
\text{First FOC} \quad \frac{\partial u}{\partial \Delta_1} &= u_{\hat{y}}(\frac{d\hat{y}}{d\Delta} - p_s) = 0 \\
\text{Rearranging} \quad \frac{d\hat{y}}{d\Delta}(\Delta_1^*) &= p_s \qquad (13.11) \\
\text{Second FOC} \quad \frac{\partial u}{\partial \Delta_2} &= u_\Delta + u_{\hat{y}} p_s = 0 \\
\text{Rearranging} \quad -\frac{u_\Delta}{u_{\hat{y}}}(\Delta_2^*) &= p_s \qquad (13.12)
\end{aligned}$$

Thus, combining Equations 13.11 and 13.12 we have:

$$\overbrace{\frac{d\hat{y}}{d\Delta}(\Delta_1^*) = p_s}^{\text{slope of risk-return schedule = slope of insurance line}} \underbrace{= -\frac{u_\Delta}{u_{\hat{y}}}(\Delta_2^*)}_{\text{slope of insurance line = slope of indifference curve}}$$

This tells us, referring to Figure 13.14, that Juliana will choose:

- *First tangency, point **c***: the risk level Δ_1 such that the slope of the risk-return schedule equals the slope of the insurance line
- *Second tangency, point **d***: the risk level Δ_2 after choice of insurance such that the slope of the utility function equals the slope of the insurance line

Checkpoint 13.4: Insurance and risk taking

Return to the planting in Palanpur problem and consider what would happen if an insurance company offered for a fee to fully compensate any farmer who planted early and lost his seeds to the birds (paying the value of the crop that would have resulted had the birds not intervened), charging a premium (fee) for this insurance policy sufficient to cover the insurance company's expected payments to the farmers.

- Is there some level of the premium at which the farmers would buy the insurance and plant early?
- Could the insurance company make a profit on this offer?
- If all the farmers had purchased insurance and planted early, would they purchase insurance again the next year?

13.9 *Buying and selling risk: Two sides of an insurance market*

We now consider Juliana's interaction with a wealthy person, Konstantin, who initially is exposed to zero risk (maybe his only form of wealth is U.S. T-bills). He is therefore not very risk averse; he is close to risk neutral. Juliana is poor and exposed to risk, with a good and a bad state affecting her realized income and occurring with equal probability. As a result of both limited income and a high level of risk exposure Juliana is highly risk averse.

Insurance: Buying and selling risk

The two may engage in an exchange to alter the distribution of expected income and risk between them. We can represent their interaction using the Edgeworth box that you encountered in Chapter 4 . But there is a difference: in the Edgeworth boxes you previously studied, what was being allocated was two goods – coffee and data. In the case we now consider two people - Konstantin and Juliana - who are exchanging a good - expected income - and a "bad" - risk. In Figure 13.15 we present the Edgeworth box with Juliana's and Konstantin's indifference curves.

In the initial state with no exchange between the two (point **z**) the (expected and realized) income of Konstantin is $\hat{y}_z^K = y_z^K$ (he faces no risk, so his expected income and his realized income are equal). Juliana's expected income is $\hat{y}_z^J$ and her realized income is $y^J = y_z^J + 0.5\Delta_z$ in the good state and $y^J = y_z^J - 0.5\Delta_z$ in the bad state. Recall, that the good and bad states are equally likely, so her expected income is $\hat{y}_z^J$.

The dimensions of the box are the total amount of expected income that the two will jointly experience ($\bar{\hat{y}} = \hat{y}^J + \hat{y}^K$) and the total amount of risk to which the two will be exposed ($\bar{\Delta} = \Delta^J + \Delta^K$ which is $\bar{\Delta} = \Delta_z^J$ at Juliana's endowment (at point **z**) because prior to their exchange she bears all the risk and Konstantin bears none.).

Because as before risk is a bad and expected income is a good, the indifference curves slope upwards. For example Juliana is indifferent between two possible allocations indicated by point **b**, namely, exposed to a lower level of risk (Δ_b^J) along with a lower level of expected income ($\hat{y}_b^J$) and point **c**, being exposed to more risk (Δ_c^J) and a higher level of expected income ($\hat{y}_c^J$).

The slope of each of her indifference curves is as before a measure of her degree of risk aversion. The steeper the indifference curve, the more expected income she is willing to give up to reduce the amount of risk to which she is exposed. She would prefer any point above and to the left of point **z** – less risk and more expected income – but that would make Konstantin worse off (less utility than his participation constraint u_z^K), so she will not have that opportunity.

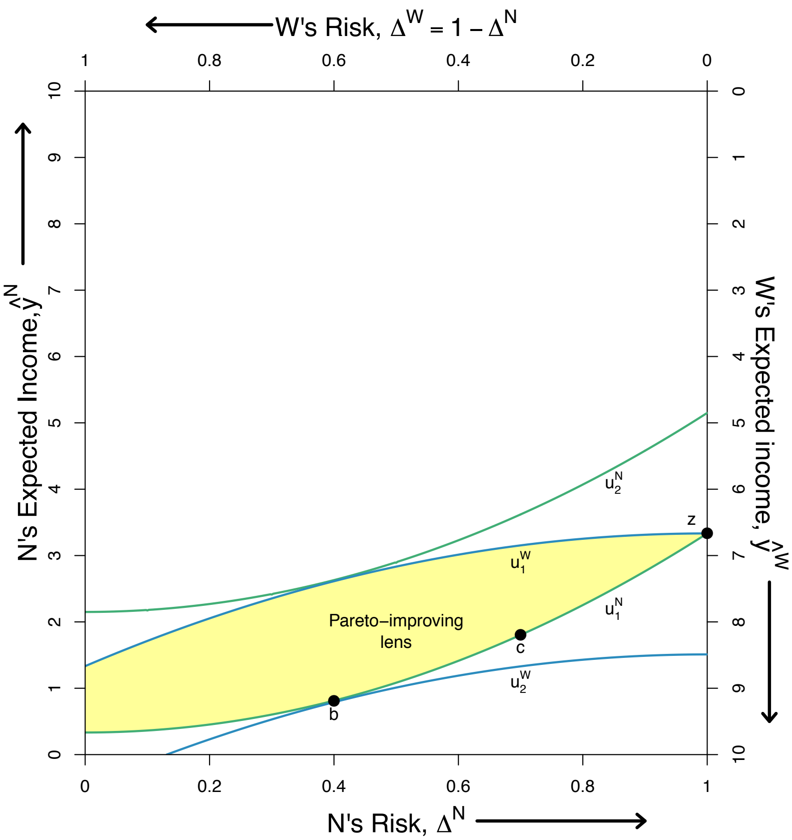


Figure 13.15: **Feasible allocations of expected income and risk and indifference curves over these allocations.** Every point in the Edgeworth box is a possible allocation that divides the total amount of expected income ($\bar{y}$) and risk ($\bar{\Delta}$) among the two people. Konstantin (blue indifference curves) prefers allocations lower and to the right; Juliana (green indifference curves) prefers allocations higher and to the left. So there are conflicts of interest in comparing higher points on the left with lower points on the right. But there are some mutually beneficial re-allocations, starting at point **z** and comparing points down and to the left as indicated by the shaded Pareto-improving lens.

A Pareto-improving insurance contract

There are, however, opportunities for a mutually beneficial bargain. In Figure 13.15 we can see that the endowment point **z** is not Pareto-efficient because at that point the indifference curves of the two are not tangent, they intersect. Therefore, there are other allocations of risk and expected income that both would prefer to point **z**. The shaded area indicates all of these win-win allocations. In them Konstantin takes over some of the risk exposure in return for Juliana transferring to him some of her expected income. In other words Konstantin sells Juliana some insurance.

REMINDER Remember from Chapters 4 and 5 an allocation is Pareto efficient (and therefore will be a point on the Pareto-efficient curve) if at that point the two participants' marginal rates of substitution are equal,meaning that their indifference curves are tangent to each other.

The agreement between the two to move to an allocation in the shaded area would take the form of an insurance contract: Juliana would give up some of her income and Konstantin would take on some of her risk exposure. Figure 13.16 provides the game tree illustrating the interaction.

The game begins with Konstintin offering a price per unit of risk reduction that Juliana wishes to purchase (we do not analyse why he offers this particular price.) Prior to the good or bad state having been realized, Juliana either rejects the offer (she takes the right branch of the tree) or accepts and (she

REMINDER: PRICE SETTING To determine the price he would set, we could use the models of first mover advantage (price setting) and "take it or leave it power" studied in Chapter 4 or the Cournot model of competition among those wishing to sell insurance to Juliana studied in Chapter 9 figures/risk/sequential_risk.pdf

Figure 13.16: **A game tree explaining the sequence of the insurance contract**. Konstantin offers a price for insurance, p_s. Juliana can accept or reject that price of insurance. In either case of purchasing the insurance or not, she will face a good or bad state. The expressions at the bottom nodes of the tree are the realized outcomes in the bad and good states for Konstantin (top row) and Juliana (bottom row.)

takes the left branch.)

In the latter case she buys her chosen amount of insurance $\bar{\Delta} - \Delta^J$, where Δ^J is the risk exposure that Juliana will be subjected to under this contract, paying a total of $p(\bar{\Delta} - \Delta^J)$. Because the total amount of risk exposure of the two is $\bar{\Delta}$, it follows that Konstantin's risk exposure at the post-exchange allocation is $\Delta^K = \bar{\Delta} - \Delta^J$. Δ^K is the amount of insurance that Juliana receives from Konstantin – it is risk to which she initially was (but is no longer) exposed. That part of her initial risk is now his (as shown in M-Note 13.5).

After the state is revealed, Juliana pays $0.5(\bar{\Delta} - \Delta^J)$ to Konstantin if the good state has occurred while Konstantin pays $0.5(\bar{\Delta} - \Delta^J)$ to Juliana if instead the bad state has occurred.

Table **??** summarizes the realized income in the good and bad states for the two people when the insurance contract is implemented. Table 13.1 gives the expected income and risk exposure in the initial situation and the situation following the implementation of the insurance contract.

Figure 13.17 captures the interaction. Dashed line $p_s z$ illustrates Konstantin's price offer (we do not analyse why he offers this particular price.) Juliana is constrained to a point somewhere along the purple insurance line, including point **z**, meaning reject the offer. The point **s** is on the highest indifference curve that is available to her at this price. Her indifference curve at point **s** is tangent to the insurance line, meaning that her marginal rate of substitution is equal to the the marginal rate of transformation (the slope of the insurance line, or p_s) She picks point **s**, pays $p_s(\bar{\Delta} - \Delta_s^J)$ and receives $\Delta_s^K = \bar{\Delta} - \Delta_s^J$ insurance.

Allocation **s** represents a Pareto-improvement over initial endowment **z**. But, while both Juliana and Konstantin are better of as a result off exchange, the exchange has increased inequality of expected income. On the other hand, the exchange reduced inequality between the two in risk exposure.

Checkpoint 13.5: Risk preferences and the price of insurance

In Figure 13.17 we found a price of insurance p_s and a post-exchange allocation **a** at which Juliana bought an amount of insurance $s = p_s \Delta_a$.

Consider the shape of indifference curves and re-draw the Edgeworth box while considering two other actors (different to Juliana):

1. Vladimir is *more risk averse* than Juliana. What will his indifference curves look like? At the same price of insurance, p_s, will he have a higher or lower cost $(p_s(\bar{\Delta} - \Delta_s^V)$ of insurance than Juliana? Explain.
2. Carla is *less risk averse* than Juliana. What will her indifference curves look like? At the same price of insurance, p_s, will she have a higher or lower cost of insurance $(p_s(\bar{\Delta} - \Delta_s^C)$ than Juliana? Explain.

M-Note 13.5: Exchanging risk

Having purchased insurance, Juliana will experience a good state and a bad state as follows. We first consider the good state, which occurs with a 50% probability and experi-

Person	*Before insurance: risk and expected income*	*After insurance: risk and expected income*
Konstantin: rich not risk exposed, less risk averse,	Risk: $\Delta_z^K = 0$	$\Delta_s^K = \overline{\Delta} - \Delta_s^J$
	Expected income: y_K^z	$\hat{y}_z^K + p_s \Delta_s^K$
Juliana: poor risk exposed, more risk averse,	Risk: $\Delta_z^J = \overline{\Delta}$	Δ_s^J
	Expected income: $\overline{y}_z^J$	$\hat{y}_z^J - p_s \Delta_s^K$

Table 13.1: **Risk exposure and expected income in the initial state and following implementation of the insurance contract**.

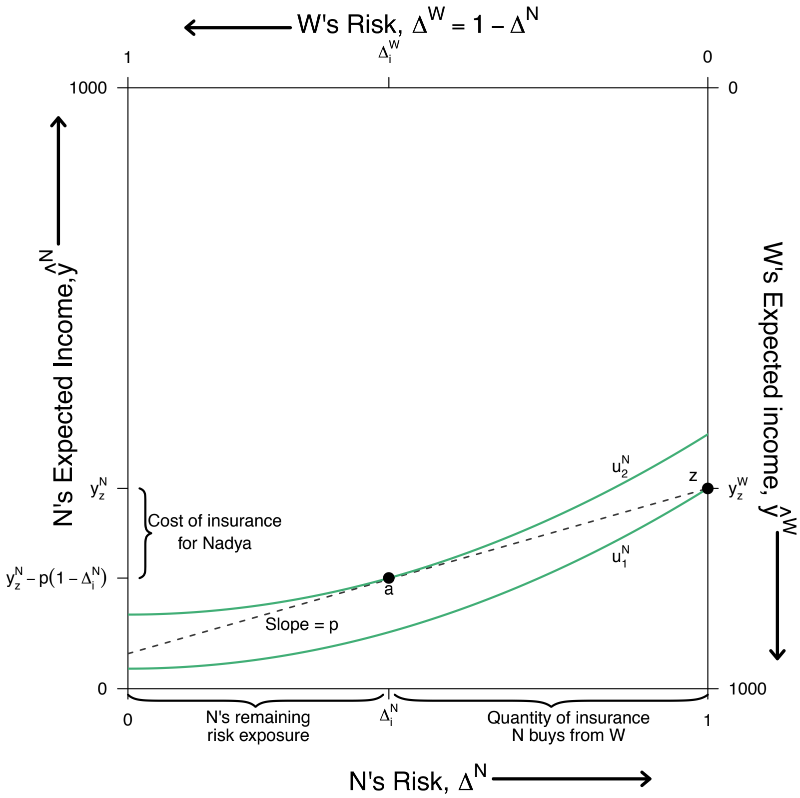


Figure 13.17: **Buying and selling risk.** Konstantin has price-setting power, sets a price equal to the slope of the line $p_s z$. Juliana, who may choose any point on that line, picks **s** which is on the most preferred of her indifference curves that is feasible: it is the highest indifference curve (u_2) she can obtain given the price line which constrains her given the price Konstantin stipulates (p_s). Juliana has given up some expected income to purchase insurance in the form of reduced risk exposure.

encing good state income y^{JG}:

$$\begin{aligned} y^{JG} &= y_z^J - \overbrace{p_s(\bar{\Delta} - \Delta_s^J)}^{\text{Insurance cost}} \overbrace{+0.5\bar{\Delta}}^{\text{Gain from good state}} \overbrace{-0.5(\bar{\Delta} - \Delta_s^J)}^{\text{Payment to K in good state}} \\ &= y_z^J - p_s(\bar{\Delta} - \Delta_s^J) + 0.5\bar{\Delta} - 0.5\bar{\Delta} + 0.5\Delta_s^J \\ &= y_z^J - p_s(\bar{\Delta} - \Delta_s^J) + 0.5\Delta_s^J \end{aligned}$$

But she also has a 50% chance of having a bad state and experiencing bad state income y^{JB}:

$$\begin{aligned} y^{JB} &= y_z^J - \overbrace{p_s(\bar{\Delta} - \Delta_s^J)}^{\text{Insurance cost}} \overbrace{-0.5\bar{\Delta}}^{\text{Loss from bad state}} + \overbrace{0.5(\bar{\Delta} - \Delta_s^J)}^{\text{Insurance payment from K in bad state}} \\ &= y_z^J - p_s(\bar{\Delta} - \Delta_s^J) - 0.5\bar{\Delta} + 0.5\bar{\Delta} - 0.5\Delta_s^J \\ &= y_z^J - p_s(\bar{\Delta} - \Delta_s^J) - 0.5\Delta_s^J \end{aligned}$$

The risk that she experiences is $y^{JG} - y^{JB}$ (the difference in her income between the good and bad states):

$$\begin{aligned} y^{JG} - y^{JB} &= y_z^J - p_s(\bar{\Delta} - \Delta_s^J) + 0.5\Delta_s^J - \left(y_z^J - p_s(\bar{\Delta} - \Delta_s^J) - 0.5\Delta_s^J\right) \\ &= 0.5\Delta_s^J + 0.5\Delta_s^J = \Delta_s^J \end{aligned}$$

For Konstantin, we can also consider what his income looks like in the good and bad states when he sells insurance to Juliana:

$$y^{KG} = y_z^K + \overbrace{p_s(\bar{\Delta} - \Delta_s^J)}^{\text{Insurance payment received}} + \overbrace{0.5(\bar{\Delta} - \Delta_s^J)}^{\text{Payment from J in good state}}$$

And his income in the bad state is the following:

$$y^{KB} = y_z^K + \overbrace{p_s(\bar{\Delta} - \Delta_s^J)}^{\text{Insurance payment received}} - \overbrace{0.5(\bar{\Delta} - \Delta_s^J)}^{\text{Payment to J in bad state}}$$

As a result of the insurance, Konstantin will be exposed to the following risk (the difference between his income in the good and bad states, $y^{KG} - y^{KB}$):

$$\begin{aligned} y^{KG} - y^{KB} &= y_z^K + p_s(\bar{\Delta} - \Delta_s^J) + 0.5(\bar{\Delta} - \Delta_s^J) - \left(y_z^K + p_s(\bar{\Delta} - \Delta_s^J) - 0.5(\bar{\Delta} - \Delta_s^J)\right) \\ &= 0.5(\bar{\Delta} - \Delta_s^J) + 0.5(\bar{\Delta} - \Delta_s^J) = (\bar{\Delta} - \Delta_s^J) = \Delta_s^K \end{aligned}$$

13.10 Application: Free tuition with an income-contingent tax on graduates

HISTORY Bill Gates, who is said to have risked it all by dropping out of Harvard to found Microsoft, explained that he actually did not quit Harvard but took a leave of absence so that he could return if necessary. He also had enough money so it would not be a problem if Microsoft had failed. Think how much innovation we might get if everyone had such a good fallback position.

The term "risky investment" is associated with headline grabbing disasters like the insurance giant AIG's (American International Group) business model prior to its crash in 2008. But one of the riskiest investments of all is investing in yourself.

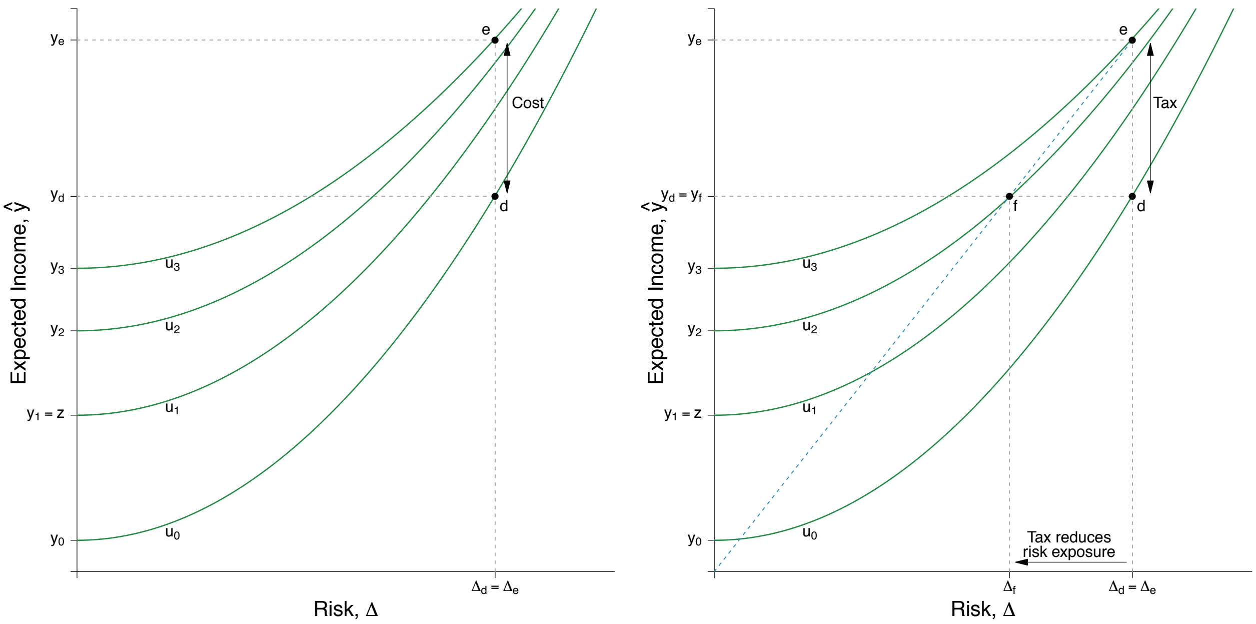
Yourself: A risky investment

The decision to attend a particular university and to study a particular subject is made while uncertain about:

- Do I have the talent and discipline to do well studying this subject?
- Will I find it interesting enough to pursue a career in a related area?
- Will instructors at the university I attend teach me well, and certify that I am competent in my field?
- Will there be a demand for the skills (and credentials of them) I will acquire as a result of my decisions?

Then there is the cost, both directly to pay tuition and accommodation at the university and the opportunity cost of being a student, that is, forgoing the salary or other income that one would have gained if you had not attended university. Most people are not rich enough to be even close to risk neutral, and as we saw in Chapter 12 a great many families, even in high income countries are unable to borrow the substantial sums required to pay for higher education privately. So unless higher education were provided or subsidized by government or private philanthropic contributions few people would attend university.

To understand why, consider Sofia, who has the opportunity to attend university without cost. Sofia has completed a two-year degree as a medical



(a) Education that is paid privately or free

(b) Education Choice with a Tax

Figure 13.18: **Effect of income-contingent taxation of graduates on choices concerning further education.** In the left panel we show Sofia's indifference curves and her possible choices: taking a zero-risk job now (no further education) at income $\hat{y}_1 = y_a$, point **a** or undertaking additional education along with additional risk exposure, shown if she pays no costs by point **e** , and by point **d** if she has to pay the cost of her education. From the indifference curves you can see that she prefers taking the job now (no further education) if she has to pay the cost, that is, $u_1 > u_0$. But she would prefer to continue with her education if it were free because $u_3 > u_1$ In the right panel, we show the case in which tuition is free but, following graduation, she will pay a tax proportional to her income. Under the income-contingent tax her expected income and risk exposure are reduced in the same proportion (along the blue line through point **f**). So instead of being at the simple free tuition (no tax) outcome **e** she is now at point **f** on indifference curve u_2. Because $u_2 > u_1$ she will continue her education.

technician and is considering two options:

- *No risk, certain job and salary*: taking a job with a certain salary, which would be supplemented by the annual interest on a modest financial asset that she has, giving her a total income of $\hat{y}_1 = y_a$ (**a** for the certain option available to her now and it does not have a hat because it is certain), or
- *Continue education, bear risk, higher expected income*: continuing her education and as a result increasing her expected income to $\hat{y}_e$ while also being exposed to risk, the amount of which is Δ, the difference between her income in the good and the bad state.

If she chooses to get more education, the good state might be described by a positive answer to the four bulleted questions about yourself as a risky investment. The bad state could be a negative answer to the questions. As before, we let the good and the bad state be equally likely to occur. We also assume that the realized income in the bad state is less than the certain income she would have were she to decide not to continue her education (otherwise she might just ignore the risk, as she could be certain to have a higher income – even in the bad state – by going to university.)

In Figure 13.18, point **e** shows her expected income and risk exposure if she continues her education when education is free (paid for by the government from general tax revenues), point **d** corresponds to when she pays for her education privately. If instead she chooses *not* to continue her education, she would receive a certain income of $y_1 = y_a$ and be exposed to no risk, indicated by point **a** in the figure. Her choices are mutually exclusive. If she chooses to

Outcome (Point in Figure 13.18)	**Free tuition (e)**	**Private cost (d)**	**Free tuition & income-contingent tax on graduates (f)**
Good outcome	$y_e^G = y_e + 0.5\Delta_0$	$y_d^G = y_e + 0.5\Delta_0 - c$ $= y_d + 0.5\Delta_0$	$y_T^G = (\hat{y}_e + 0.5\Delta_0)(1-\tau)$
Bad outcome	$y_e^B = y_e - 0.5\Delta_0$	$y_d^B = y_e - 0.5\Delta_0 - c$ $= y_d - 0.5\Delta_0$	$y_T^B = (\hat{y}_e - 0.5\Delta_0)(1-\tau)$
Expected income	$\hat{y}_e = \frac{y_e^G + y_e^B}{2} = \frac{\hat{y}_e + 0.5\Delta_0 + \hat{y}_e - 0.5\Delta_0}{2}$	$\hat{y}_d = \frac{y_d^G + y_d^B}{2} = \frac{\hat{y}_d + 0.5\Delta_0 + \hat{y}_d - 0.5\Delta_0}{2}$	$\hat{y}_f = \frac{y_T^G - y_T^B}{2} = \frac{(\hat{y}_e + 0.5\Delta_0)(1-\tau) + (\hat{y}_e - 0.5\Delta_0)(1-\tau)}{2} = \hat{y}_e(1-\tau)$
Experienced risk exposure	$y_e^G - y_e^B = \hat{y}_e + 0.5\Delta_0 - (\hat{y}_e - 0.5\Delta_0)$ $\Delta_0 = \Delta_e$	$y_d^G - y_d^B = \hat{y}_d + 0.5\Delta_0 - (\hat{y}_e - 0.5\Delta_0)$ $\Delta_0 = \Delta_d$	$y_f^G - y_f^B = \hat{y}_e + 0.5\Delta_0(1-\tau) - (\hat{y}_e - 0.5\Delta_0)(1-\tau)$ $= \Delta_0(1-\tau)$ $= \Delta_f$

Table 13.2: **Comparison of the free, privately paid, and income-contingent tax policies for education**. Point **a** is not shown in the table as the risk (Δ_a) is zero and the only state is the certain income y_a. We use the y-subscript T to indicate that results under the income contingent tax when she has continued her education. G and B refer to the good and bad states respectively.

continue her education, then point **a** – the job that she has currently been offered – is no longer available to her. (Her options as just described are summarized in Table 13.2.)

How would Sofia compare the outcomes available to her? This is where the indifference curves come in. If Sofia were wealthy, she might not be concerned about risk exposure – like Bill Gates deciding to leave Harvard – but having a modest income and being limited in how much she can borrow, she is risk averse, as the upward slope of her indifference curves indicate.

Her indifference curve u_1 through point **a** gives the combinations of risk exposure and expected income that Sofia prefers equally to the certain income of y_a. Therefore, y_a is the certainty equivalent of the combination of expected income and risk given by every point on the indifference curve labeled u_1. Comparing points **a** and **e** and the indifference curves on which they appear, we can see that if tuition is free she will choose to undertake further education ($u_3 > u_1$).

But other citizens might object: it was they who payed the taxes that allowed the government to provide Sofia's education for free. Why should taxpayers subsidize Sofia's investment in herself? As a result of higher education, Sofia would have a higher expected income, but taxpayers who had not attended university would have a lower income (having subsidized Sofia's education). It does not appear to be fair to those not as fortunate as Sofia who has already attended 14 years of schooling.

Let us therefore consider the case in which Sofia pays the cost of her education. As a result her expected income is now: expected income with free tuition minus the costs of her education. The difference between her incomes in the good and bad state is not affected – her realized income in the two states is just reduced by the private cost of her education, as shown by point **d**. So her exposure to risk is unaffected: it is still $\Delta_e = \Delta_0 = \Delta_d$. Point **d**, however, lies on indifference curve u_0. If she has to pay the cost of her tuition she would be better off not pursuing further education because u_0 her utility if she pays for her education is less than u_1, her level of utility when she takes the risk free job.

Income-contingent taxation of graduates

An alternative way of funding higher education that would reduce the risk exposure of those pursuing further studies, while at the same time addressing the unfairness of general tax payers subsidizing the advancement and higher income prospects of university graduates has been proposed: Here is the idea: let tuition be free when the student attends university, but then impose a tax on graduates that is based on the income that they actually receive (their *realized* income not their *expected* income) after they graduate.

One version of this idea would have the tax be proportional to income and set at a rate such that the taxes collected from graduates would on average pay for their education. This would require that graduates who experienced a bad state (meaning a low income) would pay in taxes an amount less than the cost of their education, while those who experienced a good state would pay more than what their education cost.

Adding the taxes paid by the unlucky and the lucky graduates, the total revenue collected would totally cover the cost of their education. We call this kind of policy *income-contingent taxation of graduates*.

INCOME-CONTINGENT STUDENT LOANS
In many countries students can finance the expenses of higher education by borrowing money with the unusual feature that the extent of repayment will be less, the lower is the income of the graduate, so that only those with substantial incomes fully repay. The shortfall due to less than full repayment by lower income graduates is made up by tax revenues.

Under the income-contingent tax, her income in each state, and her expected income would be as indicated in the third column of Table 13.2. The policy allows Sofia to trade away some of her gain in expected income were she to continue her education in order to reduce the degree of risk to which she would be exposed if she chose to invest in herself.

This is shown in the right panel in Figure 13.18. The government would set the the level of the tax to collect total revenues sufficient to cover the full costs of the education that people had decided to pursue. The tax rate as a fraction of income, τ, that would accomplish this is equal to the costs of her education as a fraction of her before-tax expected income as shown in M-Note 13.6.

The blue line through point **e** shows the effects of differing tax rates on reducing both the after tax expected income and risk exposure of the graduates. For any tax rate we know from Table 13.2 that

$$\text{Risk exposure with tax:} \quad \Delta_0(1-\tau)$$
$$\text{Expected income with tax:} \quad \hat{y}_e(1-\tau)$$

Because the tax reduces both expected income and risk by the same proportion this means that increasing the tax rate will move the resulting allocation downward to the left of point **e** along the blue line. The tax rate the government will implement is the one that collects in taxes an amount equal to the cost of education. This is indicted by point **f**.

With the income-contingent taxation of graduates, Sofia now has a choice between point **a** and point **f**. Because **f** is on a higher indifference curve than point **a** Sofia will continue her education, Under the income-contingent tax plan, Sofia "purchases" a bundle that includes two years additional of higher education, expected income before taxes of $\hat{y}_e$, an amount $\tau\hat{y}_e = c$ in expected taxes (meaning what she would pay averaged over the bad and good state) and a reduced level of risk exposure $\Delta_0(1-\tau) = \Delta_f$. She is be better off than had she taken her risk-free job instead.

M-Note 13.6: The tax rate and after tax income in the good and bad

The tax collected from graduates on the average will be $\tau\hat{y}_e$ and this amount will have to cover the cost of their education c. So $\tau\hat{y}_e = c$ which means that $\tau = \frac{c}{\hat{y}_e}$.

Sofia gets tuition-free education and in the the good state she obtains income $y_e + 0.5\Delta$ minus the tax she must pay on her income which is $\tau(y_e + 0.5\Delta)$ as in Table 13.2:

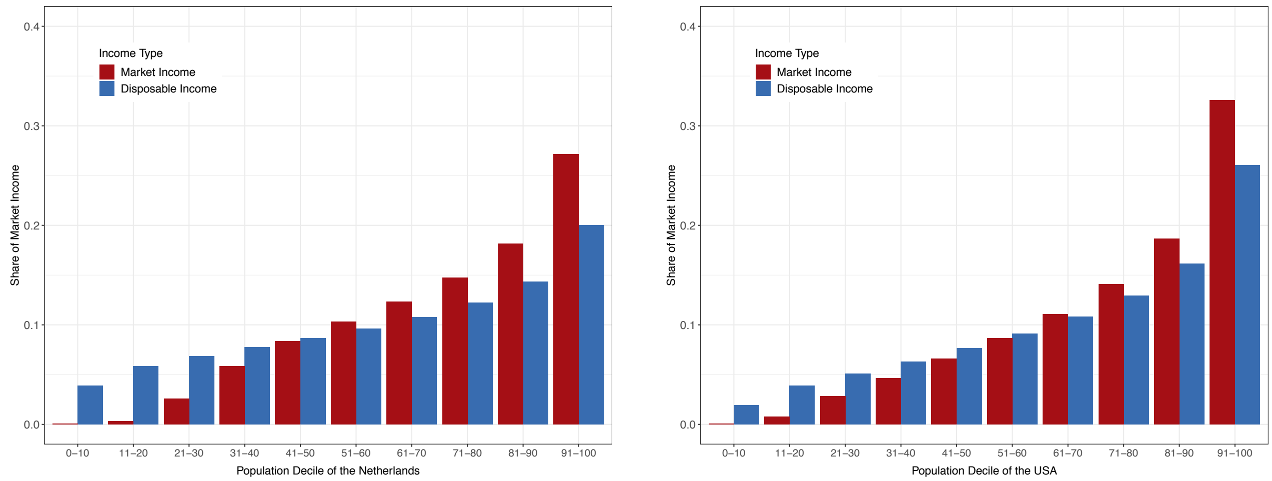
$$\begin{aligned} y_T^G &= \overbrace{y_e + 0.5\Delta}^{\text{Income in the good state}} - \overbrace{\tau(y_e + 0.5\Delta)}^{\text{Tax payment in good state}} \\ y_T^G &= (y_e + 0.5\Delta)(1-\tau) \end{aligned} \tag{13.13}$$

Similarly, if she experiences the bad state, then her income will be the following:

$$\begin{aligned} y_T^B &= \overbrace{y_e - 0.5\Delta}^{\text{Income in the bad state}} - \overbrace{\tau(y_e - 0.5\Delta)}^{\text{Tax payment in bad state}} \\ y_T^B &= (y_e - 0.5\Delta)(1-\tau) \end{aligned} \tag{13.14}$$

Checkpoint 13.6: But is it fair?

Imagine that it is five years after Sofia's decision to continue her education under the income contingent taxation of graduates policy. She graduated with honors and landed a well paying job. Is it fair that Sofia is now paying in taxes much more than her education cost? One way of answering this is to return to 5 years earlier, and ask Sofia's opinion of the policy before she made her choice and before she found out that she would end up in the good state (high salary) rather than the bad state. If Sofia had had the option to pay up front the entire costs of her own further education and in return been exempted from the tax, what choice would she have made? Use Figure 13.18 to answer the question, and explain how your answer contributes to a reply to the larger question: "is the income contingent tax on graduates fair?"



(a) Netherlands Income Shares by Decile

(b) US Income Shares by Decile

Figure 13.19: **Inequality of disposable income and market income in the Netherlands and the U.S.**

13.11 Another form of insurance: A linear tax and lump sum transfer

The income-contingent graduates' tax mitigated the riskiness that Sophia faced in choosing to continue her education. Like deciding whether to attend university and if so what subject to major in, or to invest in bullocks rather than irrigation pumps, deciding on a policy of taxation and spending by a government involves trade offs between expected income and risk. So, what you have learned about risk aversion and insurance gives you an insight into the sometime controversial economics and politics of taxation and redistribution.

Taxes, transfers, and redistribution: A look at the data

Countries differ in the extent to which government taxes and payments to citizens reduce the degree of inequality in what families and people are able to spend. Figure 13.19 provides an illustration using data from the Netherlands and the U.S.

The horizontal axis in both panels refer to deciles of the population from to the poorest ten per cent (on the left) through ever richer segments of the population to the richest ten per cent. The height of the bars shows what fraction of the total income of the country is received by each of these groups of 10 percent (called deciles). The blue bars show the share of each decile in what is termed market income, that is, income before taxes are paid or payments from the government received. The red bars show the income shares of the ten deciles in what is termed disposable income, that is, income after the payment of taxes and receipt of government payments.

The blue bar on the far right of the right panel, for example, shows that the top ten percent of income recipients in the U.S. receive about a third of all of the market income; the farthest left blue bar in the same panel shows that the poorest ten percent in the U.S. receives about 2 percent of all market income.

Studying the two panels you can see:

- *Redistribution*: the effect of taxes and transfers is to reduce the disposable income of the rich and to raise the disposable income of the poor: the red bars are shorter than the blue bars in the right of the figures, and the opposite is true on the left.
- *Differences in the extent of redistribution*: This redistribution to the less well off is greater in the Netherlands than in the U.S. For example in market income the poorest two deciles in the Netherlands are poorer than the bottom two deciles in the U.S. (compare the blue bars); but they receive a much larger share of disposable income (the red bars).
- *Differences in inequality of disposable income*: Inequality in disposable income is less in the Netherlands than in the U.S. The ratio of the disposable incomes of the top decile to the bottom deciles is 6.95 in the Netherlands and 16.8 in the U.S.

If the distribution of disposable income after taxes and transfers is more equal than the distribution of market income, then the tax and transfer policy is termed **progressive**. If disposable income is more unequal than market income, the tax and transfer policy is termed **regressive**. The tax and transfer systems in the Netherlands and the U.S. are progressive, but the U.S. is less progressive than the Netherlands.

PROGRESSIVE TAX AND TRANSFER If the distribution of disposable income after taxes and transfers is more equal than the distribution of market income, then the tax and transfer policy is progressive. The opposite case – disposable income more unequal than market income – is termed regressive.

The degree of risk exposure with linear taxes and lump sum distribution

To understand tax-financed redistribution, let's consider what is called a linear tax and lump-sum transfer policy.

LINEAR TAX AND LUMP SUM TRANSFER refers to a tax that is proportional to income (a linear tax), the proceeds of which are divided equally and transferred to citizens (a lump sum).

- *Linear tax*: The *linear tax* part is that each family or person pays a fixed fraction of their income in taxes, so the taxes paid are a linear function of the pretax income a person has.
- *Lump sum*: The *lump sum* part is that the total amount of taxes collected, net of the costs of collecting taxes and distributing the transfers, is divided equally among all of the citizens.

HISTORY Arthur Okun, who served on U.S. President John Kennedy's Council of Economic Advisors wrote this about tax and transfer policies: "The money must be carried from the rich to the poor in a leaky bucket. Some of it will simply disappear in transit, so the poor will not receive all the money that is taken from the rich." The leaks he had in mind included not only the administrative costs of the policies but also the possible effects of the taxes on incentives.

Because we will eventually turn to the politics of taxes and transfers, we call the tax-paying (and transfer-receiving) person or family the citizen. The citizen is exposed to some risk: they can experience either a "good state" or a "bad state," occurring as before with equal probability. To introduce a tax and transfer policy we define τ as the tax rate, that is the percentage of market

Outcome	*General result*	*Low income citizen*	*High income citizen*
	Before tax		
Good outcome before tax	$y^G = \hat{y} + \frac{\Delta}{2}$	70000	140000
Bad outcome before tax	$y^B = \hat{y} - \frac{\Delta}{2}$	30000	100000
Experienced risk before tax	$y^G - y^B = \Delta$	40000	40000
Expected income before tax	$\hat{y} = \frac{y^G + y^B}{2}$	50000	120000
	After tax		
Transfers received	$= \tau \underline{y} - \phi \tau \underline{y}$ $= \tau \underline{y}(1-\phi)$	22800	22800
Good outcome after tax	$y_T^G = (\hat{y} + \frac{\Delta}{2})(1-\tau) + \tau \underline{y}(1-\phi)$	71800	120800
Bad outcome after tax	$y_T^B = (\hat{y} - \frac{\Delta}{2})(1-\tau) + \tau \underline{y}(1-\phi)$	43800	92800
Experienced risk after tax	$y_T^G - y_T^G = \Delta(1-\tau)$	28000	28000
Expected income after tax	$\hat{y}(1-\tau) + \tau \underline{y}(1-\phi)$	57800	106800
Expected taxes	$\hat{T} = \tau \left(\frac{y^G + y^B}{2}\right) = \tau \hat{y}$	15000	36000

Table 13.3: **A comparison of the before and after tax experiences of low-income and high-income citizens.** To make this table we set a mean income of the entire economy (not just the two citizens chosen here as examples) equal to 80,000, a degree of risk pre-tax of $\Delta = 40,000$, a tax rate, $\tau = 0.3$, and cost of the tax, $\phi = 0.05$. The table shows what is also illustrated in Figure 13.20,for example, that the citizen with lower expected income receives more in transfers (22800) than she can expect to pay in taxes (15000) and that both higher and lower income citizens experience a lesser degree of risk (28,000, the difference between the good and bad outcomes after the tax) than they did before the tax (40,000).

income paid in taxes. The cost of administering the program is ϕ percent of the taxes collected. Let us now consider the citizen's after-tax and transfer disposable income.

In Table 13.3 we show the taxes paid and transfer received in the good and bad states for two citizens of differing income levels. Taxes paid depend on which of the two equally likely states occur so just averaging across these states we have

$$\text{Individual's expected taxes} \quad \hat{T} = \tau \left(\frac{y^G + y^B}{2} \right) = \tau \hat{y} \tag{13.15}$$

A citizen's expected income after taxes but before receiving the transfer is just expected income ($\hat{y}$) minus expected taxes or ($\hat{T}$): $\hat{y} - \hat{T} = \hat{y} - \tau\hat{y} = \hat{y}(1-\tau)$. Because the tax is proportional to pre-tax income, just as with the income-contingent graduates' tax, citizens pay more if they experience the good state than if the bad state occurs.

But what they receive as a transfer does not depend on the state that they experience.

$$\text{Transfer received} = \frac{\text{total taxes collected - cost of admin}}{\text{number of citizens}}$$

Figure 13.20: **The relationship between income and taxes paid & transfers received** Someone who has low pre-tax expected income ($y_L = 50,000$) receives *more* in lump sum transfers (22,800) than the taxes they pay (15,000). Someone who has high pre-tax expected income ($y_H = 120,000$) receives *less* in lump sum (22,800) transfers than the taxes they pay (30,000). To make this figure we have used $\tau = 0.3$, $\phi = 0.05$, and $\underline{y} = 80,000$ as in Table 13.3. Remember that $\underline{y}$ is the mean income of the *entire* economy, *not* the average income of the two citizens included in the example.

The total taxes collected divided by the number of citizens is just the tax rate τ time the average pretax income of citizens ($\underline{y}$) and so is equal to $\tau\underline{y}$. The cost of administering the tax and transfer policy is ϕ times that amount. So we have

$$\begin{aligned} \text{Transfer received} &= \tau\underline{y} - \tau\underline{y}\phi \\ &= \tau\underline{y}(1-\phi) \end{aligned} \qquad (13.16)$$

And the citizens' expected income after taxes and transfers is thus

$$\text{Expected after-tax and transfer income} \quad \hat{y}_T = \hat{y}(1-\tau) + \tau\underline{y}(1-\phi) \qquad (13.17)$$

In Figure 13.20 we show the expected taxes and transfer levels for citizens of all income levels. The first term in Equation 13.17 is the upwards sloping line (it's a line not a curve because the tax is linear). The horizontal blue line is the second term in Equation 13.17, the transfers received by all citizens.

The vertical distance between the blue and the green lines shows the difference between the expected taxes paid and transfer received, positive for higher income citizens and negative for lower income citizens. For a person whose income $y = \underline{y}(1-\phi)$, the expected taxes paid equals the transfer received. This can be seen from the equation above rewritten as follows:

$$\hat{y}_T = \hat{y}(1-\tau) + \tau\underline{y}(1-\phi) = \hat{y} - \tau(\hat{y} - \underline{y}(1-\phi)) \qquad (13.18)$$

We can conclude two things about the linear tax and lump sum transfer policy.

- the policy is *progressive*: it redistributes income from higher to lower income citizens. Expected disposable income – that is $\hat{y}_T$ – is more equal than before tax and transfer income. This is because (from Figure 13.20) expected taxes exceed the transfer for citizens with higher expected income and the reverse is true for people with lower expected income.
- the policy is a form of *insurance*: it redistributes income from the lucky to the unlucky.This is because independently of whether their expected incomes are high or low, people experiencing the good state pay more in expected taxes than do people experiencing the bad state.

Concerning the second bullet, from Equation 13.20 (repeated below) we can see that the policy *reduces the risk exposure* of both high income and low income citizens.

$$\text{Expected disposable income} \quad \hat{y}_T = \hat{y}(1-\tau) + \tau\underline{y}(1-\phi)$$

Expected disposable income = (1-τ)(the risky part)$+\tau$(the certain part)

The reason why the tax and transfer policy serves as a kind of insurance is that the transfer received —that is $\tau\underline{y}(1-\phi)$ — is certain: unlike before tax income, it is independent of the state that the individual experiences, whether good or bad. This is because there are a very large number of citizens paying taxes, some experiencing good states some experiencing bad states. As you can see from the equation, the larger is the tax, the less weight does the risky part of her income have in determining in the citizen's expected disposable income. And correspondingly, the larger is the weight of the certain part.

The effect, as was the case with the income-contingent graduates tax, is to reduce the difference between disposable income in the good and the bad state from $\underline{\Delta}$ to $\underline{\Delta}(1-\tau)$. (This is illustrated for the two citizens in Table 13.3.)

FACT CHECK If families are limited in how much they can borrow and do not have any other way of insuring against income shocks, the equation for expected income with taxes in Table 13.3 tells us that where taxes are more redistributive from the higher to lower income families we expect to see also see smaller differences in consumption between families (some lucky, some unlucky). Using the fact that tax rates and transfer policies differ greatly from one state to the other in the US, economists have determined that this is exactly the case. Consumption is more equally distributed in states with more progressive tax and transfer systems.

Checkpoint 13.7: The citizen with income equal to $\underline{y}(1-\phi)$

Consider the values in Table 13.3. The mean income is \$80,000 with $\tau = 0.3$ and $\phi = 0.05$:

1. Calculate the expected income of the citizen with income $\underline{y}(1-\phi)$.
2. Show that the citizen with this income pays an amount in taxes exactly equal to the transfers they receive as a lump sum.

13.12 A citizen's preferred level of tax and transfers

Even though most people place some value on the well-being of people other than themselves, citizens will obviously differ in the level of the linear tax that they would prefer unless they are perfect altruists (value others' income as much as their own). Of course people do not get to pick their preferred tax rate – the same rate applies to all. But political parties propose differing levels of taxation, and citizens do pick which party to support.

Think about a particular citizen, Helmut, who in the absence of the tax and transfer policy would experience risk level Δ_f and expected income $\hat{y}_f$, as indicated by point **f** in Figure 13.21. Helmut's expected income is substantially above $\underline{y}(1-\phi)$ and so like the person indicated by point **c** in Figure 13.20, his expected income after taxation will be less than in the absence of taxation.

But even if he is entirely self-regarding (cares only about his own income and risk, not that of others) he may favor a tax on his income along with a lump sum transfer of the average tax revenues. The reason is that, like the free tuition and graduates' income contingent tax policy that made it advantageous for Sofia to continue her education, the linear tax and lump sum transfer will reduce the difference between his realized income in the good and bad state. A reduction in his expected after tax income may be a cost he is willing to pay for the risk reduction that the tax and transfer policy implements.

In the right panel, the blue line through point **f** – called the the **tax and transfer line** – shows how various levels of taxation could transform his combination of after tax and transfer expected income and risk exposure. The options range from $\tau = 0$ in which case he would remain at point **f** to $\tau = 1$. In this case, that is with a 100 per cent tax rate, the only after-tax income a person received would be the transfer so he and everyone else would have an income after taxes and transfers of $\underline{y}(1-\phi)$. Points on the line closer to $\underline{y}(1-\phi)$ on the vertical axis represent greater risk reduction by means of higher taxes and transfers.

Even though Helmut can expect to pay more in taxes then he will receive in the transfer, will he nonetheless prefer some positive level of taxes? To answer this we need to identify the benefits and costs to Helmut of increasing the tax.

TAX AND TRANSFER LINE is a line that shows the feasible combination of reduced risk exposure and reduced expected income associated with various levels of taxation.

As you can see from Figure 13.21 (and as is shown in M-Note 13.7) the slope of the tax and transfer line is:

$$\text{Slope of the tax and transfer line} = \frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} \tag{13.19}$$

We can see from Equation 13.18 that the numerator of Equation 13.19 is the reduction in disposable income associated with the tax and transfer policy. This is the cost of the policy. We know from the fact that the experienced level of risk under the tax and transfer program is $\Delta(1-\tau)$ that the denominator is the effect of the reduction in risk resulting from the tax and transfer policy. This is the benefit of the policy. The ratio of the two – how much the benefit of risk reduction costs in terms of the expected disposable income forgone – is the opportunity cost of risk reduction by this policy.

As in the previous cases studied, the slope of his indifference curves, a measure of his risk aversion, is $\frac{-u_\Delta}{u_{\hat{y}}}$, his marginal disutility of risk divided by the marginal utility of expected income when experiencing the indicated level of risk and expected income. The slope of an indifference curve indicates how much he values risk reduction relative to disposable income. So we have:

$$\text{Slope of an indifference curve} \quad = \frac{-u_\Delta}{u_{\hat{y}}} = \quad \frac{\text{marginal disutility of risk}}{\text{marginal utility of } \hat{y}}$$

We can see that at point **f** – the status quo that Helmut would experience in the absence of any tax – the following inequality holds:

$$\begin{aligned}
\text{Slope of tax and transfer line} \quad &< \quad \text{Slope of the indifference curve} \\
\frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} \quad &< \quad \frac{-u_\Delta}{u_{\hat{y}}} && (13.20) \\
\text{which, rearranged, reads} & && (13.21) \\
u_{\hat{y}} \cdot (\hat{y} - \underline{y}(1-\phi)) \quad &< \quad -u_\Delta \cdot \Delta && (13.22) \\
\text{Marginal costs of risk reduction} \quad &< \quad \text{Marginal benefits of risk reduction}
\end{aligned}$$

This means that that at Helmut's status quo (point **f** that is, $\tau = 0$) the reduction in income associated with a tax times the marginal utility of expected income is less that the reduction in risk times the disutility of risk. Because the costs are less than the benefits, Helmut wold prefer some positive level of taxation.

To find the level of τ that he would prefer, we use the analytical framework above, but we require that the marginal costs and benefits of risk reduction be equal, which means that the slopes of the tax and transfer line and the indifference curve are equal. The level of taxation that implements this rule is shown by point **f'** in Figure 13.21.

His utility at that point – $u(\hat{y}_T, \Delta_T) = u_2$ – is the same as he would have experienced without the tax and transfer policy if he had been substantially richer shown as point **j** also on the indifference curve u_2 in the figure. At that point he has an expected before tax and transfer income of $\hat{y}_j$ with his initial level of risk exposure(Δ_f). The difference between this hypothetical income at point **j** in the figure and his actual expected before tax and transfer income $\hat{y}_j - \hat{y}_f$ – is a measure in income units of how valuable the reduction in risk exposure accomplished by the tax and transfer policy is to him.

You think of point **f** as the citizen's fallback option: no tax and transfer policy. Then the amount $\hat{y}_j - \hat{y}_f$ is the rent he would receive if the tax and transfer policy indicated by the point **f'** were implemented. We call these **political rents** because they are the result of public policies that are implemented as the result of a political process.

M-Note 13.7: The opportunity cost of reduced risk exposure

Summarizing the results so far, the citizen's expected disposable income (after taxes and transfers) is (see Equation 13.18):

$$\hat{y}_T = y(1-\tau) + t\underline{y}(1-\phi) = y - \tau(y - \underline{y}(1-\phi)) \quad (13.23)$$

And her exposure to risk under the tax and transfer policy is (see Equation **??**):

$$y_T^G - y_T^B = \Delta_T = \Delta(1-\tau)$$

We can use these two equations to find the equation for the tax and transfer line, expressing $\hat{y}^T$ as a function of Δ^T. First we rearrange the equation immediately above to get an

(a) Indifference curves and allocation without taxes & transfers

(b) Post-tax and transfer income curves and allocation with reduced income and reduced risk exposure

Figure 13.21: **A citizen's preferred tax and transfer policy.** Helmut's potential choices about risk and expected income. His bundle of risk and expected income puts him on indifference curve u_0 at point **f** in panel a. Taxes that he pays and which are distributed to him (and others) as a lump sum could move him to u_1, with decreased risk exposure and lower average income at point **f'**.

expression for the tax rate, or

$$\tau = 1 - \frac{\Delta^T}{}$$

Then we substitute this expression into the top equation for $\hat{y}^T$:

$$\hat{y}^T = \hat{y} - (1 - \frac{\Delta^T}{\Delta})(\hat{y} - \underline{y}(1-\phi))$$

This is the equation for the tax and transfer line for a person with market expected income of $\hat{y}$ and market risk exposure of Δ. A tax and transfer policy transforms a reduction in expected income (a cost) into a reduction if risk exposure (a benefit). We find the marginal rate of transformation by differentiating $\hat{y}_T$ with respect to Δ_T or

$$\frac{d\hat{y}_T}{d\Delta_T} = \frac{y - \underline{y}(1-\phi)}{\Delta}$$

This is the marginal rate of transformation of reduced expected income into reduced risk, which is the opportunity cost of reducing risk, or the slope of the tax and transfer line. The numerator is $-\frac{d\hat{y}_T}{d\tau}$ the marginal cost of higher taxes in terms of expected income foregone. The denominator is $-\frac{d\Delta_T}{d\tau}$, the marginal benefit of higher taxes in terms of reduced risk exposure.

Equation 13.24 shows that the tax and transfer line will be flat (mrt = slope = zero, meaning no cost) for the citizen with an income equal to $\underline{y}(1 - \phi)$. For those with incomes less than $\underline{y}(1 - \phi)$, the tax and transfer line is downward sloping, meaning that the tax and transfer policy raises their expected disposable income.

Equation 13.24 says that a tax and transfer policy that reduces risk exposure to zero ($\Delta_T = 0$) means that everyone will have the same expected disposable income, because if we substitute in $\Delta_T = 0$ we would then have:

$$\begin{aligned} \hat{y}_T &= \hat{y} - (1 - \frac{\Delta_T}{\Delta})(\hat{y} - \underline{y}(1-\phi)) \\ &= \hat{y} - (\hat{y} - \underline{y}(1-\phi)) \\ &= \underline{y}(1-\phi) \end{aligned}$$

irrespective of the person's level of market expected income.

M-Note 13.8: Helmut identifies his preferred tax rate

Helmut wants to determine which tax rate (including zero) would be the best for him. We first write his utility function

$$u(\hat{y}_T, \Delta_T) = u(\hat{y}(1-\tau) + \tau\underline{\hat{y}}(1-\phi), \Delta(1-\tau)) \tag{13.24}$$

Equation 13.24 says that utility depends on a person's *after tax* and transfer income and risk exposure. Now, we want to see the tax rate under which his utility is maximized. Differentiating with respect to τ and equating the result to 0 so as to find the maximum we have:

$$\frac{\partial u}{\partial \tau} = u_{\hat{y}} \cdot (-\hat{y} + \underline{\hat{y}}(1-\phi)) + u_\Delta \cdot (-\Delta) = 0$$

Simplifying:

$$\frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} = -\frac{u_\Delta}{u_{\hat{y}}} \tag{13.25}$$

In order to see what this point means for the figure, let's understand Equation 13.25. The right side of the equation you are familiar with. It is the marginal rate of substitution. The left side of the equation you also know (from equation 13.24) is the marginal rate of transformation of reduced expected income into reduced risk. In other words, the left side of the equation is the slope of the tax and transfer line in Figure 13.21.

Helmut's utility is maximized when the slope of the tax and transfer line is equal to the slope of the highest indifference curve that is feasible for Helmut, that is, when:

$$mrt(\Delta, \hat{y}) = \frac{\hat{y} - \underline{y}(1-\phi)}{\Delta} = -\frac{u_\Delta}{u_{\hat{y}}} = mrs(\Delta, \hat{y})$$

Checkpoint 13.8: The "leaky bucket" problem

The cost of administering the tax and transfer policy represented by ϕ is sometimes referred to as the fraction of tax revenues that "leaks away" before being making it back to citizens as transfers. Redraw the figure of the tax and transfer lines with a value of ϕ

- higher than that shown and
- lower than that shown

Provide a few reasons why ϕ would be higher in some societies than in others.

13.13 Political rents: Conflicts of interest over taxes and transfers

So far the tax and transfer policy resembled the insurance policy that Juliana purchased. Helmut was willing to trade a reduction in expected income (the taxes he paid minus the transfer he received) for a reduction in risk exposure, just as Juliana paid an insurance premium so as to reduce her risk exposure. Their risk mitigation opportunities were represented by tax and transfer line, and a virtually identical insurance line.

But there is a big difference. Unlike Juliana who was free to choose any point along the insurance line, citizens cannot choose *individually* how much risk reduction they will "purchase": the amount will be determined by a tax and transfer policy adopted by the government applied to all citizens. And this means that citizens will *disagree* about the policy to adopt, due to their differing levels of expected market income and risk exposure.

FACT CHECK Political scientists Torben Iversen and David Soskice and explored the relationship between support for taxation and redistribution policies and the degree of risk exposure associated with of a person's skills in 11 high-income democracies. For example, a skilled operator of a specific kind of machine would be more risk exposed than a person with coding skills. Taking account of other influences on political preferences (income, sex, employment status, party affiliation, and age) the degree of a person's skill risk exposure predicts support for tax and transfer policies, equal in the size of this effect to income.

To see why think about person with the much less expected income than Helmut and the same level of risk exposure, shown by point **e** in Figure 13.22. Her expected income is just below the mean income of the economy such that the expected amount amount she will pay in taxes is exactly offset by the amount she will receive in transfers.

So for her there is no cost to risk reduction. This is why her tax and transfer line is horizontal. Her utility will be maximized if $\tau = 100$ per cent. She and Helmut have a conflict of interest, as it is clear that Helmut's preferred level of taxation is much less.

Now consider another citizen much richer than Helmut and exposed to the same level of risk, shown as point **i** in Figure 13.22. You can see that all the points on his tax and transfer line are on indifference curves with a lower level than u_4 which is what he will experience in the absence of any tax and tax and transfer policy. The tax that Helmut prefers (which would bring Helmut to point **f'**) would inflict a substantial loss in utility on this richer individual. So the political rent that Helmut would receive under his preferred tax (the quantity $\hat{y}_j - \hat{y}_f$ in Figure 13.21 has a counterpart in the higher income individual's loss if Helmut's plan were adopted.

These differences in expected income and risk exposure levels are the basis of conflicts of interest, in which the stakes for some of the citizens are substantial. You can think of the political rent that Helmut would receive were his preferred tax implemented as the maximum amount he would be willing to pay in order to ensure the election of a party that would implement that plan, if the alternative were no tax and transfer policy at all.

In the right panel of Figure 13.22 we can see that even citizens with the same level of expected income will differ in their preferred level of taxation if they have differing levels of risk exposure. To see this think about the preferred tax level for Citizen M and Citizen K whose expected income and levels of risk exposure are shown by points **m** and **k** in the figure. Citizen K prefers a substantial reduction in risk (shown by point **k'**) even thought it would cost her a major loss in expected income. By contrast, from his status quo point, the tax and transfer line provides Citizen M with no opportunities for a gain in utility (if he moved down and the the left along his tax and transfer line he would be crossing even *lower* indifference curves).

The differences between what different citizens would choose shows the figures/risk/tax_transfers_insurance_initial.pdf

(a) Conflicts of interest between rich and poor

figures/risk/tax_transfers_insurance_horizontal.pdf

(b) Conflicts of interest between more and less risk exposed

Figure 13.22: **Conflicts of interest over the level of taxation**. Citizens with pre tax and transfer expected incomes and risk exposure given by points **f, e,** and **i** differ in the level of taxes they would prefer because they differ in their level of expected income. Citizens with risk exposure and expected inocmes indicated by points **m** and **k** differ in their preferred level of taxation because they differ in the extent of their risk exposure.

conflicts at the heart of many policy debates in contemporary societies with citizens' preferences mirroring their choices of the political outcomes they would prefer. The model we have presented can help us understand these debates and conflicts over the extent of tax-based redistribution from the lucky to the unlucky and from the well off to the less well off.

13.14 Application: Choosing justice, a question of ethics

So far we have asked questions about the level of taxes and redistribution that people would prefer if they cared only about themselves – their own expected income and risk exposure, not that of their fellow citizens. But we know from Chapter 2 that people do care about others and sometimes dislike inequality, even if their own income is higher as a result.

The same model about risky decisions that we have been using, with some important additions, can clarify how people might reason not about their own self interest, but about an ethical question: how much inequality is just and how much redistribution a government should do.

The common element in questions about risk on the one hand and economic injustice on the other is that both concern differences in income. Risky decisions are about actions that result in differences in one's own realized income in good and bad states. Judgement about the justice or injustice of income inequality is also about differences, but in this case it is differences between particular people in their income (averaged over good and bad states if they are risk exposed).

REMINDER *Inequality aversion* is a preference for more equal outcomes and a dislike for either *disadvantageous inequality* that occurs when others have more than the actor and and (to a lesser extent typically) *advantageous inequality* that occurs when the actor has more than others, or both.

Just as people have preferences about risk – ranging from strong risk aversion to risk neutrality – people have preferences about inequality: in behavioral experiments, for example, many subjects are *inequality averse*. People can be both risk averse and inequality averse. Or risk neutral and inequality averse.

Feasible choices of the extent of inequality and the level of average income

We now ask you to imagine that you are Adam Smith's "impartial spectator" and you are asked to design your ideal society. In the society you're considering, it has already been determined that there will be two groups of equal size: the first called "richer" and the second called "poorer." Your job is to answer the following question: *how much* richer than the poorer people should the richer people be?

REMINDER Adam Smith in his *Theory of Moral Sentiments* wrote "We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it."

Maybe you would question why there should be any rich and poor at all. But you are also told that perfect equality in the society would mean that there were insufficient incentives for people to work hard, take risks and study. As a result some inequality would be better for everyone, *even for the poor*, because the poor would have more in a world in which most people work hard, study and are willing to take some risks than in a world of no inequality.

But too much inequality can also be a problem, lowering average income. For example where high income people are much richer than low income people, social conflicts – ranging from labor strikes to property theft may be more common. In all countries a substantial fraction of of the labor force work as police, private security workers and others whose job it is to maintain order and protect property. But the ranks of what is called "**guard labor**" are substantially larger in highly unequal countries. People guarding things are not producing things, so inequality may reduce average incomes by diverting a nations productive potential from producing to guarding.

GUARD LABOR are those employed as police, private security personnel, the armed forces and others whose job is maintaining order.

The comparison of the Netherlands and the U.S. in Figure 13.19 provides an illustration. Disposable income is much more equally distributed in the Netherlands than in the U.S. And the fraction of the Dutch labor engaged in what are termed "protective services" (literally guards) is less than a third of the fraction of the US. labor force engaged as guards. Countries even more equal than the Netherlands – Denmark and Sweden – employ even fewer guards.

FACT CHECK We provide more evidence about inequality and guard labor and illustrate how policies that reduce inequality could increase average incomes in Chapter 16. There is some evidence that the most unequal of the high income economies (the U.K and the U.S.) are to the right of point **m** in the figure, meaning that there exist policies that could both reduce inequality and raise average incomes.

Figure 13.23 illustrates an economy in which average income first rises, reaches a maximum and then falls as inequality increases, as depicted by $\hat{y}(\Delta)$. The $\hat{y}(\Delta)$ function is similar to the risk-return schedule. It gives the feasible combinations of average income and inequality. The negative of the slope of $\hat{y}(\Delta)$ is *the marginal rate of transformation of inequality into* *average income*, or the opportunity cost of greater average income in terms of greater inequality. For low levels of inequality the opportunity cost of greater inequality is negative because greater inequality results in higher average income.

Because the classes are of equal size, average income is the **mid-point** between the income of the rich and the income of the poor. So the rich get average income plus one half of Δ and the poor get average income minus one half of Δ. Recall that the average income, $\bar{y}$, is determined by the level of inequality using the function $\hat{y}(\Delta)$. Therefore:

$$\text{Income of the poor} \quad y^P = \hat{y}(\Delta) - \frac{1}{2}\Delta \tag{13.26}$$

$$\text{Income of the rich} \quad y^R = \hat{y}(\Delta) + \frac{1}{2}\Delta \tag{13.27}$$

Between 0 inequality and point **P** in the figure 13.23, not only the rich but even the poor would benefit from greater inequality. (We show in Mnote 13.9 that **p** is the point on the average income function where the slope is equal to $\frac{1}{2}$.) So increased inequality would result in Pareto improvements over an economy in which average income and inequality are both very low.

Inequality beyond **P** would benefit the rich and hurt the poor, but average incomes would still rise: the rich would be receiving a larger piece of a larger pie. Beyond point **m**, however, with greater inequality, average income would fall, and the rich would be getting a larger slice of a smaller pie. There would eventually be some level of inequality so extreme that for inequality greater than this, even the income of the rich would suffer. This is indicated by point **R**.

M-Note 13.9: How much inequality would maximize the income of the poor?

To answer the question we have to vary Δ to maximize Equation 13.26.

$$\hat{y}(\Delta) - \frac{1}{2}\Delta$$

So we differentiate Equation 13.26 with respect to Δ

$$\frac{dy_P}{d\Delta} = \hat{y}_\Delta(\Delta) - \frac{1}{2}$$

We set the result equal to zero:

$$\frac{dy_P}{d\Delta} = 0$$

$$\therefore \hat{y}_\Delta = \frac{1}{2}$$

So, the income of the poor is maximized when the slope of the average income function equals one half. A line segment with slope $\frac{1}{2}$ that is tangent to the average income curve

./figures/risk/justice_inequality_feasible.pdf

Figure 13.23: **Average income and income inequality.** How much equality is too much, to little, just right? Point **m** indicates the level of inequality at which average income is the greatest. **P** and **R** respectively are the inequality levels that maximize the income of the poor and the rich.

will indicate the level of inequality that would maximize the income of the poor, which is shown by Δ^P in Figure 13.23.

We leave it to you to show that differentiating the equation giving the income of the rich equation 13.27 with respect to Δ and setting the result equal to zero shows that the income of the rich is maximized when the slope of the average income function equals minus one half. A line segment with slope $-\frac{1}{2}$ that is tangent to the average income curve will indicate the level of inequality the rich would prefer, which is shown by Δ^R in Figure 13.23.

An Impartial Spectator's preferences about inequality

What level of inequality would an Impartial Spectator choose?

The Impartial Spectator would not choose a level of inequality less than Δ^P, because point **P** would be a Pareto-improvement over any outcome for which $\Delta < \Delta^P$(both the rich and the poor would benefit if the economy were organized at point **P** rather any less unequal outcome. Similarly, the Spectator would be unlikely to select a level of inequality greater than Δ^R because such extreme levels of inequality would be Pareto inefficient: both the rich and the poor would be better off by a reduction in equality to Δ^R.

To consider the possible levels of inequality between points **P** and **R** we need

./figures/risk/justice_inequality_neutral.pdf

(a) Indifference curves of an inequality-neutral Spectator

./figures/risk/justice_inequality_averse.pdf

(b) Indifference curves of an inequality-averse Spectator

Figure 13.24: **The degree of income inequality chosen by two Impartial Spectators: an inequality averse Spectator and an inequality-neutral Spectator**. The curved green line $\underline{y}(\Delta)$ is the set of feasible average income choices the person faces when choosing how to design her society. Δ_m is the inequality choice of a person who is not concerned inequality at all and wants simply to maximize average income – they are inequality-neutral (as shown in panel a.). The inequality-averse person, shown in panel b, chooses a point like (Δ^a, y^a).

to know more about the Impartial Spectator's preferences. Figure 13.24 shows average income as a function of the degree of inequality as in the previous figure, along with two sets of indifference curves that the Impartial Spectator might have, an inequality-averse Spectator on the right and a inequality-neutral Spectator on the left. The slope of these indifference curves: it is a measure of inequality aversion.

$$\begin{aligned} \text{Slope of an indifference curve} &= \text{A measure of inequality aversion} \\ \frac{-u_\Delta}{u_{\underline{y}}} &= \frac{\text{Marginal disutility of inequality}}{\text{Marginal utility of average income}} \end{aligned}$$

With preferences indicated by these indifference curves, the Spectator will seek a point on the highest possible indifference curve that is also feasible, which is where the feasible frontier is tangent to an indifference curve.

Panel a. of Figure 13.23 shows that an inequality-neutral Spectator would choose the level of inequality that maximizes average incomes. Panel b. shows that an inequality-averse Spectator would choose a lower level of inequality.

13.15 Risk, uncertainty and loss aversion: Evaluation of the model

The model of doing the best you can in risky situations clarifies important aspects of economic behavior: investment in risky assets, what methods of production to use, the value of continuing one's education, how much insurance to purchase, citizens' evaluations of alternative tax and transfer policies. A feature of the model is that it bases the study of risky decisions on familiar analytical tools: indifference curves and feasible sets.

But the model is limited in a number of ways.

Uncertainty: Not knowing the probabilities of the relevant contingencies

We have assumed that our decision makers – the student considering higher education, the citizens, the Indian farmers – knew the probability that the good and bad state would occur (in most of the cases we assumed that the two states are equally probable). But in many situations the probabilities of the various contingencies affecting the outcomes of people's actions are not known. So people face not the problem of risk , but the much more difficult problem of *uncertainty*. People typically have a pretty good idea about the risk of rain (you can get an estimate of that online, so this is risky situation) but no clue about the likelihood of an earthquake (so this is a case of uncertainty).

In our utility function we have let Δ, the difference between the equally probable good and bad outcomes, represent what the risk averse person considers to be the "bad" – namely risk. But not knowing the probabilities of the relevant contingencies there is no way to assign a particular level of risk to a choice, even if we know the difference in expected incomes between the good and bad state. To see this suppose that in the good state you gain 10,000 Euros and in the bad state you loose 10,000 Euros. How risky is your decision?

It is obviously a lot more risky if the probability of the bad outcome is one-half than if it is, say one-in-a-hundred. The same would be true if the probability of the bad outcome were 99 in a hundred: the expected income associated with this choice would be very low and most decision makers whether risk averse or not would want to avoid it. But the reason would not be its riskiness: the choice would not be as risky as the fifty-fifty probability.

So not knowing the probabilities we cannot say how risky a choice is, making it impossible to use our risk averse utility function. The same problem arises using another way that economists treat risk: by assuming that decision makers make choices to maximize their expected utility. If the probabilities of the contingencies are not known, expected utility cannot be computed.

Often the best we can do is to identify the consequences of the one or more "bad states" and adopt policies to reduce the likelihood of their occurring. The term used to describe these policies is commonly "prudence" (meaning, roughly "caution" in the face of uncertainty). But prudence does not deliver the prescriptions for action such as those that are possible when the probabilities of contingent events are known.

The limits of the model are especially clear when applied to the problem of climate change. The reason is that while many of the relevant facts are reasonably well established – that human activity contributes to climate change, for example, we really are not able to assign well-informed probabilities or even guesses to some critical contingencies. For example we cannot know or even intelligently guess how probable human extinction is over the range of relevant earth surface temperatures.

Loss aversion as an alternative reason for avoiding risks or uncertainty

Loss aversion is a well-documented aspect of human behavior according to which the loss of some given amount – say, a Euro – reduces our utility by more than a gain of the same amount would have raised our utility. A loss-averse person would refuse a coin flip in which they stood to gain one Euro if the coin came up heads, and lose one Euro if it came up tails. Loss aversion differs from risk aversion because it captures the fact that we treat losses and gains differently even if they are very small. This differs from risk aversion in that a person who was highly risk averse in a situation involving a substantial risk (a large difference between the payoff at the good and bad state) would be approximately risk neutral for very small risks.

LOSS AVERSION is present when the loss of some given amount – say, a Euro – reduces our utility by more than a gain of the same amount would have raised our utility.

Suppose you have an income of $1000 per month. At the beginning of one month, you receive the news that you will receive a positive income shock, but you don't know exactly the amount. It could be either $100 or $300, each with a probability of 0.5. Now, suppose that you have an income of $1200 per month, but a given month you know that you will face either a positive income shock of $100 or a negative shock of $100, each equally likely. The two situations are analogous. The expected income is $1200, the good and bad states exactly the same. However, we treat gains and losses differently. We prefer a gain of $100 to achieve an income of $1100 than a loss of $100 that results in the same income. Our behavior would probably differ in both situations. For example, many people would want to buy insurance to avoid the loss of $100 in the second scenario, but it is not likely we would buy it in the first situation.

No single model is entirely adequate to cover all of the relevant cases ranging from standard problems in finance, to loss averse behaviors, to the absence of information about the probabilities of the relevant contingencies concerning climate change. A combination of models on a case by case basis, with contributions from psychology, biology and other sciences seems the best way of understanding how we behave in situations in which we lack information on the outcomes of our actions.

13.16 Conclusion

When it comes to how we live our ordinary lives and the decisions we have to take, risk and uncertainty are the rule, not the exception. Richard Feynman, Nobel Laureate in physics – a field widely known as an 'exact science' – had this to say: "I have approximate answers and possible beliefs and different degrees of uncertainty about different things, but I am not absolutely sure of anything. "

In this chapter we have adapted and extended the model of constrained optimization that you have studied so far to analyze risky decisions. The main new element is risk aversion, which builds on the idea that just as we may have likes and dislikes when it comes to different foods, or engaging in hard work, we also have preferences about the degree of risk to which we are exposed. Indifference curves based on risk preferences allow us to study the trade-offs that we face when the opportunities we have to increase our expected income also expose us to more risk .

We will see in Chapter 15 that the model of decision-making in the presence of risk will provide important insights into why capitalism is such a dynamic economic system. And we will continue a theme introduced in this chapter: how risk aversion may also contribute to a second attribute of capitalism, namely elevated levels of economic inequality.

Making connections

Constrained optimization Treating expected income as a "good" and risk as a "bad" allows us to extend the familiar analytical tools – indifference curves and feasible sets – for use in understanding decisions making over risky options.

Mutual gains and conflicts over their distribution Insurance allows a valuable risk reduction to the insured and profits to insurance providers, thus implementing a Pareto-improvement over the 'no insurance' default option. How the mutual benefits made possible by insurance are divided between the insurance provider and the insured is a matter of conflict.

Heterogeneity: Wealth differences. Differences in risk aversion reflect not only personality differences among people but also differing situations in which people find themselves. Having substantial wealth and (as a result) having access to credit reduces risk exposure and hence risk aversion. Lesser income and wealth limits a person's ability to borrow and increases people's risk exposure and risk aversion.

Inequality and poverty traps The risk aversion of people without access to credit will motivate them to avoid making high risk choices with high expected returns (such as changing ones occupation, starting a business, or relocating), thereby perpetuating their limited income.

Policies Tax and transfer policies by governments are a kind of insurance that distributes income not simply from higher income to lower income citizens, but also from the lucky to the unlucky; this can reduce the extent of risks to which a person is exposed and thereby allow people with limited wealth to make less risk averse choices with higher expected incomes.

Important ideas

contingency	lottery	expected value/payoff
expected utility	fair insurance	risk
risk neutrality	uncertainty	risk aversion
risk premium	certainty equivalent	insurance
risk-return schedule	loss aversion	risk-loving
inequality aversion	investment	Gini coefficient
Lorenz curve	utility a concave function of income	degree of risk
disposable income	market (pre-tax) income	tax and transfer line
no-tax curve	specific/general asset	risk pooling
progressive/regressive taxation		

Mathematical Notation

Notation	**Definition**
$u(\Delta, \hat{y})$	a utility function (for the study of risk as a "bad")
Δ	risk, the difference between the good and (equally probable) bad outcome
p	probability of an event
y	realized income
$\hat{y}$	expected income
$\underline{y}$	average income
$\hat{y}(\Delta)$	risk and return schedule
$\underline{y}(\Delta)$	inequality and average income schedule
s	insurance (amount of risk reduction)
p_s	price of insurance
τ	tax rate
ϕ	cost of administering taxes
$\hat{T}$	expected taxes paid

Note: We use the superscripts G and B to indicate realized values in the good and bad states respectively. The subscript T refers to an after tax and transfer value. A variable with a "hat" (such as $\hat{y}$) means "expected" and an underlined variable (such as $\underline{y}$) means average.

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

See Reference list.

14

Perfect Competition & the Invisible Hand

> . . . he intends only his own gain, and he is in this, as in many other cases, led by an *invisible hand* to promote an end which was no part of his intention. Nor is it always worse for the society that it was not part of it. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it.
>
> Adam Smith, Wealth of Nations, (1776) Book 4, chapter 2.

DOING ECONOMICS This chapter will enable you to do the following:

- Understand what economists mean by a perfectly competitive equilibrium among price-taking buyers and sellers.
- Explain why if all contracts are complete then the allocation observed in such an equilibrium will be Pareto efficient and show why perfect competition is necessary for this to be the case.
- Give the main cases in which the complete contracts assumption does not hold.
- Explain how an imaginary "auctioneer" hypothetically discovers and and sets equilibrium prices under which all markets would clear.
- Describe a more realistic process by which competition among buyers and sellers could arrive at a Pareto-efficient allocation as long as contracts are complete.
- Understand the conditions under which bargaining among private economics actors can implement a Pareto-efficient outcome.
- Know why – when more than one of the conditions necessary for a Pareto-efficient outcome are violated – then addressing one but not all of the violations may worsen rather than improve the outcome.

14.1 Introduction

At the height of the Cold War, in a model kitchen installed as an exhibit at the American embassy in Moscow to impress Russians with the living standards of the American people, two aggressive debaters faced off: Nikita Khrushchev, Premier of the Soviet Union, and Richard Nixon, Vice President of the United States (and later to be President). The date was July 24, 1959.

Nixon: "I want to show you this kitchen. It is like those of our houses in California." [points to the dishwasher, see the picture]

Khrushchev: "We have such things."

Nixon: "This is our newest model . . . Our steel workers as you know are now on strike. But any steel worker could buy this house."

Khrushchev: "In Russia, all you have to do to get a house is to be born in the Soviet Union. You are entitled to housing . . . In America, if you don't have a dollar you have a right to choose . . . sleeping on the pavement."

Nixon: "We have 1,000 builders building 1,000 different houses. We don't have one decision made at the top by one government official. This is the difference.

Khrushchev: ". . . in another 7 years, we'll be at the level of America, and after that we'll go farther. As we pass you by, we'll wave "hi" to you."

Figure 14.1: **Nikita Kruschev and Richard Nixon at the so-called Kitchen Debate, 24 July 1959.**

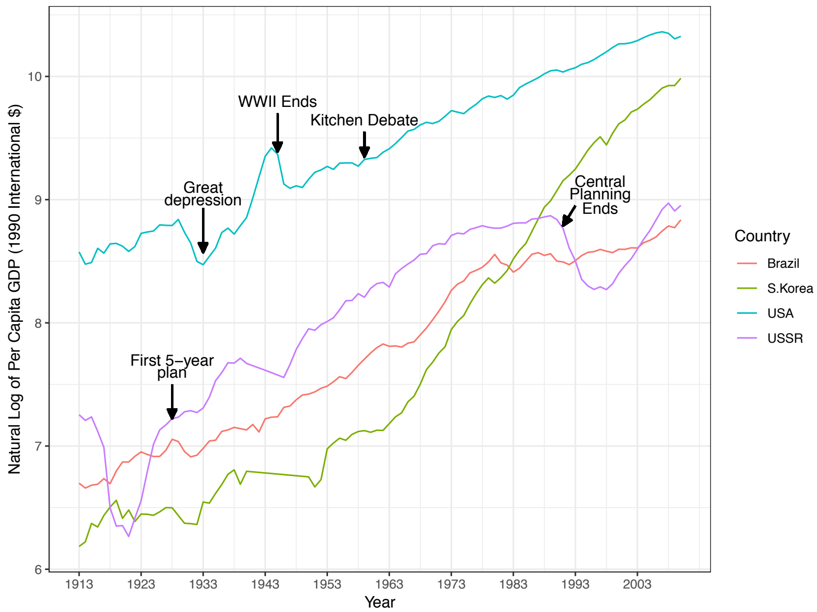


Figure 14.2: **The natural logarithm of real per-capita GDP of the USA, the USSR, Brazil and Argentina over the 20th century.** When the natural logarithm is plotted against time, the slope of the curve is the rate of growth, so aside from the Great depression and the Second World War, the rate of growth of per capital GDP in the U.S. has been fairly constant over the period. The rate of growth in the USSR exceeded that of the U.S. during the 1920s and 1930s but then declined to less than the U.S. in the decade prior to the end of Communist Party rule there. South Korea's growth dramatically increased in the mid 1960s and has remained very high since. Source: The Maddison-Project, http://www.ggdc.net/maddison/maddison-project/home.html, 2013 version

The "kitchen debate" was a surprisingly genial episode in the competition between the Soviet Union and its allies on the one hand and the United States and other capitalist nations on the other. The conflict would bring humanity to the brink of global nuclear war on more than one occasion. At issue was the superiority, even survival, of two competing economic systems: central economic planning and capitalism.

Four decades before Khrushchev and Nixon met in the model kitchen in Moscow, economists had debated whether or not a *government* could do a better job of allocating society's resources than the *market*. At issue was whether the economy should be guided by the "one decision . . . by one government official" that Nixon mentioned (a **centralized** economy), or instead by the countless decisions of the buyers and sellers, investors and workers and others whose actions are coordinated by markets in a **decentralized** economy.

A *centralized* economy (also called a "centrally planned" economy) is one in which the government (the central planner) decides what should be produced, where, by whom, and when, and how the resulting goods should be distributed among the population. In a more decentralized or "market" economy, on the other hand, individual firms, consumers and other economics actors make choices about production and consumption with only limited coordination by a government. No economy is entirely centralized, as private decisions of people are never entirely controlled by a government. Similarly no economy is entirely decentralized, as government policies limit the feasible actions that people may take, and alter the benefits and costs of particular actions.

CENTRALIZED VS. DECENTRALIZED ECONOMY A *centralized* economy is one in which the government decides what should be produced, where, by whom, and when, and how should the resulting goods be distributed among the population. In a decentralized economy, on the other hand, who produces what, when, how, and for whom is determined by the uncoordinated decisions of owners of individual firms consumers , and other economic actors.

The sometimes academic debate (which we return to at the end of this chapter) took a decidedly practical turn when, in 1928, the Soviet Union became the first-ever country to centralize economic decision-making. With their first five-year plan the government of the Soviet Union replaced markets as the main mechanism for determining the priorities and functioning of the economy. They substituted instead the decisions of government officials.

Just a year later the capitalist world plunged into the Great Depression, a cataclysm of economic insecurity and unemployment that barely affected the Soviet Union. Shortly thereafter, Adolf Hitler's rise to power in Germany — in part a result of the high levels of unemployment — raised the stakes of the ongoing debate. Germany's economy under fascist rule also for the most part avoided the Great Depression. These events and the idea that the centrally planned economy could avoid the boom and bust dynamic of the capitalist economies buttressed the case against capitalism.

Opinions differ on who won the kitchen debate, but Figure 14.2 – depicting the GDP per capita over time of several countries, including the US and USSR – shows that Khrushchev's claim that the Soviet Union would overtake the U.S simply did not happen. Notice from Figure 14.2, too, that at the height of the debate on planning versus the market the Soviet economy was growing rapidly, under the first and second five year plans (1928-1938), while the U.S. economy was struggling to recover from the Crash of 1929. But in the long run, the centrally planned economy about which Khrushchev had boasted was outstripped by the U.S. especially in the three decades following World War II, termed the "golden age of capitalism."

REMINDER In Chapter 1 what we called an invisible hand game was one in which entirely self-regarding players achieve a Pareto-efficient outcome even though they interact non-cooperatively, that is without coordinating their actions in any way.

For the half century following the introduction of the five year plans, the Soviet economy did well compared to some capitalist countries – Brazil and Argentina, for example – despite the devastation of the German invasion and occupation during World War II.

14.2 A general competitive equilibrium

Long before the 1930s debate on centralization vs decentralization, economists have studied what we call a perfectly competitive general equilibrium system. A key objective has been to determine the conditions under which set of interrelated markets (that is, a market system) that allow unimpeded entry and exit to a large number of entirely self-regarding buyers and producers could implement a Pareto-efficient allocation.

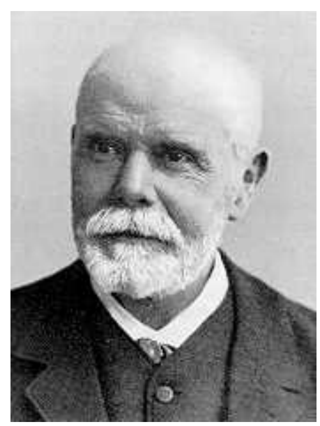


Figure 14.3: The idea of a perfectly competitive general equilibrium was developed by Léon Walras (1834-1910) a French economist with a passion for social justice and for mathematics. His father had been a schoolmate of Augustin Cournot, whose model of competition among firms you studied in Chapter 9. Along with Alfred Marshal, Walras (pronounced val-RAS) is considered the founder of the "neoclassical school" of economics that in most countries was the predominant approach in economics during the 20th century.

The invisible hand and general equilibrium analysis

The question dates back to Adam Smith and his, at the time, shocking idea that the uncoordinated individual pursuit of self-interest in a competitive economy would be guided by an "invisible hand" to serve the public interest. To better understand how Smith's invisible hand might work (and why it might not) economists have modeled the economy as a whole, rather than focusing on a single market.

- *Partial equilibrium:* When we analyze the equilibrium of a single market, say the market for fish populated by fish buyers and fish sellers, then we are engaged in *partial* equilibrium analysis.
- *General equilibrium:* When we analyze the equilibria of all markets simultaneously, then we engage in *general equilibrium* analysis, where we try to understand the ways in which markets affect each other.

HISTORY In the century after Adam Smith wrote, the idea of the invisible hand made its way into popular culture, as this snip from Lewis Carroll's 1865 *Alice in Wonderland* shows:"The game seems to be going on rather better now," Alice said. "Tis so," said the Duchess: " and the moral of it is – 'Oh, 'tis love, 'tis love, that makes the world go round."' "Somebody said," whispered Alice, "that it's done by everyone minding their own business."

Our analysis of the market for Benneton's clothing items in Chapter 10 and credit in Chapter 12 and buying and selling risk in Chapter 13 are examples of partial equilibrium economics. Our "whole economy" model in Chapter 11 is an example of general equilibrium analysis because it brings together the equilibrium number of firms and prices in the goods market with the equilibrium wages and employment levels in the labor market. Our models using the Edgeworth box – buying and selling coffee and data for example in Chapter 1 – are also simple examples of general equilibrium analysis (and we will use that method here.)

Attempts to clarify the conditions under which Adam Smith's radical claims for the invisible hand might be true have occupied some of the best minds in economics since the origin of our discipline. What they found out is of some interest for that reason alone. Kenneth Arrow and Frank Hahn put it this way:

PARTIAL AND GENERAL EQUILIBRIUM Partial equilibrium analysis is the study of single markets while general equilibrium analysis is a study of the entire economy.

> There is by now a long and ... imposing line of economists from Adam Smith to the present who have sought to show that a decentralized economy motivated by self interest and guided by price signals would be compatible with a coherent disposition of economic resources that could be regarded in a well-defined sense as superior to a large class of possible alternative dispositions.

They were not doing this as advocates of the invisible hand but more out of scientific curiosity, they explained:

REMINDER Remember the *classical institutional challenge* from Chapter 1? Consider how remarkable it is if a set of institutions as simple as those regulating a competitive market can address the challenge, even if only in special circumstances.

> It is important to understand how surprising this claim must be to anyone not exposed to the tradition ... That [this claim] has permeated the economic thinking of a large number of people who are in no way economists is itself sufficient grounds for investigating it seriously ... it is important to know not only whether it is true but whether it *could be true.* (original emphasis)

Almost two and a half centuries after Smith – in part through the pioneering research of Arrow himself – we now know that there are indeed conditions under which this remarkable conjecture could be true. But we will also see that, except in very special cases, market competition does *not* lead to a Pareto-efficient outcome. We begin with an illustration of the model of general competitive equilibrium.

figures/UtopianCapitalism/markets_tioli_logs.pdf

Figure 14.4: **The Edgeworth box, participation constraints, and the Pareto-improving lens**. Each trader – Ayanda and Biko or Adamo and Beatriz – has an endowment shown by point **z** with a corresponding participation constraint u_z^A and u_z^B. Each trader also has a best-response function to every potential price ratio that they could be offered that will take them from their endowment at **z** to another endowment along their best-response function (also called their incentive compatibility constraint for any potential voluntary trade).

A general equilibrium model of the exchange of two goods

We return to the model that you studied in in Chapter 4, illustrated in the Edgeworth boxes shown in Figure 14.4. The total amount of the two goods available ($\bar{x}$ and $\bar{y}$) determine the dimensions of the box: $\bar{x} \times \bar{y}$ (in this case (10×15). The initial holdings of the two goods are indicated by point **z**: Ayanda has a lot of coffee and Biko has a lot of data. These two bundles of coffee and data are referred to as Ayanda and Bikos endowments. We term the endowment the the fallback option for two traders because it is the outcome that they will experience if they do not trade. The participation constraints of the two are the indifference curves u_z^A and u_z^B which show all the allocations resulting from a trade that each values as highly as not trading at all.

REMINDER The rectangle in Figure **??** is called an Edgeworth box and its dimensions give the total amount of the two goods available ($\bar{x}$ and $\bar{y}$). The coordinates of any point in the box describes an allocation, that is, the amount of each good that each person has. Point **z** is the initial allocation – their endowments – while point **n** is the allocation following exchange.

They both can be better off if Ayanda exchanges some of her coffee for some of Biko's data. This is shown by the Pareto-improving lens bounded by their participation constraints. The lens is the set of all allocations that are Pareto-superior to their endowments at **z**.

REMINDER Given two arbitrary allocations R and Q, allocation Q is Pareto superior to allocation R if at allocation Q at least one person is better off and no person is made worse off than at R. The term "Pareto improvement over" is equivalent to "Pareto superior to."

The Pareto-improving lens narrows down the set of possible trades that A and

B might make. Remembering that because their exchange must be voluntary, points not in the Pareto-improving lens cannot occur. The reason is that at least one of the two would be worse off as a result, and so would refuse to trade. Every point in the Pareto-improving lens is an allocation that both A and B could agree to if the only alternative was no trade at all. The reason is that moving from point **z** to that point would make both better off (or at least one of them better off and the other not worse off).

To understand the trades that each might actually make, we need to consider the prices at which the two might exchange their goods. To do this we adopt an indirect strategy We *do not* ask where the prices come from, as we did in Chapter 4, when we gave one of the actors price setting power or take it or leave it power (to set both a price and a quantity traded). Instead we work backwards from what is required for an outcome of the traders' interactions to be a Nash equilibrium, specifically, the properties that an equilibrium price must have if it is to be consistent with the rules of the game of a perfectly competitive market.. So we:

- consider all possible hypothetical prices,
- assume that the traders take each price as exogenously given (that is, they act as price takers)
- analyse the traders' best responses to each of these prices,
- for one of these hypothetical prices find pair of mutually consistent best responses (the buying and selling intentions if the two traders) .
- this is a general equilibrium because, as a mutual best response, neither trader could benefit by offering to change the price or the quantity of goods that they are transacting.

Because best response functions play a central role in this line of reasoning Figure 14.5 shows A's best response function. The difference between Ayanda's endowment **z** and any point on her best-response function represents a rearrangement of her ownership of goods (giving up some coffee, getting more data) that she would be *willing* to implement through voluntary exchange. For example point **j** indicates that in order to get 4 gb of data (so she would then have 5 rather than just one in her endowment), she would at most be willing to give up 4 kilograms of coffee (she would as a result have 5 kilograms of coffee rather than the 9 kilos at her endowment).

This information tells us how much Ayanda will offer to sell (and be willing to buy) for each of the hypothetical prices. There is no money in this model so a price is a ratio of one good to the other, such as how many many gb of data Ayanda will receive in return for one kg of coffee is given by the (negative of the) slope of a line passing through point **z**. One of these possible prices for the two goods is indicated by the price line p_j in Figure 14.5 at which Ayanda

REMINDER A trader's best-response function in an Edgeworth box (also an *incentive compatibility constraint* or *offer curve*) shows the least amount that a person will give up in return for each amount of another good (or money) offered by another person. In the Edgeworth box, the price means the price of the x good in terms of how many y goods are required to purchase one unit of the x good. This means that a "price" is is the price of coffee. So a price like $p_j = 1$ means that 1 kilogram of coffee costs 1 gb of data. A higher relative price of coffee would be 2 gb of data for 1 kg of coffee (a steeper line), a lower price would be 0.5 gb of data for 1 kg of coffee (a flatter line).

gets one gb of data in exchange for one kg of coffee, so the "price of coffee" is one (gb of data) and the price line p_j has a slope of -1.

Recall from Chapter 4 a best-response function is constructed by treating the price line as a constraint on the trader's utility-maximizing process. So in Figure 14.5 think if each of the three price lines p_k, p_j and p_n as alternative budget constraints corresponding to the three prices. We then find a point of tangency between each of the given price lines and one of Ayanda's indifference curves. This means that each point on her best-response function – such as **n, j,** and **k** – is a tangency between the price line and an indifference curve.

14.3 Market clearing and Pareto-efficiency

How will the two traders respond to some hypothetical price? To answer the question we need to refer also to Biko's best-response function. In Figure 14.6 we show the desired trades of the two when the price is p_j. Here is what we find out:

figures/UtopianCapitalism/markets_tioli_logs_brf.pdf

Figure 14.5: A's best response to three different prices. The price of coffee (gb of data in return for one kilogram of coffee) is the (negative of) the slope of each of the three price lines p_k, p_j and p_n. Given some low price price such as $p_k = 0.58$ A will sell less coffee (meaning that she will buy fewer gb of data) indicated by point **k** than at a higher price $p_j = 1$ (point **j**).

- Ayanda would not want to sell much coffee (indicated by point **j**, the intersection of the price line and her best response), while

- Biko would want to purchase more coffee (indicated by point **h** the intersection of the same price line and his best response).

This means that at price p_j the markets *do not clear*:

- At the price p_j, the demand for coffee (by Biko) exceeds the supply (Ayanda's offer), or, what is the same thing,
- Biko's *supply* of data exceeds Ayanda's *demand* for data.

At the price p_j, Ayanda is on the short side of a non-clearing market, the side for which the number of desired transactions is least. Biko, who wants to exchange more goods than Ayanda, is on the long side. As you know saw in Chapter 9 and as you know from the way the labor market works, the transaction that will occur is determined by the amount of trade the short-sider desires.

This is because the exchange is voluntary. The long-sider, Biko, who would like to trade more at the B-favorable price p_j, cannot require Ayanda to trade more than she would like. In this case Ayanda is like the employer, the short side in the labor market: the number of people hired is determined by the employers not by those wishing to find a job. And the amount of coffee and data that will be exchanged is determined by Ayanda.

Could point **j** be a Nash equilibrium? To see that it cannot, imagine that you are Biko. At the price p_j you would like to buy more of Ayanda's coffee but she is not willing. What would you do? You would offer her a somewhat higher price of coffee and she would agree to sell (as you can see from Ayanda's best response function, raising the price, e.g. to p_n will get Ayanda to sell more). Both would then be better off. Or Ayanda, thinking along the same lines, could have just offered to sell more coffee if Biko would pay a bit more of his data.

So allocation **j** resulting from the price p_j cannot be a Nash equilibrium. Exchanging goods at that price cannot be a mutual best response.

Checkpoint 14.1: Non clearing markets

Using figure fig:IntersectingOfferCurvesexplain what exchange would occur if the price was p_j: who trades how much of what to whom?.

.

A Pareto-efficient market equilibrium

The fact that at point **j** Biko could offer Ayanda terms under which exchanging more of the goods would be mutually beneficial means that the allocation figures/UtopianCapitalism/property_walras_baseline_pec.pdf

Figure 14.6: **A non-clearing market in an Edgeworth box** The blue and green best responses show that at the relative prices of the two goods given by the slope of the gray price line each trader will choose a corresponding allocation given by the intersection of the price line and their best-response function. The price line (p_j) results in excess demand for coffee and excess supply of data. Points j and h on the price line result in different levels of utility for each trader. Biko has a higher utility at point h, thus $u_h^B > u_j^B$. Ayanda, however, has a higher utility at point j, such that $u_j^A > u_h^A$.

at point **j** could not be Pareto efficient. The following reasoning confirms this:

- At point **j**, Ayanda's indifference curve u_3^A is tangent to the price line (that's why Ayanda wanted to sell that amount).
- At point **j**, Biko's indifference curve u_2^B is *not* tangent to the price line (that's why Biko wanted to buy more coffee than Ayanda wanted to sell).
- Therefore the two indifference curves have different slopes at point **j**, which means that they intersect and so cannot be tangent.
- But this means that the allocation at point **j** cannot be Pareto efficient: Above and to the left of point **j** you can see a small Pareto-improving lens representing allocations with more goods being exchanged.

This is why point the price p_j and the allocation **j** cannot

To find the Nash equilibrium price, think about the definition: A Nash equilibrium is a mutual best response. What this means is that it must be a point on each of their best response functions, in other words their intersection.

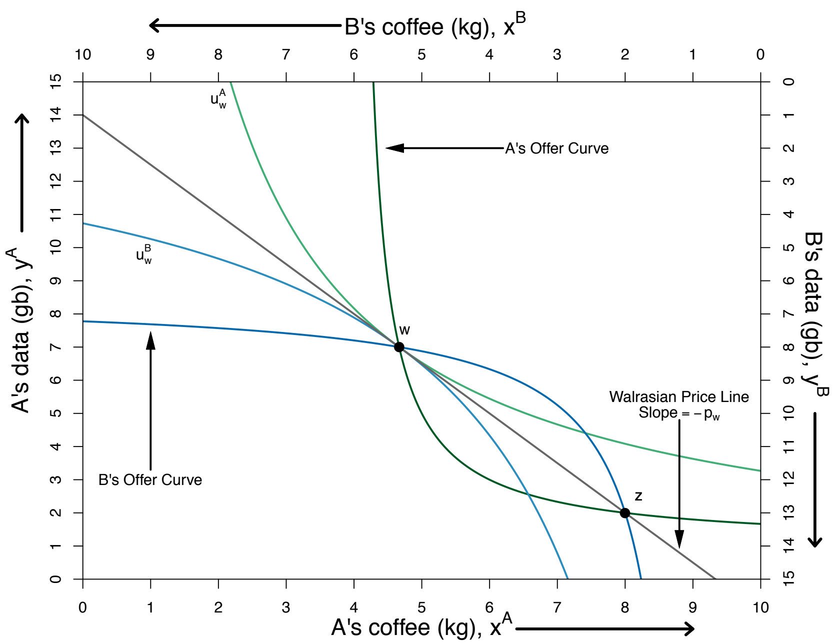


Figure 14.7: **A general competitive equilibrium.** The blue and green best responses show that at the relative prices of the two goods given by the slope of the gray price line each trader will choose a corresponding allocation given by the intersection of the price line and their best-response function. The price line, p^N, shows how Ayanda's desired sales of coffee to Biko are equal to Biko's desired purchases of coffee. At the same price Biko's desired sales of data to Ayanda are equal to Ayanda's desired purchases of data. The market clears at p^N and there is neither excess demand nor excess supply of either good.

To see this, turn to Figure 14.7, in which we show the two best response functions. You can see that there is a price of the two goods – shown by the price line p^N – under which three conditions hold.

First, *markets clear*: at point **n** Ayanda's supply of coffee is equal to Bikos demand for coffee and Bikos supply of data is equal to Ayanda's demand for data.

Second, The allocation is *Pareto efficient*. Here is why.

- At point **n**, Biko's indifference curve is tangent to the price line; it is on Biko's offer curve which are all points of tangency between a price line and the person's indifference curve.
- By the same reasoning, at point **n**, Ayanda's indifference curve is tangent to the price line
- If both people's indifference curves are tangent to the same price line at point **n**, they must have the same slope (the Law of One Price).
- Therefore the two indifference curves must be tangent to each other, meaning that the allocation at point **n** is Pareto efficient because their marginal rates of substitution are equal.

Third, the allocation is a *Nash equilibrium*: As a mutual best response, Ayanda is doing the best she can given the quantities of data and coffee Biko is willing to transact, and the Biko is doing the best he can given the transactions that

Ayanda is willing to implement. Neither can benefit by offering to trade at some price other than p_n so being a price taker is also the best they can do. The model we have used is very simple, but extensions of the model to cases with many buyers and sellers of many goods in which they are not simply "endowed" an initial allocation but *produce* the goods they exchange yields the same result.

REMINDER An allocation between two people is Pareto-efficient if (and only if) the indifference curves of the two are tangent at that point because at that point the marginal rates of substitution are equal.

Because the buyers and sellers are price-takers and we made use of the Law of One Price, the resulting allocation is sometimes referred to as a perfectly competitive general equilibrium. But the model does not address the question: how did competition determine the price p_n.

What was shown instead is that there exists a price like p_n resulting in an allocation with the three above properties: *market clearing*, *Pareto efficiency*, and *Nash equilibrium*. The gap in the model – the lack of an explanation of how the equilibrium prices come about has proven difficult to repair without dropping the assumption that people act as price-takers. We will below suggest a way that this can be done by adopting a more realistic model of how competition works.

14.4 Prices as messages, markets as information processors

Even keeping this caveat in mind the general equilibrium model supports a surprising result: there exists a set of prices such that people independently maximizing their own utility would implement a Pareto efficient allocation *if these prices were somehow the one's at which they were constrained to transact*.

The italicized passage is important, and we will later ask: can we expect markets (even if perfectly competitive) to produce prices like p_n? But for now we assume that set of relative prices like p_n is (somehow) known and are the prices at which goods transact.

To see how prices can allow the decentralized coordination of peoples actions imagine that markets were outlawed, and that you were tasked with allocating two goods, x and y, between two people Adamo (A) and Beatriz (B) in a Pareto efficient way. Think of this as an extension of the problem of allocating goods between Ayanda and Biko. We will use this two-person case as a lens for studying the relevant case, in which there are thousands of Adamos and Beatrizes.

HISTORY Looking ahead, the main theme of Friedrich Hayek's 1945 critique of the centralized economy was that it would be impossible to collect all of the information necessary to plan a centralized economy. People buying and selling on competitive markets unknowingly produce the information that they need – prices – not intentionally, but as an unintended by product of their privately motivated actions.

Consumption: Unintentionally equating marginal rates of substitution

You know that Pareto efficiency requires that you find an allocation such that the people's indifference curves are tangent (equal marginal rates of substitution) or:

$$mrs^A(x^A, y^A) = \frac{u_x^A}{u_y^A} = \frac{u_x^B}{u_y^B} = mrs^B(x^B, y^B) \tag{14.1}$$

To do this you would have to know the indifference curves of the two, a challenging task in the case of just two people, and virtually impossible for an entire economy. You would have to know (which means: devise ways of finding out) the utility functions of each person.

But if the ideal perfectly competitive general equilibrium prices were in force the problem would be a lot simpler. Each individual would maximizes utility *constrained by* the same *given* set of prices. The trader buys or sells goods until the bundle she has obtained equates her own marginal rate of substitution to the equilibrium price ratio. Because all traders are doing the same thing – constrained by the same ratio of prices– they unknowingly equate their marginal rates of substitution to the marginal rate of substitution of all the other traders. In other words:

$$mrs^A(x^A, y^A) = \frac{u_x^A}{u_y^A} = \frac{p_x}{p_y} = \frac{u_x^B}{u_y^B} = mrs^B(x^B, y^B) \tag{14.2}$$

Equation 14.2 says that even though they did not intend it, by pursuing their own interest – maximizing their utility – they implement an allocation in which their marginal rates of substitution are equal, and which, therefore is Pareto-efficient.

REMINDER Remember that a *relative price* can be thought of as the *ratio of prices* of two goods, e.g. $\frac{p_x}{p_y}$, which shows the price of one good (x) *relative to* the price of another good (y). We have also thought of relative prices as the *marginal rate of transformation* or the *opportunity cost* of choosing one good over another.

What information did Adamo and Beatriz need to have to unintentionally implement an efficient outcome? Not much. Each had to know their own preferences (not the preferences of the other) and the prices of the goods.

Production: Unintentionally equating marginal rates of transformation

We now extend the model to include production. To see what Pareto efficiency requires return to Figure 6.7. There we showed that an individual optimizes her utility by producing and consuming such that

$$\begin{aligned} \text{marginal rate of substitution} &= \text{marginal rate of transformation} \\ \text{which requires:} & \\ \text{slope of indifference curve} &= \text{slope of production possibility frontier} \end{aligned}$$

Each actor – Adamo and Beatriz – not only consumes the goods x and y, they *produce* the goods x and y. To produce the goods, they make choices about which technologies to use, the scale of output, the mix of inputs resulting in the marginal costs of producing the goods $c_x^A, c_y^A, c_x^B, c_y^B$.

REMINDER In Chapter 3 you saw how the person, doing the best she can, equated the slope of her indifference curve (the marginal rate of substitution) to the slope of the feasible frontier (the marginal rate of transformation or opportunity costs).

Remember that under perfectly competitive conditions firms act as price-takers, and as a result they maximize profits by producing up to the point that marginal cost equals the (given price of the good). So, considering each Adam and Beatriz as producers of the goods, the ratio of prices will equal the ratio of marginal costs for both. That is:

$$\text{A's mrt} \quad = \frac{c_x^A}{c_y^A} = \frac{p_x}{p_y} = \frac{c_x^B}{c_y^B} = \quad \text{B's mrt} \tag{14.3}$$

Bringing together production and consumption

Having added production by using each trader's costs of production, efficiency now requires that the marginal rate of substitution in people's indifference curves be equal to the marginal rate of transformation in the production of the two goods. We now have:

$$\frac{u_x^A}{u_y^A} = \frac{c_x^A}{c_y^A} = \frac{p_x}{p_y} = \frac{c_x^B}{c_y^B} = \frac{u_x^B}{u_y^B}. \tag{14.4}$$

REMINDER Optimization covers a trader's choice to maximize utility subject to their budget, to minimize costs subject to an output constraint, or to maximize profits given prices and costs.

All traders optimize with respect to the same relative prices. They therefore equate their own marginal rate of substitution in consumption as well as their marginal rate of transformation in production (the ratio of marginal costs) to the other trader's marginal rates of substitution and transformation. They thereby implement a Pareto-efficient allocation.

Equation 14.4 shows that prices convey two kinds of information:

- *Subjective value to consumers*: How *valuable* it is to those who consume or use it (measured by their marginal rates of substitution, which is their willingness to pay), and
- *Cost*: How *costly* it is to produce it (measured by the ratio of their marginal costs, which is the same thing as the marginal rate of transformation).

Equation 14.4 shows that the perfectly competitive equilibrium equates these two aspects of scarcity are the same: the marginal rate of transformation (the ratio of marginal costs, the relative costliness of the goods) and the marginal rate of substitution (the ratio of marginal utilities, the relative value of the goods to users) are both equal to the ratio of prices and therefore are equal to each other.

SCARCITY refers not to the *quantity* of a a good that is available but instead to the *value* of the good to a user or its *costliness* to a producer of the good. The economic term scarcity is what linguists call a *false friend*, that is the same or similar word that in different languages have entirely different meanings, for example embarrassed in English is a false friend of *embarazado*, meaning pregnant, in Spanish, and *sensible* meaning reasonable in English, means sensitive in Italian.

As a result, perfectly competitive equilibrium prices send messages to consumers and producers. If a drought in the American mid-west has decimated the wheat crop, then the price of bread will rise, and the message to the conumser is "maybe put potatoes on the table tonight instead of bread." If the price of tin has risen due to dwindling reserves of the metal, the owner of the firm hears: "consider redesigning your product using plastic."

Prices do more than convey messages about scarcity: they also provide the motivation to act on the information. If bread or tin is more expensive, the consumer or the producer using tin as an input will save money by shifting to an alternative.

To summarize:

Price = Message about scarcity + Motivation to act on the message

But there is a hitch: the prices have to be right. Suppose that instead of rising due to a drought, the price of bread had fallen because farmers are now using a new fertilizer, which, when it runs off into nearby rivers and streams, destroys the aquatic environment and the tourism or commercial fisheries that depends on it. The message sent by the lower price of bread would be "lets have bread tonight rather than pasta." But the message would be mistaken: the lower price does not measure the full cost of putting bread on the table.

PRIVATE AND SOCIAL COST The private cost (marginal or average) is the cost that the decision maker bears as a result of some action that he or she takes. The social cost is the private cost plus any costs imposed on others as negative external effects.

14.5 The Fundamental Theorems and Pareto efficiency

We can generalize from the bread example introduced earlier by returning to the distinction between **private costs** and benefits and **social costs** and benefits first introduced in Chapter 5. The private cost (marginal or average) is the cost that the decision maker bears as a result of some action that he or she takes. The social cost is the private cost plus any costs imposed on others as negative external effects. That is:

$$\text{Marginal social cost} = \text{Marginal private cost} + \text{External cost} \quad (14.5)$$

REMINDER The fishermen in Chapter 5 incurred a private cost of fishing, but they also imposed costs on the other fishermen due to depletion of fish stocks. The reason for the coordination failure in that case – over-fishing – was that the fishermen took account of *only* the private cost of each fishing more and this did not accurately measure the social cost of fishing because it did not include the external cost that one fishing imposed on the other.

REMINDERA Pareto-efficient can be implemented if policies are adopted so that actors so that actors bear the otherwise external costs that their actions impose on others. The fishermen in Chapter 5 reduced their fishing hours when they were forced to pay (in taxes) for the external costs of their actions on others: they internalized the external costs and therefore reduced their hours spent fishing. They therefore achieved a Pareto-efficient outcome.

For prices to send the right messages they must:

- measure how much goods contribute to consumer satisfaction (utility), and
- how much they cost society to produce (the marginal social costs and including the negative external effects).

We can use Equation 14.5 to understand whether an outcome is Pareto efficient or not. For the perfectly competitive general equilibrium allocation to be Pareto efficient Equation 14.4 is *not* sufficient. The reason is that what perfect competition (along with the Law of One Price) ensures is that the ratio of prices is equal to the ratio of *private marginal costs.* This is what Equation 14.4 says.

But what is required for Pareto efficiency is that the price ratio equals the ratio of marginal social costs. That is, Pareto efficiency requires that the price ratio equals the *entire* costs the producer's marginal private costs plus the external costs the producer imposes on others.

In the example of bread prices and wheat farming above, this means that the price of bread would have to include the costs that the new fertilizer imposes on fishing and tourism. In the example of "conspicuous consumption" in Chapter 7 building a luxury home in an otherwise modest neighborhood imposes a disutility on the neighbors. This disutility is an uncompensated external effect, so it would have to be included in the price of the new home for home sales to result in Pareto efficient outcomes.

For marginal private costs to equal marginal social costs it must be the case that:

- *No missing markets*: there are markets in every good and service that people value, so that everything that matters has a price.
- *No uncompensated external effects*: when people exchange goods and services any aspect of the production and use of the good that affects anyone's well-being (including those not party to the transaction) is measured in the price.

When these two conditions are both met, we say contracts are complete. Where markets are entirely missing we do not have **complete contracts** because *there are no contracts*.

COMPLETE CONTRACTS mean that there are

- *No missing markets*: there is a market in every good and service that people value, so that everything that matters has a price and
- *No uncompensated external effects*: when people exchange goods and services any aspect of the production and use of the good that affects anyone's well-being (including those not party to the transaction) is measured in the price.

Where these conditions are met, the equilibrium of a perfectly competitive economy will be Pareto-efficient. This is expressed what is termed the First Fundamental Theorem of Welfare Economics but which might better be called, honoring Adam Smith, the Invisible Hand Theorem.

First Fundamental Theorem of Welfare Economics: *A perfectly competitive equilibrium of an economy with complete contracts is Pareto-efficient*

HISTORY This result was proven independently by Kenneth Arrow and Gerard Debreu in 1951 and published in paper they co-authored three years later.

FIRST WELFARE THEOREM A perfectly competitive equilibrium of an economy with complete contracts is Pareto-efficient.

When contracts are complete, the competitive general equilibrium of the economy is efficient because prices send the right message: the marginal private cost of a good to the firm is exactly equal to the marginal social cost to society of having another unit of that good available (the marginal social cost). We can summarize the two requirements that if met ensure that an equilibrium be Pareto efficient:

- *Perfect competition:* All markets must be *perfectly competitive*, so that in equilibrium each buyer and seller take prices as given, and the Law of One Price holds. This insures that *prices will be equal to the marginal private cost* of their production.
- *Contracts must be complete*: so that there are no uncompensated external effects. This insures that *marginal private cost equals marginal social cost.*

If both conditions hold then the ratio of the prices of any two goods will be equal to the ratio of their marginal social costs.

Using the superscripts P to refer to marginal private costs and S to refer to marginal social costs we can extend Equation 14.4 to take account of the requirement that marginal private costs = marginal social costs:

$$mrs(x,y) = \overbrace{\frac{u_x}{u_y} = \frac{p_x}{p_y} = \frac{c_x^p}{c_y^p}}^{\text{Perfect competition}} = \underbrace{\frac{c_x^p}{c_y^p} = \frac{c_x^s}{c_y^s}}_{\text{Complete contracts}} = msrt(x,y) \qquad (14.6)$$

Equation 14.6 is the usual condition for Pareto efficiency stating that the marginal rate of substitution in consumption (mrs) must be equal to the marginal rate of transformation in production. But here we use $msrt$ to mean the marginal *social* rate of transformation, meaning the ratio of marginal social costs, that is, taking account of both private costs and the costs imposed on others by the uncompensated external effects of the production of the good.

Equation 14.6 also makes it clear that the equality of the mrs to the ratio of prices and the ratio of marginal private costs and marginal private rate of transformation mrt will be true under the conditions of perfect competition while the private and marginal social rate of transformation will be equal if contracts are complete. This summarizes the conditions under which the a general equilibrium of the economy will be Pareto efficient:

- perfect competition and
- complete contracts.

14.6 Perfectly competition and inequality: Distributional neutrality

Economists celebrate the First Welfare Theorem because it explains what is required for the prices to be right. Doing this clarifies the (very demanding) conditions under which a perfectly competitive market would allow buyers and sellers to realize all of the possible mutual gains from exchange. The same model also has a lot to say about how these gains from mutually beneficial exchanges will be shared.

A remarkable feature of the model is that at the competitive equilibrium prices the value of each actors' endowment wealth is the same as the value of the goods they hold in the competitive equilibrium after exchange. This means that the distribution of wealth is unchnaged by the process of exchange, a feature termed the distributional neutrality of the market.

Distributional neutrality: A graphical illustration

As in Figure 14.6, in Figure 14.8, any point in the Edgeworth box represents a feasible allocation of the entire supply of the goods ($\bar{x}$ and $\bar{y}$) between the

M-CHECK In Figure 14.8 and later figures we have given the two traders quasi-linear utility functions (linear in the good y) so that the Pareto-efficient curve is a vertical line. This occurs because changing the amount of y that the two have does not alter their marginal rates of substitution, so Pareto efficiency is a characteristic of a single distribution of the x good, but is *independent* of the distribution of the y good. This device allows us to focus here on distribution while setting aside the question of Pareto efficiency.

REMINDER Remember that $\bar{x} = x^A + x^B$, the

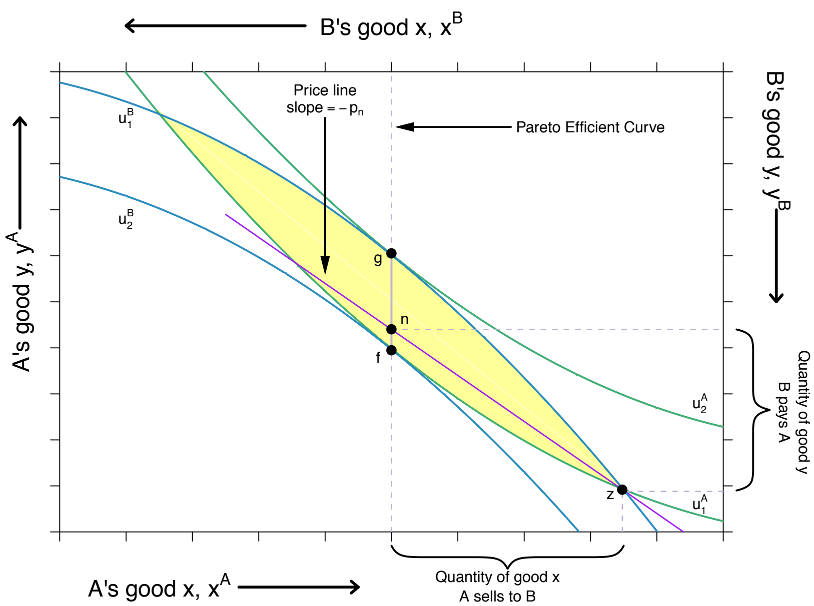


Figure 14.8: **Distributional neutrality of the perfectly competitive equilibrium.** At their endowments, point z, Adamo has 8.5 units of x and 0.9 units of y. He trades 3.5 units of x, to obtain 3.5 units of y, resulting in his competitive equilibrium allocation of 5 units of x and 4.4 units of y. At the equilibrium, Beatriz has 5 units of x and 5.6 units of y. The market-clearing price is shown by the slope of the market-clearing price line, $-p^N$ (this is the ratio of the number of units of good y exchanged for units of good x, therefore a *relative* price). The value of the traders' wealth at the equilibrium outcome (n) is the same as the the value of their wealth at their endowments (z). So with the Nash equilibrium price, $p^N = 1$, for Adamo, $p^N \cdot x_n^A + y_n^A = p^N \cdot x_z^A + y_z^A = 9.4$ and for Beatriz $p^N \cdot x_n^B + y_n^B = p^N \cdot x_z^B + y_z^B = 10.6$.

two people, that is, an allocation that exhausts the supply of both goods. We consider two people – Adamo and Beatriz – but imagine that there are a great many people similar to the two in the market. We can think of Adamo as one type of trader (an A-type) and Beatriz as another (a B-type).

The two traders each have a positive but different *initial holdings* or *endowments* of both goods (x_z^A, y_z^A) and (x_z^B, y_z^B) indicated in the figure by point **z** (as in the previous figures, and in Chapter 4). The endowment is an exogenous initial distribution of wealth, the determination of which is beyond the model.

In Figure 14.8 before exchange takes place the Adamo-types have a lot of the good x and not much of the y good, while the reverse is the case for the Beatriz-types. So it is natural to guess that Adamo will be the seller of the A good and Beatriz the buyer. We now suppose that the Adamos and Beatrizes trade on a perfectly competitive market and ask:

- How much (what quantity) of the x good will be bought and sold?
- What will be the good's price?

REMINDER Recall that with a quasi-linear utility function that is linear in the good y, the person's marginal rate of substitution does not depend on how much of the y good she has. As a result a difference in the amount of y that the person has simply shifts up or down the indifference curves without changing the slope. We call this a 'vertical displacement' or 'vertical shift.'

To do this we introduce the traders' preferences. The indifference curves for Adamo (and all the other traders like him) are measured from the lower-left (south-west) origin while the indifference curves for Beatriz (and all the other traders like her) are measured from the upper-right (north-east) origin.

The participation constraints of the two traders – u_z^A and u_z^B – are the indifference curves passing through their initial endowments (**z**). We can see from Figure 14.8 that mutual gains are possible through trade: trading from the endowment point to any point in the Pareto-improving lens will make the traders better off.

The Pareto-efficient curve in the figure shows all of the allocations of the two goods that are at the tangencies of the two traders' indifference curves, where their marginal rates of substitution are equal. The first welfare theorem states that if the two traders have the endowments given by point **z**, the perfectly competitive equilibrium will occur at some allocation between points **f** and **g**, which satisfies the participation constraints and requires that the final allocation is Pareto-efficient.

REMINDER The Pareto-efficient curve in the figure is similar to the one shown in Chapter 4, when Ayanda and Biko were engaged in the trade of work hours.

The Law of One Price means that all transactions are made at the same market equilibrium price. The Law has two far-reaching implications.

- *Equal treatment*: An implication of Walras' assumption that all transactions take place at equilibrium prices is that two consumers who have the same preferences and initial endowment ownership will end up in equilibrium with the same consumption bundle, because they face the same equilibrium prices and the same budget constraint, and will choose the same utility-maximizing consumption plan. This is the *equal treatment property* of Walrasian competitive equilibrium. No matter which Adamo-type we consider, all of their transactions will be at the same price, so all of the Adamos will end up at the same final allocation. (The same will be true of all the Beatrizes, of course.)

EQUAL TREATMENT An exchange process has the characteristic of *equal treatment* if consumers with the same preferences and endowment have the same consumption.

- *Distributional neutrality*: The final allocation reached by all the Adamos will have the same value (as total wealth) at the market equilibrium price, p^N, as the initial endowment (**z**). This follows because the value of any Adamo's holding of the commodities will not change in any transaction when the transaction takes place at the market equilibrium price. So all points on the price line – including the endowment and the equilibrium allocation – have the same wealth value (at the price given by the slope of the price line). The value of what they get at the market equilibrium prices is the same as the value of the initial endowment at those prices. The same is true of all the Beatrizes.

DISTRIBUTIONAL NEUTRALITY means that the process does not change the distribution of wealth between the traders. In the case of competitive trading, Adamo and Beatriz can evaluate their wealth at the competitive prices and their wealth before and after exchange will be the same (the quantities of each good multiplied by each good's price). In other words, at equilibrium prices the distribution of wealth is the same at point **z** (at their endowments or fallback) and point **n** (a corresponding Walrasian competitive allocation or Nash equilibrium). This is true because each party's wealth remains the same at the equilibrium price, which passes through both points (**z** and **n**). As a result, any inequality in wealth that existed before exchange will also exist after exchange. But, the traders' *utility* will differ between the pre- and post-exchange allocations.

In addition to Pareto efficiency we have a second important result about the equilibrium of a perfectly competitive economy: the degree of inequality that results after the trading process is identical to the inequality in the value of the traders' initial endowments, when the endowment and post-exchange allocations are valued at the equilibrium prices. Some economists think that this feature of the equilibrium outcome is as important as its efficiency properties.

M-Note 14.1: Wealth and distributional neutrality

At his endowment, trader A has (x_z^A, y_z^A). Distributional neutrality requires that a trader's wealth – the value of the endowment at competitive prices – be the same at their endowment as it is at the post-exchange allocation determined by market prices. There are two goods in the market, x and y. We call good y the numéraire and say that its price is $1. A numéraire is an index good against which all other goods are priced. We make y the numéraire in this context because y affects the traders' utilities linearly – it has a marginal utility of 1.

We can now consider the Nash equilibrium prices, p^N, of good x as the price at which the outcomes must be distribution-neutral.

Therefore, if a trader's wealth at their endowment is given by m_z^A and m_z^B respectively, then for distributional neutrality to be true, it must be the case that:

$$m_z^A = y_z^A + p^N x_z^A \quad = \quad y_n^A + p^N x_n^A = m_n^A \tag{14.7}$$

Equation 14.7 says that the Nash equilibrium price, p^N, is the price at which the trader's wealth at their endowment (x_z^A, y_z^A) is equal to their wealth at their post-exchange allocation (x_n^A, y_n^A). The same is true for B.

Markets, inequality and wealth redistribution

REMINDER In addition to complete contracts, the second theorem requires – for the model economy in the previous figure – that indifference curves have the usual concave (bowed in) to the origin shape. Confusingly to some, this is called the *convexity assumption*. In a model that includes production of the goods and not just their exchange, convexity requires that there not be economies of scale in the production of the goods.

The equal treatment property – equal wealth at the endowment and the post-exchange allocation – leads to an interesting challenge. How does a society deal with inequality if market processes ensure that wealth is the same before and after exchange? What is called the Second Welfare Theorem addresses how an outcome that is both fair and efficient might be implemented by the combination of a redistribution of wealth and the process of perfectly competitive market exchange.

Theorem 14.1: Second Welfare Theorem

Any Pareto-efficient allocation can be implemented by an assignment (a redistribution) of the endowments (the wealth) of all parties, followed by a perfectly competitive market exchange process.

SECOND WELFARE THEOREM Given complete contracts, any Pareto-efficient allocation can be implemented by some assignment of the endowments of all parties, followed by a perfectly competitive market exchange process.

To see the theorem's importance, suppose that the citizens of an economy wish to redistribute goods to the less well off members of society and select a particular Pareto-efficient allocation as their preferred outcome. The **Second Welfare Theorem** says that an alternative, fair outcome can be implemented by reassigning property rights among the citizens (changing who has what at the initial endowments) followed by a Walrasian exchange process. For the Second Welfare Theorem, in addition to complete contracts (no external effects), a further assumption is required. We need to rule out, for example, increasing (instead of decreasing) marginal utility in the goods people consume and any substantial degree of economies of scale in production.

The process of redistribution for the Second Welfare Theorem is illustrated

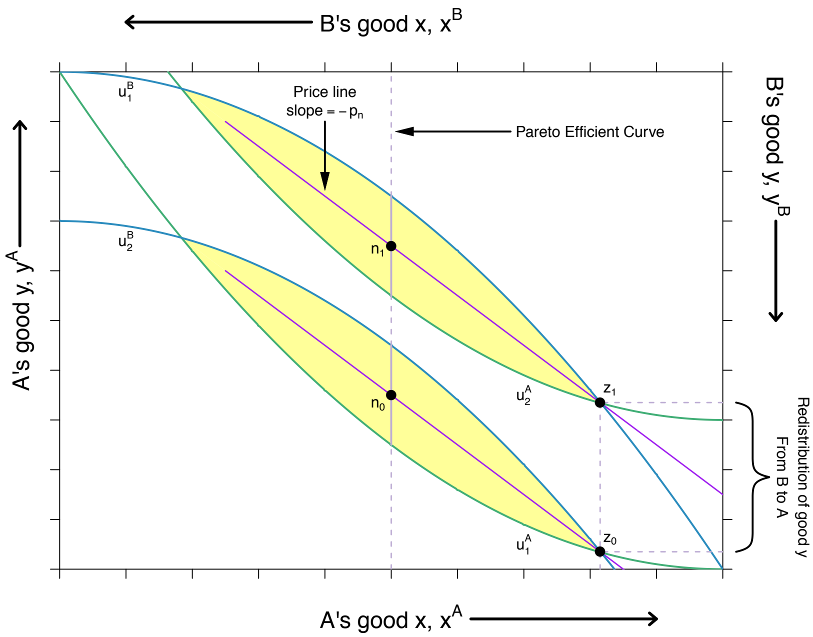


Figure 14.9: **The Second Theorem of Welfare Economics.** Initially, each Adamo has some of the good, x^A, but very little good y, y^A, as shown by point z_0. If perfect competition unfolded as usual with an imaginary auctioneer, then the competitive price would be given by p^N and the traders would end up at n_0. But, if the government – or another third party – chose to redistribute the initial endowments to z_1 by taking good y from each Beatriz and giving it to each Adamo, then trading would start at point endowment z_1. As a result, at the competitive prices p^N, the traders would arrive at the Pareto-efficient outcome n_1.

by Figure 14.9 in which the endowments are redistributed from z_0 to z_1, after which Walrasian exchange at competitive prices takes the traders to n_1 rather than the original post-exchang allocation n_0.

HISTORY Nobel Laureate Amartya Sen (1985: 11) wrote that the Second Welfare Theorem "belongs to the revolutionists' handbook" because of suggested than an efficient way to reduce inequality is by redistributing wealth.

Consider what would happen if the Adamos and Beatrizes making up the population decided that the Adamos really deserve more and they thought that this could be achieved by changing the *price* rather than by redistributing goods before market exchange. Because the Adamos are the sellers of the good, this could be done by passing a law setting a price of the good p_1 that is higher than the competitive equilibrium price p^N and then letting the two groups exchange goods in any way they wished. But, at a higher price, markets would not clear (p^N is the market clearing price) and there would be excess supply of the good. We also know that the result would not be Pareto-efficient.

Suppose instead that they decided to redistribute the *wealth*: taxing the wealth of the Bs and giving the revenue to the As. The new endowment would then be z_1 and n_1 could be reached from there by the two exchanging goods at the market clearing prices. So under the assumptions of the Second Welfare Theorem, particularly the assumption that all exchange takes place at the eventual equilibrium prices, wealth redistribution followed by Walrasian exchange represents a mechanism that can implement *any* equal-treatment Pareto-efficient allocation.

Why are the results of the Second Welfare Theorem interesting to economists?

Often, governments adopt policies to alter final allocations of goods and services by changing *prices*, for example, by placing price controls on certain goods, or taxing others, or minimum wage laws that put a floor under wages, and rent control laws that put a ceiling on rents. The Second Welfare Theorem tells us that instead of changing prices, governments seeking to address economics inequalities could do so by altering the initial distribution of endowments and then letting markets operate, rather than by changing prices. The contrast of the two approaches is illustrated to policies that helped poor farmers.

HISTORY Shortly after South Africa's first democratic election in 1994 that ended the system of racial inequality called apartheid, President Nelson Mandela asked one of the authors of this book for advice about economic policy. Bowles recommended that the land and other wealth that had been confiscated from the black African population should be returned to their former owners (or their families) and that the market exchange process would then implement a more just allocation. He was simply repeating the logic of the Second Welfare Theorem.

- *Changing prices*: The government of the Indian state of West Bengal, elected with support of less well off farmers, placed a ceiling on the share of the farmers' crops that could be claimed by the land lord. This is a policy to affect *prices*, in this case, the rent that the tenant farmers have to pay to their landlords (the owners of the farms).
- *Redistributing wealth*: Half a century ago the governments of South Korea and Taiwan adopted policies that limited the amount of land that large landlords could own and distributed land to landless farmers. This is how policy can alter *initial endowments*, in this case the *ownership* of the land through redistributing wealth.

Efficiency and neutrality: Why the theorems are important

When the two Fundamental Welfare Theorems are taken together they appear to leave little room for ethical concerns about the operation of a competitive market system except for the distribution of well-being. The distribution of well-being is determined not by markets themselves but rather the distribution of initial endowments because "markets" are distribution neutral.

Kenneth Arrow pointed out that under the conditions specified by the theorems that he first proved:

> Any complaints about [the market system's] operation can be reduced to complaints about the distribution of income . . . [but] the price system itself determines the distribution of income only in the sense of preserving the status quo.

John Roemer's, thinking about the relationship between workers, capital and the labor discipline model we explored in Chapter 11, said: "If the exploitation of the worker seems unfair, it is because one thinks the initial distribution of capital stock, which gives rise to it is unfair."

Arrow's and Roemer's observations had been anticipated by the U.S. Supreme Court in its 1915 decision Coppage v. State of Kansas

> . . . wherever the right of private property exists, there must and will be inequalities of fortune; . . . it is impossible to uphold the freedom of contract and the right of private property without at the same time recognizing as legitimate these

> inequalities of fortune that are the necessary result of the exercise of those rights.

HISTORY Even philosophers like David Gauthier have drawn conclusions from the model:

> The operation of a market cannot in itself raise any evaluative issues. Market outcomes are fair if . . . they result from fair initial distributions . . . the presumption of free activity ensures that no one is subject to any form of compulsion, or to any type of limitation not already affecting her actions as a solitary individual. . . . [Thus] morality has no application to market interaction under the conditions of perfect competition.

At first glance, the combination of the concept of perfectly competitive equilibrium and the Welfare Theorems appears to vindicate Adam Smith's conjecture that competitive markets would ensure that traders be "led by an invisible hand to promote an end which was not part of" the participants' intentions. Gerard Debreu, who along with Kenneth Arrow proved what we have called the invisible hand theorem, told the French journal *Le Figaro* "The superiority of the liberal economy [meaning unregulated competitive markets] is incontestable and can be demonstrated mathematically . . . " He was referring to his own justly famous First Welfare Theorem.

But few economists take the theorem as an exoneration of any real world market institutions. Fewer still take the Second Welfare Theorem as a prescription for wealth redistribution to implement a distributionally fair Pareto-efficient allocation. In the next section, we discuss the important reasons why considering market failures is so important to understand real competitive equilibria.

14.7 Market failures due to uncompensated external effects

There are several reasons for the limited applicability of the concept of competitive equilibrium and the Fundamental Welfare Theorems. Some of the reasons concern the lack of realism of assumptions on which the theorems are based. Others point to the inadequacy of the theorems themselves even if the assumptions were empirically valid.

Here we deal with the first concern: that contracts are complete is generally a poor assumption, so the First Welfare Theorem does not apply to most real economic interactions.

Uncompensated external effects

With the example of two fishermen over-fishing a lake in Chapter 1 we introduced coordination failures – Nash equilibria that are Pareto-inefficient due to the lack of coordination among the economic actors involved. In Chapter 5 we returned to the problem, showing how the two fishermen determine the hours they devote to fishing by balancing the marginal costs (their own disutility of effort in fishing) and the marginal benefits (how much each additional hour added to the catch). The coordination failure among the fishermen was due to the fact that each fisherman – in deciding how much to fish – did not include in the costs of fishing the negative effect of their effort on how much fish the other fisherman caught.

Another example: when fuel costs are low, more people decide to drive to work rather than taking public transport or choosing to ride their bicycles. The information conveyed by the low price does not include the environmental costs of deciding to drive. The effects on the decision-maker are termed *private* costs and benefits, while the total effects, including costs inflicted or benefits enjoyed by others, are *social* costs and benefits.

We can understand why these and other market failures are common by thinking about how they could be avoided. How could the cost of driving to work accurately reflect all of the costs incurred by anyone, not just the private costs made by the decision maker? The most obvious (if impractical) way would be to require the driver to pay everyone affected by the resulting environmental damage (or traffic congestion) an amount exactly equal to the damage inflicted.

REMINDER Costs inflicted on others (pollution and congestion that are worse because you drive to work) are termed negative external effects; while uncompensated benefits conferred on others are positive external effects.

This is impossible to do, but it sets a standard of what has to be done or approximated if the "price of driving to work" is to send the correct message. A similar approach applies if you drive recklessly on the way to work, skid off the road, and crash into somebody's house. Tort law (the law of damages) in most countries would require you to pay for the damage to the house. You are held liable for the damages so that you would pay the cost you had inflicted on another.

Knowing this, you might think twice about driving to work (or at least slow down a bit when you are late). It will change your behaviour and the allocation of resources. But while tort law in most countries covers some kinds of harm inflicted on others (reckless driving), important external effects would not be covered by tort law (adding to air pollution of congestion by driving your car).

Table 14.1 gives examples of important types of market failure. In all of these examples there are *uncompensated external effects*. For each of them, it is the case that either:

- the *private costs* to the decision-maker of an activity differ from the *social costs* including both the private costs and the negative external effects on others, in this case the private costs are less than the social costs; or
- the *private benefits* to the decision maker differ from the **social benefits**, in this case the social benefits exceeding the private benefits; or of course
- both of the above

Consider an example from Table 14.1. A farm uses pesticides that contaminate the local water supply. The firm's private costs for its pesticide use do not include the external cost the farm's owners impose on fishermen and private citizens who use the river water or other water sources contaminated by the use of the pesticides. Because the farmers do not pay the full cost

Type of problem	*The decision*	*Uncompensated external cost or benefit*	*Market failure*
Public good	A firm invests in R & D	Other firms can use the innovation	Too little R & D
Public bad	You take an international flight	You increase carbon emissions	Overuse of airplanes
Negative external effect	A farm uses pesticides that contaminates water	Damage to fishing industry	Overuse of pesticide
Positive external effect	A firm trains a worker	Another firm benefits if the worker quits	Too little worker training
Common property resource	You travel to work by car	Congestion for other road users	Overuse of roads and highways
Moral hazard (labor)	An employee on a fixed wage decides how hard to work	Hard work increases her employer's profits	On the job effort is too low
Moral hazard (credit)	A borrower decides how much risk to take	Lender is exposed to default risk	Excessive risk
Adverse selection (insurance)	A person with an undisclosed illness buys insurance	Insurer exposed to risk, insurance costs rise	Insurance is excessively costly or non-existent

Table 14.1: **Examples of uncompensated external effects and market failures**.

of the pesticides (social cost = private cost + external cost), they *over-use* pesticides.

Now consider an example not from the table, but describing the kind of neighborhood you might live in. A home-owner plants a beautiful garden that her neighbors can see. She enjoys the flowers she has planted at his house, but the social benefits include her enjoyment of the flowers as well as the enjoyment of neighbors who pass by. Because home-owners do not reap the full benefits of maintaining a beautiful home (social benefits = private benefits + external benefit), many will under-provide home maintenance and garden cultivation.

Incomplete contracts and missing markets

If the gardener somehow owned the sight of her flowers which she could sell or rent to people walking by, or the neighbors of the polluting farmers owned the clean water around their house which they could sell to the polluting farmers, the market failure could be avoided. Market failures occur because the external benefits and costs of a person's actions are not owned by anyone.

Think about waste: if you redecorate your house and you tear up the floor or knock down a wall, you own the debris and you have to dispose of it, even if you need to pay someone to take it away. But this is not the case with pollutants from a farm or loud music you might play at night. You do not have a contract with the farm company specifying at what price you are willing to accept contaminated water, or with your neighbour about the price of the right to play music after 10pm.

If the problem is uncompensated external effects, why don't countries just rewrite their laws so that benefits conferred on others must be rewarded, and costs inflicted on others be paid by the decision-maker?

In Chapter 11 we reviewed the reasons why the kinds of contracts that would enforce these objectives are incomplete or unenforceable: the necessary information is either not available or not verifiable, the external effects are too complex or difficult to measure to be written into an enforceable contract, or there may be no legal system to enforce the contract (as in pollution that crosses national borders).

For these and other reasons, in most cases it is impractical to use tort law or any other body of law to make people liable for the costs they inflict on others, because we don't have the necessary information, and even if we did the legal costs of adjudicating the cases on a one by one basis would be prohibitive. It is equally infeasible to use the legal system to compensate people for the beneficial effects they have on others, for example, to pay those who keep beautiful gardens an amount equal to the pleasure this confers on those who pass their house, because a court would have to know how much that pleasure was worth to each person who walked by.

In the earlier examples of uncompensated external effects, the reason why uncompensated external costs and benefits occur is the same:

REMINDER When information is *non-verifiable* it means that it cannot be taken to a court or some other third party who will adjudicate on a contract, and the contract is therefore incomplete even if information is not asymmetrical between the parties to the contract.

- Some information about an aspect of an exchange that is of concern to someone other than the decision-maker is *asymmetric* or *non-verifiable*.
- Therefore there can be no contract or property rights ensuring that external effects will be compensated.
- As a result, some of the social costs or benefits of the decision-maker's actions will be *excluded* from (or will not be sufficiently valued in) the decision-making process.

The result is a market failure even in a perfectly competitive equilibrium. Market failures are a kind of *coordination failure*, where the broader term is sometimes used to describe a Pareto-inefficient outcome in problems like traffic congestion where the relevant social interactions are not on a market.

14.8 Market dynamics: Getting to an equilibrium and staying there

EXAMPLE Many results in economics depend on *proofs of existence*, that is, defining the conditions under which specific outcomes will occur. But, the existence of an outcome does not mean that it is meaningful.

But does price-taking competitive equilibrium help us to understand real economies? For us to answer that question, we need to know, in addition:

- Would an economy that got to a price-taking equilibrium stay there?
- How could an economy get to a price-taking competitive equilibrium to begin with?

Staying there: Is a price-taking "equilibrium" a Nash equilibrium?

In the real world, the markets that most closely approximate price-taking are also those where the commodity being traded has been standardized and certified by some third party (like the market for #2 Red Winter Wheat we introduced in Chapter 10). The standardization of commodities increases the effective number of competitive suppliers and greatly reduces the influence any one competitor can have on the market price. Large numbers of competitors make it more difficult for firms to collude.

REMINDER We explained the dynamics of the markets for commodities with complete contracts, such as wheat or corn in Chapter 10. Refresh your memory about the characteristics of these goods by reading the introduction to that chapter.

If an entire economy were like the market for #2 Red Winter Wheat with goods defined by complete contracts as enforced by the Chicago Board of Trade and the found itself in a price-taking competitive equilibrium, would it stay there? Or would buyers and sellers have an incentive to change their behavior in ways that would result in a different kind of equilibrium?

If, in a competitive price taking equilibrium, there is something a firm or a family could do to raise its profits or utility – for example acting like a price-maker rather than a price-taker – given what everyone else is doing (that is, price taking), then the price-taking "equilibrium" is not a Nash equilibrium and it will not persist.

REMINDER In Chapter 9 we discussed how firms would *differentiate* their goods to ensure that they could maintain market power.

The first thing that at least some firms would explore is finding a way to escape from the competition of other firms by making their product distinct in some way so that other firms' products are no longer perfect substitutes for the firm's differentiated output. Product redesign or advertising might accomplish this.

If successful, the firm would then face a downward-sloping demand curve, and be able to restrict output, charge a price greater than marginal cost, and make profits above the opportunity cost of capital. Even a small amount of price-making power could be sufficient to attract the firm away from the price-taking behavior.

The important conclusion is that price-taking may not be profit-maximizing behavior even in a competitive market because there are opportunities to make more profits by deliberately altering the nature of the competition that a firm faces.

Even without product differentiation, in real-world economies firms may face downward-sloping demand curves because the number of effective competitors they face is limited. One example of these limits is geography; think of restaurants, how many competitors does a mid-priced Italian restaurant have in a small city?

REMINDER In the Cournot model of Chapter 9 the typical firm sees itself as sharing the market with a limited number of competitors, and therefore as having some power as the marginal producer to influence market price.

Checkpoint 14.2: Adam Smith on competition

In *The Wealth of Nations*, Adam Smith famously wrote that "People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices." Less often cited is his next sentence: "It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty." On the basis of this passage how would Smith respond to the following questions:

- Is price taking behavior among competitors a Nash equilibrium?
- What should be the objectives and limits placed on policies to ensure competition among firms?

Getting there: The parable of the auctioneer

Walras – whose model of competition we have so far employed – failed to explain the *process* by which an economy of consumers and firms would reach the price-taking competitive equilibrium that his equations describe. In one edition of his book, Walras proposed the idea that the market-clearing equilibrium prices could be found by adding to the economy adding a fictional character, the *Auctioneer*.

The Auctioneer, Walras wrote, "cries out a system of prices at random" and asks that all the consumers and firms report their price-taking profit- and utility-maximizing supplies and demands of all the commodities at those random prices. Given that the Auctioneer is just randomly choosing prices, for any given set of prices suppliers would typically produce:

- *excess demand*: too little of some commodities (so that demand exceeds supply), and
- *excess supply*: too much of others (supply exceeding demand).

But, excess supply and demand are not a problem for the auctioneer because no one actually produces or exchanges at these random prices in Walras's parable. Walras thought that the Auctioneer could by trial-and-error (*tâttonement* in French) find market-clearing price-taking competitive equilibrium prices. Actual production, exchange, and consumption would take place, according to Walras's scheme, only once the equilibrium prices were known to all traders and everyone could restrict their transactions to the equilibrium prices.

Neither Walras nor anyone since has imagined such experimentation by an Auctioneer is how prices and market clearing come about. Walras created the parable as a device to demonstrate that market-clearing prices could exist and could somehow be discovered. The lack of a convincing account of how an economy might get from an initial endowment to a price-taking competitive equilibrium challenges a common interpretation of the Second Welfare Theorem: namely that redistribution of endowments followed by market exchange can implement any Pareto-efficient allocation.

HISTORY All that Kenneth Arrow (Nobel Laureate) and Frank Hahn claimed for the Second Welfare Theorem is that "in a certain sense any desired efficient allocation can be achieved by redistribution of initial assets followed by the achievement of an equilibrium" (Arrow and Hahn (1971: 95)). They are careful not to suggest that the equilibrium can be achieved in a decentralized manner. They illustrate the Second Fundamental Welfare Theorem with an example of "an omniscient state" that "computes a price vector ... satisfying the hypotheses of the theorem."

How do buyers and sellers act out of equilibrium? Can buyers and sellers move to a Pareto-efficient equilibrium from any initial endowment? The Walrasian model does not provide answers to these questions. As a result, we cannot say *how* buyers and sellers can achieve any Pareto-efficient equilibrium from a given initial endowment without exploring alternatives to Walras' Auctioneer.

14.9 Bargaining and rent-seeking: A more realistic model of market dynamics

But traders might get to an equilibrium without the services of the Auctioneer. How? Traders can reach market equilibrium through a decentralized process involving actual trades at disequilibrium prices. To do this we return to thinking about traders as perfect competitors (introduced in Chapter 9), not as price takers as in the Walrasian perfectly competitive general equilibrium. A perfect competitor is an economic actor who seeks out each and every opportunity for gains through exchange, driving the economy to an equilibrium by a process of rent-seeking that continues until there are no rents left to seek.

REMINDER Remember (from Chapter 9 that price-taking is not an *assumption* about how people behave in general. Price-taking is the *result* of a perfectly competitive market equilibrium. Price takers at the market equilibrium become *price makers* when the market is out of equilibrium. Competitive markets reach an equilibrium through some of the buyers or sellers changing their prices in order to capture the rents that exist when the market is not in equilibrium. Price-taking does not mean that buyers or sellers cannot set a price different from what others are setting; they can set any price they wish. Price-taking means they cannot benefit from doing so.

More realistic assumptions for exchange

What does an adequate account of such a system require?

Let us start with basic facts about the *people trading*: They are empirically plausible in what they are attempting to do and in what they know.

- people differ in their preferences and, because of specialization and the division of labor, people differ too in their endowments
- people trade voluntarily and refuse exchanges that make them worse off, so the prices at which they exchange are mutually agreed upon
- people know their own preferences but not those of (most) other people

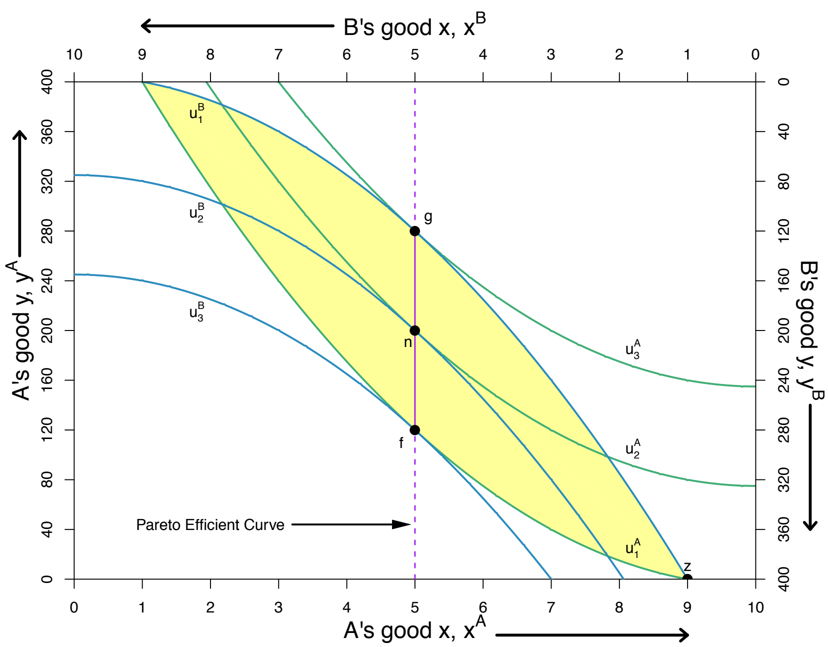


Figure 14.10: **The basic Edgeworth box that two traders would confront in the simulation.** The traders have identical quadratic, quasi-linear preferences with $\bar{p} = 100$ and $\bar{x} = 10$. All of the Type A traders – the Adamos – have several items of the good ($x^A = 9$) to sell, but no good y ($y^A = 0$). We could think of them as the *producers* or *sellers* in this economy. All of the Type B traders – the Beatrizes – have very little of the good ($x^B = 1$), but a significant endowment of cash to purchase the good ($y^B = 400$). We could think of them as the *buyers* in the economy. These simulation models are based on similar simulations from Foley (1994).

Next are the facts about the institutions that provide the rules of the game for trading: the process of exchange is truly *decentralized*.

- *Decentralized trade:* trading is arranged by traders not by a government or some other centralized body
- *No auctioneer*: there is no imaginary Auctioneer saying which trades are allowed
- *Trade is voluntary*: a trade takes place if the trade is mutually beneficial to the two traders involved, and does not take place if it is not mutually beneficial.

Duncan Foley, the economist who pioneered this approach, described a set up like this:

> [traders] enter the market knowing only the transactions they view as improving their condition given their endowments, preferences, technology, and expectations [they] encounter other [traders]; and make mutually advantageous transactions with them in a disorderly and random fashion.

We can build a theory of real-world decentralized allocation of resources from these basic facts to provide an account of how uncoordinated actions of self-interested consumers and firms would systematically lead to a Pareto-efficient outcome. We do not need the Auctioneer to show that the equilibrium of a competitive market exchange process can be efficient (assuming complete contracts so as to preclude difficulties that arise from incompleteness).

HISTORY Stephen Smale, a mathematician, introduced an element of market realism by abandoning the Auctioneer and allowing transactions take place at non-equilibrium prices. In his model, similar to Duncan Foley's, starting from an initial endowment, people participate in a series of exchanges consistent only with the requirements that the transaction increase the satisfaction of the parties to the exchange and that no such mutually beneficial exchanges remain unexploited. Traders reach an equilibrium that is Pareto-efficient in this model, similar to the scenario illustrated in Figure 14.11. Smale comments: "The exact equilibrium depends on factors such as which agents first encounter each other." As in Foley's model, final wealth and utility therefore come down to an individual trader's trading history and luck.

Let us suppose an initial interior endowment is represented in Figure 14.8 by point **z**, namely an allocation such that:

$$mrs_z^A = u_x^A / u_y^A \quad < \quad u_x^B / u_y^B = mrs_z^B \tag{14.8}$$

So the condition for a Pareto-efficient allocation is *not* met at the endowment because the marginal rate of substitution of Beatriz-types exceeds the Adamo-type's marginal rate of substitution. The traders recognize the differences in willingness to pay and willingness to accept, so each Adamo might wish to exchange some of his z for some of a Beatriz's y and each Beatriz would conversely wish to trade some of her y for some of an Adamo's x, so mutually advantageous trades will be possible. But at what price will each pair of traders transact?

Traders engaging in an exchange that results in an allocation in the lens formed by the two indifference curves u_z^A and u_z^B find that their trade is both feasible and represents a Pareto-improvement over their initial endowments. Any exchange process in which agents seek out mutually advantageous trades as long as any are possible has to lead them eventually to a Pareto-efficient allocation (as the First Welfare Theorem asserts). But which allocation on the Pareto-efficient curve and in particular what distribution of utility they will reach depends on the details of the institutions and preferences governing the interaction. The resulting final allocation will be **path-dependent** on the exact pattern of transactions.

PATH DEPENDENT We spoke about path dependence in previous chapters, starting in Chapter 1. In basic terms, path dependence means that "history matters", that is, the sequence in which trades occur matters for which traders outperform other traders, or which equilibrium of a game is more likely to occur.

Each trader might engage in a series of trades with different partners, always implementing Pareto-improvements (remember trade is voluntary!). This process will continue until *no one* can find a partner with whom a mutually advantageous trade was possible. When this is the case, we know two things:

- *Nash equilibrium*: The outcome must be a Nash equilibrium because it is true for each trader that she cannot find any other trader with whom a mutually beneficial exchange can happen
- *Pareto efficiency*: The outcome is Pareto-efficient for the same reason that it is a Nash equilibrium: there are no mutually beneficial exchanges that could be carried out.

Without knowing more about the details of the exchange process, such as the exact order in which the agents meet and the exact trades they might make, we cannot say where on the efficient set the economy would land.

There are many trading paths that produce Pareto-improvements at each step. One Adamo may trade consistently at more favorable prices than another on some particular trading path, and get a larger share of the consumer surplus. We could conceive of other trading processes, but one cannot say much about the outcomes of the exchange process unless the institutions governing it are specified.

The market process can be thought of as a large collection of bargains over potential surpluses like the interactions we have analyzed in Chapters 4 and 9.

Computerized simulations of general equilibrium

We can illustrate how the exchange process might work using a computer simulation. We simulate disequilibrium trade using two goods (x and y) and two types of traders, A and B. Each trader has their own utility function and initial endowment of the two goods. There are an equal number (1000) of each type of trader.

In each round of trading, the traders are paired randomly (traders have the same probability of being paired with another trader of the same type as with a different type). For each pair the program computes offer prices given their utilities and their current holdings of the two goods, and chooses a random price in the interval between the two offer prices. We limit trade at the random price between the traders' offer prices by the smaller of their two utility-maximizing offers. This involves calculating the demand function for each trader and ensures that the trades are feasible and do not decrease a trader's utility. Each trader's holding of the two goods is updated by the trade, and carried forward to the next round of trading.

As this process continues, it is possible that two traders of the same type enter a round with different holdings of the two commodities and therefore have different offer prices and can make a mutually advantageous trade. (This is not possible in the first round of trading because all traders of each type start with the same original endowment.)

REMINDER Remember that "offer price" is another name for marginal rate of substitution because it is the price at which a trader is willing to give up some y in order to get an extra unit of x. Therefore when the offer prices are equal, the marginal rates of substitutions are equal and the allocation is Pareto efficient.

This process continues until there are no available mutually advantageous trades, which implies that the final allocation is Pareto-efficient, and that the offer prices of all traders are equal.

When both traders have quasi-linear utility functions, their offer prices depend only on their holding of the non-linear good. In this case, all Pareto-efficient allocations give the same amount of the non-linear good to all the traders of the same type, and the final price ratio is the same as the Walrasian equilibrium, which gives the same allocation of both commodities to each type.

REMINDER Remember that a trader's participation constraint is their fallback utility or the utility that they would receive at their endowment, **z**.

We illustrate the basics of the simulation in Figure 14.10. The two types of traders started off with asymmetrical endowments. Each trader has an indifference curve going through the endowment (**z** in the figure) and these indifference curves are the traders' *participation constraints*. The participation constraints show all of the possible outcomes at which the traders would be no better off and no worse of than not trading.

She slopes of the indifference curves at point **z** show that there are many prices at which A would like to sell some of the good and B would like to buy some of it. They have identical preferences but given their holdings at the endowment point, B has a strong preference for good x compared to good y, while A prefers y to x.

The Pareto-efficient allocation of the good is for each of the As and Bs to have five units of the good. The Pareto-efficient allocations are shown in both Figure 14.10 and 14.11. The traders will go on trading until they reach that allocation. Remember trading continues until there are no mutually beneficial exchanges possible, until a Pareto-efficient outcome has been reached.

14.10 Disequilibrium trading creates inequality

What is undetermined is how much money each will have when the trading ends.

Figure 14.11 shows trading paths for two Type A traders: one who did poorly and another who did well. Though the traders do not perform equally well, their trading implements a Pareto-efficient outcomes.

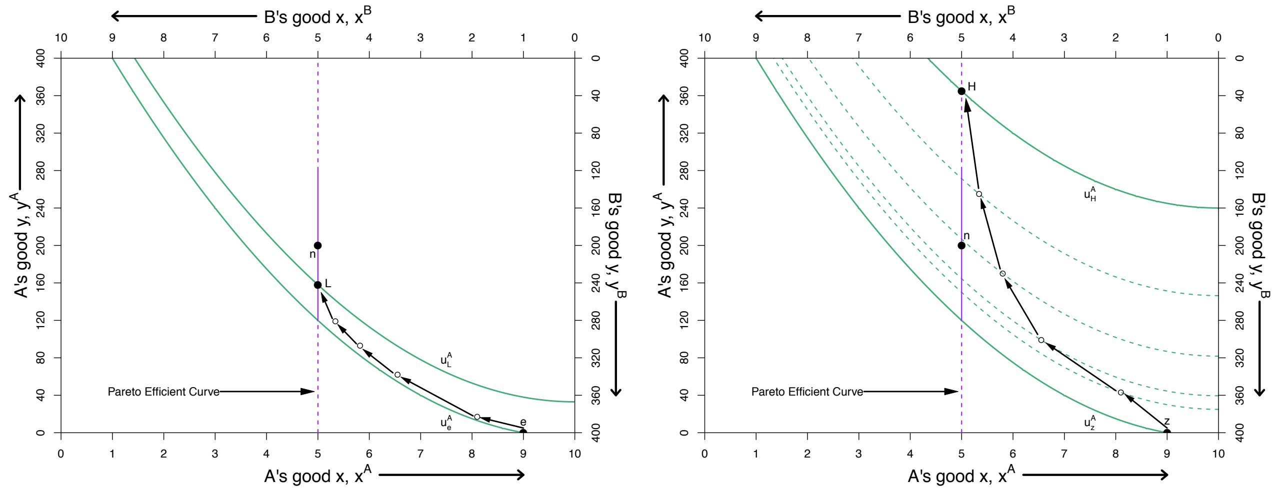
Why did one do so much better than the other? Talent? Bargaining power? Mistakes?

No, there are no mistakes: each trade made by both traders made them better off as the right panel in the figure shows, where the trader moves to ever higher indifference curves.

Moreover, the model does not advantage any one player over any other. The pairing to trade is random, and so is the selection of prices (from the mutually beneficial range at which they both wish to trade). The difference between the two is a matter of *luck*. In our simulated model, we show that the inequality in final allocations occurs as a result of randomness or luck. In real-world economies institutions, norms, and history may affect these trades, as would bargaining power, discrimination, or social networks.

Here is what we know about the outcome:

- *Pareto efficiency*: All of the traders end up on the Pareto-efficient curve.
- *Gains from trade*: Both A's and B's benefited from trading (no A is below his participation constraint, and no B is below hers).
- *Inequality*: One of the A's barely improved over his participation constraint, whereas another A did substantially better (as shown in Figure 14.11). The same inequalities are also evident among the B's in the simulation. The



(a) A low-performing Type A trader

(b) A high-performing Type A trader

Figure 14.11: **Bargaining to a Pareto-efficient and unequal outcome.** In both panels, n is the Walrasian equilibrium outcome as would be implemented for all traders by the fictive Auctioneer. Panel a shows a simulated path for a Type A trader who does not obtain a significant increase in utility as a consequence of the trades that they make in the simulation. He traded several times, but his trades only resulted in small increases in utility, with him eventually arriving at a final allocation (L for low) on the Pareto-efficient curve. Panel b show a simulated path for a Type A trader who obtains a significant increase in his utility, resulting in a final allocation (H for high) on the Pareto-efficient curve. Each intermediate trade is shown by a hollow circle. The endowments are shown by point **z**, and the final allocations are shown by the black dots L and H.

differences in welfare are demonstrated further by the distributions of utility shown in Figure 14.12: some A's do really well and others do poorly.

- *Wealth inequality.* The two types of traders started off with identical wealth within their types, but they ended up with very different levels of wealth. The left panel of Figure 14.12 shows that a great many of the B's and a few of the A's suffered a loss in wealth as result of the market exchange.

Figure 14.11 shows that when there is trade at disequilibrium prices different traders with the same initial endowment and indifference curves can wind up in equilibrium with different amounts of good y, and hence different utility levels and differing amounts of wealth. This occurs even if the traders have identical preferences. In this case, unequal wealth and unequally shared gains from trade occur simply because some traders are luckier than others in which other traders they meet and the prices at which they transact. As Figure 14.11 demonstrates, one of the Type A traders did incredibly well because of the fortunate trades he made (right panel), whereas another trader did quite poorly as a result of the trades he made (left panel). Such inequality can emerge as the result of *decentralized* market processes even when the market implements a Pareto efficient outcome.

Simulated actors vs. rent-seeking

You can think about the matching and trading process that we simulated as if each trader is a rent-seeker looking around for a good deal, and only stops trading when there are no more deals to be found. But the Adamo and Beatriz-types were not really as aggressive as real rent-seekers. A rent-seeker tries to obtain as much rent as possible given the next best alternative of their counterpart, whereas each trader in the simulations agreed to *any* trade that was greater than their own participation constraint, they were not

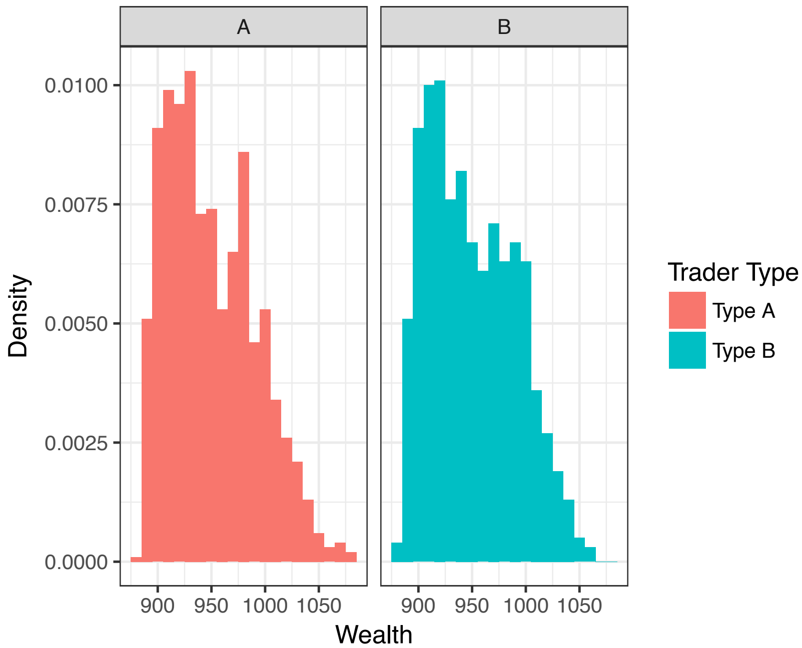


Figure 14.12: **The wealth distributions of the traders from the simulations.** The figure shows the distribution of wealth at the final allocations for all traders. As the distributions demonstrate, within the types of traders, utility and wealth are unequally distributed with some traders achieving much higher levels of utility and wealth than other, less fortunate, traders.

trying to get their counterpart to have a lower share of the rent. A more realistic model could include these and other features that would allow the actors to take up all of the rent seeking activities of the perfect competitor.

Adding such realism to our simulation would introduce even more inequality.

- *Bargaining power*: If we also introduced some differences in bargaining power, then the price at which a trader settles may be more or less favorable than some other trader; so inequality would no longer just be a matter of luck.

- *Discrimination*: Or, if some traders were the targets of racial, religious, gender, or other discrimination, then they would tend to face less favorable prices, introducing yet more inequality.

- *Non-random matching*: Finally perfect competitors in the real world would not settle for being randomly matched: those with more information would seek out and find trading counterparts holding very different quantities of the goods than they had, and with whom the range for mutually beneficial trades was therefore especially large.

Our model has affirmed that competitive exchange can implement Pareto-efficient outcomes, even without the Auctioneer. But there has been some collateral damage to the other results of the Walrasian model of perfect competition.

REMINDER You have seen that unequal treatment is a result of a competitive labor market in Chapter 11. Among a group of identical workers (with identical skills and other endowments) some workers will be employed and others not employed. In other words, some are renting their endowments for a wage w while others are not transacting at all (they are *labor market-excluded* workers).

Gauthier's claim that in a competitive market "no one is subject . . . to any type of limitation not already affecting her actions as a solitary individual" is no longer true: trading on the market generated substantial differences in wealth and utility among traders with identical initial endowments. We also have to give up Arrow's insistence on the distributional neutrality of markets, namely that markets merely preserve the status quo distribution of wealth at the initial endowment. Whether the inequalities emerging in the trading process among identical people are of significant magnitude remains an open question. The distribution-neutrality of the perfectly competitive market depends on the Law of One Price which we now see, like the Auctioneer called in to enforce the Law, is a fiction.

14.11 Bargaining to an efficient outcome: The Coase Theorem

Our model of traders with limited information, buying and selling at whatever prices they can agree upon, until there are no more mutually beneficial trades to be made, suggests that bilateral bargaining may play a key role in achieving a Pareto-efficient outcome for an economy.

Ronald Coase (1910 – 2013) observed that even if a market for some external effect does not exist, the economy can still reach a Pareto-efficient outcome through bilateral exchanges as long as traders can bargain efficiently, including over the uncompensated external effects. Our model of competitive exchange is an example of Coase's reasoning, as we have assumed that any possible exchange with the potential to benefit both traders will be implemented.

Coase versus standard approaches in economics and law

In the field called **welfare economics**, the standard approach to coordination failures of the type illustrated in Table 14.1 is that the government should impose taxes or provide subsidies designed so that private economic actors internalize the external benefits and costs of their actions on others. The tax or subsidy transforms each person's objective function and hence, their utility- or profit-maximizing first-order conditions, so that each will act as if they were taking account of the effects of their actions on others. We saw an example of how taxes do this in Chapter 5 when each fisherman was taxed by the government for each hour that they fished. The tax reduced their fishing to the Pareto-efficient level.

Compelling arguments for "green taxes" and subsidization of schooling are routinely made on these grounds, invoking reasoning originating with Alfred Marshall and A.C. Pigou (1877–1959) early in the twentieth century. In legal theory, standard approaches to activities generating external costs are to prohibit them or make those generating the external costs legally responsible

REMINDER For the fishermen, taxes change the cost of exerting effort, and forced them to incorporate the cost of their negative external effect on their counterpart by including that cost in their utility function. For firms, taxes result in higher market prices with firms and consumers responding accordingly (typically reducing the market-clearing quantity bought and sold). Taxes that take account of external effects are called Pigouvian taxes for A. C. Pigou. In Chapter 1 we showed how legal systems ("tort" or "liability" law) can accomplish similar results by requiring people to compensate for harms done to others.

WELFARE ECONOMICS is a branch of economics that studies the effect of economic policies and institutions on societal well-being ("welfare") measured, for example by total utility, a weighted sum of individual utilities, or consumer and producer surplus.

Table 14.2: **Three approaches to internalizing external costs and benefits**.

Approach	*Ways to internalize or limit external effects*	*Examples*
Welfare economics	(Pigouvian) taxes and subsidies by governments	Green taxes, public education
Legal theory	Prohibitions and liability (tort) law by courts	Prohibition of incandescent bulbs
Coase	Bargaining by private parties	Cap and trade environmental policies

for ("liable for") compensating those harmed for the external costs inflicted on them. An example of a prohibition would be requiring the replacement of incandescent light bulbs with LED lighting to reduce greenhouse gas emissions. We used a liability rule in Chapter 1 to get the fishermen to a cooperative outcome by making each liable for the damage they caused the other.

Coase challenged the need for intervention through taxes, subsidies, or prohibitions to address the problem of external costs and benefits. He demonstrated that, under specific conditions, entirely *private* transactions motivated by self-interest could accomplish the same objectives.

Table 14.2 summarizes approaches to coordination failures and demonstrates Coase's contribution.

Coase re-considered Pigou's theory and ideas. In 1920, Pigou had explained the role of external costs using the example of a railroad where the trains that used the railroad could create sparks that would cause sparks to light fires in the farmland through which the farm passed. The fires damages the crops and cost the farmers valuable profits. Pigou had asserted, conventionally, that to internalize the external cost imposed on the farmers the owners of the railroad should be liable for the damage caused by the trains. If the owners of the railroad anticipated liability, the liability would induce the owners to account for the costs of their actions on others.

Coase responded that "if the railroad could make a bargain with everyone having property adjoining the railway line and there were no costs involved in making such bargains, it would not matter whether the railway was liable for damages caused by fires or not." (31) This surprising conclusion is motivated by the observation that if the costs of the fires exceeded the cost of preventing the sparks (say, by redesigning the engines) then those harmed could simply pay the railroad a sufficiently large sum to induce them to agree to eliminate the fire damage. Conversely, if the cost to the railroad of controlling sparks is very high, the railroad could pay the neighboring farmers not to cultivate the

figures/UtopianCapitalism/cloquet_fire.jpg

Figure 14.13: In 1918, the Cloquet-Moose Lake Fire in Minnesota, USA, was started by sparks from a railroad. 453 people died and over 50,000 were injured. More than 250,000 acres of land burned and the fire caused over $73 million in damage to properties. During the expansion of railways internationally fires occurred both as a result of passing trains and as a result of building railways.

vulnerable land near the tracks.

You have already encountered Coase's proviso – *costless bargaining* – in the model of competitive exchange where we assumed that traders exchanged goods if there was a mutual benefit to be had. If executing a trade involved, for example, substantial legal expenses — that is **transaction costs** — then this would not be the case.

TRANSACTION COSTS occur when contracts are incomplete and include legal costs for litigation about inferior quality goods, the costs of monitoring and disciplining employees, the costs of screening loan applicants and the losses from unpaid loans.

COSTLESS BARGAINING Bargaining is *costless* when the parties to the bargain, such as traders in an Edgeworth box interaction, do not incur costs in executing a trade other than the price of the good exchanged. Costs of trading or transacting are often called *transaction costs*.

Bargaining is *costless* when the parties to the bargain, such as traders in an Edgeworth box interaction, do not incur costs in executing a trade other than the price of the good exchanged. Costs of trading or transacting are often called *transaction costs*. **Costless bargaining** — or as Coase sometimes put it the lack of "impediments to bargaining" — is important, and, unlike many who have invoked Coase against governmental regulation, Coase himself stressed it: ". . . if market transactions were costless all that matters (questions of equity aside) is that the rights of the various parties should be well defined and the results of legal actions easy to forecast. But . . . the situation is quite different when market transactions are so costly as to make it difficult to change the arrangement of rights established by the law."

What came to be called the *Coase Theorem* achieves a seemingly dramatic extension of the Fundamental Theorems of Welfare economics. Even where contracts are incomplete and, as a result, uncompensated external effects like those illustrated in Table 14.1 occur, efficient allocations can result from bargains. Efficient allocations occur if the people affected can bargain efficiently over the rights governing the actions that result in the external costs or benefits.

Because there is some controversy about what the "theorem" means (as Coase himself said, there is no explicit "Coase Theorem"), it may be useful to consult its author. In his Nobel lecture, Coase (1992: 717) wrote:

> What I showed . . . was that in a regime of zero transactions costs, an assumption of standard economic theory, negotiations between the parties would lead to those arrangements being made which would maximize wealth, and this irrespective of the initial assignment of rights.

14.12 An example: How Coasean bargaining works

To understand the conditions under which Coase's argument would be true, think about a situation similar to one of the examples that Coase used when he first introduced the idea.

Let's consider two people: Anders and Bianca. Bianca operates a small metalworking shop next to Anders's yoga studio and when Bianca cuts metal the loud noise disturbs the people practicing yoga next door. The loud noise is a negative external effect (a cost) that Bianca imposes on Anders. To explain what occurs we shall describe a set of payoffs for Anders and Bianca depending on the actions they are able to take and the outcomes that occur. This will allow us to construct Game Trees to explain what the Nash equilibrium outcomes of the interaction will be. We consider two cases:

- *Bianca has the initial rights*: Bianca can *use* her loud machinery – imposing a cost on Anders – as much as she likes; he can bargain with her to get her to restrict her usage.
- *Anders has the initial rights*: Anders has the right to *restrict* Bianca's usage of the machinery; she can bargain with him to allow her to use it.

Bargaining when Bianca has initial rights

Lets say that if there is no bargaining, and if Bianca uses her metal-cutting saw during the hours of operation of Anders yoga studio, then Bianca receives a payoff of 5 and Anders 1. If she is restricted to not use the saw when the yoga studio is open, then both get payoffs of 4.

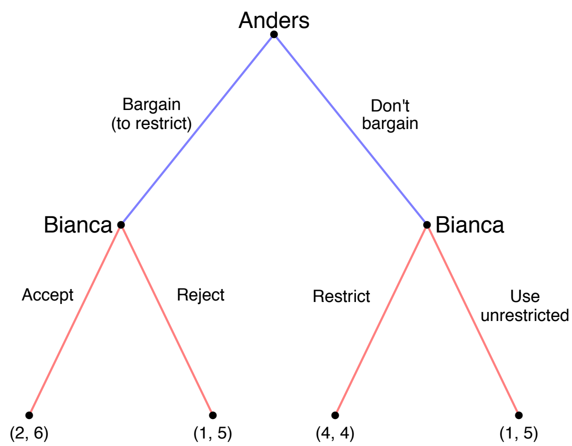
Problems like this are sometimes addressed by what are called zoning laws, which for example *exclude* "nuisance activities" such as garbage incinerators or pig farms from residential neighborhoods or professional office locations. Coase's idea is that private bargaining rather than government policies could get Bianca to *internalize* the external cost she imposes on Anders, giving her the private incentive to operate her shop in a way that does not disturb Anders's yoga sessions.

The initial assignment of rights in this case is whether or not Bianca has the right to generate as much noise as she wants, whenever she wants. The possible bargains between her and Anders include a payment from one to the other along with the transfer of the initial rights from one to the other. For example, if Bianca initially had the right to operate her machinery in an unrestricted way, she might be willing give up that right – giving Anders the right to restrict her use of the saw – if Anders paid her enough.

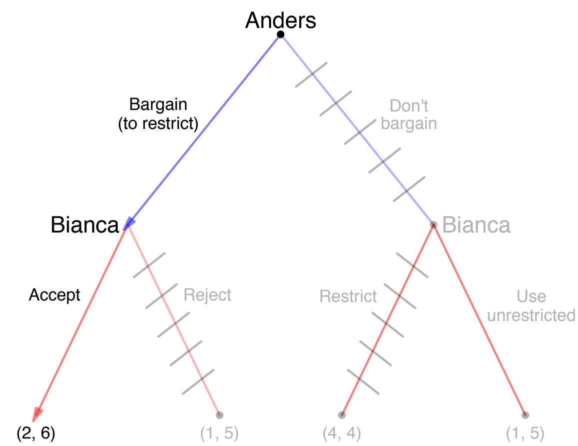
Remember, like any other exchange, for a bargain to be implemented it must result in both parties being better off (or at least one of them being better off and the other not worse off). This is because exchange is *voluntary*. Both parties must have an incentive to accept the deal. So the result of the bargain must be a Pareto improvement over the result that would occur without the bargain.

The game tree in Figure 14.14 illustrates the case in which Bianca has the right to unrestricted use of her metal saw. At the top of the tree Anders can choose either:

- *Don't bargain*: such that he does *not* bargain with Bianca to get her to restrict her usage of the machinery, or



(a) Bargain to restrict



(b) Solved game tree

Figure 14.14: **The case of unrestricted bargaining between Anders and Bianca.** The payoffs at the terminal nodes are listed in the order (Anders, Bianca). Anders is the first-mover with his decision branches in blue. Bianca is the second-mover with her decision branches in red. Anders can Bargain (to restrict) or not. If Anders bargains, Bianca can accept or reject the bargain. If Anders does not bargain, Bianca can Restrict voluntarily or choose to Use unrestricted. Panel a. shows the game without the solution, panel b. shows the solution to the game with the pruned branches faded out and the Nash equilibrium shown as (Bargain (to restrict), Accept).

- *Bargain (to restrict)* and offer her 2 to restrict her use of the machinery.

If he offers this bargain, then, moving down the game tree, Bianca has the choice between:

- *Reject* (meaning reject Anders's offer and exercise her right to unrestricted use of her machinery), or
- *Accept* the bargain and Restrict.

On the right branch of the tree, Anders does not bargain, and Bianca then chooses between:

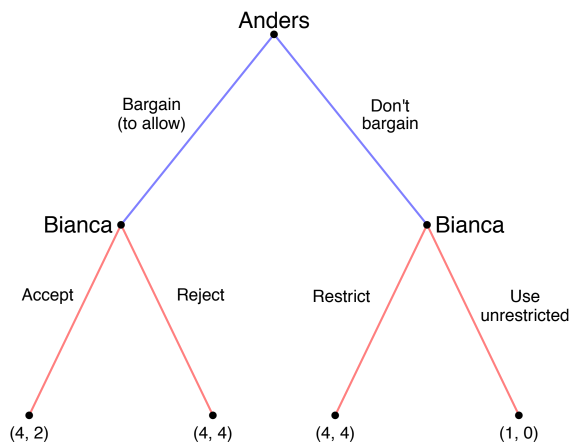
- *Restrict*: to limit her usage of the metal-cutting saw, or
- *Use unrestricted*: meaning use her metal-cutting saw whenever she pleases.

The numbers at the end of the branches are the payoffs to the two people: Anders's first, Bianca's second. You already know that if Anders does not bargain and Bianca chooses Leave unrestricted then Anders gets 1 and Bianca gets 5, while if Bianca restricts they both get 4.

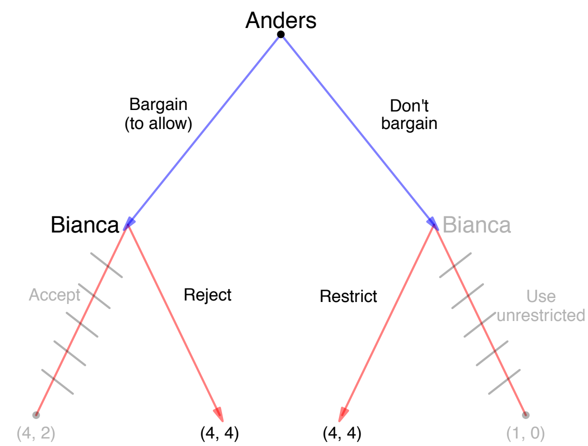
On the left branch of the tree, that is if Anders chooses Bargain (to restrict) offering Bianca 2 to *not* use her saw the outcomes are as follows:

- If Bianca rejects the offer and retains unrestricted rights to use her machine, then she gets 5 and Anders gets 1
- But if she accepts his offer, then before being paid by Anders Bianca gets 4 and then with the payment from Anders she gets $4+2=6$; as a result, Anders gets $4-2=2$.

How would this game be played, if both Anders and Bianca cared only about



(a) Bargain to allow



(b) Solved game tree

Figure 14.15: To be completed.

their own payoffs. To find the outcome of the game, look for the Nash equilibrium, that is an outcome resulting from Bianca's choice being a best response to Anders's choice and Anders's choice also being a best response to what he can anticipate will be Bianca's choice.

Anders is the first mover. To decide what to do anticipates what Bianca will do in response. He goes directly to the payoffs. The best he could do is to get 4, that is, if he did not bargain and Bianca choose Restrict. But, then, looking at Bianca's payoffs, he can see that if he does not bargain, she will choose Leave unrestricted, because she will then get 5 rather than 4.

So Anders now considers the bargaining option. The best he can do is 2, that is if Bianca accepts his bargain. To figure out if she will, he needs to consider how Bianca will make her choice. Bianca would reason that she can get 5 if she chooses Unrestricted. But she could get 6 if Anders offers her 2 to secure an agreement to Restrict (that is, her payoff of 4 if she restricts plus the 2 that Anders offers in this case).

Anders will be better off bargaining with Bianca, giving her 2, to Restrict and ending up with the 4 he gets if Bianca chooses Restrict, minus the 2 he pays to Bianca, or 2 for himself, which is better than the 1 he would get if Bianca chose not to restrict.

So the outcome of the game will be (Bargain to restrict, Accept) with the total payoffs of the two being $6 + 2 = 8$, which is larger than the total payoffs of the two , or $5 + 1 = 6$, without the bargain. Coase's point, "negotiations between the parties would lead to those arrangements being made which would maximize wealth" applied here means that the noise problem would be solved by the bargaining process providing Bianca's incentives for adopting Restrict even if she has the right to choose “Unrestricted.”

Bargaining when Anders has initial rights

Now consider the alternative case in which Bianca initially does *not* have the unrestricted right to to use her metal saw. Could she bargain with Anders to get that right, just as Anders had bargained with Bianca to get her to adopt Restrict?

Figure 14.15 shows that the answer is "no." If Bianca initially did *not* have the right, we assume that if she plays Unrestricted anyway, she has a payoff of zero (due to incurring legal costs or fines imposed by the courts).

Alternatively, Anders could propose to transfer the right to play Unrestricted to Bianca in return for a payment from Bianca to Anders (she would not incur legal costs or fines in this case!). But how big a payment would she have to offer him?

Consider Anders's fallback option: he can get a payoff of 4 if he lets the initial assignment of rights do its work and Bianca has to play Restrict. That is, he must be at least as well off as when he did not bargain. He therefore needs a payoff of at least 4 when he transfers the right to her.

We saw in Figure 14.14 that if Bianca Rejects when Anders offered a Bargain, Anders would get a payoff of 1 (he has fewer clients when there's a lot of noise next to his yoga studio). The difference between 1 and 4 tells us how much Bianca would have to pay him: $4 - 1 = 3$! An amount of 3 is the *least* amount Bianca could pay.

But if Bianca paid Anders 3 to be allowed to choose Leave unrestricted, she would end up with $5 - 3 = 2$. A payoff of 2 is less than if she had voluntarily chosen to restrict, where she gets 4 and Anders gets 4. So if Anders offers this bargain (knowing the least amount he would accept), Bianca would choose Reject, and play Restrict, with payoffs (4, 4).

The result is the same outcome as would occur if Anders had chosen *Not* to bargain: Bianca would play Restrict, because she does not have the right to play Leave unrestricted.

Checkpoint 14.3: Explaining the payoffs

Explain:

a. How the payoff of 2,6 in Figure 14.14 comes about.
b. What is the smallest payment that Anders could have offered to Bianca that would have led her to agree to the bargain and choose Restrict?

Coasean bargaining: Why it works and why it might not

Looking at the two cases in which Bianca had or did not have unrestricted rights, the outcome – Restrict – occurs independently of who had the initial rights. Notice that the outcome when she *did* have the rights, with payoffs (2,6) is Pareto superior to the outcome in that same case without bargaining with payoffs (1,5). But the payoffs when Bianca has the right to choose Unrestricted (2,6) are not Pareto superior to the outcome in which he did not have that right (4,4).

The Pareto improvement shows the Coase Theorem in action. As long as the two parties can bargain costlessly:

- *Rights do not affect whether a bargain occurs*: who has the rights does not affect whether the external costs get internalized and the inefficiency addressed.
- *Rights do affect distribution*: who has the rights will affect the distribution of payoffs; Bianca does better and Anders does worse when Bianca has the right to the unrestricted use of the machine.

As Coase said, the reason why bargaining will implement the Restrict outcome is that the total payoffs under restrict (8) exceed the payoffs when Bianca plays Unrestricted (6). The difference between the two is the sum of two effects:

- The *external cost* to Anders of unrestricted use of Bianca's machine, that is, the difference in Anders's payoffs under Restrict and Unrestricted $(4-1=3)$ and
- The *opportunity cost* of Bianca restricting his use of the machine, that is, the difference between Bianca's payoffs under Unrestricted and Restrict $(5-4=1)$.

Restricting Bianca's use of the machine – eliminating the external cost in the first bullet above – will benefit Anders by 3 but cost Bianca just 1. Whoever has the initial rights, there will be some bargain that will result in the use being restricted.

HISTORY In a paper published in 1960 (one of the two that won him the Nobel prize in economics three decades later) Coase wrote that the idea on which his theorem is based, namely, that there were no costs of bargaining, was "a very unrealistic assumption. To carry out a market transaction people need to discover who it is that one wishes to deal with, to inform people that one wishes to deal and on what terms, to conduct negotiations leading up to a bargain, to draw up the contract, to undertake the inspection needed to make sure that the terms of the contract are being observed, and so on.

But in stating his "theorem" Coase was careful to assume that "there are no costs involved in making such bargains." Letting the example be a little bit more realistic makes it clear why.

- Bargaining is costly.
- The number of people bearing the external costs affects the cost of a bargain.
- External costs are often borne by many people, the benefits of the action are often reaped by only one person, so those adversely affected often face a coordination problem.

Bargaining is costly

Imagine that Anders and Bianca are not neighbors able to bargain informally with one another at virtually zero cost but instead have to hire teams of lawyers to bargain on their behalf. Legal fees make could be 3 times the amount actually transferred between the two in the bargain.

Return to the left panel of the figure and ask if bargaining will succeed in implementing the outcome (Bargain, Restrict). Remember this occurred because Anders paid Bianca 2 to secure her agreement to Restrict. Studying the payoffs, you will see that he could have offered Bianca just a little more than 1 (so that Bianca would bet a bit more than 5 if she chose Restrict). But even paying this lesser amount to Bianca, along with the legal fees would cost Anders a total of 4, leaving her with a payoff of zero, which is worse than what he would get if Bianca chose Unrestricted. So there would be no agreement that Anders could have offered that would benefit both and so be both proposed and accepted.

People adversely affected are many and diverse
A second step in the direction of realism provides another reason why Coase's zero bargaining costs assumption is not generally applicable. Pigou's example of the railroad company and the farmers along the train's route is a good one because at least on one side of the interaction there are a large number of people (in this case the farmers).

Before bargaining with the railroad company, the farmers would have to agree with one another, perhaps bargaining about which farmers should pay more or less of the costs of securing the railroad's agreement to redesign their engines, if that is the solution. Even if the external costs imposed on the farmers exceeded the opportunity cost to the railroad of redesigning the engines to avoid the fires, the costs in legal fees and the farmers' own time could be large enough so that no bargain could be struck, and the railroad continued causing the farmers' crops to burn.

People adversely affected face a coordination problem
Related to the previous problem, as the number of people bearing an external cost increases, the likelihood that they can coordinate to bargain with the party imposing the costs decreases. This problem is sometimes called the problem of concentrated benefits and diffuse costs. The person imposing the cost on others – Bianca in this case – received concentrated benefits from being able to run her machine (the profit she makes from doing so). But, if instead of affecting Anders alone, Bianca's actions affected Anders, Caroline, Deepal, Erkan, Friederike, and many others each of whom bore some costs, but not that great a cost individually, then the costs are *diffused* among many people.

If one or some of them were able to reach a bargain with Bianca to restrict her activities, the result of the bargain itself would be a public good (non-rival and non-excludable, as discussed in Chapters 1 and 5). Each person would therefore have an incentive to free ride on the efforts of others to bargain with Bianca: obtaining the benefits without paying the costs. This means bargaining itself is a coordination problem. We explore this dynamic further in Chapter 16 and explain why it is hard to design policies to overcome such problems of coordinating collective action among many people who bear costs they'd rather others incur.

Checkpoint 14.4: Bargaining to *allow* external effects

Think about a case different from that shown in Figure 14.15 in which the payoff to Bianca from Unrestricted is much larger, 10, which means that the opportunity cost of restricting (that is 6, or 10 minus the 4 she gets if she chooses Restrict) is much greater. The payoff to Anders if Bianca is unrestricted is, as before, 1. Using the reasoning above but with these different payoffs:

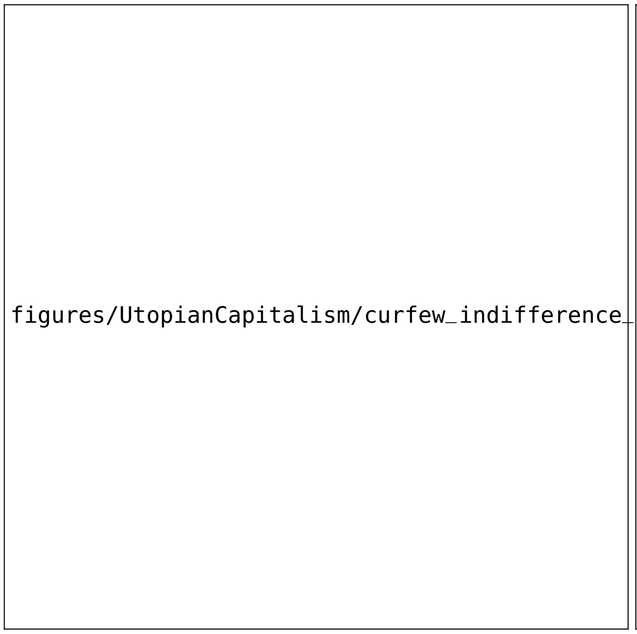
a. show that if the two can bargain, then Bianca will make unrestricted use of her machinery whether she initially has the right to do so or not;
b. if Bianca did not have the unrestricted right to use the machine and if Anders could make a take it or leave it (TIOLI) offer to Bianca to grant her permission to unrestricted use of the machine, what bargain would he offer
c. if Bianca did have the unrestricted right to use her machine explain why Anders cannot bargain with Bianca to secure her agreement to Restrict.

14.13 Application: Bargaining over a curfew

Having used game trees to understand how Coasean Bargaining might work, we can use another set of familiar tools to understand how Coasean bargaining might work: utilities, indifference curves, and Edgeworth boxes. Using these tools allows us to explore in greater depth how Coasean bargaining can achieve Pareto-efficient outcomes, that is, outcome on the Pareto-efficient curve in an Edgeworth box, or outcomes at which players achieve outcomes from which neither of them would be interested in unilaterally deviating. As in the previous case, though, even if the bargaining parties achieve an efficient outcome, the results may be unequal with rents being distributed unequally among them depending on the assignment of initial rights.

The neighbors' utilities and indifference curves

Anna (A) and Bertolt (B) are two neighbors. Bertolt is a night-owl who plays loud music late into the night, while Anna worships the rising sun, and hence wants to go to sleep early. A curfew is proposed specifying the time of night, T, after which no music is to be played. If A could determine the curfew she would set $T = T^A = 9pm$, while B would select $T = T^B = 3am$. $T^B > T^A$: Bertolt prefers to get to bed at 3am and Anna at 9pm. For either of them, going to bed earlier or later than their preferred time results in disutility.



(a) A's indifference curves

(b) B's indifference curves

Figure 14.16: **Indifference curves for Anna and Bertolt over payments and the curfew time.** For Anna, a later curfew time (T) is a bad and payments from Bertolt (y) are a good. So her indifference curves slope upward. For Bertolt, payments from Anna ($-y$) are a good and a later curfew is a good (T), so his indifference curves slope downward. (note: payments *to* Anna would be a *bad* for Bertolt).

To see these preferences, let the utility functions of Anna and Bertolt be:

$$\begin{aligned} \text{Anna's utility: } & u^A(T,y) = y - \alpha(T^A - T)^2 \\ \text{Bertolt's utility: } & u^B(T,y) = -y - \beta(T^B - T)^2 \end{aligned} \tag{14.9}$$

α and β are positive constants indicating the importance of the curfew time relative to income in the well-being of each (the relative intensity of their preferences). For simplicity, let $\alpha + \beta = 1$. It is important for what follows that the two utility functions are quasi-linear, so that transfers of y are equivalent to transfers of utility, and that they exhibit a constant marginal utility of income. Anna and Bertolt's utilities should be interpreted in such a way that the y and $-y$ are transfers between them or differences in their incomes y^B and y^A (as shown in the indifference curves in Figures 14.16 a and b). Therefore if y^A is positive then y^B is negative and vice versa.

The neighbors' indifference shown in Figures 14.16 a and b demonstrate the following. For Anna:

- She prefers an earlier curfew time, so later curfew times are a *bad*.
- Anna's utility given by Equation 14.9 includes the term y so she likes to receive payments from Bertolt, so payments received from him are a good
- Alternatively, paying Bertolt would reduce her utility

- Her indifference curves therefore slope upward from 9pm to 3am (before 9pm they slope downard; she would prefer a later curfew than 8pm).

For Bertolt:

- He likes a later curfew time (T) so later curfew times (up to 3am) are a good.
- Bertolt's utility given by Equation 14.9 includes the term $-y$.
- That is, if he *receives* payments from Anna that would *increase* his utility ($y^B > 0, y^A < 0$).
- Alternatively, *paying* Anna, meaning a larger *positive* rather than negative y would *reduce* his utility ($y^B < 0, y^A > 0$).
- His indifference curves therefore slope *downward* up to 3am (after which they slope upward; he prefers an earlier curfew than 4am).

As Figure 14.16 shows, for Anna, her indifference curves slope downward until they reach her preferred curfew time at 9pm. After her preferred curfew time, her indifference curves slope upward. For Bertolt, his indifference curves slope downward until they reach his preferred curfew time at 3am. After his preferred curfew time, his indifference curves slope upward as a later curfew is a bad.

Given the conflict of interest between the neighbors about the curfew time the best an impartial spectator, such as the mayor of the town in which Anna and Bertolt live, can do is to minimize the level of disutility each incurs, as we now show.

A mayor proposes a curfew between two conflicting neighbors

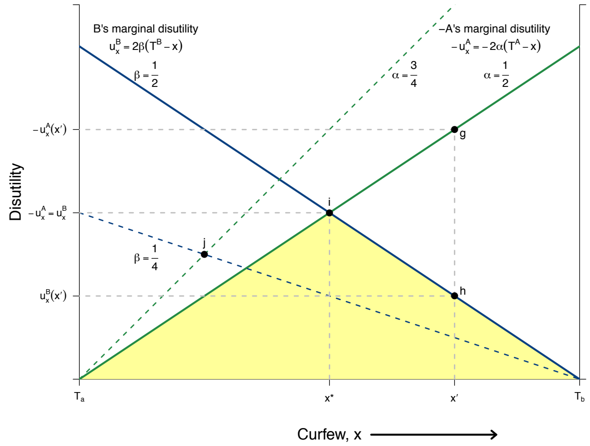
Suppose you are the mayor of the town and, knowing the neighbors' utility functions, you wish to set T to maximize total social utility, $W = u^A + u^B$ (or, equivalently, to minimize disutility). The mayor will choose the a curfew, T^i, as follows:

$$T^i = \alpha T^A + \beta T^B \tag{14.10}$$

We show why this is the socially optimal curfew in M-Note 14.2.

The curfew, T^i, that the impartial spectator would choose is a weighted sum of the two preferred curfew times. We'll call this the *socially optimal outcome*, and relate it later to the whole class of Pareto-efficient outcomes. If $\alpha = \beta$, the socially optimal curfew is half-way between the two neighbors' preferred times because:

- Each neighbor experiences rising marginal disutility as the curfew time diverges from his or her preferred time.



(a) Curfew when neighbors have the same intensity of preferences

(b) Curfew when Anna feels more intensely about her curfew than Bertolt does

- The sum of the disutility is minimized by balancing the players' marginal disutilities.

Figure 14.17 illustrates this.

The horizontal axis is the time of the curfew, ranging from early (T^A) to late (T^B). In both panels, the area under the two marginal disutility curves is the sum of disutilities. The sum of the disutilities is *minimized* by a curfew set at T^i, the Pareto-efficient curfew.

For $\alpha = \frac{1}{2}$ and $\beta = \frac{1}{2}$, a later curfew, e.g. $T' = 1:30am$ would result in *higher* disutility for A and lower disutility for B, but it would not be a social optimum.

Total welfare is optimized by setting $T = T^i$. If $T = T' > T^i$ the marginal disutility to A ($-u^A(T')$) of a later curfew (closer to T^B than the curfew at T^i) exceeds the marginal disutility to B ($u^B(T')$) of the time T'. It would *lower* total disutility if the curfew time were made earlier and set closer to T^i.

In panel b. two different disutility curves are shown with a higher α for A and a

Figure 14.17: **The socially efficient curfew.** The horizontal axis is the time of the curfew, ranging from early (T^A) to late (T^B). In both panels, the area under the two marginal disutility curves is the sum of disutilities. The sum of the disutilities is *minimized* by a curfew set at T^i, the Pareto-efficient curfew in panel a at point **i**. In panel b. two disutility curves are shown with a higher α for A and a lower β for B, with the intersection resulting in correspondingly earlier curfew, illustrated by point **j**.

lower β for B, with the intersection resulting in correspondingly earlier curfew, illustrated by point **j** at 10:30pm. That is, if A feels much more strongly about her 9pm curfew than Bertolt feels about his 3am curfew, then the mayor – as the impartial spectator – will choose a curfew closer to Anna's preferred curfew to minimize disutility.

M-Note 14.2: The socially optimal curfew

The mayor wishes to impose the socially optimal curfew, taking account of both Anna's and Bertolt's preferences in her social welfare function, W.

$$\begin{aligned} W &= u^A + u^B \\ \Rightarrow W &= y - \alpha(T^A - T)^2 - y - \beta(T^B - T)^2 \\ \Rightarrow W &= -\alpha(T^A - T)^2 - \beta(T^B - T)^2 \end{aligned} \tag{14.11}$$

To maximize total social welfare, the mayor needs to differentiate Equation 14.11 with respect to T and impose first order conditions (set the derivative equal to zero).

$$\begin{aligned} \frac{dW}{dT} &= 2\alpha(T^A - T) + 2\beta(T^B - T) = 0 \\ \Rightarrow \underbrace{2\alpha(T^A - T)}_{u_T^A} &= -\underbrace{2\beta(T^B - T)}_{u_T^B} \end{aligned} \tag{14.12}$$

Therefore, the mayor finds the curfew at which A's marginal utility of T (u_T^A) equal to the negative of B's marginal utility of T (u_T^B). Or, equivalently, the negative of A's marginal utility to the positive of B's marginal utility:

$$\Rightarrow -\underbrace{2\alpha(T^A - T)}_{u_T^A} = \underbrace{2\beta(T^B - T)}_{u_T^B} \tag{14.13}$$

We showed the equality given by Equation 14.13 in Figure 14.17 with two sets of values for α and β.

What level of T does this result in? To find this, we have to isolate T in Equation 14.12

$$\begin{aligned} 2\alpha(T^A - T) + 2\beta(T^B - T) &= 0 \\ 2\alpha T^A - 2\alpha x + 2\beta T^B - 2\beta x &= 0 \end{aligned} \tag{14.14}$$

Divide by 2 and isolate T terms

$$\begin{aligned} \alpha T + \beta T &= \alpha T^A + \beta T^B \\ T(\alpha + \beta) &= \alpha T^A + \beta T^B \\ T^i &= \frac{\alpha T^A + \beta T^B}{\alpha + \beta} \end{aligned} \tag{14.15}$$

At this stage, we have two potential assumptions about the parameters α and β that could apply. If $\alpha = \beta$ then, Equation 14.15 simplifies to:

$$\begin{aligned} T^i &= \frac{\alpha T^A + \alpha T^B}{\alpha + \alpha} \\ T^i &= \frac{\alpha(T^A + T^B)}{2\alpha} \\ \alpha \text{ cancels} \qquad \therefore T^i &= \frac{(T^A + T^B)}{2} \end{aligned} \tag{14.16}$$

In which case, the mayor picks the midpoint of the two neighbors' curfews (as in Figure 14.17). On the other hand, if $\alpha+\beta=1$, then Equation 14.15 simplifies to:

$$\begin{aligned} T^i &= \frac{\alpha T^A+\beta T^B}{1} \\ \therefore T^i &= \alpha T^A+\beta T^B \end{aligned} \tag{14.17}$$

Equation 14.17 says that the mayor will choose a weighted average of the two neighbors' curfew times depending on the *intensity* of each neighbor's preference (given by their utility functions).

Checkpoint 14.5: The logic of the curfew

Be sure you can apply the same logic to a curfew set earlier than T^i.

a. What would B's disutility be?
b. What would A's disutility be?
c. What could improve total welfare?

Private bargaining

The Coase Theorem says that it doesn't matter for Pareto efficiency which of the two determines the curfew or even if some third party determines it as long as the two can efficiently bargain to rearrange the relevant property rights, meaning in this case, the curfew itself. Bargaining is efficient if the outcome is on the utility possibilities frontier (and hence is Pareto-efficient). Suppose the bargaining takes the form of either neighbor paying the other an amount of money. For example B pays y to A in return for A agreeing to a later curfew than whatever is initially announced (with $y>0$ being a payment from B to A, $y<0$ would be a payment from A to B for an earlier curfew).

REMINDER In Chapter 4 we used the utility possibilities frontier when discussing the choices an impartial spectator might make about which Pareto-efficient outcome to choose depending on their social welfare function. Those outcomes that were Pareto-efficient lay on the utility possibilities frontier (feasible utility frontier).

Would private bargaining achieve the same result as the curfew chosen by the impartial spectator?

Figure 14.18 provides the basic infrastructure for thinking about the bargain between the two. The solid line at $y^A=y^B=0$ is the case when they do not bargain. Points **a** and **b** correspond to each neighbor's preferred curfew time: Anna's at 9pm (with indifference curves u_a^A for her and u_a^B for Bertolt) and Bertolt's at 3am (with indifference curves u_b^B for him and u_b^A for Anna). The impartial spectator's choices is shown by point **i** on this same line where neither can transfer money to the other and each has a utility $u_4^A=u_4^B=4.5$.

Consider what would appear to be the worst case, no outside-imposed curfew at all, which means that in the absence of any bargaining between the two, B will impose loud music on A until T^B o'clock every night (3am). That is, Bertolt has the initial rights just like Bianca did in the previous example to run her machine whenever she wanted to unless Anders struck a bargain with her.

To see if Bertolt and Anna might strike a bargain, consider the interaction between the two as illustrated in Figure 14.19. As in the previous figure, the time of the curfew (T) is on the horizontal axis and the difference in incomes for A and B are measured vertically. If there is no rule about a curfew, then this means that B simply plays the music as late as he likes (3am), shown by point **b**. The curves u^A_b and u^B_b are combinations of curfew times and payments which, for each, are as good as B's preferred curfew time with no payments. That is, they are the participation constraints going through point **b** where B has initial rights. Both preferred and inferior combinations are indicated by the other indifference curves.

REMINDER Because the marginal utility of y is $u^A_y = 1$ and $u^B_y = -1$, their marginal rates of substitution correspond to the ratios $\frac{u_T}{u_y}$ with which you are familiar. Remember that because $mrs(T,y)$ is defined as the negative of the slope of the indifference curve, and A's indifference curves have a positive slope, $mrs^A(T,y)$ is negative.

The Pareto efficient curve is at a bargain where the curfew is at midnight and and where the neighbors' indifference curves are tangent. That is, where:

$$\begin{aligned} mrs^A(T,y) &= mrs^B(T,y) \\ -2\alpha(T^A - T) &= 2\beta(T^B - T) \end{aligned} \tag{14.18}$$

As in Figure 14.16, because the marginal utility of income is constant for both people quasi-linear utility functions, the indifference curves are vertical displacements of one another (notice that y does not appear in the expression for the marginal rates of substitution of the indifference curves). So, other tangencies can be found along a vertical line through midnight, giving the Pareto-efficient curve. Pareto-efficient outcomes will set the curfew at midnight, but will differ in the payments between the neighbors.

If the status quo curfew is T^B, B gets utility 0 while A gets -18. Both would prefer any point in the Pareto-improving lens formed by the indifference curves corresponding to their participation constraints.

We do not know what bargain the two traders will agree to. We know that the bargain will depend on the institutions and norms governing the bargaining process. We assume that any outcome must be *voluntary* and hence cannot be worse for either party than the curfew of T^B with zero payments between the two. Referring to Figure 14.19:

REMINDER A negative y means A transfers to B, whereas a positive y would mean B transfers to A.

- if B can make a take it-or-leave-it offer to A the outcome will be $u^A_{t^B} = -18, u^B_{t^B} = 9$
- A pays B the amount $y = -15$ (A has a negative income difference $y^A = -15$ and B a positive with $y^B = 15$)

- The curfew is set at $T^i = 4$ (midnight), with A gaining a utility greater than -18 by an arbitrarily small amount.

REMINDER The Pareto-improving lens must exist because at $T = T^B$, $du^B/dT = 0$ (T^B is *B*'s preferred curfew time) while $du^A/dT < 0$ so there will exist some $dT < 0$ and some payment from *A* to *B* which will make both better off. This lens in (T, y) space gives us the bargaining set you've seen in Chapter 5, which corresponds to the idea of the utility possibilities frontier limited by the participation constraints.

Efficient Coasean bargaining and the bargaining set

What we do know – here is the Coasean condition for an agreement – is that if the *institutions* and *norms* governing the bargaining process allow Pareto-efficient bargains, then the outcome will be Pareto-efficient. That is, the allocation of goods and income resulting from traders bargaining with each other will be an allocation along the utility possibilities frontier, or equivalently, along the Pareto-efficient curve within the Pareto-improving lens.

Figure 14.20 shows the utility possibilities frontier and the bargaining set for the allocations of time and money, doing so in terms of utility. Because Figure 14.20 is in term of utility it therefore depicts the surplus that the players can obtain as a consequence of trading: the gains from exchange. The Pareto-improving lens in Figure 14.19 in terms of allocations of the curfew and money corresponds to the utilities that the players can obtain in Figure 14.20.

To see why this has to be true, imagine that the two had settled on some curfew other than T^i, say T', in Figure 14.17. A would be willing to pay B to further reduce T, the maximum payment offered by A ($u^A(T')$) exceeding the minimum acceptable to B ($u^B(T')$). The outcomes consistent with efficient Coasean bargaining differ from the standpoint of distribution, but all are efficient.

So Coase is right: who holds the property rights does not matter ("questions of equity aside"). In this case, the Coasean formulation that "the outcome will be same regardless of the assignment of property rights" is exactly correct as to the curfew time, which is the same at all Pareto-efficient allocations. If we take "questions of equity" into account, then the outcome is emphatically *not* the same regardless of the assignment of property rights because the distribution of income depends crucially on the bargaining power of the neighbors.

Constraints on Coasean bargaining: wealth, credit, and power

Often, though, people cannot borrow money easily. Recall how we spoke about people being *credit-constrained* or *credit-market excluded* in Chapter 12. Sometimes people are unable to borrow *enough* money to finance projects they want to finance, or they are unable to borrow money altogether. So it may be the case that A is not wealthy and does not have (and cannot borrow) the funds necessary to compensate B in the curfew example. Assume, for concreteness that A cannot pay more than \$4.50. If A cannot pay more than \$4.50, A's constraint limits the Pareto-improving lens in Figure 14.19 to the lens above A's indifference curve $u_2^A = -6.75$ and below B's indifference curve $u^B = 0$. Because A is *credit-constrained* or *poor* (has low wealth) these additional constraints limit the extent to which A can engage in mutually beneficial exchange with other traders. So A can voluntarily exchange with other people, but A's share of the rents from exchange are decreased by her low wealth and inability to access credit.

Checkpoint 14.6: Switching the initial rights

If instead of Bertolt having initial rights as in Figure 14.17, assume instead that Anna had initial rights to impose a curfew at 9pm and that Bertolt could offer her a bargain to extend the curfew to midnight.

a. Show on Figure 14.18 what the Pareto-improving lens would be.
b. Draw indifference curves that would lead them to choose a Pareto-efficient outcome.
c. Show the take-it-or-leave-it (TIOLI) offers each would make to the other were a bargain to be made.
d. Indicate an arbitrary, voluntary bargain when neither of them has TIOLI power.

14.14 Bargaining, markets, and public policy

Coase's contribution proved controversial, in large measure because it was used to advocate a limited role for government in addressing market failures. Buchanan and Tullock wrote:

> If the costs of organizing decisions should be zero, all externalities would be eliminated by voluntary private behavior regardless of the initial structure of property rights. There would, in this case, be no rational basis for state or collective action beyond the initial minimum delineation of the power of individual disposition over resources.

Among the more surprising claims said to be based on Coase's reasoning is that the assignment of property rights is efficient in actual economies, and that transitions from one economic system to another could be seen as the outcome of efficiency-enhancing Coasean bargaining.

However, when the Coase theorem is presented sufficiently precisely to be correct, all it says is that *if* there are no impediments for traders to bargain efficiently, *then* outcomes will be Pareto-efficient. This seems disappointingly similar to the Fundamental Welfare Theorems. As Henry Farrell pointed out, the information conditions under which the Coase theorem holds – no impediments to efficient bargaining – are exactly those that would allow complete contracting. This has led to a concern that:

- Where the Coase theorem works, the Welfare Theorems also hold, and the Coase theorem is unnecessary.

- Where the Welfare Theorems fail (due to contractual incompleteness), the zero bargaining costs assumed by the Coase Theorem will also not hold.

Some have concluded on this basis that when the Coase theorem is needed it fails, and is therefore of little relevance. But this interpretation misunderstands the significance of Coase's analysis. What he pointed out is that trading on competitive markets is *not* the only way to get from an inefficient initial endowment to a point on the Pareto-efficient curve or at least closer to it (in the Pareto sense): people can bargain bilaterally *without* an Auctioneer and still obtain Pareto-efficient outcomes.

We therefore need not understand the Coase theorem as a case *against* the Pigouvian tax-and-subsidy welfare economics tradition. Intead, where neither markets nor governments succeed, we can understand The Coase Theorem as a specification of the conditions under which private rearrangements of property rights may overcome – or lessen the effects of – coordination failures.

By indicating what is required – efficient bargaining – the Coase theorem makes clear just how improbable it is in many situations that private decentralized allocations will be Pareto-efficient. In this respect, it may resemble the Fundamental Welfare Theorems: it neither advocates nor opposes decentralized solutions, rather it clarifies what is required for the results to be Pareto-efficient.

The Coase Theorem also underlines the value of distinguishing between efficiency arguments and fairness arguments concerning policies for coping with coordination failures. The Pigouvian position, – for example, that polluters should pay for the harm they do – is often voiced by environmentalists. But many would object to this principle on grounds of fairness if the polluters in question were not carbon-based energy companies but instead people who live in cities in poor nations who cook and heat by means of fires because they lack the income to purchase more environmentally-friendly stoves. It is not clear, however, how a Coase-inspired solution to this problem – those who care about air quality bargaining with those cooking over wood fires – could work, given the extraordinary transaction costs of arranging a private bargain among all of those involved.

FACT CHECK The World Health Organization reported in 2018 that about 3 billion people worldwide cook using polluting open fires or stoves with kerosene, dung, wood, and other biomass. The external cost? About 4 million people a year die of respiratory diseases related to air pollution caused by air pollutants from such cooking and heating. Finding a way to shift these households to electric cooking and heating – perhaps even solar-powered – would result in significant improvements to health and life expectancy.

Coase's contribution was to point out the following. Let people start from an arbitrary endowment and allow them to bargaining. They will use local information and attain a Pareto-efficient outcome under conditions less stringent than those required by the Fundamental Welfare Theorems. Coase did not need the assistance of an imaginary and all-powerful Auctioneer.

But the idea that distribution of property rights does not matter for efficiency is mistaken. Barriers to efficient bargaining are common, credit constraints limit the resources people may deploy in Coasean bargaining, and the distribution of wealth influences both the barriers and the credit constraints.

Checkpoint 14.7: Pigou and external effects caused by people living in poverty

Answer the following questions:

1. Explain why the example above – asking that a person living in poverty who causes pollution should bear the costs of that pollution – is consistent with the Pigouvian position outlined earlier in the chapter.
2. What would a Coasean bargain potentially look like in this case between another party and the poor person who cooks with their polluting stove?

14.15 Application: Planning vs the market in the history of economics

HISTORY The Soviet economy in practice was a much more complicated social phenomenon than von Mises described. Though "top-down" central direction through rough measures of planning and strong political supervision of productive managers by commissars who represented the Communist Party played a significant role in the Soviet system, it also depended to a great extent on "bottom up" decentralized interactions of productive enterprises and the improvisation of immediate producers.

The fierce debates over capitalism vs. socialism in the nineteen-twenties and thirties – known as the *Socialist Calculation Debate* – were about ideal systems, that is ways of organizing an economy in which troublesome details like incomplete contracts and information were simply ignored.

The socialist calculation debate

After the Bolshevik Revolution in 1917 that brought the Communist Party to power in Russia, the Austrian philosopher Ludwig von Mises (1881–1973) predicted the failure of the Soviet attempt to institute a "planned" economy. He reasoned that an efficient government plan needs to mimic the efficiency properties of an ideal perfectly competitive economy. To do this this it would be necessary to solve hundreds (or thousands, or tens of thousands) of equations describing the equality of marginal rates of transformation and marginal rates of substitution for each of the huge number of produced goods and services that make up a modern economy.

Von Mises thought that decentralized interaction of buyers and sellers in competitive markets could practically solve these equations under conditions of capitalist commodity production, but that it would be impossible for a centralized planning agency to solve them, and the result would be chaos.

Theoretical economists interested in socialism, including Pareto and his student Enrico Barone (1859–1924), had explored exactly these questions in depth before the First World War and the Russian Revolution. Barone concluded that the job of the "ministry of production" of a socialist regime would be to *mimic* the conditions of price-taking competitive equilibrium by equating marginal rates of transformation and marginal rates of substitution across the different sectors of the economy. Because any profits or other surpluses or rents that might accrue to socialist enterprises would be controlled by the socialist state, Barone envisioned the state as redistributing the economic surplus in line with its political preferences over income distribution.

Barone and later advocates of came to be called “market socialism” such as Oskar Lange (1904–1965) and Abba Lerner (1903–1982) noted that even if the direct solution of the numerous equations describing the conditions for Pareto-efficiency proved to be beyond the capacity of existing mathematical methods, a socialist economy could in principle instruct its managers to act *as if* they were price-taking profit-maximizing capitalists to achieve the same efficient outcomes.

By the 1940’s the debate was over. Even the arch-opponent of socialism, Joseph Schumpeter, had conceded: “Can socialism work? Of course it can. . . . There is nothing wrong with the pure theory of socialism.”

Pareto had similarly concluded that “pure economics does not give us a truly decisive criterion for choosing between the organization of society based on private property and a socialist organization.”

What then was wrong with centralized planning? And what was wrong with the economic theory that so inadequately captured the economic shortcomings of centralized allocations and vindicated socialist planning?

The problem of information and price signals

A striking feature of the calculation debate had been that both sides deployed the Walrasian model on behalf of their arguments. Hayek soon appreciated the error and counter-attacked on stronger grounds. In his 1945 paper “The Uses of Information in Society” he re-framed the debate in terms of the costs and limited availability of information, ideas the Walrasian model of competition left out.

The problem with socialism, according to Hayek, is that the information needed by the planner is *privately* held by millions of economic actors. The actors lack the will and the means to transfer their information to a central authority.

By contrast, according to Hayek, *decentralized* markets make effective use of dispersed information. People know their own preferences and respond to prices. Under ideal conditions, people observe prices and behave as if the prices reflect the scarcity of the goods in question.

We now know that there is no realistic model of market competition for which these ideal conditions hold:

- Many relevant prices do not exist because the markets are missing.
- Prices do not reflect scarcity because of external costs and benefits.

HISTORY One of the leaders of the Bolshevik Revolution Leon Trotsky, explained the essentials of Hayek’s reasoning many years earlier: "If a universal mind existed, such a mind, of course, could *a priori* draw up a faultless and exhaustive economic plan, beginning with the number of acres of wheat down to the last button for a vest. The bureaucracy often imagines that just such a mind is at its disposal; that is why it so easily frees itself from the control of the market . . ."

HISTORY In the lecture commemorating his Nobel Prize, Coase commented on the role of neoclassical economic advisors and the costly transition from a centrally planned to a capitalist economy in the 1990’s: “Without the appropriate institutions no market economy of any significance is possible. If we knew more about our own economy, we would be in a better position to advise them.”

- Prices may be unknowable, prices of future goods for example.

By focusing attention on which institutions more effectively use the available information, Hayek's paper counts as a landmark work in the theory of economic institutions.

In formalizing a major shortcoming of centralized planning, Hayek also pointed to the deficiencies of the Walrasian paradigm, namely the assumption of complete contracts.

For many economists, the Walrasian model of perfect competition had provided a strong justification to leave the allocation of resources to the market. But most economists, including Arrow , held that the empirical implausibility of the First Welfare Theorem's assumptions – especially complete contracts – disqualified it as a defense of *laissez faire*. The irony is that the perfectly competitive model neither justified an unregulated market economy nor provided the basis for understanding the shortcomings of a centrally planned economy. Critics of central planning complained that many socialists argued on "**utopian**" grounds – for example assuming that the central planner had all the information necessary. In response, economists commonly insist on studying what they called "actually existing socialism" meaning the anything-but-utopian realities of the Soviet economy and society. Because the assumptions of complete contracts and price-taking equilibrium abstract from the commonplace observations of the day-to-day operation of capitalist economies, these assumptions also have a *utopian* character.

HISTORY A utopia is a world in which everything is "perfect." It literally translates as "no-place" (because no such perfect place actually exists). The model of Walrasian competitive markets is utopian because it assumes that there is perfect competition and contracts are complete (neither of which are true in the real world). Thinking about and defining utopias has been a mainstay of political philosophy going back to Plato whose "Republic" was a form of utopia or perfect society. Plato did not allow poets and actors in his utopia, which the authors of this book would object to.

We can surely imagine an economy in which the assumptions of perfect competition and complete contracts hold and it may be insightful to reason about life in that imaginary economy, just as Arrow and Debreu did when proving the First Welfare Theorem. But, can feasible changes in real-world markets bring real-world, "actually existing" capitalist economies closer to the ideal the assumptions of price-taking competitive equilibrium that complete contracts represent?

14.16 Perfect competition or the perfect competitor

Walrasian competitive markets are not really about capitalism, or any other market system. Nor do they capture even the idealized logic of a system of decentralized allocation among people with limited information.

An irony of the planning vs. the market debate is that the Walrasian competitive markets are highly centralized, requiring the Auctioneer to find equilibrium prices.

Markets play no real role in this model, nor is the model consistent with any plausible *process* of how equilibrium is reached. The reason is that buyers and sellers do not set prices (they are "price takers").

HISTORY Arrow and Hahn drew attention to this gap: "If we did not stipulate ... an auctioneer, we would have to describe how it comes about that at any moment of time two goods exchange on the same terms wherever such an exchange takes place and how these terms come to change under market pressure."

	Rent-seeking competition (Chapter 9)	*"Perfect competition" (Chapter 14)*
Actors	Price makers	Price takers
Law of One Price	No	Yes
External effects	Present	Absent
Equilibrium	Not necessarily market clearing	Market clearing
Distribution	Markets affect inequality	Distribution neutrality
Pareto efficiency	No	Yes
How prices change	Rent seeking, bargaining	Fictional Auctioneer

Table 14.3: **Comparing rent-seeking competition and perfect competition.**

Fortunately economics has a lot more to offer about how markets work than the perfectly competitive market equilibrium, which, as Table 14.3 shows, differs in many ways from the markets you have studied.

Given that coordination failures due to incomplete contracting – in greater or lesser degree – that are absent from the model of "perfect competition" and yet are endemic to most social interactions, one may wonder why the Fundamental Welfare Theorems have attracted such attention. No doubt some of the interest in the theorems comes from a misreading of them. In this misreading, people argue that the theorems demonstrate the desirability of limiting government's role in the economy to the definition and enforcement of property rights, as advocated by people favoring "laissez-faire" economic policy.

But the question of the Pareto-efficiency of competitive equilibrium outcomes now plays virtually no role in scholarly discussions of economic policy and institutions. Instead people have re-focused on the practical question of choices among feasible institutions, mechanisms, and policies supporting real-world outcomes that improve welfare, but may not be Pareto efficient.

REMINDER Recall the constitutional challenge from Chapter 1. By talking about a *mechanism* we are talking about the *institutions* that would permit people to engage in decentralized actions and avoid outcomes that none of them would have chosen, that is, avoiding Pareto-inefficient outcomes.

In this practical task, however, the lessons of the Fundamental Welfare Theorems remain important. Under the right conditions, people acting autonomously in pursuit of their own interests may implement socially desirable outcomes. Enhancing the capacity of private actions – either buying and selling on markets, or Coasean bargaining – to accomplish these ends remains an important aim of policy.

14.17 Conclusion: Ideal systems in an imperfect world

Economists – often assisted by mathematical reasoning – have been able to contribute some light to the often heated debates about the invisible hand and its policy implications.

Two conclusions stand out:

- The Walrasian model of perfect competition on which the Fundamental Theorems are based does *not* provide a decentralized model of how the process of competition among firm owners and people would reach and remain at an equilibrium.
- The assumptions under which a perfectly competitive equilibrium – if reached and sustained – would be Pareto-efficient are unlikely to hold in any real economy.

These somewhat disappointing results do not detract from the contribution of the invisible hand and the debates surrounding it, and clarify the economics of the world which we live in. In this world, a process approximating Smith's invisible hand reasoning can sometime actually operate, providing a a set of institutions that is superior to some of the alternatives, including a highly centralized economy, like the one that Nikita Khrushchev advocated in the kitchen debate with Richard Nixon.

The lessons of the perfectly competitive model and the invisible hand motivate the final two topics we will address.

First, in the next chapter we turn away from the abstract and idealized world of the perfectly competitive equilibrium to what might be called "actually existing capitalism" and its economic constitution. The inhabitants of this world are not the undifferentiated price-taking "traders" who we have considered here but instead a much more interesting and lifelike cast of characters: employers and employees, lenders and borrowers, the wealthy and the property-less, the included and the excluded, the price-making first-movers and the second-movers.

Second, in the final chapter we will study governmental policies and how they might improve the functioning of a capitalist economy. The demonstration that the conditions under which Adam Smith's invisible hand will work are not realized in any economy is not a sufficient basis for concluding that government interventions in the economy will improve economic outcomes. While well-designed policies can play an essential role in sustaining both more fair and more efficient outcomes, idealized models of the government (like the idealized model of the perfectly competitive economy studied here) are inadequate.

Making connections

Rents Many traders engaged in exchange in perfect competition or bilateral bargaining get gains from trade or rents that take them above their participation constraints. They will voluntarily choose to trade rather than remaining at their endowments.

Distribution The rents created by exchange can be equally distributed or not, depending on the initial endowments or redistribution subsequent to market operations (the Second Welfare Theorem). With bilateral bargaining and random encounters, inequality emerges from market operations. The inequality would not have existed in the fictional world of the Walrasian Auctioneer. In Coasean bilateral trading, the distribution of rents will depend on the status quo institutions and who weights their preferences more (who is willing to pay more for the allocation they want).

Efficiency Traders can obtain Pareto-efficient outcomes from perfect competition with a Walrasian Auctioneer, from biltareal bargaining with random pricing, or from Coasean (bilateral) bargaining among perfect competitors who exploit all opportunities to obtain rents.

Institutions The institutions for market operation (complete contracts, Walrasian Auctioneer) are required for perfect competition to achieve Pareto-efficient outcomes (the First Welfare Theorem). For Coasean bargaining to result in Pareto-efficient outcomes, traders need to be able to make side payments to each other, they cannot incur transaction costs from trading, and they need access to savings (liquid capital) or credit (they cannot be excluded from the credit market) to make the side payments.

Important Ideas

decentralized exchange	central planning	agent based simulation
double auction	socialist calculation debate	invisible hand
utility function	marginal rate of substitution	Edgeworth box
Pareto-improving lens	Pareto efficiency	Pareto-efficient curve
First welfare theorem	utility possibilities frontier	endowment
post-exchange allocation	Social Planner	participation constraint
Second welfare theorem	Auctioneer	excess supply
excess demand	market-clearing	institutions
Walrasian model of exchange	gains from trade	inequality
economic rent	socially efficient	negative external effect
positive external effect	incomplete contract	
social costs & benefits	deadweight loss	
equal treatment	distributional neutrality	path dependence
alloction & distribution	non-verifiable information	asymmetric information

Mathematical Notation

Notation	Interpretation
x, y	goods to be allocated among people
$\bar{x}$, $\bar{y}$	total amounts of x and y available for trade
	also the dimensions of the Edgeworth box
p_x, p_y	price of goods x, y
p^N	market-clearing price
c	marginal cost of production
T	time of curfew
T^i	socially optimal time of curfew
T^A, T^B	preferred time of curfew for A and B
α and β	intensity of dislike to the difference between the preferred and actual time of curfew for A and B
W	social utility function

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

figures/UtopianCapitalism/edgeworthbox_curfew_initial.pdf

Figure 14.18: **Indifference curves at different endowments and at the allocation chosen by the mayor (impartial spectator).** T^i is the social optimum irrespective of the initial allocation of rights over the curfew. The horizontal axis is the time of the curfew (T), with T^A and T^B indicating A's and B's preferred curfews. The vertical axis shows the transfers between A and B in terms of money (y). If B has the right to play music as long as he wants, then the initial endowment is point **z** where there is no curfew and no transfer, so B simply plays music until his preferred time $T^B = 3am$, shown by point **z** and therefore resulting in the fallbacks or participation constraints given by $u_1^A = u_z^A$ and $u_5^B = u_z^B$. If A has the right to restrict music playing, then the initial endowment is point **z'** where there is a curfew at 9pm and no transfer,

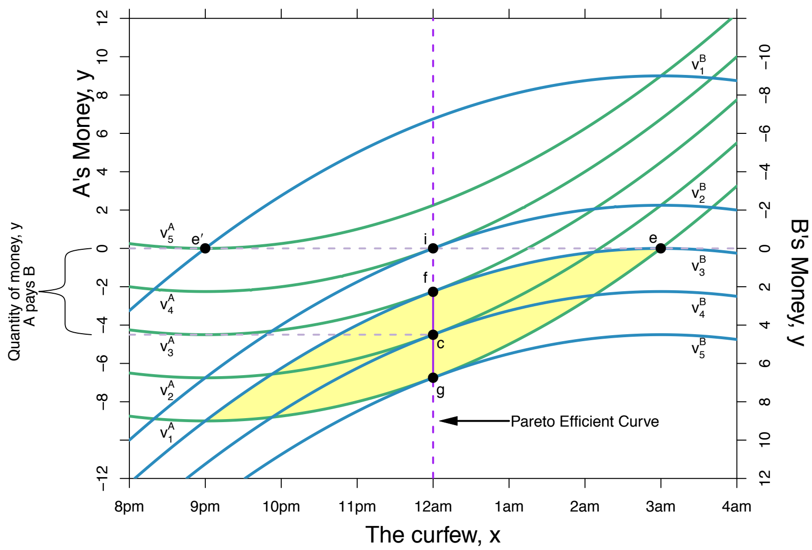


Figure 14.19: **Optimal Coasean bargaining.** The horizontal axis is the time of the curfew (T), with T^A and T^B indicating A's and B's preferred curfews. T^i is the social optimum irrespective of the initial allocation of rights over the curfew. The vertical axis shows the transfers between A and B in terms of money (y). If the players can transfer cash to each other, then they can arrive at a bargain in the Pareto-improving lens and on the Pareto-efficient curve. If A has take-it-or-leave-it power over B, then she will choose point $\mathbf{t^A}$. If B has take-it-or-leave-it power over A, then he will choose point $\mathbf{t^B}$. If they arrive at a negotiated solution to split the rents 50-50, then they will arrive at point **v** (a **v**oluntary, yet arbitrary, Pareto-efficient Coasean bargain). We assumed $T^A = 1(9pm), T^B = 7(3am), \alpha = \frac{1}{2}, \beta = \frac{1}{2}$.

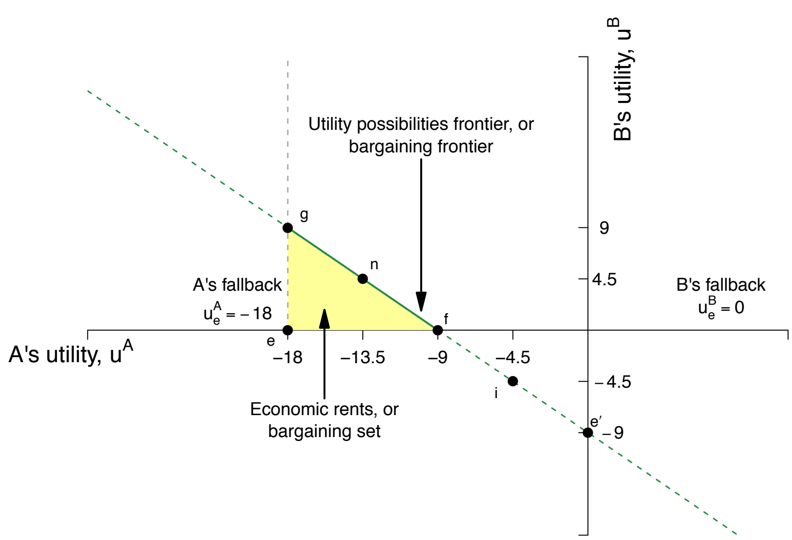


Figure 14.20: **The utility possibilities frontier and bargaining set that results from the Coasean bargaining over the curfew between the players**. The triangle in yellow is the bargaining set and corresponds to the Pareto-improving lens in Figure 14.19. Points **a**, **b**, $\mathbf{t^A}$, $\mathbf{t^B}$, **i** and **v** correspond to the same points in the Figures 14.18 and 14.19, but now represent the relevant utilities at each bundle rather than the allocations of curfew (T) and differences in income (y^A and y^B).

15

Capitalism: Innovation & Inequality

> . . . the proprietors of . . . establishments and their operatives do not stand on an equality, . . . their interests are, to a certain extent, conflicting. The former naturally desire to obtain as much labor as possible from their employees, while the latter are often induced by the fear of discharge to conform to the regulations which [are] detrimental to their health or strength. In other words, the proprietors lay down the rules and the laborers are practically constrained to obey them.
>
> United States Supreme Court, 1898

15.1 Introduction

In the 14th century, Ibn Battuta (who you met earlier in Chapter 4), one of the leading geographers and explorers of his age, traveled widely in Asia, Africa, the Middle East, Russia, and Spain. In 1347, he visited the land we now call Bangladesh. “This is a ... country... abounding in rice,” he wrote. He described traveling along its waterways passing “between villages and orchards, just as if we were going through a bazaar.” Six and a half centuries later, a third of the people of Bangladesh are undernourished, and the country is among the world’s poorest.

At the time of Ibn Battuta’s visit to Bangladesh, Europe was reeling under the impact of the bubonic plague, which took the lives of a quarter or more of the residents in many cities. Manual workers in London, probably among the better off anywhere on the continent, consumed less than 2000 calories per day (a physically active man needs close to 3000 calories per day to maintain his health).

The shortage of labor following the plague boosted real wages somewhat through the middle of the next century. But over the next four centuries real wages of laborers did not rise in any European city for which records exist: in most cities wages fell by substantial amounts and in Northern Italy wages fell to half their earlier level.

DOING ECONOMICS

This chapter will enable you to do the following:

- Understand why in many countries "the capitalist revolution" was associated with rapid increases in living standards.
- Understand how the fact that less wealthy people are both more risk averse and face limited access to credit may explain why the conventional firm – private owners of capital goods hiring labor – is the most prevalent economic organization in modern economies.
- Explain how, by placing decision-making power in the hands of wealthy and hence not very risk-averse owners and managers, the capitalist firm may promote innovation.
- Use the Lorenz curve and Gini coefficient – measures of inequality – and models of credit, labor, and product markets to explain how limited competition in product markets affects inequality.
- Use the model of product, credit, and labor markets to explain the effects of public policies can affect the extent of unemployment, wages, and inequality.
- Explain how principals and agents in labor and credit markets are connected not only by exchange, but also politically, with employers and lenders exercising power over employees and borrowers.

./figures/Capitalism/londonwage.pdf

Figure 15.1: **London Craftsmen’s Real Wages, with 1850 as the base year, from 1264 to 2001.** Source: Robert Allen, 2001

Since around 1800 real wages rose dramatically, first in England and later by even greater amounts in other European cities. Figure 15.1 illustrates the increase, showing the real wages of London craftsmen from 1264 to 2001 including the recent twenty-fold increase.

The astounding increase in wages in London was replicated in similar increases in average living standards (measured by per capita income) in Great Britain, followed by even more dramatic increases in Japan and Italy. Some of the evidence can be seen in what we call the hockey stick of history, shown in Figure 15.2. Nothing similar happened in Bangladesh.

At the time of Ibn Batutta's visit to Bangladesh the world was flat, economically speaking: countries and regions did not differ much in their average living standards. The main inequalities were between the rich and the poor within a country: land lords and farmers, masters and slaves, men and women.

CAPITALISM is an economic system in which most production takes place in privately-owned firms that employ labor in return for wages to produce goods and services to be sold on markets to make a profit.

Starting around the middle of the 18th century this pattern of within-country inequality began changing as vast differences in average incomes began to develop between the rich countries and regions and the poor.

How did this happen?

The answer, very briefly, is **capitalism**. Capitalism is an economic system in which most production takes place in privately-owned firms that employ labor in return for wages to produce goods and services to be sold on markets to make a profit. There are two fundamental characteristics of a capitalist firm. They concern:

REMINDER Institutions are the laws, social norms, and "rules of the game" that govern how people behave and interact in social interactions.

- *Control rights*: Who makes *decisions* about how the firm is run, what it produces, what technologies are used and so on?

- *Residual claimancy*: Who *owns* the *net revenues* (profits) from the sale of what is produced after the payments for the inputs have been made?

In a capitalist firm, the *owners* of the capital goods used in the production process have the control rights and are residual claimants on the revenues of the firm. This is what ownership of a firm means.

The emergence and diffusion of a the novel set of institutions that came to be called "capitalism" in many countries brought about a vast expansion in the productivity of human labor. Higher productivity eventually led to higher wages, especially in those countries where workers' bargaining power was augmented by the expansion of workers' political rights.

This happened in Europe and not in Bangladesh.

./figures/Capitalism/gdp.pdf

Figure 15.2: **GDP per capita in China, Britain, India, Japan and Italy from the year 1000 to 2004.** Source: The Maddison-Project, http://www.ggdc.net/maddison/maddison-project/home.htm, 2013 version

15.2 Capitalism's success: The hockey stick of history

Capitalism is an economic system of recent origin, having its roots in the urban economies of Northern Italy, England, Belgium and the Netherlands starting around 500 years ago. Capitalism expanded rapidly, first in Europe, later in the places where European migrants located, and eventually to most economies in the world.

Capitalism inaugurated a new economic era as different from what preceded it as did the emergence of agriculture and the spread of the new institutions associated with settled agriculture roughly 11,000 years before. One of the most striking outcomes of the "capitalist revolution" was the rapid increase in the productivity of labor, making possible an extraordinary and prolonged increase in people's material living standards. This accomplishment of productivity and improved well-being is not controversial even among the most severe critics of capitalism – Marx and Engels stressed it in their 1848 *Communist Manifesto*. But what aspects of capitalism's success might be controversial?

HISTORY The use of the word "capitalism" has had its booms and busts over the course of history; and it is often, but far from always associated – in newspaper articles, for example – with the word "democracy." Cognitive scientist Simon De Deo tracks what he calls "the marriage (and divorce) of capitalism and democracy" using computer science methods applied to centuries of newspaper text. See the video here https://cmu-lib.github.io/dhlg/project-videos/dedeo/

- Not all capitalist economies have prospered in all centuries, for example, many Latin American economies over the twentieth century failed to grow or failed distribute the benefits of their growth equitably.
- Some other economic systems have also fostered sustained rapid economic growth. For example, the Soviet Union under centralized economic

planning, from the Great Depression until the 1970s (admittedly followed by slump), or Vietnam and China under a mixture of markets and planning since the 1980s.

But these examples do not diminish capitalism's record over centuries as a uniquely productive system.

What capitalism accomplished, and what accounts for much of its productive success, is that capitalism allowed some individuals to innovate. When people *innovate* they introduce new technologies, new products, and new ways of organizing production and marketing. But innovation alone isn't enough. Capitalism created incentives for people to innovate by giving people:

- A reasonable expectation of reaping the rewards if they successfully innovated,
- while bearing the costs if they failed.

But innovation also ties to inequality. Inequalities in wealth combined with credit and other financial markets (aided by the introduction of limited liability, which we explained in Chapter 12) allowed a single individual or a small group of people to amass substantial resources under unified control and take risks on a grand scale. Labor markets allowed these material resources to be put to use to employ vast numbers of workers, so the owners of capital could reap the rewards of technological innovation and economies of scale.

INNOVATION The key to the success of capitalism in raising living standards in the countries shown in the figure is that it brought about a permanent technological revolution, a never ending process whereby new products, new technologies, and new forms of organization would be introduced and widely diffused, resulting in a long run increase in amount of output produced in an hour of labor.

The wealth and ability to access credit of those people directing these business projects made the risks of innovation tolerable. It also allowed them to offer *de facto* insurance to those people they employed, in the form of the wage contract. The result was that those people with nothing to supply other than their labor could be employed by people with capital initiating new and innovative projects. These workers would never have been willing to assume these risks themselves, but were comfortable working for wages on behalf of others who could bear these risks.

For the first time in history, surviving in the competition among members of the economic elite depended on one's success in introducing unprecedented ways of organizing production and sales, new technologies, and novel products. Joseph Schumpeter referred to the process of innovation as "creative destruction." The creative part was the new products, new employment opportunities, and increased productivity associated with the innovation. The destruction part was the fate of the losers in this process: those whose jobs are lost, or businesses bankrupted.

HISTORY Joseph Schumpeter (1883-1950) was Finance Minister of his native Austria before becoming a professor of economics at Harvard University. Like Karl Marx, he described capitalism as an economic system that was constantly in motion that would eventually succumb to what he called "the march into socialism." Unlike Marx, he was no friend of socialism: his most famous book *Capitalism, Socialism and Democracy* (1942) celebrated the successes of capitalism.

The success of these arrangements of work and innovation hinged critically on the relative security of possession associated with the rule of law (including private property rights), accomplished in large part by the increasingly powerful nation-states that grew in conjunction with capitalist economic institutions.

But capitalism's success did not hinge on contracts being complete. Quite the contrary, capitalism fostered the rapid diffusion of new techniques through a competitive process where firms that imitated the innovators captured much of the increased economic rents generated by innovators. This was possible because patents, copyrights, and trademark law provided little protection for the intellectual property rights of the innovators. As a result, as innovators they could not effectively monopolize their innovation or prevent imitators for adopting their innovations.

HISTORY Petra Moser, an economic historian, studied the engineering and other innovations on display at Chrystal Palace Exhibition known as The Great Exhibition or The Great Exhibition of the Works of Industry of All Nations in London in 1851. She found that countries with no or weak intellectual property rights were at least as innovative as those with stronger patent laws and other intellectual property rights.

The process of innovation and imitation implemented allocations that diverged from those implied by the static efficiency conditions characterizing the competitive equilibrium of the idealized economy that we saw in Chapter 14.

Capitalism further enhanced productivity by greatly expanding the scope of both labor and financial markets. Financial and labor markets were notorious for their incomplete contracts and resulting market failures. The secret of capitalism's success was not that it avoided market failures and allowed an efficient allocation of resources. Instead, capitalism promoted innovations and investments in capital goods that radically increased the productivity of labor from one year to the next.

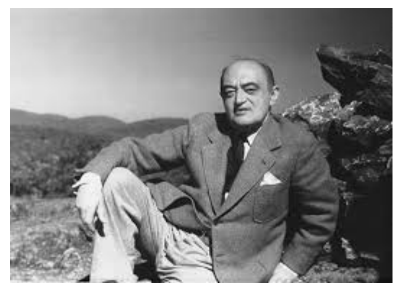


Figure 15.3: **Joseph Schumpeter (1883-1950)**, professor of economics and author of *Capitalism, Socialism and Democracy*.

As in Chapter 14, in this chapter we represent the economy as an entire system, with many markets and actors buying and selling, hiring and working, saving and investing. But the system we describe here – capitalism – bears little resemblance to the perfectly competitive equilibrium of Chapter 14.
To understand capitalism we will draw upon what you have learned about risk and about how markets work, especially the labor market and the credit market, as these are keys to understanding how different capitalism is from the model of perfect competition.

A key to the success of capitalism is that it concentrated economic decision-making power in the hands of wealthy employers – capitalists – who had both the incentives and the resources necessary to undertake the risky investments which contribute to the process of innovation.

15.3 Capitalism and inequality

Closely connected to the innovation that it supports, a second feature of capitalist economies is economic inequality.

Figure 15.4 presents data on wealth inequality for the past century. Here you see something very different from Figure 15.2. When it comes to inequality of

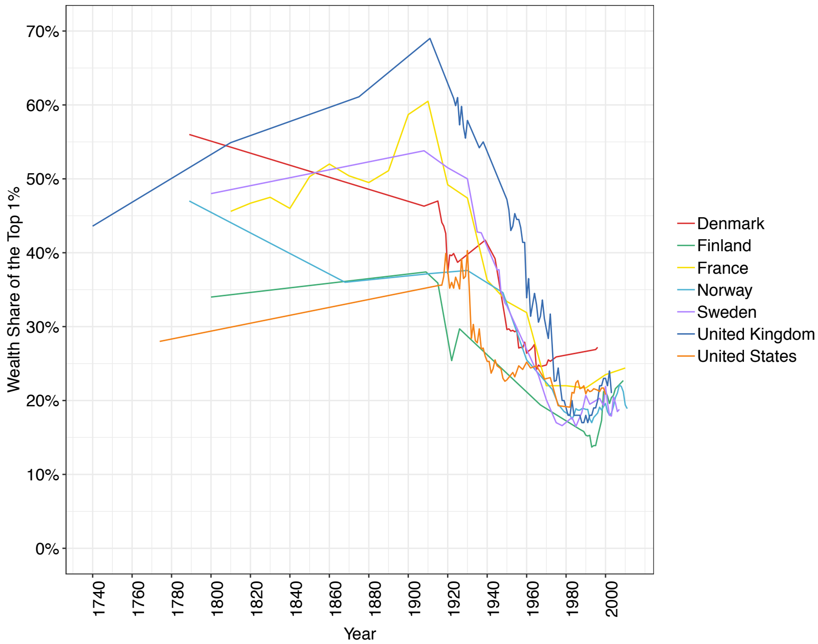


Figure 15.4: **The wealth shares of the top 1% of the wealth distribution.** A sample of countries the period 1910-2015. Source: The CORE Project, Unit 19.

wealth the modern capitalist and democratic societies are strikingly similar to those of the past.

The only exceptions are the early small-scale societies that did not have governments – which were much more equal than the other societies – and the societies of the past in which slavery was present – in which inequality was somewhat greater. Aside from these exceptions, all social systems on which we have data experience substantial wealth inequality. By these measures, the democratic and capitalist economies are slightly *more* unequal than the other societies.

In these modern societies inequality in disposable income is substantially less than inequality in **market income** because taxes and transfers redistribute income.

MARKET INCOME Market income includes all income *before* the payment of taxes or the receipt of transfers from the government; it includes earnings (wages and salaries from employment) as well as income from self-employment and from the ownership of assets (interest, rents, or dividends).

To understand both the innovation and the inequality that are characteristic of a capitalist economy, we will return to the principal-agent models that we used to study the markets for labor and credit, extending the analysis in two ways:

FACT CHECK In a YouGov survey in 2016, barely half of Americans expressed a favorable view of "capitalism." Among those polled who were under 30 years of age "socialism" was preferred by a wide margin (42 percent to 32 percent). We do not know, of course, what the survey respondents meant by these words.

- We will ask: why is the firm directed by the owners of the capital goods used in the firms production rather than by the employees that work there? Or as Paul Samuelson wrote why not "have labor hire capital"?

- We will put the credit market and the labor market together to study the relationship among borrowers and lenders, employees and employers across the whole economy.

15.4 Employment as insurance

Capital hires labor in the sense that the owners of the capital goods of the firm – its buildings, machinery, and intangible property (such as patents and trademarks) – decide who to employ in the firm and on what terms.

A useful starting point in providing an answer to why this is such a common form of organization in modern economies is the employment relationship that you studied in Chapter 11. A key feature of the labor discipline model is that the firm is represented as a group of suppliers of inputs to a common production process. The firm's activities are coordinated by means of an authority structure – owners and managers giving orders – rather than by market exchanges coordinated by complete contracts and prices.

EXAMPLE Firms do *not* use prices internally to determine which manager allocates which worker to which task. Managers and executives give instructions and workers follow them. But firms do use prices to determine which goods get sold to whom outside of the firm.

Because labor, credit, and other contracts are incomplete, those with decision making authority in the firm have power over other people's money, other people's assets, and other people's labor. The authority structure of the firm addresses the conflicts and incentive problems that arise from the decision-makers having the power they have.

An important question, then, is why is it generally the case that control rights are not held those people who *work* in the firms, but rather held by the people who supply capital to the firm, or their representatives, such as the managers and executives in a firm?

We will see that there are two reasons why this is the case, both arising from the fact that most workers do *not* own substantial amounts of wealth:

- *Risk aversion:* Workers are often risk averse and they may view wage employment as a kind of insurance.
- *Credit constraints:* Workers are not able to borrow substantial sums of money at interest rates as favorable as those available to wealthier borrowers, or potentially are not able to borrow at all, so it is difficult for them to become owners of the firm's assets.

Wage employment as insurance for risk-averse workers

Working for a wage can be considered a kind of insurance if the alternative is that one's income depends entirely on the revenues resulting from the sale of the products one produces.

This is because:

- *Income variation:* the income flowing from any production process varies as the result of shocks to costs and to market demand, unlike a fixed wage contract.

- *Risk preferences:* the cost of bearing the risk of shocks is greater for the suppliers of labor who are poorer and hence more risk-averse than for the suppliers who are wealthier and often more nearly risk-neutral.

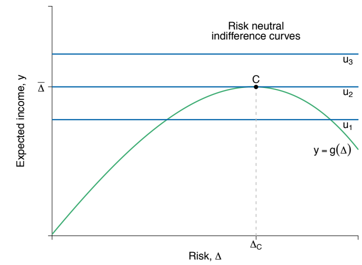


Figure 15.5: The indifference curves of a risk-neutral actor tangent to a risk-return schedule

The fixed-wage contract provides a worker insurance against variations in earnings and this insurance is more valuable to the suppliers of labor than it is costly for the suppliers of capital to provide. The fixed wage contract makes the suppliers of capital the residual claimants on the income from production. The owners of the firm therefore also 'own' the risks associated with shocks to costs and prices.

This being the case, an arrangement in which the capital suppliers also exercise control over the relevant assets reduces the cost of attracting capital to the project. The reason is that being residual claimant on the income stream of an asset that one *does not control* will be unattractive to investors if how the asset is used is not subject to complete contracts.

Capitalist firm are therefore an accommodation to differing levels of risk aversion among the input suppliers.

- Workers are more risk-averse than the owners of capital, so workers end up subject to the authority of owners.

- Owners are less risk-averse or risk-neutral than workers, so they hire capital inputs, pay for the monitoring costs of workers, make the decisions about how their capital is used, and profit from the effort of workers and their investment of capital.

Employees' limited wealth and credit market exclusion

The second reason why capital hires labor is that workers typically are not wealthy enough to acquire the funds necessary to purchase and own the capital goods with which they work. Therefore, the assignment of rights of residual claimancy and control to the suppliers of capital is the standard case because the cost of capital supplied to a firm controlled by its employees will be higher than the cost of capital faced by an otherwise identical firm controlled by its capital suppliers.

The result of differing costs of capital follows from two facts:

- *Inequality in wealth:* Labor suppliers generally have limited wealth, whereas capital suppliers tend to be wealthy

- *Inequality in prices:* As you know from chapter 12, for a project of given size and quality, if credit markets are competitive, the interest cost of a loan decreases as the wealth of the borrower increases and for any given interest rate the wealthier person can borrow more.

15.5 Explaining the hockey stick: Capitalist firms share risks & promote innovation

The model will show why the ownership of firms by the wealthy may promote innovation.

To see why, we need to combine the credit market model of Chapter 12 and the analysis of decision-making with risk from Chapter 13.

Combining these models will explain why suppliers of labor who have less wealth may prefer share-cropping or wage employment *because* these contracts shield them from risk, even if their expected incomes would be higher were they suppliers of capital, meaning they have control rights and are residual claimants. The model will also show that if sufficient assets were redistributed to the less wealthy, they might no longer wish to avoid the risk exposure of being an owner.

REMINDER Take a look back at Chapter 12 and the "You WIN!" competition in Nigeria: previously poor entrepreneurs who had access to significant wealth were much more willing to take on risks and innovate in their business practices and investments.

This section addresses three questions:

- Under what conditions will the less well off prefer to hold productive assets exposed to risk?
- Why will a conventional firm (with control and residual claimancy rights held by the suppliers of capital) be more innovative than a worker-owned firm?

In the next section we will pose a further question:

- Is it possible to redistribute wealth to the less well off in such a way that it will be sustainable once implemented?

Risk and expected profit

As in our model of choice in risky situations in Chapter 13, risk arises because there is some outcome that will be either good or bad with equal probability, and the extent of the risk is how different the two are. If the good outcome is much better than the bad outcome (if Δ is large) then there is a lot of risk.

We consider a project in which the owner of some capital goods may:

- hire an employee or
- operate the production process herself (be an owner-operator).

If an employee is hired, he is paid at the end of the period, so the value of the total capital invested (k) in the project is the following:

- the amount produced (x) times
- the number of units of the capital good required per unit of output (a_k) times

- the price per unit of the capital good (p_k)

$$\text{Value of total capital invested} \qquad k = p_k a_k x \tag{15.1}$$

To determine how much risk will be chosen under different scenarios we introduce a risk-return schedule and indifference curves based on expected income ($\hat{y}$) being a good, and risk (Δ) being a bad. The risk-return schedule is similar to the one you studied in Chapter 13: it measures the expected income ($\hat{y}$) or expected profit $(\hat{\pi})$ of the owner of the capital goods used in the production process, over a given period (say, a week, or a year) and how this varies with the degree of risk (Δ). Here we take account of:

- The *opportunity cost* of the capital invested in the project (ρk) where ρ is the opportunity cost of capital and k is the value of the capital goods used and, if an employee is hired.
- The total *wages* paid over some given period of time.
- The total costs of *monitoring* the operator of the capital goods if that person is an employee rather than the owner of the capital goods.

Whether the good or bad state occurs, the owner pays a wage w to an employee, who, using the capital goods will produce some goods which will be sold. The owner also devotes some resources m per period to monitoring the worker to ensure that she provides adequate effort to the job, yielding the owner an expected profit of

$$\text{Expected Economic Profit} \qquad \hat{\pi}^E(\Delta) = \hat{y}(\Delta) - \rho k - m - w \tag{15.2}$$

Now suppose the project is organized like a conventional firm: the capital goods are owned by a wealthy individual.

- *Control rights:* The owner has control rights over the production process: he chooses the risk level.
- *Residual claimancy:* The owners is the residual claimant on the income of the project (so he also bears all of the risk).

We now ask: what level of risk will the owner choose? To answer this question recall that the risk-return curve is the feasible frontier – that is, the constraint – that he faces, and his indifference curves represent his preferences for expected income and risk. So this is another familiar constrained optimization problem: he will seek the tangency of the risk-return curve with the highest feasible indifference curve.

REMINDER From Chapter 13, we know that the slopes of the indifference curves – with risk on the horizontal axis and expected return on the vertical – are a measure of risk aversion: steeper means greater risk aversion. For a risk-averse person indifference curves are a) upward-sloping because expected return is a good and risk is a bad; b) are steeper for greater risk exposure (farther to the right); c) become flatter for at higher income levels (higher up in the figure, due to decreasing risk aversion) and d) are flat for zero risk exposure. The y-axis intercept of an indifference curve is the "certainty equivalent" of all other points on the indifference curve, that is the amount of return which, if received with absolute certainty, would be as valuable as all of the combinations of risk and return on the same indifference curve.

figures/Capitalism/risk_contrast_monitoring_indiff.pdf

Figure 15.6: **Choice of risk level and wage chosen by a wealthy owner** The figure shows the choice of risk, $D\bar{el}ta$, for a wealthy owner who is risk neutral. Given this risk level and given the monitoring costs, m, to pay a worker the figure also shows the wage that is consistent with the competition condition (w^c). That is, the wage that is consistent with a given level of ρ given the entry and exit of firms.

If the employer is risk-neutral, we know that he will select $\Delta = \bar{\Delta}$, the maximum of the risk-return schedule $\hat{y}(\Delta)$.

How much will he then pay the worker?

To answer this question, we introduce competition among many similar firms whose owners are also risk-neutral. There are many other firms that will enter this industry if the profits are sufficiently high. Incumbent firms will leave if profits are too low. We assume that there are no barriers to entry so that in equilibrium the expected profit rate on the project must be equal to the opportunity cost of capital for the wealthy individual ρ. Or, what is the same thing, the expected economic profits from equation 15.2 on the project $(\hat{\pi}^E(\Delta))$ must be zero.

To find the wage consistent with the zero expected profits condition we do the following:

- Set $\hat{\pi}^E(\Delta) = 0$,
- Re-arrange equation 15.2 and solve for w.

So we have:

$$\begin{aligned} \text{Equilibrium wage} &= \text{Expected Profit} - \text{Opportunity cost of Capital} - \text{Monitoring Costs} \\ w^c(\bar{\Delta}) &= \hat{y}(\bar{\Delta}) - \rho k - m \end{aligned} \tag{15.3}$$

The c superscript for the wage (w) indicates that this is an equilibrium value reflecting the degree of competition as represented by the competition condition. The c superscript for the wage (w) indicates that this is an equilibrium value reflecting the degree of competition as represented by the competition condition.

REMINDER This is the same as the case of unlimited competition in Chapter 12 in which the owner had zero economic profits at the competition condition when competition was unlimited ($b = 0$) therefore $\pi^A = \rho$ and $\pi^E = \pi^A - \rho = 0$.

REMINDER This is the same as the case of unlimited competition in Chapter 12 in which the owner had zero economic profits at the competition condition when competition was unlimited ($b = 0$) therefore $\pi^A = \rho$ and $\pi^E = \pi^A - \rho = 0$.

We depict Equation 15.3 in Figure 15.6, where the lower risk-return schedule shows the wage which would be consistent with the zero expected economic profit condition for each risk level that the owner might take. We showed above that the risk-neutral owner will choose $\bar{\Delta}$, so the wage that will be offered given the zero expected economic profits condition is $w^c(\bar{\Delta})$.

15.6 How can more equal societies also be innovative?

When we turn our model to questions of public policy in real economies, we need to broaden the concept of innovation. We have identified innovation with risk-taking by a firm in the form of introducing a new product, technology, or form of organization.

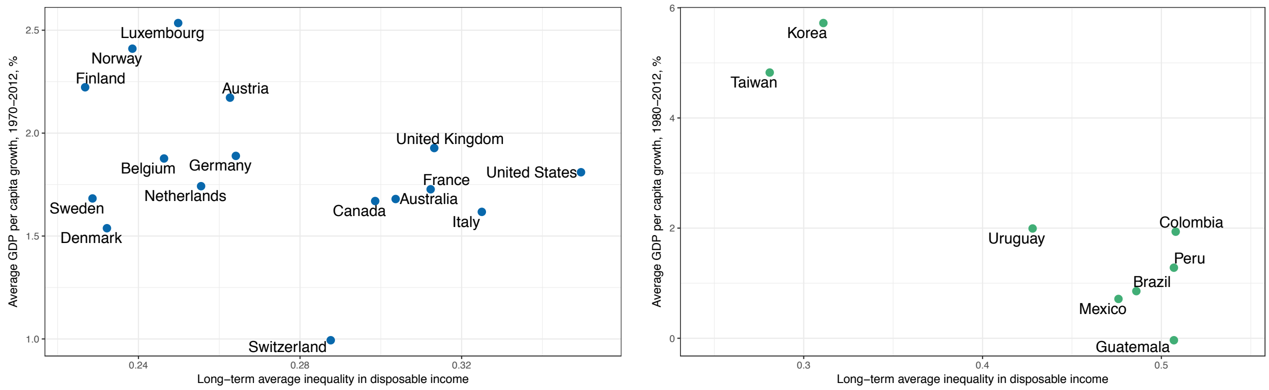
But individuals face a far wider range of choices in which they must consider trade-offs between expected income and risk. We can include many examples of this trade-off:

- The question of whether to attend higher education or what subjects to study in university, which we studied in Chapter 13
- If one's employer has gone out of business, to search for a job in the local labor market (which may be depressed because other similar firms, too, have failed), or relocate to a different part of the country – or even abroad – in search of work. The former option may have less risk but also lower expected income than the latter.
- If you are stuck in a dead-end job, do you stick with it or quit and take your chances on finding a new employer?

An innovative society is one in which the riskier but higher expected income options are more likely to be taken.

An important take-away from our model is the following:

- The wealthy are less risk-averse and also able to borrow larger amounts of money at lower interest rates than those with limited wealth, which is a reason why



(a) Growth and inequality among wealthy countries

(b) Growth and inequality among middle and low income countries

Figure 15.7: **Inequality and GDP per capita growth for various wealthy and middle-income countries.** GDP per capita growth rate per year over a long period is measure from low to high on the vertical axis. Inequality, measured by the disposable income Gini coefficient, is measure low to high along the horizontal axis. The high growth rates achieved by Taiwan and Korea were in part made possible because they could borrow new technologies originating in the richer nations. The Latin American countries shown did not manage to benefit from these "catch up" opportunities.

- firms tend to be owned and controlled by wealthy and less risk-averse individuals, which
- promotes higher levels of innovation than would be the case if the less wealthy individuals owned the firm or had a voice in its decision-making process.

We would therefore expect to find a trade-off between the degree of equality and the extent of innovation in an economy. Economies like Germany, where employees have a substantial consultative voice in the management of large firms, should be laggards in technology and innovation. But comparing across a large number of capitalist economies, this does not appear to be the case.

- Sweden, Finland, Denmark and Germany all ranked ahead of the United States and the UK in Bloomberg's 2017 index of the world most innovative economies. South Korea topped the list.
- Another commonly used measure of innovation is the number of "triadic" patents (those filed for the same innovation in the U.S., the European Community, and Japan) per head of population. On this measure Sweden and Finland outrank the U.S. with Denmark not far behind.
- All three of these Nordic countries with modest inequality outrank the U.S. on R & D by businesses as a percentage of GDP, researchers per 1000 employees, and venture capital as a percentage of GDP.
- Comparing another measure influenced by innovation across the high-income countries, the rate of growth of labor productivity shows that the more equal countries have a slight advantage, as can be seen in Figure 15.7 panel a. The same is true if we focus on middle-income economies (panel b.) where labor productivity may grow rapidly through a process of technological borrowing from world leaders.

REMINDER In Chapter 13 you saw that the elimination of tuition for higher education combined after leaving university with a progressive "graduates income tax" would reduce the risk of investing in ones own higher education, and make it possible for less wealthy individuals to continue their education.

REMINDER Recall that in Chapter 11 we defined the average labor productivity as γ, which represented how much of the total output could be attributed to each unit of labor, on average.

We do not conclude from these data that more equal countries are more innovative than less equal ones but simply that there are a good number of countries with modest levels of inequality that are highly innovative. Partly as a result, these economies have experienced rapid growth in living standards made possible by impressive improvements in labor productivity. Commonly mentioned explanations of this – particularly in light of the expectation from our model that this would not be the case – include:

- high quality education for virtually all citizens (as in Finland and Korea)
- substantial governmental support for basic research and communications infrastructure
- high wage policies that force low productivity businesses to close along with retraining and reemployment for displaced workers, as in Sweden.

Our model itself may provide some further clues. Return to Figure **??** to confirm that for a risk-averse person, the degree of risk-aversion (the steepness of the indifference curve) is greater the more risk the person is exposed to. Citizens of most of the more equal countries we have mentioned are exposed to less risk because of two reasons:

- The lesser inequality in disposable incomes, so that the difference in living standards between winners and losers, lucky and unlucky are less pronounced than in more unequal countries.
- The public provision of basic goods and services such as preschool through secondary education and health services.

We turn now from innovation to the second characteristic of the capitalist economy, inequality.

EXAMPLE The United States has a larger population than Norway and many wealthy people willing to invest capital to generate new ideas, to innovate. But that is not the same as talking about the average willingness to innovate or bear risks. For example, recent evidence in the US from the Affordable Care Act (ACA) showed that soldiers whose parents got access to health insurance through the ACA (and who could therefore insure their adult children under the age of 26) were less likely to re-enlist in the army and more likely to go back to school to get further education. The social cushion of the insurance allowed them to take on the risks of activities they wouldn't otherwise have done.

15.7 Measuring economic inequality: The Gini coefficient and the Lorenz curve

Because people care about fairness, and policies to address unfair inequality are often controversial, comparisons of economic inequality are important pieces of information. But what does it mean to say, for example, that economically speaking Germany is more equal than the U.S. or that since 1980 economic inequality has increased in India and China?

To describe the *amount* of inequality we need to ask:

- Inequality of what?
- Among whom? And...
- Find a way to *summarize* the amount of inequality.

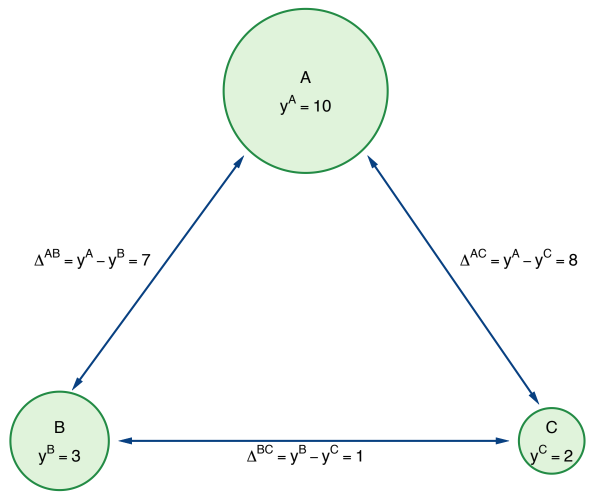


Figure 15.8: **Inequality measured as differences between pairs of people.** There are three households: the Ali, Brown and Cohen families. Each household is represented by a circle: the larger the circle, the more income (y) the household has ($y^A = 10$, $y^B = 3$ and $y^C = 2$ units of income). Each of the double-headed arrows indicates a unique pair of households: A and B, A and C, and C and B. The numbers on the arrows show the income difference difference between the indicated households (Δ^{ij}). In the figure, $n = 3$, the number of pairs is 3, the total differences between pairs is 16 (8 + 7 + 1), therefore the average difference is $16/3 = 5.33$. The total income is 15 (10 + 3 + 2) and therefore mean income is $15/3 = 5$. The relative mean difference is the average difference divided by the mean income, and one half of that gives us the Gini coefficient: $(16/3)(3/15)(1/2) = 0.53$

Economic inequality is most commonly measured by differences in wealth or income either among members of a population, for example, a nation, or between distinct groups of members of a population, for example, between men and women, or people of different religions or ethnic identity.

Two widely used ways of measuring the extent of inequality are:

- the **Gini coefficient** that provides a single number measuring how much disparity there is among the members of a the population.
- the **Lorenz curve** that represents the entire distribution of income or wealth in the population.

LORENZ CURVE The Lorenz curve summarizes the distribution of income or some other measure across a population, mapping the cumulative (poorest to richest) population shares and corresponding income shares. It was invented in 1905 by Max Lorenz (1876-1959), an American economist, while he was still a student.

GINI COEFFICIENT This measure of inequality using income as an illustration is the average difference in income between every pair of individuals in a population relative to mean income, multiplied by one-half. The Gini coefficient is usually calculated as the area between the Lorenz curve and the perfect equality line, divided by the total area under the perfect equality line. It is named after the Italian statistician Corrado Gini (1884-1965) who developed the idea.

The Gini coefficient

A simple and informative way to understand the Gini coeffient is to think about a population of people (or families) and then ask about each of the possible pairwise comparisons among them: how different are these two individuals, relative to the mean income. This way of seeing the Gini coefficient is shown for a three-person population in Figure 15.8.

We represent the population as a network. The circles (called the "nodes" of the network) are individuals or families – Ali (A), Brown (B), and Cohen (C) – and the size of the circle is proportional to the amount of income they have. One of them, Mr. Ali, might be the employer whose income is profits (after paying taxes) made by hiring a worker from the Brown family whose income (also after taxes) is the wages that Mr. Ali pays her, and a member of the Cohen family is the unemployed person receiving some kind of government assistance (financed from the taxes the others pay).

The important information about the network is on the arrows (called the "edges") between the circles: the labels on the arrows indicate the difference in income between the two individuals connected by the arrows. The Gini coefficient is based on the sum of the income differences among all of the pairs of households relative to the average income, as shown in the formula in the figure, explained in M-Note 15.1.

An advantage of this representation of the Gini coefficient is that it allows very intuitive inferences about the meaning of any particular value of this measure of inequality. Two examples illustrate this.

- from the equation above we see that the relative mean difference is $2 \times G$; so, for example, a Gini coefficient of 0.41 means that the average difference between all pairs in the population is 0.82 times the mean income $(0.82 \times \bar{y})$.
- Consider what happens if there are only two people in a population and they are dividing a 'pie' representing total income. We shall use households A and B as our two-household economy. The portion received by the disadvantaged household of the pair (Household B with $y^B = 4$) is $\frac{4}{14} = 0.28$; so using the same Gini coefficient as above and given that the smaller slice is 28 percent of the total, the richer of the pair receives the rest of the pie or 72 per cent.

M-Note 15.1: Inequality as differences between people

Figure 15.8 represents a population of just 3 people. But the Gini coefficient, like the Lorenz curve, is used to measure inequality in populations of millions. To see how this is done, we take the following steps.

- If there are n members of the population then the total number of pairs is $\frac{n(n-1)}{2}$, shown as the three edges among the n = 3 families in the figure.
- Let Δ^{ij} be the absolute difference in income between family i and family j meaning the income of the richer family minus the income of the poorer family.
- We then define the sum of the differences between all pairs as $\Delta = \sum \Delta^{ij}$, and
- the average difference as this number divided by the number of pairs or $\frac{\Delta}{(n^2-n)/2}$.
- If we then let $\bar{y}$ be the average income

Then we have the following measure of the Gini coefficient:

$$\text{Gini coefficient} \equiv G = \left(\frac{\Delta}{(n^2-n)/2}\right)\left(\frac{1}{\bar{y}}\right)\frac{1}{2} \qquad (15.4)$$

This means the Gini coefficient is the mean difference among all pairs (the first term: total differences divided by total number of pairs) relative to (divided by) the mean value of y (the "relative mean difference") times one half. The Gini coefficient for the three-person economy shown is 0.53 as shown in Figure 15.8.

Checkpoint 15.1: Understanding the Gini coefficient

- Using equation 15.4 for the Gini coefficient in the M-Note confirm that if A has all the income, and B and C none, then the Gini coefficient is equal to one.
- The Gini coefficient (like the Lorenz curve) can be used to represent inequality in any dimension, not just income, or wealth. Collect information from your friends or the class you are in on some individual characteristics – height, for example – and using equation 15.4, calculate the Gini coefficient for height.
- In the original article explaining his famous coefficient, Gini measured inequality in both economic measures (e.g. wealth) and physiological characteristics (e.g. height, right hand strength). Looking at the data he commented "Notice how the degree of inequality is different for economic and for physiological measures: for most of the physiological characteristics the value (of the Gini coefficient) is smaller than ten percent; for most of the economic measures it is greater than half." Research since has confirmed this observation. Why do you think this would be the case?

The Lorenz curve

Unlike the Gini coefficient which summarizes inquality among a group a people with a single number, the Lorenz curve gives us a picture of the disparity of income across the whole population. The Lorenz curve shows the entire population lined up along the horizontal axis from the poorest to the richest. At any point on the horizontal axis, the height of the curve indicates the fraction of total income received by the fraction of the population given by that point on the horizontal axis.

M-CHECK The representation of the Gini coefficient using the Lorenz curve and equation 15.5 does not work for very small populations while the network representation in Figure 15.8 using Equation 15.4 in the M-Note can be used on populations of any size.

Two Lorenz curves (in green) for the Netherlands are shown in Figure 15.9. In the figure the diagonal blue 45-degree line is how the Lorenz curve would look hypothetically if everyone had the same income: for example, it shows that ten percent of the population receive ten percent of the total income, fifty percent of the population receive fifty percent of the income and so on. This is called the *perfect equality line*. The perfect equality line is the Lorenz curve if everyone in the Netherlands had the same income.

For any Lorenz curve we can calculate the Gini coefficient as the area between the Lorenz curve and the perfect equality line divided by the area under the perfect equality curve or from the figure:

$$\text{Gini} = \frac{A}{A+B} \tag{15.5}$$

The Lorenz curve shows how far a real distribution of income departs from this line of perfect equality. The two panels in the figure differ in their answer to the "inequality of what?" question. On the left we have the distribution of total market income before payment of taxes or receipt of government transfer. So this includes all income including earnings (wages and salaries from employment as well as income from self-employment, and from the ownership of assets – interest, rents, or dividends). The Lorenz curve indicates that the poorest 25 per cent of the population (0.25 on the horizontal axis) receives about 2 percent of total income (0.02 on the vertical axis), and the lower-earning half of the population (0.50) has less than 20 percent of income.

DISPOSABLE INCOME Disposable income is the maximum a household can spend ('dispose of') without borrowing, after paying tax and receiving transfers (such as unemployment insurance and pensions) from the government.

Along with the distribution of market income, taxes and transfers – illustrated by the linear tax and lump sum transfer policy you have just studied — we show the Lorenz curve for *disposable income*, which better captures living standards. Disposable income is the maximum a household can spend ('dispose of') without borrowing, after paying tax and receiving transfers (such as unemployment insurance and pensions) from the government.

PROGRESSIVE AND REGRESSIVE POLICIES A system of taxes and transfers or other policies that make inequality in disposable income less than market income inequality are called "progressive." Policies that have the opposite effect are called "regressive."

Notice from the figure that in the Netherlands, almost one-fifth of the households have a near-zero market income, but most nonetheless have enough disposable income to survive, or even live comfortably: the poorest one-fifth of the population receives about 10 percent of all disposable income. The Lorenz curve for disposable income is much closer to the perfect equality line than is the Lorenz curve for market income, meaning that the system of taxes and transfers in the Netherlands reduces income inequality.

Policies that have this effect are called "**progressive**." Policies that benefit the rich at the expense of the poor are called "**regressive**." To see how progressive government policy is in the Netherlands, notice that the bottom 20 percent receive only 0.53 percent of market income but 10.41 percent of disposable income.

15.8 Inequality and the macro-economy: A micro-economic explanation

The labor-discipline model provides an analysis of the relationship between wages and the levels of employment and unemployment in the entire economy. But employment, unemployment and wages are not the only macroeconomic outcome we might wish to study. We might want to understand, for example, the relationship between the levels of employment, unemployment, and wages on the one hand and the extent of inequality in an economy on the other. The micro-economic model already showed us that some workers are unemployed perhaps receiving some income from government transfers such as unemployment insurance, other workers are employed and receiving a wage that includes an employment rent, so they are better off than the unemployed, and employers make profit as a consequence of hiring workers.

REMINDER The labor discipline model (the Ford model) explains how wages, work effort and the number of workers hired are determined in the interaction of a profit-maximizing employer and a team of utility-maximizing employees.

How do these results translate into income inequalities in the aggregate economy? To answer the question we need to use the two measures of inequality: the Lorenz curve and the Gini coefficient.

Recall:

- The Lorenz curve for disposable income – that is income after government transfers and taxes – is closer to the perfect equality line, indicating less inequality, and that
- The extent to which government taxes and transfers reduces inequality differs greatly between countries, as the comparison of the Netherlands and the US in Chapter 13 showed.

Here we look at inequality in market incomes, that is before taxes and transfers, and we ask: what attributes of an economy make market incomes more or less unequal?

As an illustration, think about an economy in which there are no self-employed people and nobody works for the government. Also there are no taxes or government expenditures, so the only income is either wages or profits. As a result everyone in the economy is included in the the three groups of people in the labor discipline model and we assume that they have the following incomes:

- *the unemployed*: they receive nothing (income is zero)
- *the employed workers:* workers who receive in wages or salaries some share of the value of the goods they produce, σ, called the **wage share** and
- *employers*: who receive the complementary share of the value of goods produced, that is one minus the wage share or $1-\sigma$.

The wage share, σ, is:

$$\begin{aligned}\sigma &= \text{wage share}\\ &= \frac{\text{real hourly wage}}{\text{real value of output produced by a worker in an hour}}\\ &= \frac{w}{\gamma}\end{aligned}$$

The Lorenz curve shows the entire population lined up along the horizontal axis from the poorest to the richest. At any point on the horizontal axis, the height of the curve indicates the fraction of total income received by the fraction of the population given by that point on the horizontal axis.

THE WAGE SHARE is the fraction of total income that goes to workers in the form of wages. We can also think of the *profit share*, which is the fraction of income that instead goes to the owners of capital goods in the form of profits.

The Lorenz curve for income in this economy is depicted in Figure 15.10.

The Lorenz curve for this economy is made up of three line segments with the beginning point (at the lower left of the figure) having coordinates of (0, 0) and the endpoint (at the upper right) having the coordinates (100, 100). The first line segment is a portion of the horizontal axis because the poorest segment of the population (the unemployed) have no income at all. The first kink in the curve occurs when we have counted all the unemployed people, so everyone else has some income.

The second kink in the Lorenz curve is the interior point, whose coordinates are (fraction of total number of the population, the wage share). The curve between the interior point and the upper right corner is steeper than the other two segments of the curve because the employers make more than their employees, so adding a given population fraction along the horizontal axis accounts for a larger increase in the share of income accounted for.

You know from the discussion of the Lorenz curve in Chapter 13 that the blue shaded area in Figure 15.10 is the difference between the actual income distribution and the hypothetical one in there is no inequality. So this area is a measure of inequality. From the figure you can see that it will increase if:

- A larger fraction of the employees are without work (higher unemployment rate) – the first kink shifts right
- The real wage falls and nothing else changes – reducing the wage share so the second kink shifts down
- Labor productivity rises and nothing else changes so output rises but wages do not, reducing the wage share – this implies that the profit share rises, so again the second kink shifts down

M-Note 15.2 provides us with an equation that translates data on the share of population unemployed, working, and employers and their shares of total income into the resulting Gini coefficient. You can use that equation to confirm that the statements in the above three bullets are true.

The "kinked" curve in the figure is a simplification designed to show how the model of the labor market with its 3 groups of actors influences the level of inequality. Lorenz curves based on actual data are smooth like the one shown in Chapter 13 for the Netherlands and the US. This reflects three facts:

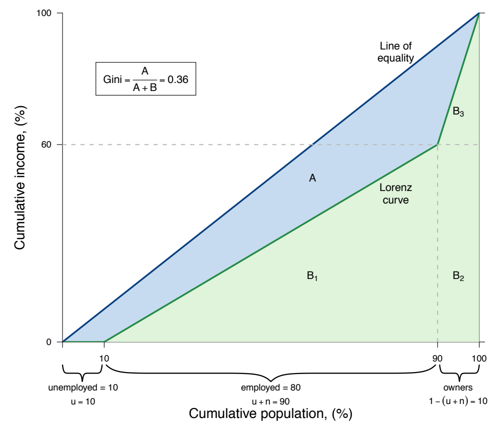
- employees are *not* identical, they are *diverse* or *"heterogeneous."* They have more or less valuable skills, more or less marketable education or credentials, and they differ in race, gender, where they live, and many other ways that affect their pay
- employers too are heterogeneous; they differ in the extent of their wealth and in whether assets they own produce goods in growing or declining demand, whether they are an innovation leader or follower, their management skills, and other things affecting the rate of profit they earn.
- even if employees and employers were identical, they would end up with differing wages and rates of return simply by the luck of the market (see Chapter 14).

	Unemployed	*Employed*	*Employers*
Population Share	u	n	$1-u-n$
Income Share	σ_u	σ_n	$1-\sigma_u-\sigma_n$

Table 15.1: **Data for the Lorenz curve.**

./figures/Capitalism/inequality_nl.pdf

Figure 15.9: **Distribution of market and disposable income in the Netherlands (2010)**. Source: LIS. Cross National Data Center. Calculations were made for the CORE Project by Stefan Thewissen (University of Oxford) in April 2015. Household market (labour and capital) income and disposable income are made comparable through PPP comparisons and are top- and bottom-coded.



(b) The Lorenz Curve and Gini coefficient

(a) The wage curve, competition condition, and number of workers employed

Figure 15.10: **The wage curve and the Lorenz curve.** In panel a. we show the wage curve and the wage for the labor market of an economy determined by the competition condition in which there are 90 workers of whom 80 are employed, receiving 60% of total income (that is the wage share) and 10 are unemployed, receiving no income. There are also 10 employers who are owners of 10 firms each of which employs 8 of the workers ($8 \times 10 = 80$). Together, these 100 people form an economy and have access to income from work that can be depicted using a Lorenz curve (panel b) with its corresponding Gini coefficient.

In the next M-Note we derive a relationship between inequality and the labor market:

$$G = u + n - (1-u)\sigma_n - (1+n)\sigma_u \qquad (15.6)$$

Equation 15.6 says that the Gini coefficient (G) depends on the proportion of the population that is unemployed (u), the proportion of the population that is employed (n), the proportion of the population that employ workers $(1-u-n)$ and the income shares of the unemployed(σ_u) and the employed (that is, the wage share, σ_n).

./figures/Capitalism/lorenz_gini.pdf

Figure 15.11: **Lorenz curve given the data in Table** 15.1.

Equation 15.6 therefore allows us to study how changes in the structure of the economy will affect the degree of inequality as measured by the Gini coefficient. We do this in Table 15.2. The first line, for example, says that the effect of an increase in the fraction of the population that is unemployed can be determined by differentiating G with respect to u. We find that this derivative is positive, so higher unemployment raises the Gini coefficient.

It is important to understand what this derivative means: it show the effect of increased unemployment, *holding constant the other parameters*. Notice that as employment increases (with unemployment constant) this means that fewer high income employers hire a larger number of workers. This increases inequality other things being unchanged. An increase in employment with no change in the number of employers would mean a decrease in unemployment, which would reduce the Gini coefficient.

Effect of an increase in	$G = u+n-(1-u)\sigma_n-(1+n)\sigma_u$	*Effect on Gini*
% **Unemployed**	$\frac{\partial G}{\partial u} = 1+\sigma_n > 0$	increase
% **Employed (fewer employers)**	$\frac{\partial G}{\partial n} = 1-\sigma_u > 0$	increase
% **of income to unemployed**	$\frac{\partial G}{\partial \sigma_u} = -(1+n) < 0$	decrease
% **of income to employed (wage share)**	$\frac{\partial G}{\partial \sigma_n} = -(1-u) < 0$	decrease

Table 15.2: The Gini coefficient and its determinants.

M-Note 15.2: The Gini coefficient in a class divided economy

Given the data in Table 15.1, we can draw the Lorenz curve as show in Figure 15.11. To calculate the Gini coefficient from these data, notice that the blue shaded area of the polygon OACE between the (green) Lorenz curve and the (blue) perfect equality line is made up of four triangles.

We derive an expression for the Gini coefficient by summing the area of these four triangles and then dividing by the area of the large triangle under the perfect equality line.

Then the distance between the perfect equality line and the points A and C is

$$\begin{aligned} |AB| &= u - \sigma_u \\ |CD| &= u + n - (\sigma_u + \sigma_n) \end{aligned}$$

You can see that the area of the triangle on the lower left

$$S(\triangle_{OAB}) = \frac{1}{2}(u^2 - u\sigma_u) = \frac{1}{2}|AB|u$$

By analogous reasoning about the other triangles, the area of the polygon $OACE$ is

$$\begin{aligned} S &= S(\triangle_{OAB}) + S(\triangle_{ABC}) + S(\triangle_{BCD}) + S(\triangle_{CDE}) \\ &= \frac{1}{2}|AB|u + \frac{1}{2}|AB|n + \frac{1}{2}|CD|n + \frac{1}{2}|CD|(1-u-n) \\ &= \frac{1}{2}\left(|AB|(u+n) + |CD|(1-u)\right) \end{aligned}$$

Therefore, the Gini coefficient is

$$\begin{aligned} G = 2S &= |AB|(u+n) + |CD|(1-u) \\ &= (u-\sigma_u)(u+n) + (u+n) - (\sigma_u + \sigma_n)(1-u) \\ &= u(u+n) - \sigma_u(u+n) + (u+n)(1-u) - (\sigma_u + \sigma_n)(1-u) \\ &= u + n - (1-u)\sigma_n - (1+n)\sigma_u \end{aligned}$$

In some applications it will be the case that the share of income going to the unemployed (σ_u)or to the poorest segment of the population will be zero (as is the case in Figure 15.10), so that we will have

$$G = u + n - (1-u)\sigma_n$$

15.9 Market power and the distribution of income

An important trend in many of the high income economies of the world in recent decades has been the increase in market income inequality. As you can see from the table, the increase in inequality could have occurred because of an increase in unemployment. But in many countries – the US and the UK are examples – unemployment at the end of the second decade of the 21st century has not increased. What could account for these developments?

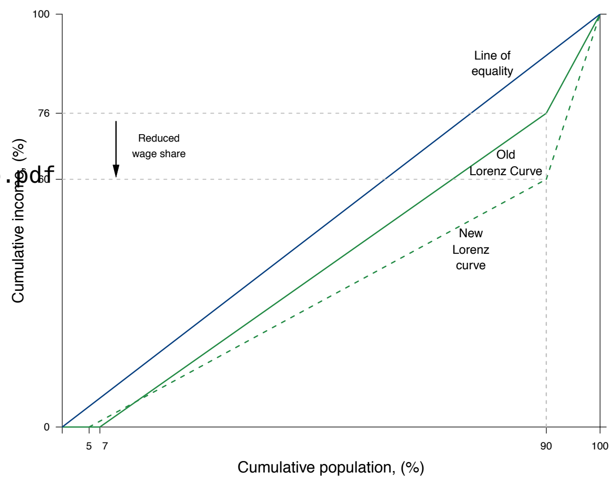
We can use equations 11.23 and 15.8 to suggest a two-part answer.

- *Less competition*: A decline in competition shifted the competition curve downwards.
- *Less bargaining power for workers*: A decrease in workers' access to government transfers when unemployed and fewer impediments to workers being fired resulted in a downward shift of the wage curve.

First, the emergence of modern monopoly and the decline in the degree of competition discussed in Chapter 9 raised the mark-up ratio, and this pushed down the competition condition (Equation 11.23) as shown in Figure 15.12. Decreased competition would account for an increase in inequality, but by itself this should have also decreased employment and increased unemployment (the shift from point **a** to point **b** in the Figure 15.12). So something else must have changed too.

This is the second part of the answer. In many countries trade unions became weaker, it became easier to terminate workers, and harder for the unemployed to gain access to government payments. This lowered the unemployment figures/Capitalism/wage_share_unemployment_prop.pdf

(a) Shifts in the competition condition and the wage curve



(b) Changes in the Lorenz Curve

Figure 15.12: **Effect of shifts in the wage curve and the competition condition on the Lorenz curve.** Panel a. depicts a shift in the wage curve and a shift in the competition condition. The shift in the competition condition results in a movement from the original equilibrium at point **a** to a new equilibrium at point **b** with a lower wage and lower level of employment. The subsequent shift in the wage curve results in a transition from point **b** to point **c** with higher unemployment at the new real wage.

The consequences of these changes are shown in panel **b** in the Lorenz curve. First of all, there is a decrease in unemployment at point **c** in the right panel which is reflected in the Lorenz curve on the horizontal axis: more workers have jobs (85 workers are employed so 5 are unemployed). These employed workers, however, have a lower real wage and therefore the wage share is lower, so the Lorenz curve has pivoted slightly inwards and inequality is higher. In comparing the two different outcomes it is important to identify that with more workers, there is more output, but the workers' cumulative share of income is their wage share. We have assumed that worker productivity is $\gamma = 1$, so output is 83 before the change at point **a** and output is 85 at point **c**. At **a**, workers receive in total $0.76 \times 83 = 63$ as a share of the total income 83, which is exactly the wage share $0.76 = \frac{63}{83}$ shown with the old Lorenz curve. At **c**, workers receive in total $0.6 \times 85 = 51$ as a share of the total income 85, which is exactly the wage share $0.6 = \frac{51}{85}$ shown on the new Lorenz curve.

benefit (B) and raised the probability (t) that a worker caught shirking would be fired. The effect was to shift downward the wage curve, as shown in Figure 15.12. So the economy shifted from point **a** not to point **b** but to point **c** with somewhat lower unemployment and a lower wage share.

The effect of these changes on the distribution of income is illustrated in panel b of Figure 15.12.

To understand how competition in markets and inequality interact, we use Figure 15.12. The initial outcome is shown by the original wage curve intersecting the original competition condition at point **a** in the right-hand panel with corresponding level of inequality shown by the Lorenz curve in the right-hand panel. Using Equation 15.6, the resulting Gini coefficient is $G = 0.83 + 0.07 - (1 - 0.07)(0.76) = 0.19$.

As a consequence of a reduction in the level of competition, however, the competition condition shifts downwards with a new intersection with the wage curve at point **b**. If this were the only change, this would reduce unemployment to 78 workers at the intersection of the new competition condition with the old wage curve. However, when the wage curve shifts downwards to a new wage curve (as was the case with the worsening fallback position of workers since the 1970s), then the new competition condition and the new wage curve intersect at point **c** with the lower wage and higher level of employment of 85 workers.

Point **c** corresponds to this new level of inequality in the right-hand panel shown by the new wage share (60) and the lower unemployment (only 5 work-figures/Capitalism/markup_and_gini.pdf

Figure 15.13: **The mark-up and the Gini coefficient over time in the United States over the period 1970 to 2015.** Source:

ers unemployed vs. 7 previously). The slope of the Lorenz curve is flatter, however, because workers now receive a smaller share of income. Using Equation 15.6 again, the level of inequality will therefore provide a Gini coefficient of $G = 0.85 + 0.05 - (1 - 0.05)(0.6) = 0.33$, which is significantly higher than the previous Gini coefficient. The dynamics of the model therefore help us to understand empirical facts we have observed in the world: from increasing market concentration and a higher mark-up with decreased competition, to lower unemployment alongside higher inequality.

Recent trends in the markup and the Gini coefficient in the U.S. are shown in Figure 15.13. The blue line, economic profits as a fraction of total output, presents data you have already seen in Chapter 9. The orange line shows the Gini coefficient for market incomes over the same period.

15.10 Modern monopoly, winners-take-all and public policy

But there is an important difference between Standard Oil along with other giant firms such as U.S. Steel and British Petroleum on the one hand, and a novel market structure exemplified by Amazon, Microsoft and Facebook.

The term "modern monopoly" has been used to describe them but like the "conventional monopolies" with which they are contrasted they aggressively compete other companies selling similar products.

What is "modern" about them are are extraordinary competitive advantages of large scale. .

These advantages occur for two reasons:

- First, many modern monopolies may have substantial fixed costs and very low marginal cost. Examples are companies providing information in which there may be substantial costs to produce or acquire rights to the information (this is called a "first copy cost"), but zero costs to making it available to consumers after that. Think Spotify and other streaming music services.

- Second, as the examples of Amazon and Facebook illustrate, modern monopolies benefit from what we have called "network-based economies of scale in demand." One of the benefits of buying on Amazon is that you can read reviews of the product you are interested in by hundreds of other Amazon users. People looking for an apartment for a weekend go to AirBnB.com because many apartment owners have posted their apartments there, which they would not have done unless lots of apartment seekers were logging onto AirBnB.

REMINDER Economies of scale in demand are sometimes called "network economies of scale" because the value of being part of the network (e.g. of Facebook or Instagram users) increases the more other users there are.

A result of these novel forms of economies of scale is winner-take-all competition, which is to say, a process of competition which results in a monopoly or near monopoly. As countless companies that challenged Amazon have discovered, the advantages of size are decisive. In modeling modern monopoly, we have to reject our assumption (in the Cournot model) that average and marginal cost are independent of the output.

Let's consider the extreme case of a firm or firms which have a substantial fixed cost c_0. Total costs (irrespective of output) are then c_0 and marginal costs are zero. Therefore, average costs are $\frac{c_0}{X}$. A single firm with this cost structure would behave just as the monopoly in the Cournot model, producing a quantity such that the marginal revenue equaled the marginal cost (namely zero). As with the monopoly in Figure 9.10, the monopoly would produce less output than if price were equal to marginal cost (a price of zero).

EXAMPLE Think about policies and practices that might make it easier for firms to perfectly price discriminate and the markup ration increases approximate the conditions for a Pareto-efficient outcome. These would involve the ability to collect accurate information on the willingness to pay of each buyer, and to make credible take-it-or-leave-it offers to each. How would you evaluate these policies, weighing the possibly competing objectives of Pareto-efficiency, privacy, and fairness in the distribution of income?

What would be an appropriate policy to remedy this loss in consumer surplus?

The first thing that would come to mind is the traditional antitrust remedy: break up the monopoly into a sufficiently large number of firms so that prices and deadweight loss would fall, while consumer surplus would increase.

Type	*Examples*	*Characterization*	*Harm*	*Remedy*
Old	GM, U.S. Steel Tata Industries BASF British Petroleum	Downward-sloping demand curve; stable competition among the few.	Inefficiently limited buyer access to goods and services; Foregone consumer surplus survival of high-cost firms	Objective: Restore competition (in the market) by Breakup of large units; Price caps; Public regulation/ownership of natural monopolies
Modern	Amazon Facebook Tencent/WeChat Alibaba Uber/Lyft Bristol-Myers-Squibb (Pharmaceuticals)	New natural monopoly High first copy costs zero or low marginal costs; demand-side increasing returns; Serial monopoly	Foregone consumer surplus when $p > mc$; Centralized control over essential social infrastructure; limited competitor access to market; winner-take-all firm which may not be most efficient or highest value provider	Public regulation or ownership; Support competition for the market & facilitate take-overs; Auction of market dominance; Corporate governance reforms

Table 15.3: **Conventional and Modern Monopoly**

But this would make no sense. Why? Instead of a total cost of providing the goods equal to c_0 by a single firm, the costs would be the number of firms multiplied by the fixed costs, or nc_0. Therefore, the average cost of the goods or services provided would be nc_0/X rather than c_0/X. The average cost would be the number of firms, n, times the cost under monopoly.

The challenge for public policy is that in the case of modern monopoly small firms are inefficient firms. Large size is socially beneficial because it spreads a single fixed cost over all the units produced. The logic of policy for conventional monopolies is to promote "competition in the market" through the increased number of firms. But this logic cannot make sense for modern monopoly. In the case of modern monopoly two approaches have been proposed.

- Public regulation or ownership: First, recognize the production of this good or service is unavoidably a monopoly and let the government regulate the prices it can charge and other aspects of its business. This has been the conventional approach to what are called "natural monopolies" such as railroad and electricity networks.

- Competition *for* the market rather than *in* the market: The objective of "competition for the market" policies is to seek to ensure that natural monopolies are owned by people who will produce the best products at the lowest cost.

To see how this might work, think of an industry in which the sharply declining costs per unit of output sustains a winner-take-all kind of competition. Suppose a single firm is dominant in producing some good or service whose owners are incompetent but are able to deter entry by competitors through predatory pricing and other strategies. Then another group of potential owners, who would run the firm more profitably might consider buying a controlling

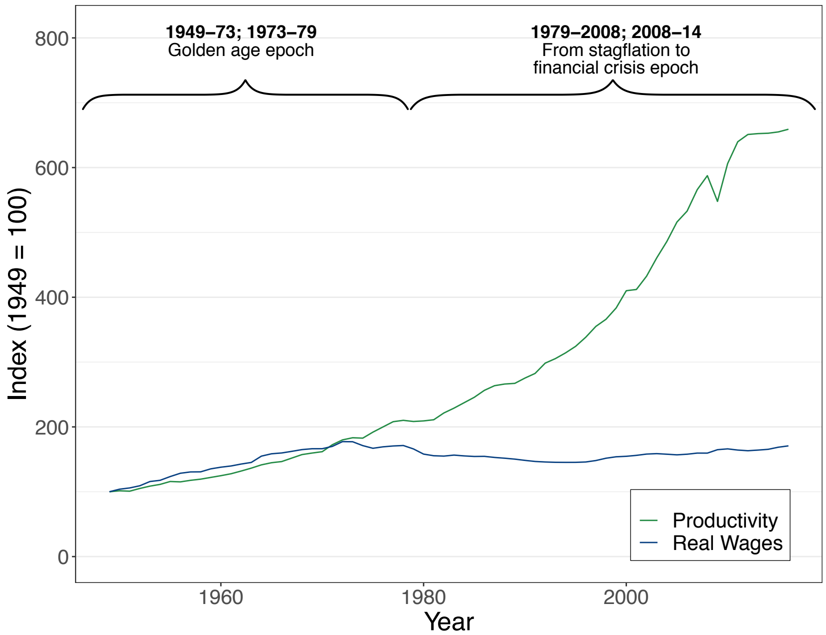


Figure 15.14: **Real wages and productivity of manufacturing workers over time in the United States 1949 to 2016.** Real wages and productivity increased in tandem during the period 1949 to 1979. Productivity is an index measuring real output per hour of all persons.

share of the stock of the incumbent firm. Facilitating this kind of firm buy-out is an example of competition for (rather than in) the market. Table 15.3 summarizes differences between conventional and modern monopoly.

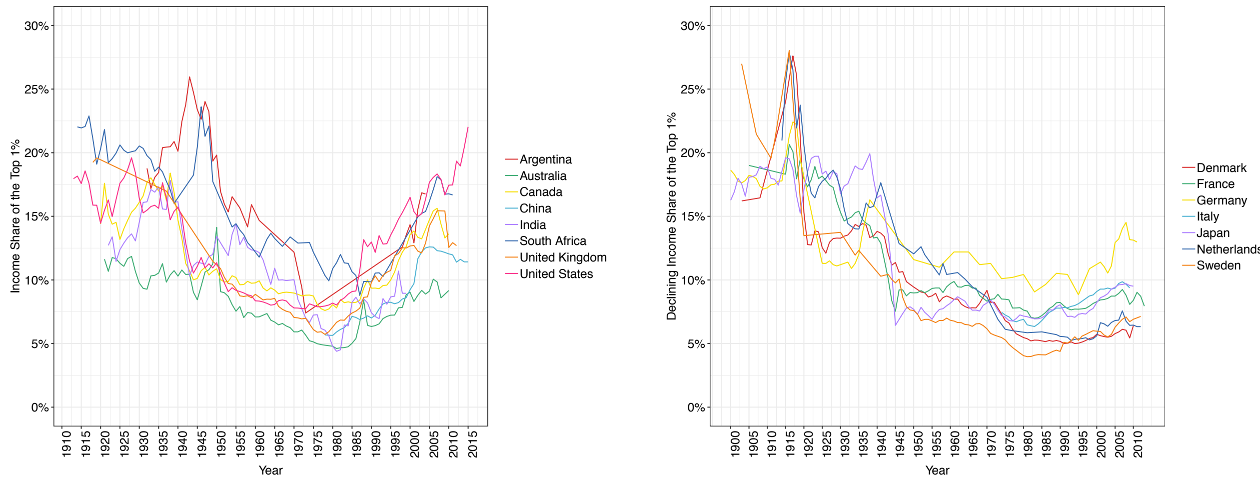
15.11 *Application: Public policy to raise wages and reduce unemployment and inequality*

The models you have learned provide some insight about why some societies would be more equal than others, and why in some periods of time inequality would increase. Figure 15.14 provides an example. From the end of the Second World War to the end of the 1970s, labor productivity and real wages in manufacturing in the U.S. grew approximately in tandem (real wages actually grew slightly faster than productivity over much of that period).

But since the 1980s, in the U.S. the real wages of employees have been roughly constant, while productivity has much more than doubled. The result predicted by our models is the following:

- increased labor productivity results in an increase in the share of total income (the economic profits) received by wealthy employers
- an associated increase in market income inequality when wages remain constant.

We will see later in this chapter that this is exactly what happened.



(a) Income share of the top 1%

(b) Declining share of top 1%

Figure 15.15: **Income share and declining income shares of top 1 percent**.
Source:

We can use the model to study how public policy could alter the wage rate, the profit rate, and level of employment in a way that is sustainable in the long run.

Sustainability here means that the desired policy outcome – higher real wages, or less unemployment – is consistent with maintaining the employees' motivation to work hard and well and the owners' motivation to invest in the firms' capital goods sufficiently and to hire additional workers even at a higher real wage.

Sustainable policy therefore requires that the Nash equilibrium be moved, which can only happen if one or both of the wage curve and the profit curve are shifted.

The Nash equilibrium of the model of the whole economy is determined by two equations (the notation is shown in Table 15.4):

$$\text{Competition condition} \qquad w^c = \frac{(1-b)\gamma}{(1+\rho)} \tag{15.7}$$

And the

$$\text{No-shirking wage} \qquad w^N = \frac{u}{tj} + B \tag{15.8}$$

Changing the Nash equilibrium requires changing one of the variables in one of these two equations. We will consider four ways that the Nash equilibrium might be changed by public policies or by labor unions.

Policy Target	*Effect*	*Shift*	*Wage*	*Unemployment*	*Gini*
Competition Policy	Reduce b	w^c up	$+$	$-$	$-$
Opportunity Cost of Capital	Reduce ρ	w^c up	$+$	$-$	$-$
Research, Education	Raise γ	w^c up	$+$	$-$	$-$
Disutility of Effort	Reduce $\underline{u}$	w^* down	None	$-$	$-$
Unemployment Insurance	Raise B	w^* up	None	$+$	$-$
Barriers to Job Termination	Increase t	w^* down	None	$-$	$-$

Table 15.4: **Policy effects on Equilibrium.** Each line shows a policy, the parameter that the policy targets, which curve it shifts and the direction of the shift, and the effect on the real wage, unemployment, and the Gini coefficient.

Competition policy to reduce barriers to entry (b)

The competition condition immediately above includes the term b, which is a measure of barriers to entry in an industry and therefore reflects how uncompetitive an industry is. This can be a target of economic policy.

As we saw in Chapter 9 the Anti-Trust Division of the Department of Justice in the United States and Competition Commissions in other countries pursue policies to reduce barriers to entry in an industry by means of legislation or legal action. The effect, as the equation shows, is to raise the real wage that is consistent with the competition condition.

Here is how that works. As you know from Chapter 9 lower barriers to entry means that more firms will enter the economy, and the effect will be to reduce the price that will maximize profits for the firms' owners. The resulting reduction in the profit-maximizing markup ratio over cost means that the share of profits in total income falls, and the share of wages increases. This is why the competition condition would shift upwards.

At a higher real wage, the level of employment that is consistent with sustaining the incentive to work (shown by the wage curve) increases. The resulting decrease in the level of unemployment along with the shift in the labor share of income then has the effect of reducing the level of inequality in the economy, as shown by the Gini coefficient, as shown in Table 15.4.

Monetary policy to reduce the opportunity cost of capital (ρ)

Equation 15.7 also allows us to see the effect of a central bank's monetary policy. If the central bank (for example, the Federal Reserve Bank in the U.S.) wishes to promote higher levels of investment, it may reduce the borrowing costs of commercial banks and hence lower the interest rates at which other firms and individuals can borrow. This lowers the opportunity cost of capital ρ.

The decrease in ρ means that the wage that is consistent with the competition condition (expected profit rate equal to the opportunity cost of capital) is now higher. This shifts up the profit curve. And just as was the case with the upward shift in the competition condition resulting from a more competitive economy, this will allow a higher level of employment and higher wages.

The resulting decrease in the Gini coefficient occurs because the level of employment has increased. The share of wages in total income is unaffected, so unlike the case of increased competition, this is not among the reasons for the reduction in the Gini coefficient.

Research, education, and training to raise labor productivity (γ)

Consider a government policy that supports basic research resulting in improved technologies and an increase in the output produced by a worker per hour of work at the required effort level (γ).

Equation 15.7 shows that an increase in γ will increase the highest wage that the firm could pay consistent with its profit rate not falling below the opportunity cost of capital. Education and training that increased the productivity of an hour of labor would have the same effect. The result will be a shift upwards in the competition condition and a new Nash equilibrium with higher wages and employment.

Improvement in working conditions to reduce the disutility of effort ($\underline{u}$)

Employers may adopt policies to improve working conditions by providing:

- Material amenities such as a safer workplace, air conditioning, and flextime scheduling, or
- Social amenities such as a voice in company decision making and a respectful and fair minded approach to their workers.

Either will reduce the disutility of effort: working hard for a kind and respectful boss is a lot less unpleasant than working for an indifferent and insulting one. The material amenities make going to work more pleasant (or less unpleasant), whereas the social amenities may stimulate motives of reciprocity towards the employer and finding some pleasure in putting in a hard day's work for him.

As you can see from Equation 15.7 the reduction in the disutility of effort $\underline{u}$ will shift down the wage curve, shifting the Nash equilibrium to the right so that employment rises (wages remaining unchanged).

Checkpoint 15.2: Putting the models to work

Re-draw the two figures – the whole economy model and the associated Lorenz curve comparing the status quo and the effect of

- an increase in labor productivity

- an improvement in working conditions that reduces the disutility of effort

Increase in the unemployment benefit (B)

For simplicity, think about an increase in the unemployment benefit that is financed by government borrowing rather than through taxation as often occurs during downturns in the business cycle when governments extend the period over which the unemployed can receive benefits. Raising the unemployment benefit would improve the employee's reservation option. Return to the equation for the no-shirking wage (w^N), and think what effect this will have on the wage curve.

In order to motivate the employee to work, the firm now will have to pay more, which shifts up the wage curve. What then occurs?

- At the initial employment level H^* firms raise the wage, which is now above the profit curve which
- leads the firms to cut back employment or to close
- the process continues until the employment level has fallen to the new intersection of the now higher wage curve with the unaltered profit curve

The effect is to reduce employment, increase unemployment, and to leave the wage unchanged. You can see from Table 15.4 the effect taken in isolation, would be to raise the Gini coefficient. But there is a second effect: an increase the share of total income going to the unemployed. This occurs for three reasons:

- more people are unemployed
- each of them is receiving more; and
- because fewer are employed, less is being produced

The two effects – reduced employment (at an unchanged wage) and increase share of income going to the unemployed (the poorest group) – work in opposite directions. But the second effect – reducing inequality and the Gini coefficient – is always larger.

As a long term strategy for reducing inequality, however, raising the unemployment benefit is likely to be infeasible. The reason is that its effect is to reduce employment and output, and its funding by counter cyclical government borrowing (which we assumed for simplicity) would have to be replaced by a permanent tax revenue source.

15.12 Application: Trade unions, inequality, and economic performance

LABOR OR TRADE UNION A labor/trade union is an organization made up of employees who together bargain with one or more employers about wages and working conditions. The term trade union is also used, the word 'trade' meaning a particular kind of skill, such as "the building trades." A union may negotiate a contract for an entire industry, or for a single firm. This process is called collective bargaining. The contracts typically cover such things as hours, wages, and working conditions; they do not specify the amount of work to be done by employees.

In many industries, the contract under which employees work is not between individual workers and the employer, but instead between a labor union representing the workers and the employer. In this case the firm is not a wage setter as we described in Chapter 11, but instead engages with the union to bargain for the wage rate and other conditions of work. Countries and industries differ markedly in the fraction of all employed workers that are working under a collectively bargained contract. Figure **??** shows this statistic for twenty-nine countries ranging from France, Sweden, Denmark, and Belgium at the high end, to U.S., Japan, and U.K at the low end.

If the economy is not at the Nash equilibrium, n, in Figure 11.13, and the wage is below the competition condition line, a trade union may bargain employers to raise the wage. But Figure 11.13 shows that if trade unions succeed in bargaining a wage higher then w^C, then firms' profits will be insufficient to justify their hiring labor and they will either close down or relocate.

In the long run, the effect of labor unions on the labor market occurs because of effects on the Nash equilibrium, that is, by shifting either the wage curve or the competition condition.

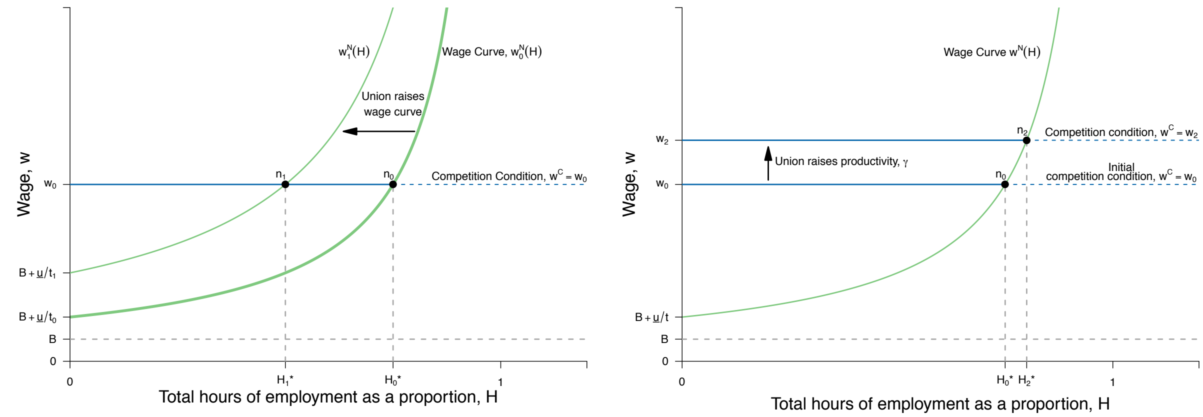
For example, based on Equation **??**, the union might reason that if it were more difficult for the firm to dismiss shirking workers, they would be forced to pay higher wages to motivate employees to actually work. Reducing t, the equation shows, will raise the "no shirking wage".

Equation **??** shows that this will raise the wage curve for all levels of employment. Figure 11.13 shows that making termination more difficult (lowering t) does not have the effect the trade union intended. On the contrary, making it harder for employers to fire workers would shift the Nash equilibrium to the left, reducing employment and leaving the wage unaffected. A more promising strategy by the union would be to bargain with the firm for better working conditions or an environment in which workers feel more respected and not subject to arbitrary treatment by supervisors.

This is called the **union voice effect**. Voice effects could reduce the disutility of working, which from Equation **??** would lower the wage curve. Why would *lowering* the wage curve be in the interest of workers? From the figure you can see that increasing voice (and decreasing the disutility of work) would not affect the wage but it would increase the level of employment.

THE UNION VOICE EFFECT occurs when a trade union, by providing a 'voice' to otherwise unheard employees, improves their treatment by employers and their job satisfaction (therefore decreasing the disutility of work).

Trade unions can also shift up the competition condition by providing a means of resolving conflicts about work rules and other practical aspects of the



(a) Negative effect of unions

(b) Positive effect of unions

Figure 15.17: **Effects of labor unions on employment and wages.** Left-hand side: The union decreases t from 1 (the shirker is sure to be dismissed) to $t_1 < 1$. From equation **??**, you can see that as a result, the wage curve shifts upwards and there is a new Nash equilibrium, n_1 with the same wage and lower employment. Right-hand side: When a union increases the productivity of workers (γ) the profit curve shifts upwards because now a profit rate equal to the opportunity cost of capital is achieved with a higher real wage. The result is a new Nash equilibrium, n_2, with higher wages (w_2) and higher total employment (H_2^*) than the previous Nash equilibrium n_0.

Figure 15.18: **Wage growth and unemployment in high income economies.** Source of data: US Bureau of Labor Statistics and OECD Labor Statistics.

production process between owners and employees in a cooperative manner. The effect is to increase the productivity of the workers' effort by reducing wasted effort, and this raises the wage that can be offered consistent with the competition condition. This outcome is illustrated in Figure 15.17.

In this case, as the figure shows, the effect on the union's negotiation has been to change the Nash equilibrium such that it raises both the level of employment and the wage rate.

The conclusion is that trade unions can affect the functioning of the labor market both positively and negatively. Where trade unions have pursued positive policies – the union voice effect and working with employers to raise productivity – countries in which a majority of employees work under collective bargaining agreements such as Norway, Germany, Denmark Finland and Sweden, have outperformed countries with weak trade unions, such as the U.S., Canada, and the U.K, as Figure 15.18 shows. The figure also shows, however, that some countries where labor union negotiated contracts cover a small fraction of the labor force – Japan is an illustration – perform well compared to countries with stronger unions – such as Spain and Belgium.

Checkpoint 15.3: Policy-making with labor markets

Imagine you are a policy-maker and could pick which country's labor market performance you would most like to emulate in your country. Rank the countries shown in the order of their performance, by your own values. What difficulties might you encounter in trying to emulate, say, Japan or Norway in your own country?

15.13 Capitalism as an economic and social order: Disparities in wealth and power

The Firm as a Political Institution

xxx In 1951 a paper by Herbert Simon pioneered the study of exchanges with incomplete contracting. Forty years later, he imagined a mythical visitor from Mars approaching earth in a spaceship:

> . . . equipped with a telescope that reveals social structures. The firms reveal themselves, say, as solid green areas . . . market transactions show as red lines connecting the firms forming a network in the spaces between them. . . .

What would the Martian see, mused Simon?

> No matter whether our visitor approached the United States or . . . urban China, or the European Community, the greater part of the space below it would be within the green areas, for almost all of the inhabitants would be employees, hence inside the firm boundaries. Organizations would be the dominant feature of the landscape.

The moral of the story, for Simon, is about the proper subject matter of economics:

> A message sent back home, describing the scene would speak of large green areas interconnected by red lines. It would not speak of a network of red lines connecting green spots. (Simon 1991:27)

xxxx Until recently power has been considered an alien concept by economists, something exercised by governments but not by private economic actors, at least not since the end of slavery and feudalism. These were both economic systems in which one party – the slave owner, the feudal lord – could threaten dire consequences to any of "their" slaves or serfs who did not do their bidding. But in a capitalist economy exchanges are voluntary, not coerced at gunpoint (or swordpoint), and parties to any exchange are free to walk away. That is why the participation constraint must be satisfied for any exchange to take place: as we saw in Chapter 5, this is because participation in an exchange is voluntary.

But even in a voluntary exchange among private parties, power can be exercised. We have seen that where contracts are incomplete exchange one of the actors, the principal, acting as a first-mover offers terms that induce the agent to do something in the principal's interest that could not secured by enforcing the terms of a contract. The principal does this not by threatening physical harm to the agent but by committing to terminate the interaction if the agent does not do what the principal asks. The harm threatened is economic: if the relationship ends the agent loses the enforcement rent that she received as part of the interaction.

We can begin to piece together an explanation, drawing on an economic

HISTORY Ronald Coase won the Nobel prize for his contributions to the economics of institutions, particularly the study of firms as political as well as economic organizations as well as the process of exchange when contracts are incomplete and what he termed the "transaction costs" associated with endogenous enforcement. Upon his death (at the age of 102) *Forbes* magazine called him "the greatest of the many great University of Chicago economists."

HISTORY Simon was a Nobel Laureate in economics though his undergraduate and PhD degrees were in political science. He was a pioneer in fields as diverse as artificial intelligence and organizational theory and is best known for stressing people's limited cognitive capacities and incomplete information when making decisions, what he termed "bounded rationality." He favored replacing taxes on wages and salaries by a tax the value of land.

model of the firm that we owe to an unlikely pair: Karl Marx the 19th century socialist revolutionary and Ronald Coase a leading 20th century economist whose work is often used in advocating lesser role for the government in addressing economic problems.

REMINDER So far, in Chapters 7 and 8, we have assumed that labor is simply a factor of production – or input – the amount and quality of which are what produces output and which the firm can simply purchase much as they purchase kilowatt hours of electric power. Here we replace that assumption with a more realistic model that takes account of the fact that work is done by people who, having their own ideas about how they would like to spend their day, need to be motivated by some combination of carrots or sticks to do the work on which the employer's profits depend.

Marx was the first to stress the fact that the employment contract did not cover such things as the *amount* or *quality* of work done. Rather, the employment contract specified the *hours* during which the employee agreed to submit to the *authority* of the employer. Under the employment contract, the employer does not purchase the employee's work, he rents the employee's time.

According to Marx the employee's actual supply of effort to the production process was not secured by contract but was rather an extraction that "only by misuse could . . . have been called any kind of exchange at all." Anticipating Ford, as well as late 20th century developments in economic theory, Marx pointed out that an increase in the wage might *reduce* the cost of labor per unit of output.

THE SUPPRESSION OF THE PRICE MECHANISM sounds like something that was done in the Soviet Union when the economy was centrally planned by government officials, a so-called *command* economy *without* market-determined prices. But Coase pointed out that the private owners of firms in a capitalist economy also suppress the price mechanism by not having prices *within* the firm determine which employee does what task: in a firm, workers do what managers or supervisors tell them to, not prices. Keep this curious fact in mind for later when we study the pros and cons of markets as a system of economic governance.

Like Marx, Coase (1910-2013) stressed the central role of authority in the firm's contractual relations: "note the character of the contract into which a factor enters that is employed within a firm. . . . [T]he factor . . . for certain remuneration agrees to obey the directions of the entrepreneur." Indeed, Coase (1937: 387,389) defined the firm by its *political* structure:

> If a workman moves from department Y to department X, he does not go because of a change in prices but because he is ordered to do so . . . the distinguishing mark of the firm is the suppression of the price mechanism. (387, 389)

Coase sought to understand why firms exist at all, and what determines the extent of what he called these "islands of conscious power in this ocean of unconscious cooperation." The size of the firm is determined by the decisions of its owners and managers when confronted with the question: should we purchase this input from another supplier or should we make it in-house. The more the firm produces in-house, the larger will the firm be, for a given level of final sales of the product. The reason why we have large firms according to Coase, is that for many inputs *the suppression of the price mechanism* within the firm in favor of a centralized system of control makes in-house production more cost effective than acquiring the same input on the market.

EXAMPLE For example, should a car manufacturer also manufacture tires for the car (produce them in-house) or should they purchase the tires that another manufacturer produces? This is true, too, of many aspects of the car, such as the seats, steering wheel, radio, etc.

A key idea in the chapter is this. Given the benefits of specialization and economies of scale, economic activity is necessarily social rather than individual, and the types of institutional arrangements governing production and exchange reflect the fact that the conflicts of interest among the participants are governed by incomplete contracts. The combined effect of incomplete contracts and conflicts of interest is that the determination of outcomes depends on who exercises power in the transaction. Power is generally exercised by those who hold what are called the residual rights of control, meaning the right to determine what is not specified contractually.

Control over assets and power over people

REMINDER For B to have power over A, it is sufficient that, by imposing or threatening to impose sanctions on A, B is capable of affecting A's actions in ways that further B's interests, while A lacks this capacity with respect to B. Where B has this capacity because B is on the short side of a not clearing market, we say that B has short side power (Bowles and Gintis, 1992).

What we have explained is why control rights over assets confers power over people. Samuelson's claim in the epigraph of this chapter asserts the contrary – it does not matter who hires whom. This is true in the Walrasian perfectly competitive model because in that framework the labor contract is assumed to be complete, so the notion of "hiring" simply means "buying." "What does it mean," Oliver Hart (1995) asked, "to put someone "in charge" of an action or decision if all actions can be specified in a contract?"

This basic point also explains why, in Marx's terms, contractual transactions on competitive markets appear to be a free exchange among equals ("a very Eden of the innate rights of man"), while in the workplace the two parties to the employment contract take on a different appearance: the employer is boss and the employee is "his laborer."

REMINDER Recall that in a market with incomplete contracts one side of the market will be the *short side* because it is the side on which the desired number of contracts is lowest, whereas the other side of the market is the *long side* because it is the side of the market on which the the desired number of contracts is greatest. For example, in employment markets, employers desire fewer contracts than workers or employees, and so employers have short-side power over the long-side workers.

Samuelson's and Marx's picture of the economy as a social system could not be more different. For Samuelson, the economy is a level playing field politically speaking, no actor has any power or authority over any other. For Marx the economy is also a political system in which power is exercised by those with wealth.

Expressed in modern economic terms, we update Marx to say that those on the short side of a market – employers in the labor market, lenders in the credit market for example – exercise short-side power over those on the long side of the market with whom they transact – employees and borrowers.

HISTORY Adam Smith foresaw this very problem in *The Wealth of Nations* (V. 1, 107) "The directors of such [joint stock] companies, however, being the managers rather of other people's money than of their own, it cannot well be expected that they should watch over it with the same anxious vigilance with which the partners in a private copartnery frequently watch over their own. Like the stewards of a rich man, they are apt to consider attention to small matters as not for their master's honour, and very easily give themselves a dispensation from having it. Negligence and profusion, therefore, must always prevail, more or less, in the management of the affairs of such a company."

The incomplete contracts model of the political and economic structure of a capitalist economy is summarized in Figure 15.19, which introduces the market for managers as a distinct kind of contingent renewal contract (following the structure of the same contract between employers and employees in Chapter 11).

The owners of a firm are the principals and the managers are the agents. The agent skillfully manages the owner's assets to maximize to owner's wealth, but the manager's action are a set of behaviours that cannot be subject to a complete contract. The other necessary ingredient of a principal-agent relationship – conflict of interest – is also present, because the manager has interests other than to maximize the owner's wealth. These interests include the manager's own leisure, frequent first class air travel and luxurious accommodation, and self-promotion activities that will improve his fallback position (that is, his prospects for employment in a different firm).

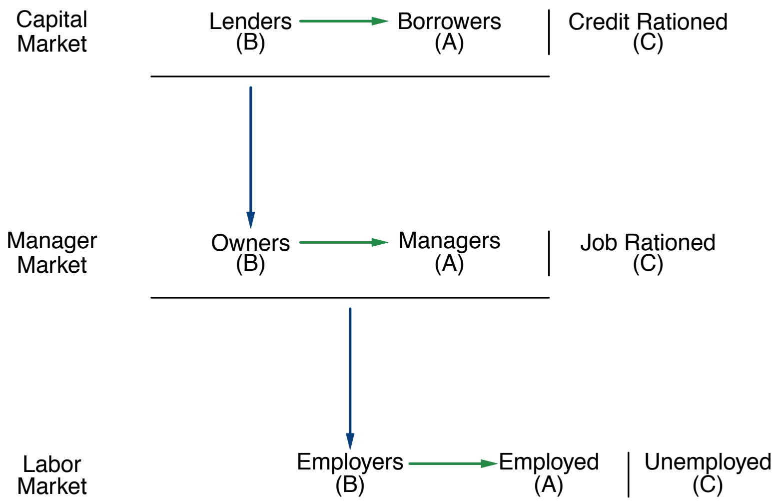


Figure 15.19: **The incomplete contracts model of the the economic and political structure of a capitalist economy.** The Bs are short-side principals exercising short-side power over the As (the long-side agents with whom they transact). The Cs are quantity-constrained long-siders: the unemployed, job-rationed or credit market excluded.

To reduce the conflict of interest, owners typically compensate managers not only with a salary but also with payments or stock options that will increase with the value of the firm.But unless the manager owns a very substantial fraction of the entire firm, the conflict of interest between principal and agent will remain an important aspect of the owner-manager relationship. The only way to eliminate the conflict of interest would be to make the manager the sole residual claimant on the value of the firm's assets, while paying to the owners some contractually fixed amount independent of the firm's value.

FACT CHECK This is often the model followed by franchising: managers of a franchise pay the owners of a brand a fixed amount for use of the brand, but are residual claimaints on the profits of their franchise.

In Figure 15.19, the short-siders (B) exercise short-side power over the long-siders with whom they transact (A) while the excluded long-siders (C) are quantity-constrained.

Notice in Figure 15.19 that the differences among the economic actors are both horizontal and vertical. The horizontal differences (indicated by the vertical lines) are among people who could be identical, the long-siders in the figure among whom the A's made the transaction they desired and the C's failed to do so.

The vertical dimension (indicated by the horizontal arrows) is between the those who have short-side power (the principals) and those over whom this power is exercised (the agents). That is, between the lenders and borrowers, the owners and the managers, and the managers and the workers.

Notice also that the people appearing as lenders and successful borrowers in the capital market appear as owners in the manager market, while owners and those who succeed in securing a position in the manager market appear as employers in the labor market. The political dimension of the economy is depicted in Figure 15.19 as a downward cascade of short-side power beginning with wealthy lenders who exercise power over borrowers wealthy enough to secure a transaction. The wealthy and the successful borrowers, then exercise power over managers (those who secured employment), who in turn, along with owners, exercise short-side power over employees.

The idea that the political structure of markets and firms should play a central role in the analysis of the economy, a recent development in economics, is in fact far from new. The U.S. Supreme Court stated it quite clearly in the headquote to this chapter.

CONTEXT Responding to Marx, Joseph Schumpeter sought to dispel the idea that the relationship of employer to employee was political in any sense or that the employer had power over the employee: "What distinguishes directing and directed labor appears at first sight to be very fundamental," he wrote, but in reality the difference "constitutes no essential economic distinction ... the conduct of the former is subject to the same rules as that of the latter ... and to establish this regularity ... is a fundamental task of economic theory" (Schumpeter, 1934: 20-21).

But many economists have considered the exercise of power by employers to be illusory. An important mid-20th century microeconomics text (Alchian and Allen, 1969: 320) may have surprised some students with the following:

> Calling the employer the boss is a custom derived from the fact that the "boss" specifies the particular task. One could have called the employee the boss because he orders the employer to pay him a specific sum if he wants services performed. But words are words.

We doubt very much that the authors would disagree with the Supreme Court's assessment as an empirical account. Like Samuelson (in the head quote), they were describing the logic of a model, not an empirical aspect of the economy.

In the approach modeled here, Alchian and Allen's example would look quite different. The employer would simply refuse any pay demanded by the worker unless it happened to be w^*, namely the wage that maximized the employer's profits.

HISTORY The seventeenth century philosopher, Thomas Hobbes wrote "... Riches joyned with liberality [generosity] is Power; because it procureth friends, and servants ..." (1651/1968:150.) In Hobbes' day the terms "servant" referred to any employee.

We have therefore come to three important conclusions:

- First, differences in wealth will be reflected in differences in the feasible set of contracts and the individual's choice of contracts from the feasible set.
- Second, some of these contractual arrangements will include a structure of authority such that participants on one side of the transaction have *power* over the others, even when the setting is competitive in the standard sense that costs of entry and exit are small. So, the fact that participation is voluntary does not prevent power from being exercised.
- Third, those exercising power will typically be those with greater wealth, so that the political-economic structure of a capitalist economy is one in which both wealth and power are concentrated at the top, and the exercise of power – commands that will in all probability be obeyed – cascades downward through this structure.

The sociology and psychology of short-side power

The model of power developed here provides a reason to doubt the old adage: "The wealthy are different from everybody else; they have more money." Wealth does indeed determine the position of one's budget constraint and wealth commands more goods and services. Substantial wealth gives a person a large feasible set – for say consumption, free time, and other valued things. Having a wider range of choice because of an enlarged feasible set, we can say that wealthy people have more freedom.

But those wealthy enough to engage in their own projects or to borrow large amounts at the going rate of interest enjoy more than superior purchasing power. They may command people as well as goods. Their access to capital allows them, but not others, to become employers of labor and as such to occupy positions of short-side power in non-clearing markets.

Those without wealth tend to be constrained not only by a more limited feasible set of consumption choices, but also by the fact that as employees or borrowers they are subject to the exercise of short-side power by others.

These disparities in power show up in many non-economic realms of our lives. The employment relationship persists over many years and the workplace is a cultural environment in which employees' and employers' preferences and beliefs evolve. Workplaces are no different from schools or neighborhoods, for they influence how we develop as human beings, and how we raise our children.

An empirical example will suggest the importance of these effects. Over a period of three decades the social psychologist Melvin Kohn and his collaborators have studied the relationship between one's position in the authority structure of one's workplace – giving as opposed to taking orders – and the individual's valuation of self-direction and independence in one's children, as well as one's own intellectual flexibility, and personal self-directedness. They concluded that "... the experience of occupational self-direction has a profound effect on people's values, orientation, and cognitive functioning."

His collaborative study of Japan, the U.S. and Poland (when it was still under Communist Party rule) yielded cross-culturally consistent findings: people who exercise self-direction on the job also value self-direction more in other realms of their life (including child-rearing and leisure activities) and are less likely to exhibit fatalism, distrust, and self-deprecation. Kohn and his co-authors reason that "... social structure affects individual psychological functioning mainly by affecting the conditions of people's own lives." Kohn concludes that:

> The simple explanation that accounts for virtually all that is known about the effects of job on personality ... is that the processes are direct: learning from the job and extending those lessons to off-the-job realities. (1990a: 59)

(a) Ana chooses to work for a wage

(b) Beata chooses to be self-employed

Figure 15.20: **The risk and income choices of two owner-operators.** Ana is very risk-averse and has a choice between point **a** on indifference curve u_1^A and a certain wage w^c on indifference curve u_2^A with $u_2^A > u_1^A$. Her utility is higher at the wage than at the combination of expected income and risk $(\hat{y}^A, \Delta^A)$, so she would prefer to be employed – the less risky option – rather than to be an owner-operator. Beata is less risk-averse and has a choice between point **b** on indifference curve u_2^B and a certain wage w^c on indifference curve u_1^B with $u_2^B > u_1^B$. Her utility is higher when she is an owner-operator with expected income and risk $(\hat{y}^B, \Delta^B)$ than at the certain wage, so she would prefer to hold the risky asset and be an owner-operator to being a wage worker.

As the personality dimensions mentioned by Kohn are part of individuals' preferences explaining how they raise their children, what kind of leisure activities they engage in and the like, this is strong evidence for the effects of workplace organization on preferences that we pointed to in Chapter 10.

15.14 Would a wealth-poor person want to hold a risky asset?

Would the employee, who could now receive w^c with certainty prefer instead to be the owner of the capital goods?

An employee who owned the capital goods would change the picture in the following way:

- The employee would now become the *residual claimant* on the income (y) resulting from the project.
- She would also have *control rights* about how the project was conducted, that is, she would choose the risk level.
- She would bear all of the *risk* of the project.
- As owner of the capital goods she would have to consider the opportunity cost of the capital goods (ρ) that she uses.

If the former employee has limited wealth the opportunity cost of capital for her may be substantially more than it was for the previous wealthy owner, namely ρ. To own the capital goods she might, for example, have to borrow at an interest rate $\bar{\rho} > \rho$.

For simplicity we assume that the amount of work she does on the project is the same as an employee or as an owner-operator. As we are considering projects in which capital goods are owned by the worker, who therefore owns the income resulting from her work, there is no need for monitoring.

As an owner-operator the following would be true:

- Her expected income is therefore the expected profit from the project minus the opportunity cost of capital or $\hat{y} = \hat{y}(\Delta) - \bar{\rho}K$ and
- Her utility would be $u(\hat{y}, \Delta) = u((\hat{y}(\Delta) - \bar{\rho}K), \Delta)$.

Her risk-return schedule is shown in Figure 15.20.

We want to know how her utility as an employee – namely $u(w^c, 0)$ – compares to her utility as an owner operator, who selects risk level Δ, that is $u((\hat{y}(\Delta) - \bar{\rho}K), \Delta)$.

To answer the question we need to proceed in two steps:

1. from the tangency of her indifference curves and the risk-return schedule, we determine which level of risk she would take under each circumstance and then

2. we ask is she better off with the resulting expected income and risk level than she would be with a certain wage equal to w^c

We do this by looking at Figure 15.20.

Figure 15.20 presents indifference curves for each of two possible owner-operators, who we shall call Ana (A) and Beata (B). Because her indifference curves are steep, we know that Ana is very risk-averse while Beata is only modestly risk-averse (so her indifference curves are flatter).

Given the risk-return schedule $\hat{y}(\Delta)$, they will choose points **a** and **b**, with risk levels Δ^A and Δ^B and corresponding expected incomes $\hat{y}^A$ and $\hat{y}^B$. Risk-averse Ana will choose less risk and as a result have a lower expected income than will less risk-averse Beata ($\hat{y}^A < \hat{y}^B$).

To compare how these workers would evaluate the prospect of being a wage worker rather than an owner-operator, we use the fact that the wage in the conventional “firm" set up is received with certainty (all of the risk is borne by the owner). We can therefore compare the certainty equivalent of points **a** and **b** (chosen by the two owner-operators) with the utility of the certain wage that each would receive as an employee.

We see that risk-averse Ana would prefer to be a wage worker: the certain wage (w^c) as an employee is higher than the certainty equivalent of the best she could do as owner-operator, namely choosing point **a** with certainty equivalent $\bar{w}^A$. You can see this in Figure 15.20, which shows that Ana's indifference curve going through w^c, where she works for a wage, provides her a higher utility (on indifference curve u_2^A) than if she were an owner-operator using capital goods at point **a** (on indifference curve u_1^A). Beata, though, is on a higher indifference curve (u_2^B) at point **b** as an owner-operator than she would be as a wage worker receiving w^c (on indifference curve u_1^B).

Less risk-averse Beata would therefore prefer to be the owner of the capital goods. So, if she were initially a wage worker and could borrow funds to purchase the capital goods at rate $\bar{\rho}$, she would do so.

But if the prospective owner-operators lack wealth and as a result the rate at which they could borrow $\bar{\rho}$ is sufficiently great, or if they were entirely excluded from credit markets, then most potential owner-operators would end up as wage workers. Or if for some reason they owned the capital good they would prefer to sell it and become an employee. This is true even though this form of production – the conventional firm – incurs costs of monitoring that are unnecessary (or at least reduced) when workers are also owners. (Because there is only a single worker owner in our example, no monitoring would be necessary – she is the residual claimant on her own work effort.)

This provides our answer to the question we asked at the beginning of this section: if the prospective owner-operators had the option to borrow the funds to buy the capital good would they buy it and become owner-operators or would they rather sell the asset and become wage workers?

15.15 Risk, redistribution and innovation

This suggests a new question. Suppose the ownership of the capital goods were simply transferred to the prospective owner operators, so that they were as a result much richer than before: would this wealth redistribution make it attractive to retain ownership rather than selling?

The fact that she now owns the capital goods does not mean that the opportunity cost of capital is irrelevant to her (she could, for example, sell the capital good, and pay off some of the debt she has incurred). Let us assume that if the worker owned the capital goods she would be wealthy enough to borrow at the same rate that was available to the rich person, namely ρ. This is her opportunity cost of capital.

The fact that she is wealthier thus has two effects, both shown in Figure 15.21:

- *Shift of the risk-return schedule:* The risk-return schedule is net of the opportunity cost of capital (ρ); ρ has now decreased so the risk-return schedule shifts upwards.
- *Greater wealth and lower risk aversion*: Because the worker is now richer, she is also less risk averse: for any given level of Δ, her indifference curves are flatter. So as an owner-operator she will be willing to take more risk, and therefore will be able to enjoy a higher level of expected income.

In Figure 15.21 the status quo before the ownership of the capital good was transferred to Ana, is shown as point a^0. Recall that she would have chosen as her risk level Δ^{A0} as an owner-operator, enjoying a certainty equivalent utility of w^{A0}. As a result, she preferred to be a wage worker receiving w^c rather than to purchase and own the capital goods.

But due to her reduced risk aversion and reduced opportunity cost of capital, the now much wealthier Ana chooses point a^1, meaning a risk level of Δ^{A1}. At this risk level and the resulting expected income she is better off than being a

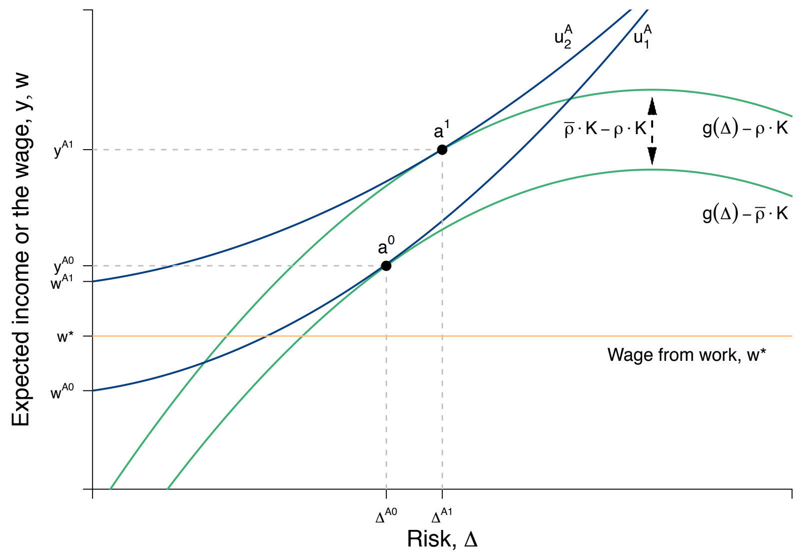


Figure 15.21: **Effects of redistribution of wealth to one owner-operator.** A redistribution of wealth makes ownership with risk exposure preferable to employment at a certain wage, w^c. The figure shows effects of the greater wealth of the owner-operator. First is the effect of the reduction in the opportunity cost of capital. Reducing ρ from $\bar{\rho}$ to ρ^* shifts up the green risk-return schedule. Second, because Ana the owner-operator is richer she is less risk-averse (flatter blue indifference curve). Both the shift in the risk-return schedule and the change in the slope of the indifference curves make ownership more attractive than wage employment, and the certainty equivalent of her chosen risk level, and the resulting expected income, exceeds the wage.

wage worker. You can see this because the certainty equivalent of point a^1 is greater than the (certain) wage she would receive as a wage worker.

What this means is that a redistribution of wealth (by a government for example) that made Ana wealthier would result in her remaining an owner-operator. She would have no incentive to sell the capital good and become a wage worker.

The hypothetical redistribution of assets is a vehicle for exploring the interaction of credit constraints, risk aversion and ownership. It is *not* a policy design. Design of actual policies of asset distribution would need to address both the policy's administrative aspects as well as general equilibrium and long-term dynamic effects not considered here. For example, whether the once-poor would adopt savings and investment strategies which would preserve, enhance, or consume their assets would need to be considered.

Also, we have focused on the relationship between a single owner and a single worker. But, due to economies of scale, a firm will typically employ a large number of workers. These workers could not become owners of their own personal firms with similar economies of scale because small owner-operated firms are much smaller than large firms that are able to exploit economies of scale.

Or, as a team of worker-owners, they would have to find a way to motivate each team worker to work hard and well, rather than simply free riding on the efforts of fellow team members.

HISTORY In 1944, when the welfare state was in its infancy, Richard Musgrave and Evsy Domar showed that by redistributing income from the lucky to the unlucky, progressive taxation and transfer policies would reduce risk exposure and thus promote greater risk taking.

15.16 Conclusion

With few exceptions (Cuba, North Korea) and many variants (U.S., Germany, China) capitalism is the economic system of the world today. Knowing how the institutions of capitalism work – promoting innovation, sustaining substantial economic inequalities, transforming our biosphere, and its other consequences – is an essential starting point for modifying the rules of the game so that the economy better serves the needs and interests of all. In our final chapter we will use the understanding you have gained to explain how well designed policies can promote this objective.

Making connections

Gains from exchange and conflicts over their distribution The institutions of the capitalist economy both facilitate the exploitation of mutual gains based on specialization and exchange, and influence how these gains are shared among the people making up the economy.

Institutions of capitalism Models of behavior under risk and models of the labor, credit, and other markets with incomplete contracts provide frameworks for understanding both the dynamism of capitalism and its characteristic forms of economic inequality.

Risk and risk aversion The wealthy and not-very-risk-averse owners of the capitalist firm make it an effective risk taking 'innovation machine.' By reducing risk exposure, insurance can promote risk taking among the less wealthy.

Distribution The concentrated wealth and economic power that facilitates innovation also implies substantial inequalities in market incomes among participants in a capitalist economy.

Efficiency The success of capitalism in raising living standards (the hockey stick) is explained by the way that this economic system (when combined with the rule of law) promotes innovation in the long run, not by its success in implementing Pareto-efficient outcomes in any given period.

Incomplete contracts, social preferences power The incomplete contracts that give rise to the principal agent relationships characteristic of the capitalist economy – especially the employment of managers and workers – mean that both social norms and the exercise of power are important in determining economic outcomes, providing a political and social dimension to this economic system.

Important Ideas

cooperative	ownership	public good
incomplete contracts	enforceability	control rights
coordination failure	specialization	economies of scale
short side power	hockey stick	insurance
incomplete markets	credit/labor market excluded	credit-rationed
home ownership and tenancy	efficient design	allocative efficiency
mechanism design	optimal contract	team production
monitoring	risk	risk-sharing
risk-return schedule	expected income	risk-return indifference curve
risk averse	risk neutral	inequality
wealth redistribution	innovation	short-side/long-side
endogenous preferences	residual claimancy	Gini coefficient
profit curve	zero-profit condition	wage curve
trade union	public policy	unemployment benefit
wage share	profit share	no-shirking condition
collective bargaining	sanctions	peer monitoring

Mathematical Notation

Notation	Interpretation
K	level of capital
p_k	price per unit of capital good
a_k	units of capital good required per unit of output
X	units of output of a project
y	income
$\hat{y}$	expected income
$\bar{y}$	average income
Δ	difference in income
$\overline{\Delta}$	amount of risk that maximizes the expected income
ρ	opportunity cost of capital
$\bar{\rho}$	interest rate that a poor person face
$\hat{\pi}$	expected profit
$\hat{\pi}^E$	expected economic profit
$\hat{\pi^A}$	expected accounting profit of an entering firm
π^A	accounting rate of profit
π^E	economic rate of profit
m	resources devoted to monitor the workers
w	wage
w^n	equilibrium nominal wage consistent with the competition condition
w^c	real wage consistent with the competition condition
w^N	no-shirking wage
n	number of people
G	Gini coefficient
b	probability o failure of an entering firm
p	price of a good
a_L	amount of labor required per unit of output
γ	productivity of labor
H	level of employment
H^N	Nash equilibrium level of employment
σ	wage share
σ_n	wage share
σ_u	income share of the unemployed
n	proportion of the population that is employed
u	proportion of the population that is unemployed
B	unemployment benefit
t	probability of being fired if shirking
j	probability that a "terminated" worker will not find a job
$\underline{u}$	disuyility of effort

Discussion Questions

See supplementary materials.

Problems

See supplementary materials.

Works cited

16

Public policy and mechanism design

Mechanism-design theory aims to give the invisible hand a helping hand.

The Economist explaining the Nobel Prizes in economics in 2007.

Seat belts in cars are now standard equipment. But they are a recent addition to the safety features of an automobile. The Australian state of Victoria was the first to introduce a mandatory seat belt law, in 1971. Over the next two decades, spurred by claims that seat belts would save thousands of lives per year (even "10,000 to 20,000" in the U.S. according to advocates), over 80 jurisdictions implemented similar laws applying to the vast majority of automobiles in the world.

They claimed life-saving effects of seat belts were based on simulated crashes in which – to take one example –"for belted occupants the deaths were reduced by 77 percent in full frontal crashes and 91 percent in rollovers." In the U.K., the Royal Society for the Prevention of Accidents summarized the evidence: "no other single practical piece of legislation could achieve such dramatic savings of lives and serious injuries."

The truly global spread of mandatory seat belt laws seemed like a case of evidence-based public policy at its best. But was it?

In Figure 16.1, we can look at the evidence on road fatalities in 17 countries during the 1970s, 13 of which passed mandatory seat belt legislation and the other four of which did not. At the time the legislation was passed, together these 17 countries accounted for 80 percent of the world's cars. In almost all of the countries road deaths fell, in part due to the substantial increase in the price of gas, and the reduction in both legally permitted and actual speeds of driving. But the drop in fatalities was much greater in the countries that had *not* passed the seat belt laws.

How could this have occurred?

DOING ECONOMICS

By the end of this chapter you will be able to do the following:

- Explain how mechanism design develops new rules of the game that will support Nash equilibria that are improvements over the status quo.
- See this process as an inversion of standard economic practice, reverse engineering the set of institutions and other mechanisms that will achieve some desired social objective rather than the more standard approach: predicting outcomes based some given rules of the game.
- Explain how a mechanism designer accomplishes this task in the case of public goods provision and other societal objectives by internalizing uncompensated external effects and other means.
- Recognize that the kinds of preferences people have – other regarding and self regarding –matter and be able to explain pitfalls of policies based on the assumption that citizens all resemble *Homo economicus*.
- Understand the limits of mechanism design and why economics and policy makers cannot escape from the second best world of imperfect solutions to societal problems.
- See that the mechanism designer today is carrying on the tradition of the 18th century philosopher-economists who sought to develop institutions that would allow societal coordination such that the voluntary actions of free people would result in socially desirable outcomes.
- Write blog post or an op ed (opinion piece for the editorial pages) for your university or local newspaper or do a Ted Talk using what you have learned to explain the pros or cons of some economic policy idea about which you are passionate.

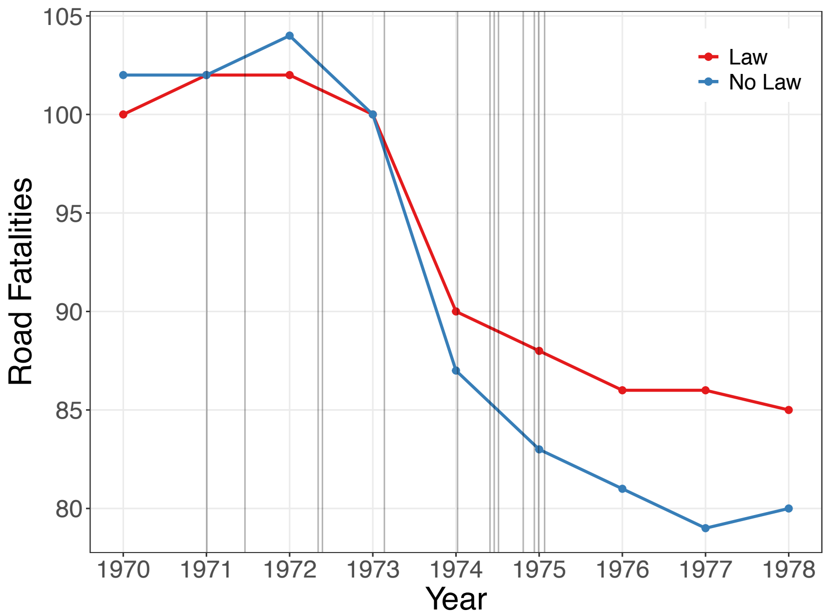


Figure 16.1: **Seat belts and road accident deaths.** Shown are indices of the number of road traffic deaths in 17 countries, with a value of 100 for 1973 in each case. The average of the indices for the 13 countries that passed seat belt laws are shown by the red line, the vertical lines representing the dates at which each of these countries introduced (and enforced) a seat belt law. The blue line represents the average indices for the 4 countries which did not impose mandatory use of seat belts. The decline in road deaths in both sets of countries is due in part to the substantial increase in the price of gasoline in 1973 (the "first oil price shock"), and the imposition of lower speed limits in many countries.

The answer, it seems, is that people drive faster when they are wearing seat belts. The result was that accidents are more frequent, and while the occupants of cars are more likely to survive an accident if one occurs, the greater frequency of accidents results in greater fatalities including among pedestrians and non-occupants.

In some cases the compensating effect of the seat belts – driving faster – was even enacted into law. In Germany buses with seat belts were allowed a top speed of 100km/h (60 m/h) while those without belts were restricted to 80km/h. Curiously the Royal Society for the Prevention of Accidents whose advocacy of the seat belt laws in the U.K was quoted above seemed to endorse this not as a bug but as a feature of the safety devices that "allowed [buses]to travel faster ...thus allowing drivers to cover more miles in the hours they are allowed."

An experiment provided evidence consistent with the increased driving speed explanation. In the Netherlands, before belt use was mandatory, a group of people who had never used seat belts participated in an experiment. The drivers were randomly selected to either wear seat belts or not. They were told that the purpose was to judge the comfort of seat belts, but in fact the experimenter measured the speeds at which they covered a specific 105 km course. Drivers wearing seat belts drove faster.

MECHANISM A mechanism (in the field of mechanism design) is a set of property rights (including control and residual claimancy rights), decision making protocols, market structure or other rules of the game.

16.1 Mechanism design: Policy implementation by Nash equilibrium

We do not conclude that seat belt laws should be abolished: mandatory seat belts and enforced moderate speed limits will surely reduce road fatalities. But the surprising results of the seat belt introduction teach an important lesson: effective policy design should take account of the diverse *unintended effects* of the policy (in this case, including greater speed), not simply the intended effects (in this case fewer fatalities among car occupants in a crash).

Adam Smith wrote (quoted in the introduction of Part I of this book): that the policy maker

> . . . enamored of his own ideal plan of government, . . . seems to imagine that he can arrange the different members of a great society with as much ease as the hand arranges the different pieces upon a chess-board . . . but . . . in the great chess-board of human society, each single piece [on the chess board of society] has a principle of motion of its own, altogether different from that which the legislature might choose to impress upon it.

The unexpected consequences of seat belt laws and Adam Smith's warning illustrate some challenges in designing public policy and also provide guidelines for effective policy interventions, the subject of this final chapter.

We can group government activities by their intended purposes:

- *Providing the economic framework* that governs how people and economic organizations interact, including the judicial system that interprets and enforces contracts and property rights, policies regulating how firms and other economic entities compete and the conditions under which people are employed, and reducing uncertainty about the level of aggregate demand and the value of the currency.
- *Addressing market failures* including through the provision or subsidization of public goods (including schooling and basic research), competition policies, and internalizing the harmful external effects of economic activity on the biosphere.
- *Addressing unfairness* in the distribution of income, wealth, or some other valued aspect of our living standards that arises due to the working of the economic framework adopted, including though taxes and transfers, direct government provision of some services (schooling, fire and police protection) and setting prices (rent control, minimum wages).

REMINDER Recall that if one party can set both the price and the quantity transacted, then this person – with take it or leave it power – is constrained by the other person's participation constraint, so the resulting allocation must be Pareto-efficient.

Thinking back to the allocation of goods between Ayanda and Bongani in Chapter 10, the above "providing the economic framework" would include determining why Ayanda—not Bongani—was the first mover, why Bongani could not simply take Ayanda's goods by force, and whether when Ayanda set a high price for her good Bongani had any other ways of acquiring the good (from a competing supplier, for example). In that example, "addressing market failure" could take the form of making it easier for Ayanda to impose both a price and a quantity on Bongani, giving her 'take it or leave it power' rather than just price setting power. If Ayanda's advantages were deemed to be unfair, then "addressing unfairness" could take the form of the government imposing a maximum price that Ayanda could charge, or possibly redistributing some of Ayanda's endowment to Bongani.

Reverse engineering good outcomes

Mechanism design is a branch of economics that seeks to design policies, decision-making protocols (like majority rule), laws, property rights, and other so called mechanisms that will implement outcomes that are judged to be desirable.

The primary focus of the field has been to provide novel rules of the game that will result in Pareto efficient Nash equilibria in cases where market exchange or bargaining among private parties fails to accomplish this result.

In previous chapters we have followed the common practice in economics, namely, start with the rules of the game for some economic interaction and a description of people's objectives and knowledge and then figure out what allocation will result. In other words, take the following steps:

- Start with a set of rules of the game or some other description of the institutional setting in which an interaction will take place, for example, the principal-agent game describing an employer and worker interacting about the wage and level of effort performed by the worker.
- Describe people's objectives (by their utility functions), constraints (by the set of feasible actions or strategies open to them), and beliefs.
- Use concepts like best response and Nash equilibrium to determine the outcome that will result from the interaction.

Mechanism design reverses these steps, starting with a desired outcome and then working back to see what rules of the game would bring that outcome about, given people's preferences and beliefs. So the steps above are taken up in reverse order:

- Start with an outcome that the mechanism designer (or the voters who elected her) would like to implement.
- Then devise a set of rules of the game that will lead individuals pursuing their own private objectives to implement the desired outcome as a Nash equilibrium.

The logic of mechanism design is simple: if the status quo is a market failure, or some other coordination failure, there must be some other allocation that

EXAMPLE: A MECHANISM FOR THE FAIR DIVISION OF A CAKE Suppose two people are going to share a small cake. What would be a good rule for making sure that the cake is divided fairly? A mechanism that is widely used is to let one person cut the cake and the other person choose which piece to take. Because the person cutting the cake will get the second piece, that is, the piece not chosen by the other, she will have an incentive to try to make both pieces be equally desirable.

EXAMPLE: KING SOLOMON AND BUDDHA AS MECHANISM DESIGNERS King Solomon in the Old Testament (the Hebrew Bible) was asked to determine which of two women should be awarded a young boy, who both claimed was their child. He asked for a sword and said he would award each women half of the infant. The first woman who spoke agreed with the proposal. But the second, in tears, objected and told Solomon to stop, and to give the child to the first woman. Solomon then knew that the first woman, who agreed to dividing the child, could not possibly be the true mother, and awarded the child to the second woman. In a an almost identical story from India, Buddha makes a similar ruling (the two women would have a tug of war over the child). Solomon and Buddha are wise mechanism designers in these stories, implementing a good outcome (the child goes to the true mother) by devising a way to get the information necessary to do this (who is the true mother).

is Pareto-superior to the status quo. So the mechanism designer then gets to work. She specifies the characteristics of an alternative feasible and superior allocation, for example that it must be Pareto efficient and meet some standards of fairness.

Her job is then to reverse engineer that desired allocation, finding a set of rules of the game or other mechanisms under which the desired allocation will be a Nash equilibrium, and therefore could be implemented by introducing the mechanisms she has discovered. The creation of a new superior equilibrium by a change in the rules of the game is called implementation by Nash equilibrium.

The policy maker who practices implementation by Nash equilibrium is respecting Adam Smith's dictum that the government cannot simply order people how to act (like moving chess pieces around on the chess board). Instead the government can alter economic outcomes by changing the circumstances under which people themselves decide what to do.

You have studied another case of implementation by Nash equilibrium in Chapter 5 where the problem faced by the policy maker was to design a tax that would deter over-harvesting fish from a lake. As in the prisoners' dilemma game, the fundamental problem was that actors did not take account of the effect of their decisions on others. In the case of fishing, this so called external effect occurred because each person fishing more meant that the other caught fewer fish. The policy that addressed this problem was the obligation to pay a tax equal to the costs that each of their fishing time imposed on the other.

This is termed "internalizing the external effects" of an individual's actions, and it is the key idea of mechanism design. Internalize means to make acting to produce a good outcome for society something that the individual wants to do. This is accomplished not by changing people's preferences, but instead by devising rules so that each person pays or is rewarded for the social costs or benefits to others resulting from his actions, that is not only the private costs and benefits, but also the external costs and benefits.

REMINDER You have seen this idea in action right from Chapter 1 where you learned how public policy could convert a prisoners dilemma game, with its undesirable mutual defect outcome, to an assurance game with a Pareto-superior Nash equilibrium. The 'mechanism' was a liability law that required the player to pay for the external costs that she imposed on another person, in this case the costs imposed on a cooperator by defecting on them.

Mechanism design has come into prominence in recent decades for two reasons.

- First, economists now recognize that the simple institutions underlying Adam Smith's invisible hand idea — perfect competition and well defined private property rights in anything that matters — represent ideals that are hardly ever realized in existing economies. The invisible hand and its associated laissez faire policies, as *The Economist* pointed out, need "a helping hand." Where markets fail, mechanism design provides this helping hand in suggesting more complex institutions including auctioning a limited number of permits for carbon emissions to address the challenge of climate change and designing kidney exchanges to match organ donors and those needing a kidney replacement.

- Second, with the growth of government's role in the economy the unintended consequences of policy interventions – often ignored by policy advocates – have become increasingly evident. The error in these policy designs is to ignore a fundamental precept of mechanism design. Policymakers must make sure that the desired outcome of some policy will be a Nash equilibrium once the policy is introduced. If this is not the case, the intended effects of the policy will be undone by the actions of private actors.

FACT CHECK San Francisco's rent control law passed in 1995 is estimated to have reduced citywide housing supply by 15 percent and raised average rents by 5 percent. This is definitely an unintended consequence, but it does not mean that the policy was mistaken. The gains by the mostly not very well off occupants of the rent-controlled apartments were substantial, and easily could be judged as more important than the losses to other renters and landlords. We illustrate this case in Figure 16.11

Checkpoint 16.1: Adam Smith: Mechanism designer

Explain how King Solomon (in the margin note above) or Adam Smith (in the quote above) fits the description of the mechanism designer.

16.2 Optimal contracts: internalizing external effects of public goods

REMINDER You know from Chapters 2 and 5 that a public good is one that is both *non-excludable* and *non-rival*, meaning that it is impossible or very costly to exclude anyone from access to a public good, and the consumption of the good by one citizen does not reduce the amount available to others. Return to Table 5.1 if you need to refresh your memory about what is distinctive about public goods.

Designing a policy to address the problem of private under-provision of public goods illustrates the method. Examples of public goods include the knowledge generated by basic research, weather reports or other information broadcast on an open access platform, and public safety.

Private provision of public goods: A market failure

In the absence of a subsidy or other public policy, private economic actors will typically not provide public goods at all, or will provide them in insufficient quantity. Under-provision means that there is some greater level of provision of the public good – for example with each person contributing more – such that all citizens would be better off.

The reason is that public goods are costly to provide, but the benefits that result from any individual's contribution to the public good are shared by everyone:

- There are private costs and private benefits that a decision-maker will take account of in choosing to contribute to a public good.
- But there are also external benefits, so the social (total) benefits exceed the private benefits.
- And unless contributions are subsidized, the individual is not compensated for the external benefits that her contributions create.

Here is a specific example of a public good and how it benefits a particular citizen, call her Bridget, who is one of n identical citizens. The amount of the

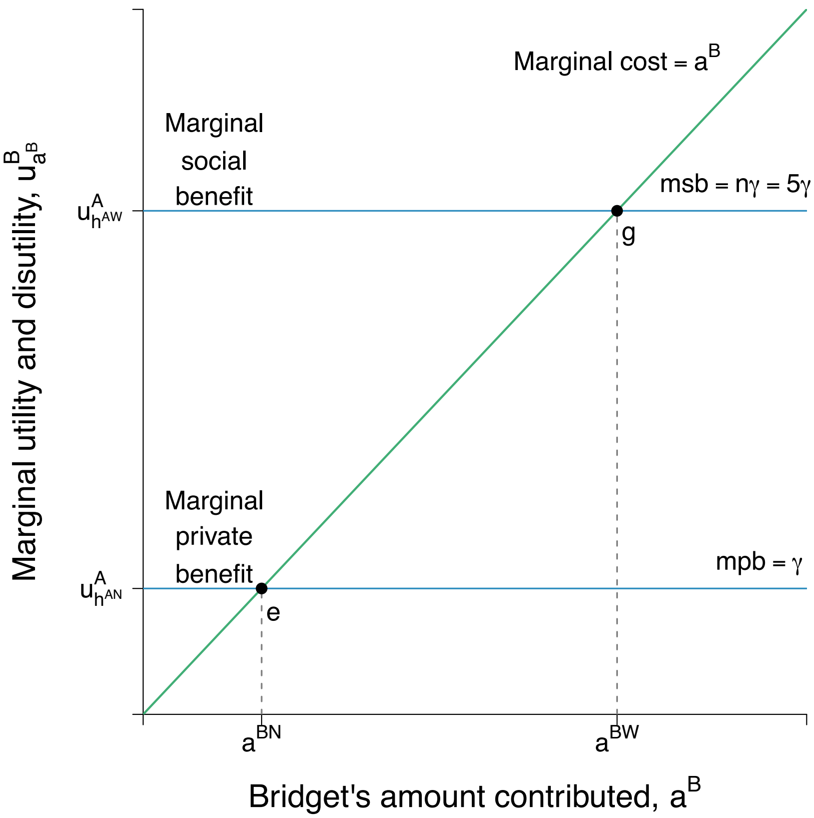


Figure 16.2: **The marginal benefits and marginal costs of contributing to a public good**. Bridget will contribute up to such point as her marginal costs, measured by a^B, equal her marginal private benefits ($mpb = \gamma$). Therefore, her Nash equilibrium contribution is a^{BN}. If there are four other citizens, then the socially optimal outcome occurs where Bridget's marginal private cost a^{BW} equals the marginal social benefit $msb = n\gamma$, which includes the benefit that Bridget confers on the 4 other citizens as a consequence of the positive external effect of her contributing to the public good. The Nash equilibrium contribution is lower than the socially optimal contribution.

public good available to Bridget and every other citizen is equal to a positive constant (γ) multiplied by the sum of all contributions or

$$B = \gamma(a^1 + a^2 + a^B + \ldots + a^n) \quad (16.1)$$

From this equation you can distinguish among these three essential concepts about the benefits of the public good.

- *Marginal private benefits* of contributing: γ. If Bridget contributes more, then the amount of the public good will increase by an amount γ times the increase in her contribution. This is her private marginal benefit.
- *Marginal external benefits* of contributing: $\gamma(n-1)$. Each of the $n-1$ other citizens receive exactly the same public goods benefit that Bridget received as the result of her contribution.
- Marginal social benefits of contributing (private plus external) $n\gamma$. The sum of the marginal private benefits γ and the the marginal external benefits $\gamma(n-1)$ is the marginal social benefit.

A fourth essential concept is the cost of contributing:

$$\text{B's disutility of contributing to the public good} = \frac{(a^B)^2}{2}$$

The marginal cost of contributing an amount a^B is the derivative of this expression with respect to a^B, which is just a^B itself. This means that the more she contributes the more costly it is to her. But when she is contributing nothing it would cost her very little to contribute a small amount. So we have

- *Marginal private cost of contributing*, a^B. There are no external costs of contributing so the social cost of contributing is the same as the private cost.

In making her decision about how much to contribute, the self-regarding citizen considers the private costs and benefits but not the external benefits. That is why the public good will be under-provided if it is left to the independent actions of self-regarding citizens. This is a standard case of market failure due to the uncompensated positive external effect that one citizen's contributing has on the well being of the other citizens.

REMINDER Social marginal costs equal private marginal costs *plus* external marginal costs.

The external benefits arise because the public good is non-rival and non-excludable. Because the good is non-excludable, I cannot exclude you from benefiting from my contribution to the good. And because the public good is non-rival, additional people can benefit from the public good without reducing its availability to others.

To see why a market failure results, consider a community of n people each of whom can contribute an amount of money to the provision of a public good. We will study the choice of one of these citizens, Bridget, and her contribution a^B. Bridget's choice is illustrated in Figure 16.2. The marginal private benefit from her contributing is the lower horizontal line. The private marginal cost is upward sloping green line.

She considers contributing nothing. But then the marginal private benefit should she contribute some small amount would exceed the cost. Similarly, were she to contribute a large amount the marginal costs would far exceed the marginal private benefits. So she will contribute up to the point that the marginal private costs equal the marginal private benefits (this is shown using math in the MNote).

REMINDER This is the Nash equilibrium of the game, and it is also the dominant strategy equilibrium; each citizen is doing the best they can and this is independent of what the other citizens are doing This is because there is a dominant strategy — to contribute γ — that is not affected by the actions of others. You can confirm this by noticing that the first order condition for the choice of the level of contribution does not involve any terms about the contributions of others.

All of the other $n-1$ members of the population, being identical to Bridget and facing the same incentives, also contribute the same amount, which we will now call simply a^N, the amount contributed by each in the Nash equilibrium without any subsidy. So the total amount contributed is na^N, and the total amount of the public good provided to each citizen is $na^N\gamma$. And because it is a public good, this amount is enjoyed by all n citizens. So the total utility experienced by the population is $n(na^N\gamma)$, which, remembering that $a^N = \gamma$, is equal to $(n\gamma)^2$.

M-Note 16.1: Public goods: A coordination failure

The amount of the public good available to Bridget and every other citizen is equal to a positive constant (γ) multiplied by the sum of all contributions or $\gamma(a^1 + a^2 + \ldots + a^B + \ldots + a^n)$. Bridget incurs a cost (in disutility) of $\frac{(a^B)^2}{2}$ when she contributes an amount a^i to the public good.

So Bridget's utility function is

$$B\text{'s utility unction} \qquad u^B = \gamma(a^1 + a^2 + a^B + \ldots + a^n) - \frac{(a^B)^2}{2} \tag{16.2}$$

To find how much Bridget will contribute we differentiate Equation 16.2 with respect to a^B, and set the result equal to zero (finding the first order condition that defines the maximum private benefit minus private cost):

$$\begin{aligned} u_{a^B} = \frac{\partial u^B}{\partial a^B} &= \gamma - a^B = 0 \\ \therefore \gamma &= a^B \\ \text{Marginal private benefit} &= \text{Marginal private cost} \end{aligned}$$

Therefore, Bridget's Nash equilibrium level of contribution is $a^{BN} = \gamma$ (N for Nash as in Chapters 5 and 9.

To see why this is a coordination failure think about the effect that Bridget contributing just a little more would have on Bridget's utility and the utility of the $n - 1$ other citizens. We know from the first order condition that she used to determine her level of contribution (just above) that

$$\frac{\partial u^B}{\partial a^B} = 0 \tag{16.3}$$

so Bridget's utility would not be affected. But her contribution and the increase in the public good that results would add an amount γ to the utility of each of the $n - 1$ other citizens. If everyone could agree to contribute a little more, everyone would be better off. Therefore the Nash equilibrium is not Pareto efficient, and the result is a coordination failure.

The socially optimal level of contribution

We now introduce an imaginary actor: the mechanism designer, charged by the citizenry with the task of devising a change in the rules of the game that will address the under-provision of public goods. The mechanism designer, we will assume, is committed to treating each citizen equally and regards the utility of each as comparable and equally worthy of being maximized. So the mechanism designer will attempt to maximize the sum of the utilities of the members of the population.

REMINDER Remember the auctioneer who regulates how markets work in the general equilibrium model in 14? We also termed him imaginary because there is no such entity. The auctioneer is a device for helping us think clearly about how markets might ideally work. We call the mechanism designer imaginary to stress that she is not intended to represent how governments or public policy makers really work, but instead to convey in simple terms how they ideally might work.

The mechanism designer's task is:

- *Objective: The desired outcome*. Determine the level of the public good that if implemented would maximize the sum of the utilities of the citizens. We call this the socially optimal level of public goods provision.

- *Implementation by a mechanism.* Devise a policy that will alter the citizens' utility functions – adding a subsidy in this case – so that the citizens acting privately will have a sufficient incentive to contribute the desired amount.

Notice how the mechanism designer inverts the process by which we have reasoned in most of the rest of this book. The economist typically takes the participation constraint and the incentive compatibility constraint as given, and then asks: given the individuals preferences and other aspects of the interaction, how much would citizens contribute in the Nash equilibrium of this interaction? The mechanism designer starts with the desired level of contributions and then works backward to discover a mechanism – a set of policies making up the participation and incentive compatibility constraints – which will make contributing this desired outcome a Nash equilibrium.

To determine the socially optimal level of public goods provision, the mechanism designer will take advantage of the fact that citizens are identical and she would like to treat them all equally, so the result to be implemented will be the same for each citizen. As a result choosing a given level that every citizen will contribute to the public good so as to maximize the utility of an individual citizen is the same thing as choosing a level for every citizen to contribute that will maximize the the total utility of all the citizens.

The optimal level of the public good provided can be determined by following these three rules (see Figure 16.2). Consider some particular level of contribution and the level of marginal social benefits (*MSB*) and marginal private costs (*MC*) when contributing this amount: then

- $n\gamma = MSB > MC = a^B$: if the marginal social benefits are greater than the marginal private costs of contributing, increase the contributions (the effect on increasing the benefits will exceed the effect on increasing costs);
- $n\gamma = MSB < MC = a^B$: if the marginal social benefit of contributing is less than the marginal cost, then reduce the contribution level;
- $n\gamma = MSB = MC = a^B$: if the marginal cost is equal to the marginal social benefit, then do not change anything: this is the socially optimal level of contributions.

The MNote explains further this result.

M-Note 16.2: The optimal contribution to the public good

Let's consider just Bridget: what level of contribution (identical across all citizens) would maximize her utility? We will refer to her utility as u^W without any B superscript because the citizens are identical, and we are considering variations in some identical level of contribution that they all will make.

Bridget's utility function is

$$u^W = \gamma na - \frac{a^2}{2}.$$

Because the contributions of the citizens will be equal we have:

$$u^W = \gamma na - \frac{a^2}{2}.$$

Differentiate the utility function with respect to a:

$$\frac{\partial u^W}{\partial a} = \gamma n - a^N = 0.$$

For the maximum, set the result equal to zero:

$$a^N = \gamma n. \tag{16.4}$$

Thus we see that the socially optimal level is that each citizen contributes $n\gamma$, not just γ which is what the individual maximizing her own utility does.

Checkpoint 16.2: Population size and optimal public goods provision.

Explain the economic reasons why the socially optimal level of the public goods provision is larger if the population is larger.

A subsidy to internalize the external benefits

The key ideas in designing the mechanism are to use economic theory to develop a diagnosis of the market failure and on the basis of this, propose a remedy:

- **Diagnosis**: the under-provision of the public good by private actors occurs because contributing produces external benefits for which the contributor is not compensated
- **Solution**: Find a way to compensate each citizen for the external benefits that their contribution confers on others.

We know that the *private* benefit that a citizen receives for each unit that Bridget contributes to the public good is γ. And each of the $n-1$ other citizens also receives the same benefit from her contribution (because it is a public good). So the external benefits conferred by Bridget contributing a^B are $(a^B)\gamma(n-1)$. Therefore, to compensate for the otherwise uncompensated external benefits, the subsidy per unit of contribution should be $\gamma(n-1)$.

Like the tax on fishing time to prevent over-harvesting fish stocks in the lake (studied in Chapter 5) the subsidy has addressed private under-provision of the public good by shifting the Nash equilibrium to a Pareto efficient outcome. The public good example, however, is a particularly simple case: the actions taken by each citizen do not depend in any way on what the other citizens did. As far as the model is concerned they could well have lived on separate islands, unknown to one another.

Most problems of mechanism design, by contrast, are complicated by the fact that how one person responds to the policy depends on what everyone else is doing, including how they respond to the policy, as was the case of over-fishing the lake. To see how the mechanism designer works in this more realistic environment, consider the case of what are called sin taxes, that is, taxes intended to raise the cost of what are widely considered to be bad habits, like excessive drinking of alcoholic beverages or smoking. The key fact here is that drinking and smoking are social activities.

M-Note 16.3: The Mechanism Designer's Socially Optimal Subsidy

In selecting the socially optimal subsidy the mechanism designer does not take account of the subsidy received as a contribution to each citizen's utility. The reason is that for every subsidy received, some citizen has paid an equivalent tax (to provide the government revenues for the subsidy). The subsidy is a transfer among citizens, and as the mechanism designer treats each of them as equivalent and comparable, it has no effect on the total utility of the citizenry.

To find the optimal subsidy we modify Bridget's utility function

$$\text{Subsidized utility} \quad u^B = \gamma(a^1 + a^2 + \ldots + a^B + \ldots + a^n) - \frac{(a^B)^2}{2} + \omega a^B$$

This modified utility function is just: Private benefits from the public good - private costs of contributing a^B + subsidy for contributing a^B.

We can now find Bridget's marginal utility with a subsidy, again by differentiating her utility function (now modified by the subsidy) with respect to her amount contributed and setting the result equal to zero. This gives us the first order condition for Bridget's contribution level when the subsidy is introduced :

$$\begin{aligned} u_{a^B} = \frac{\partial u^B}{\partial a^B} &= \gamma - a^B + \omega \\ \text{To find the optimum subsidy we set this result to zero} \therefore \gamma + \omega &= a^B \qquad (16.5) \\ \text{Private and Subsidized Marginal Benefit} &= \text{Marginal Cost} \end{aligned}$$

Equation is Bridget's best response function, giving the amount she will contribute for any value of the subsidy that the mechanism designer might choose. And we already calculated the optimal contribution (equation 16.4): $a^N = \gamma n$. Therefore, all what we have to do is to equate Bridget's best response function to the socially optimal level of contribution

$$\begin{aligned} a^N = \gamma + \omega &= \gamma n \\ \text{and then solve for the amount of subsidy: } \omega &= \gamma n - \gamma \\ \omega &= \gamma(n-1) \end{aligned}$$

The subsidy that implements the optimal level of public goods contributions exactly compensates each contributor for the external benefits that their contributions generate, that is γ for each of the other citizens, of which there are $n-1$.

Checkpoint 16.3: Internalizing external benefits

- Explain in your own words why the subsidy $\omega = \gamma(n-1)$ internalizes the external benefits of each citizen's contribution.
- If the population grew to $2n$, double its current size, explain why the optimal subsidy would also increase.
- Would the optimal subsidy double?

16.3 The social multiplier of cigarette taxes

The fact that consumption is a social activity was illustrated in Chapter 7 by a model and some evidence about how one might engage in "conspicuous consumption" as a social signal of one's own status, earning power or respectability. We return to the social nature of consumption here — smoking with friends is more enjoyable than alone — and see how this changes the effects of policies designed to discourage this unhealthy form of consumption.

The social multiplier

Suppose we want to determine the effect of a cigarette tax on the amount of smoking that people do. We know that the amount of smoking is reduced by a higher price of cigarettes and is increased by other people smoking more (or decreased by other people smoking less). The tax affects the amount of smoking *directly* because it raises the price of cigarettes. But it also affects smoking indirectly: if people smoke less after the tax is imposed, then smoking will not only be more expensive, it will also be less enjoyable because fewer people do it.

The total effect of a price increase on an individual's smoking will therefore be greater than if the price increase were experienced (hypothetically) only by a single person because of the additional indirect effect of the tax reducing how much other people smoke. The **social multiplier** measures the difference between these direct and total effects.

THE SOCIAL MULTIPLIER measures the indirect effects of social policy through its effects on other people's behavior. A tax on cigarettes reduces one's own consumption of cigarettes directly through a decrease in quantity demanded at that price and through an indirect, social effect that reduces a person's enjoyment as a consequence of fewer people smoking.

Figure 16.3 illustrates the direct and indirect effects.

Econometric estimates suggest that in other forms of "social consumption" the social multiplier may be large. A study of "heavy drinking" by Russian men, for example, stimated that the effect of a 50 percent permanent increase in the price of vodka would reduce heavy drinking by about 30 percent. The social multiplier — that is the indirect effect of reduced peers' drinking on one's own alcohol consumption — accounts for one third of this effect.

The tax on smoking alters social outcomes by altering the environment in which people decide what to do, that is, how much to smoke now that it's more expensive and others are smoking less.In this way, the design of the tax is like the constitutional challenge we discussed in Chapter 1: how do we design a tax so that people do not obtain outcomes they would rather avoid? Both the direct and indirect effects reduce the Nash equilibrium level of smoking. The model in this section is thus an application of the idea of implementation of an outcome by Nash equilibrium.

How could the mechanism designer calculate the impact of a tax on the level of smoking?

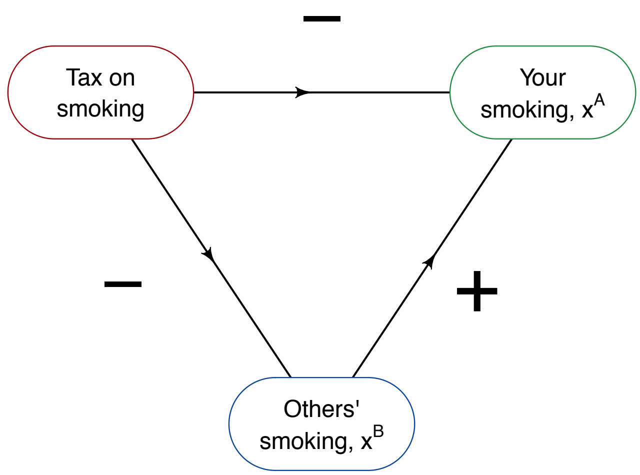


Figure 16.3: **The social multiplier of a tax on cigarettes.** The figure shows the first round effects of a tax on cigarettes smoking both directly (by making smoking more expensive) and via the social multiplier (by making smoking less enjoyable, because there will be fewer people to smoke with). The negative signs show a negative relationship between taxes and smoking: as taxes increase, smoking decreases; as taxes decrease, smoking increases. The positive sign shows that there is a positive relationship between others' smoking and your own smoking: so if other people smoke more, you smoke more, if other people smoke less, you smoke less. What is shown is just the first round effect: your smoking is now less than before, so now others will find smoking less enjoyable and reduce their level of smoking, and so on.

A model of smoking as a social activity

Let us assume, for simplicity, that there are two smokers (A and B). Each have an income of y_A and y_B. The price of each cigarette is P. A and B smoke x_A and x_B cigarettes respectively. As a social activity, the utility of smoking depends positively on the amount of smoking of the other person. Both smokers have to choose between smoking or using their money on any other good.

Figure 16.4 shows how person A will select a level of smoking. The downward-sloping solid curve is A's marginal utility of smoking cigarettes. Her marginal utility is downward-sloping because the 20th cigarette smoked in a day is less enjoyable to her than the first, or the 15th. She maximizes utility by smoking more as long as the marginal utility exceeds the price, leading her to smoke x_{A0} cigarettes. (M-Note 16.5 explains her choice using calculus and a particular utility function.)

Suppose the smoking tax raised the price of cigarettes, as shown in the figure. There would be a series of effects:

- A first round direct effect: Because smoking is more expensive, she would do less.
- A first round indirect effect: Because in response to the tax, others would smoke less too, she would enjoy smoking less.
- Subsequent indirect effects: Because as a result of others smoking less,

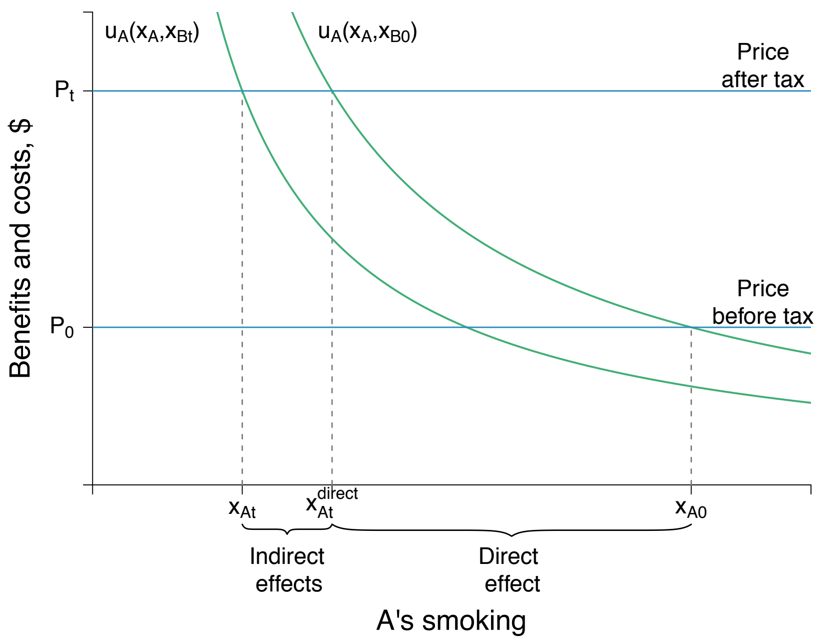


Figure 16.4: **The social multiplier of a tax on cigarettes.** The figure shows how an individual chooses a level of smoking, and how both a price increase and others' smoking less will reduce smoking. The first round and subsequent indirect effects are shown as 'indirect effects' in the figure.

she is now smoking less, too, and now others would enjoy smoking even less, and cut back, in response to which she also would reduce even further her level of smoking, and so on.

The first effect is shown by the shift up of the price line. The second and subsequent indirect effects are shown by the shift downward in her marginal utility of smoking curve: for any level of smoking that she does, the marginal utility is now less. The sum of these two effects results in her reducing her smoking from x_{A0} to x_{At}.

M-Note 16.4: A smoker's best-response function

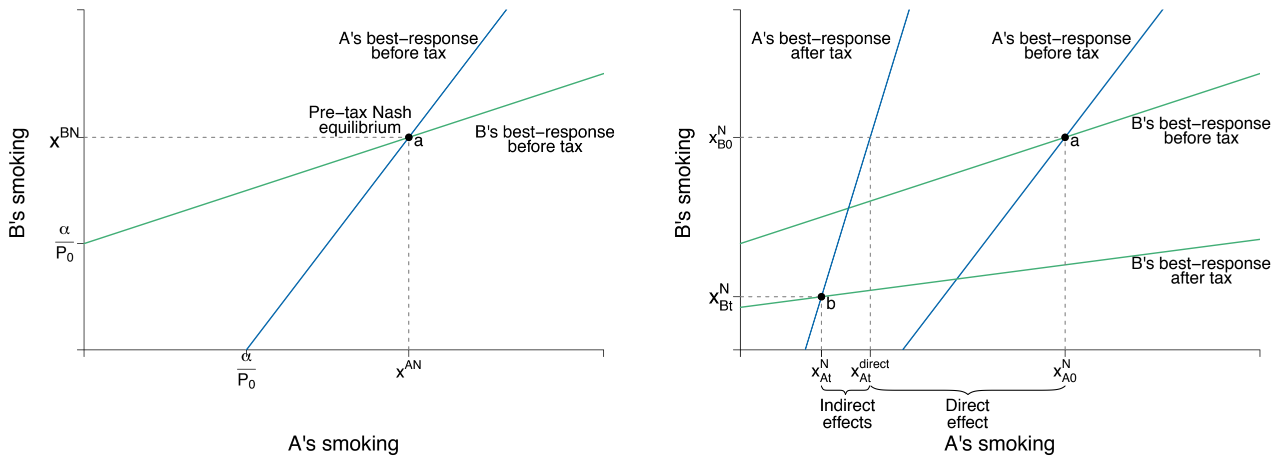
For simplicity, we assume that the utility of the money spent in other goods is proportional to the income remaining after paying for cigarettes ($Y - Px_{A,B}$). Then a utility function of A that expresses the idea that smoking is s social activity is:

$$u_A \quad = \quad (\alpha + \beta x_B) ln x_A + y_A - P x_A$$

The parameter α measures how much the smoker enjoys smoking alone (that is when $x_B = 0$) while β measures how much more pleasurable smoking is if others smoke more.

Smoker A will pick x_A to maximize her utility. Note that in this problem x_A is the only variable that A can select. As usual, to calculate x_A that the smoker will choose, we need to derive the first order condition:

$$\begin{aligned} \text{First order condition} \quad \frac{\partial u_i}{\partial x_A} &= \frac{\alpha + \beta x_B}{x_A} - P = 0 \\ \text{Marginal utility} &= \text{Marginal benefit} - \text{marginal cost} = 0 \\ \text{Rearranging to isolate } x_A \quad x_A &= \frac{\alpha + \beta x_B}{P} \end{aligned} \tag{16.6}$$



(a) Best responses before tax

(b) Best responses after tax

Equation 16.6 is the best response function of A. You can see that her smoking depends negatively on the price of cigarettes (because P is in the denominator of the expression) , and positively on the smoking of B (because the smoking by others is in the numerator with a positive sign).

Figure 16.5: **The Nash equilibrium level of smoking before (point a) and after (point b) a cigarette tax is imposed**. The effect of the tax is shown in the shift to the left of A's best-response function (the blue line) and the shift down of B's best response function (the green line).

The Nash equilibrium level of smoking

The left panel of Figure 16.4 shows how A changes her level of smoking when the price increases. But it does not show how B responds to the tax and to the fact that A now smokes less. And you may wonder if all of the subsequent rounds of people smoking less inducing others to smoke less would ever come to an end.

To find out we show in Figure 16.5 how each responds to the level of smoking of the other, assuming the price of cigarettes is fixed. The steeper of the two upward-sloping lines is person A's best-response function, showing the levels of smoking she will do in response to B's level of smoking. We derive A's best-response function in M-Note 16.5. The best-response function of A slopes upwards because, holding the price of cigarettes constant, the more B smokes, the more A will smoke. The x-axis intercept of her best response function shows how much she would smoke if B did not smoke at all.

The flatter of the upward-sloping lines is B's best response function, showing how he responds to each possible level of her smoking. The two best-response functions give us all the information we need to determine the Nash equilibrium level of smoking (x^{AN}, x^{BN}). Remember, the Nash equilibrium is a mutual best response so it must be a point on both of the smokers' best-response functions. There is only one Nash equilibrium which occurs as the intersection of the two best-response functions.

M-Note 16.5: The Nash equilibrium of an interaction between smokers

For the best-response functions to intersect we need $\frac{\beta}{P} < 1$. Otherwise, the slopes of both best-response functions are higher than one, and as long as each would smoke at least a little even if the other did not, the best-response functions do not intersect.

To derive the Nash Equilibrium of the interaction between smokers A and B we know that in the Nash equilibrium the two best response functions must intersect (as they do in Figure 16.5). Therefore, at the Nash equilibrium, $x_A = x_B$. We can therefore, substitute $x^B = x^A$ into A's best-response function to find the Nash equilibrium levels of smoking x^{AN} and x^{BN}. You may remember we derived the best-response functions the same way in Chapters 5 and 9.

Using equation 16.6, x^A and x^B to find the Nash equilibrium level of smoking:

$$\begin{aligned} x^A &= \frac{\alpha + \beta x^A}{P} \\ \left(1 - \frac{\beta}{P}\right) x^A &= \frac{\alpha}{P} \\ \left(\frac{P-\beta}{P}\right) x^A &= \frac{\alpha}{P} \\ x^{AN} &= \frac{\alpha}{P-\beta} = x^{BN} \end{aligned} \tag{16.7}$$

Equation 16.7 shows that the Nash equilibrium level of smoking will be positive because, as we have already seen, $\frac{\beta}{P} < 1$ or $\beta < P$, so the denominator of Equation 16.7 must be positive.

The direct and indirect effects of a cigarette tax

The right panel of Figure 16.5 shows how the increase in price caused by the tax changes the best response-functions of the two, and how as a result the Nash equilibrium level of smoking is reduced. There are two effects of the price increase on A's best response function

- It reduces the amount of smoking she would do if B did not smoke at all, lowering the x-intercept and
- It reduces the effect of B's smoking on her own smoking, making the best response function steeper.

The effect of the tax-induced price increase on B's level of smoking is similar. In Figure 16.5 it shown as a shift downwards of his best-response function and a flattening of the slope. The new Nash equilibrium is the intersection of the two new best-response functions, point **b** in the figure.

The total change in A's smoking x^N_{A0} to $x^N_{At}$is made up of two effects:

- The **direct effect**: shown as a reduction from x^N_{A0} to x^{direct}_{At} and
- The **indirect effect**: shown as the reduction from x^{direct}_{At} x^N_{At}

M-Note 16.6: Direct, indirect, and total effects of a cigarette tax

The direct and indirect effects of the cigarette tax shown Figure 16.3 can be identified mathematically using the *total derivative*. We used the total derivative in Chapter 3 to understand the marginal rate of substitution. You can also check that mathematical appendix if you need to review the idea of the total derivative.

To understand the total effect of the change in price, we assume that an increase in the tax impacts directly on price ($\frac{\partial P}{\partial t} = 1$). Therefore we will focus on the change on smoking given a change of the price ($\frac{\partial x^N}{\partial P}$).

Using 16.6, we calculate dx^A, the change in smoking of the individual A given a change in the price dP:

$$\begin{aligned}
\text{Change in A's smoking} &= \text{Change through price} + \text{Change through B's smoking} \\
dx^A &= \frac{\partial x^A}{\partial P} dP + \frac{\partial x^A}{\partial x^B} dx^B \\
\text{Divide through by } dP \quad \frac{dx^A}{dP} &= \frac{\partial x^A}{\partial P} + \frac{\partial x^A}{\partial x^B}\frac{dx^B}{dP} \qquad (16.8) \\
\Delta \text{ smoking for } \Delta\text{Price} &= \text{Direct effect} + \text{Indirect effect}
\end{aligned}$$

Equation 16.8 shows that we can decompose the total effect of the smoking tax into two parts:

- $\frac{\partial x^A}{\partial P}$ is the *direct* effect of the tax on smoking.
- $\frac{\partial x_A}{\partial x_B}\frac{dx_B}{dP}$ is the *indirect* effect through the change in consumption of B.

To simplify the problem further, we use the fact that the two people behave identically, as we showed in 16.5 for their Nash equilibrium levels of smoking x^{AN} and x^{BN}. Therefore, $\frac{dx_A}{dP} = \frac{dx_B}{dP} = \frac{dx^N}{dP}$:

$$\begin{aligned}
\frac{dx^N}{dP} &= \frac{\partial x_A}{\partial P} + \frac{\partial x_A}{\partial x_B}\frac{dx^N}{dP} \\
\left(1 - \frac{\partial x_A}{\partial x_B}\right)\frac{dx^N}{dP} &= \frac{\partial x_A}{\partial P} \\
\frac{dx^N}{dP} &= \frac{1}{1 - \frac{\partial x_A}{\partial x_B}}\frac{\partial x_A}{\partial P} \qquad (16.9)
\end{aligned}$$

Because the effect of one's smoking on the other's level of smoking is positive – $\frac{\partial x_A}{\partial x_B} > 0$ – we can see from Equation 16.9 that the total effect of the tax is greater than the direct effect.

The social multiplier

Considering the direct and indirect effects of the cigarette tax, the indirect effect occurs because the smokers' payoffs are interdependent and they each have negative external effects on one another. The interdependence of their outcomes makes the game *social* and leads us to consider the relative importance of the indirect effect of policy to the total effect of the policy. This relative effect of the indirect effect is called the *social multiplier* as it shows the importance of the social aspect of the interaction and how important one person's behavior is to another person mimicking or refraining from that behavior.

The social multiplier is defined as

$$\text{Social multiplier} = m \quad = \quad \frac{\text{total effect}}{\text{direct effect}} - 1 \qquad (16.10)$$

Equation 16.10 shows that if the indirect effect is absent, then the direct effect equals the total effect, and therefore the social multiplier is zero. The social multiplier is zero in a world where people's payoffs are independent of each other, that is, when each has no external effects on the other.

But in many important cases, social policy needs to consider the external effects that people have on each others' behavior. In cases where people impose negative external effects on each other – such as with smoking, eating sugary foods, and so on – taxes or fines can help to reduce behavior directly through prices, but also indirectly through changing how others behave. Alternatively, in cases where people provide positive external effects on others' behavior – such as with getting vaccinated, engaging in healthy activities, or engaging in civic behaviors – governments may choose to subsidize or directly facilitate of such behaviors to encourage them because of the effects of the social multipliers which results in more people adopting those behaviors even when the benefits are personally uncompensated.

Checkpoint 16.4: A negative social multiplier

Can you think of situation in which the social multiplier might be *negative*: the *total effect* is less than the *direct effect*. What would be the sign of the *indirect effect*? Can you think of a a real-life interaction in which we can observe a negative social multiplier?

16.4 The theory of the second best and public policy

Many public policies seek to reduce the extent of uncompensated external effects and thereby to limit the extent of market failures in an economy. Examples include imposing taxes or fines on the emission of pollutants and subsidizing contributions to public goods such as neighborhood amenities.

We encounter additional market failures when we drop the assumptions of the perfectly competitive model and recognize that while firms compete they typically also have some degree of market power, and can benefit from restricting output so as to sell at a higher price. Recall from Chapter 9 that with limited competition, the firm will select a level of output and price such that the price exceeds the marginal cost. And pricing higher than marginal costs results in a market failure.

REMINDER Recall that from Chapter 8 you know that if $p > mc$ then there will be a dead weight loss and from Chapter 14 that marginal cost pricing is also the basis of the "invisible hand" reasoning formalized in the First Welfare Theorem (see (Equation 14.4))

Government policy might wish to address both the problem of limited competition and uncompensated external effects. But, unless it can address both problems completely (totally eliminating uncompensated external effects and ensuring the prices will equal marginal costs), the government's policies may make the situation worse.

HISTORY Here is the gist of what has come to be called the *general theorem of the second best* advanced in 1956 by Richard Lipsey (b. 1928) and Kelvin Lancaster (1924-1999)): a single violation of the relevant efficiency conditions (that is equation (14.4))) means that fulfilling the remaining marginal conditions may result in an allocation that is Pareto-inferior to an allocation implementable by more extensive violations of the efficiency conditions.

The intuition behind this result is that policy makers can address the Pareto-inefficient allocations caused by the violation of one of the conditions (Equation 14.4) by violating another condition. An example: think about a firm emitting pollution when it produces a good, so that its production has negative external effects on the environment. The marginal social costs of production (the total costs to society) exceed the marginal private costs (to the firm), the firm produces more than the Pareto-efficient level of output.

But imagine that the firm is a major petroleum company. Now take account of the fact that the firm faces limited competition, so it is restricting output, and setting a price in excess of its marginal private cost. Because the firm has limited competitors, the firm restricts output and therefore partially offsets the problem of the firm's negative environmental effects (which leads it to produce more than the Pareto-efficient amount of output).

REMINDER In Chapter 9, we spoke about competition policy as a way of *increasing* the number of firms in a market that had high concentration (few firms competing). In the US, this competition policy has typically been called anti-trust policy because monopolies were called trusts. Currently, many governments – South Africa, the UK, The Philippines, Singapore, India, the EU – have a "competition commission" that administers competition policy.

Ideally a government would introduce both

- *a competition policy* ensuring that the firm faced sufficient competition so that it would produce up to the point where the price is equal to the private marginal cost and
- *an environmental policy* imposing costs on the firm equal to the negative external effects of its emissions, so that the firm's marginal private costs would now include the environmental costs imposed on others.

These dual policies would correct the market failure; this is the first best outcome. But often it is impossible to do both. In the absence of an environmental policy, a competition policy that induced this producer to choose the competitive output level – producing and emitting more – could be welfare-reducing rather than welfare-enhancing.And an environmental policy in the absence of a competition policy would impose the marginal social costs on the firm, leading the firm to restrict production even more. And this might, on balance, worsen social outcomes.

Figure 16.6 illustrates the case of a firm facing a downward sloping demand curve (due to limited competition) with marginal private costs lower than marginal social costs. In the absence of any policy, the firm will equate marginal revenue to marginal private cost, producing the amount X^{MP} and selling at the price p^{MP}. With the status quo – no policies – the consumer surplus that is lost due to limited competition is the consumer's willingness to pay minus the private marginal cost summed over the output that was not produced due to limited competition. This is the sum of areas A + D. But area D is also the external costs that would have been imposed on others, had the firm not restricted it's production. So some of the consumer surplus lost due

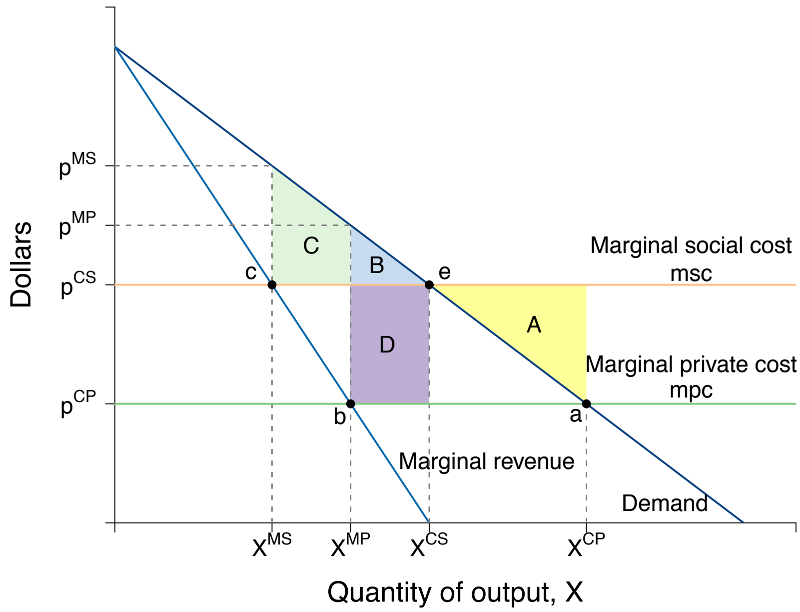


Figure 16.6: **The theory of the second best.** The theory of the second best illustrated by a comparison of competitive markets and monopoly when the marginal social costs exceed the marginal private costs because of the negative external effects imposed on others. The lower case letters indicate the outcomes under no policy (b, for bad), environmental policy (green tax)only (c), competition policy only (a), and both policies (e, for efficient.)

to restricted output is offset by the external costs avoided due to restricted output. So the net deadweight loss in Figure 16.6 is area B.

Now suppose the government were able both to

- implement taxes so that the (tax adjusted) private marginal costs equal the social marginal cost, and
- enforce competition so that the firm produces up to a point that the price at which it sells its output equals its marginal private costs (which because of the tax are now equal to its marginal social costs.)

Then the firm would choose the level of output that equates price to the marginal enforce competition so that the firm produces up to a point that the price at which it sells its output equals it marginal private costs and cost and produce at the optimal level X^{CS}. There would be no deadweight loss. This is what is called the *first-best outcome* because it is Pareto-efficient.

FIRST-BEST OUTCOME A first-best outcome is Pareto-efficient. Typically, a great many alternative Pareto-inferior outcomes may also exist, and some may be judged better than others on fairness or other grounds.

But if the government were to enforce competition in the absence of an environmental policy, firms would ignore the marginal external costs and produce up to X^{CP} where price equals marginal private cost, with corresponding deadweight loss of area A. Notice this is different from the deadweight losses studied in 8, which resulted from firms producing too little. Here the firm is producing too much, and the deadweight loss is the sum of the difference between the the marginal social costs and the (lesser) willingness to pay of the buyers.

Considering the opposite case, a monopoly that was forced by an environmental policy to pay the full marginal social costs would choose the output where its marginal revenues equal the social marginal cost, at point X^{MS}. The resulting total deadweight loss would be area B plus area C.

The four possible outcomes — no policy, both policies, and one but not the other policy — and their effects are summarized in Table **??**

Both of the single policy outcomes – the competition policy without the environmental policy and the environmental policy without the competition policy – are worse in welfare terms than the original outcome (X^{MP}, p^{MP}).

In this case imposing either policy in the absence of the other is worse than doing nothing: the area B – the doing nothing deadweight welfare loss – is smaller than either A – the loss from adopting the competition policy alone – and of course smaller than $C + B$, the loss from adopting the environmental policy alone.

Checkpoint 16.5: Second-best policies

- Explain why the points **b**,**c**,**a** and **e** in figure 16.6 indicate the results of the four sets of policies shown in the left column of Figure XXX
- It is not always true that doing nothing is better than adopting either policy in isolation. In other cases, a government should adopt one policy rather than another policy because of the relative improvements in social welfare from doing so. Can you draw the curves in Figure 16.6 so that adopting the environmental policy but not the competition policy will be better than doing nothing? Can you do the same figure for the case in which adopting a competition policy is an improvement over the do nothing policy even if it is not accompanied by the environmental policy?

The second best theorem does not question the idea that public policy ought to address market failures arising from lack of competition and uncompensated external effects. But, it does show that treating market failures in isolation rather than as a general problem can be counterproductive. It also underlines the fact that the best the policy maker can do may not be to eliminate the market failure (first best) but instead to limit its extent (second best).

16.5 Deception as an impediment to efficient exchange

In Chapter 9, we introduced an imaginary economic actor – the perfect competitor – who never misses an opportunity for private gain. And we provided conditions under which an economy made up of perfect competitors would implement Pareto-efficient outcomes.

We showed, for example, that the owner of a firm facing limited competition (and hence a downward sloping demand curve) like the firm studied in the

Policy	First order condition	Point in Fig	Price	Output	Welfare loss	Reason for loss
No policy	mr = mpc	**b**	p^{MP}	X^{MP}	B	Too little output
Green taxes only	mr = msc	**c**	P^{MS}	X^{MS}	B + C	Even less output
Competition policy only	p = mpc	**a**	P^{CP}	X^{CP}	A	Too much output
Green taxes + competition	p = msp	**e**	P^{CS}	X^{CS}	None	

Table 16.1: **Policy making in a second best world.** The rows are the four policy options discussed in the text. You can see that if introduced singly, both green taxes and competition policy result in an increase in the size of the deadweight losses over the "no policy" option.

previous section was a perfect competitor, he would not restrict output to sustain higher prices but instead would practice price discrimination. This would allow him to expand production up to the point where price = marginal cost. Thus if perfect price discrimination is possible even a monopoly would implement the conditions that we associate with perfect competition. We also used a model of bilateral exchange with rent-seeking competitors to understand how buyers and sellers could reach a Pareto-efficient outcome (recall Figure **??**.)

The perfect competitor is imaginary, like the mechanism designer and the auctioneer. The three personages are just thought experiments like mathematical models that we use to clarify our reasoning and to present in a simple form some important aspect of a more complex reality.

We gave the perfect competitor extraordinary powers, the ability of the monopolist, for example, to find out the willingness to pay of each potential buyer of her product, and to impose that price on each. But one power we did not give her: the power to lie.

Lying for profit: Why mutually beneficial trades may not occur

And so we confront a problem: Can the perfect competitor implement efficient outcomes if she (and everyone else) can lie.

When traders meet they have no incentive to report how much they value the good to be exchanged. That is, they have no incentive to report what we call their *true valuation* of the good. The reason is that a trader can profit by *misreporting* or withholding information about how much she values the good.

This behavior is quite common in bargaining situations.Think about a prospective buyer of your home. He is not going to say: "well I'd prefer a lower price, but in fact I'd be willing to pay as much as two hundred thousand dollars for it." And you, the owner, are not going to offer the information "I'd like the high-

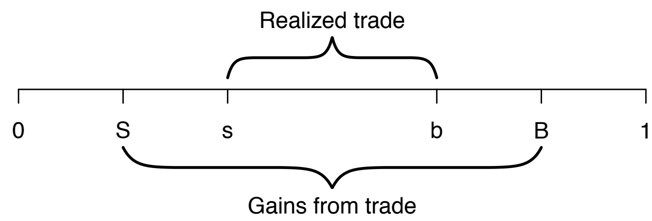


Figure 16.7: **Bargaining over the gains from trade.** For the pair of traders shown in the figure, a mutually beneficial trade is technically possible because the seller's true valuation is S and the buyer's true valuation is B. Therefore, for any price between those points the exchange is mutually beneficial. If the two announce the prices s and b, then they will trade at some price between those two numbers. If, however, s had been greater than b, no trade would take place, even though it would be Pareto-improving to execute the trade.

est price possible, but I'd be willing to part with the house for as little as one hundred thousand dollars if that is the best I can do."

If they are perfect competitors they will certainly recognize withholding or misrepresenting information about an exchange as a rent-seeking opportunity, and falsely report their valuations of the goods. The result, we will show, is that some exchanges that could have benefited both buyer and seller will not happen. The reason is that the traders' stated valuations will influence the prices at which the traders exchange. As a result, some mutually beneficial exchanges will not occur.

The problem is quite general, but it is best illustrated by an institution called a **double auction** involving goods for which the usual impediments to bargaining (such as incomplete contracts) are absent.

A particular good now in the possession of the seller, her house for example, is worth S to a seller who may sell this good to a buyer. The good is worth B to the buyer. These valuations are private information, so the other trader does not know her trading partner's valuations.

The double auction proceeds as follows:

- **Pairing buyers and sellers:** a large number of buyers and sellers for a one-shot are paired, simultaneous interaction.
- **Seller:** In each pair, the seller announces the minimum price at which she is willing to sell, s
- **Buyer:** At the same time, the buyer announce the maximum price at which he is willing to buy, b .
- **Exchange**: An exchange occurs if the buyer's offer price is greater than the seller's sell price, that is, $b > s$.
- **Price**:The price determined by the two bargaining will be between s and b: for example the bargaining rule could be that the price is midway between the two, so $p = (b+s)/2$.

Figure 16.7 depicts the valuations of the players and the potential for gains from trade.

DOUBLE AUCTION In a double auction buyers and sellers simultaneously submit to an auctioneer 'bids' and 'asks', that are the prices at which they are willing to buy and sell, respectively. An auctioneer then chooses a price that clears the market. The bargaining game covered in this section – with one buyer and one seller and no auctioneer – is a variant of the double auction.

You can see that in this exchange $B > S$ is the participation constraint: if the good is not worth more to the buyer than it is to the seller, then there is no way that the exchange could occur voluntarily. If $B > S$ and the exchange does not happen, then this result is not Pareto-efficient because then there would exist some change – the sale that did not occur – that would make both traders better off.

The incentive compatibility constraint in this exchange is $b > s$ because this inequality restricts the actual exchanges to cases where buying and selling prices that the two will voluntarily choose to announce allow a trade to occur.

HISTORY Roger Myerson was awarded a Nobel prize in Economics for demonstrating in 1983 (with his co-author Mark Satterthwaite) what is called the 'Myerson-Satterthwaite Theorem', which shows that some mutually beneficial trades will not take place in the double auction because buyers and sellers misrepresent their true values.

The reason why opportunities for beneficial trade will be missed as follows:

- **Buyers have an incentive to lie**: if a trade takes place, then the price will be more favorable to the buyer, the lower is her announced valuation, b.
- **Sellers have an incentive to lie**: if a trade takes place, then the price will be more advantageous to the buyer, the higher is his announced valuation s.

The seller therefore has an incentive to *overstate* her valuation of the good and the buyer has an incentive to *understate* his valuation of the good. By misrepresenting their valuations, if a trade nonetheless occurs, both benefit from the resulting increase in their share of the gains from trade that they will get. But this comes at the cost of reducing the probability of a transaction. As a result, when the two traders meet it may happen that $b < s$, so that trade does not occur, even though $B > S$, so the buyer valued the good more than the seller, and a mutually beneficial trade *could* have occurred were they not deceitful.

This inefficient result arises because the announced valuations of the buyers and sellers influence *both*

- the price at which the good will transact – if a transaction is concluded and
- whether a transaction will take place at all.

If the price at which the good is exchanged were determined without regard to the announced valuations, then all mutually beneficial transactions would take place.

Conflict over the distribution of the pie may mean a smaller pie

The failure of our rent-seeking perfect competitors to exploit a mutually beneficial exchange opportunity means that the outcome is Pareto-inefficient. This is another example of the general problem that the conflict over how the pie will be divided up often results in a smaller pie. Other examples you have already studied include:

- Rejections of positive offers in the Ultimatum Game
- Pareto-inefficient allocations when one actor has price setting power (but not take-it-or-leave-it power)
- Conflict over on-the-job effort and resulting unemployment and diversion of resources to monitoring, and other unproductive labor-disciplining uses
- Conflict of interest over repayment of loans and conduct of projects financed by credit leading to exclusion from the credit market of would-be borrowers seeking to finance good projects who lack sufficient wealth to post collateral.

The conclusion is that even *without* the uncompensated external effects and transactions costs that result in market failures, it is effectively impossible to achieve the Pareto-efficient outcome envisioned in either a (Coasean) rent-seeking model of perfect competitors or a perfectly competitive model (a Walrasian model) of competitive general equilibrium. We live in a second-best world.

Checkpoint 16.6: Market failures in the double auction

Explain why the following statements are true:

- if the bids and asks are submitted in sealed envelopes and the price is chosen by an assessor (a person skilled at determining the value of things) before opening the envelopes following which the buyers and sellers could then decide whether to make the exchange at the selected price, then they would have no incentive to misrepresent their true values.
- if the potential buyer truly values the good much more than the seller then even though they will misrepresent their true values, it is likely that they will execute a trade (this means that the trades that are most beneficial are likely to be take place.)

Markets would work better if Homo economicus had a conscience

But even in this second-best world, we might wonder whether an impartial Mechanism Designer or even a randomly-selected citizen might try to encourage people to change their behavior. For example, in the double auction we described, if one of the many buyers or sellers were asked to choose a strategy for announcing the buying or selling price that everyone (including her) would have to follow, what rule would she choose? She would choose telling the truth.

TRUTH-TELLING MECHANISMS are rules of the game that that would make it a best response for a self interested and amoral person (that is, *Homo economicus*) to reveal their *true* preferences (including their true value of a good that they may buy or sell).

The reason is that if everyone were to follow the “tell the truth” rule then she would do better on average because she would then be in an economy in which all mutually beneficial exchanges actually happen. So in this case $b = B$ and $s = S$ and every exchange that allowed mutual benefits would then occur, leading to a Pareto-efficient outcome.

But telling the truth is not a best response to others telling the truth in the double auction, as we have seen. So unless everyone acquired a conscience that simply banned misrepresenting their values on ethical grounds, the 'tell the truth' rule would break down. Nobody would follow it.

Economists and philosophers have long sought to devise rules of the game that would provide incentives to truthfully reveal ones preferences, called truth-telling mechanisms. But not only has the search been unsuccessful: it has shown that truth telling mechanisms cannot exist.

HISTORY In the early 1970s, philosopher Allan Gibbard and economist Mark Satterthwaite independently proved a theorem demonstrating the impossibility of mechanisms that would make honest revelation of preferences a dominant strategy.

It is clear that if all parties can agree to restrict the pursuit of their own individual advantage — forgoing the advantage of misrepresenting one's true value — all may benefit. Another example that you encountered in Chapter 1 would be in the Prisoner's Dilemma Game, to just outlaw defection.

But as Adam Smith advised us at the beginning of this chapter, there are (and should be) limits to what the government can accomplish in this respect. In many countries, for example, it is illegal to overstate the value of one's property when making an application for a loan. But neither banks nor government typically have the information or enforcement capacities required to make truth telling a best response. Many environmental policies illustrate the "outlaw defection" option. This works in cases where defection takes a specific and observable form whose banning can be easily enforced — banning the use of lead in vehicle fuel — but not in others.

16.6 When optimal contracts fail: The case of team production

Residential segregation illustrates a case where we may need to modify the logic of mechanism design: to take account of how preferences may respond to policies, and that the policy maker may have other goals than implementing a Pareto efficient outcome. In other cases the objective of internalizing external effects to implement a Pareto superior outcome (with given preferences) makes good sense, but the policies that would produce Pareto-superior outcomes are not feasible. Team production provides an illustration of why the optimal contracts a mechanism designer would introduce may not be possible.

What is team production?

Most production in a modern economy involves large numbers of people contributing to a job in ways that make it difficult to assign particular responsibility for each task to any one worker. This is called **team production**.

TEAM PRODUCTION is a form of production in which the the contribution of each individual to the output cannot be determined, either because it cannot even be *defined* or because it cannot be *measured*.

REMINDER Return to Chapter 10 to remind yourself of the aspects of contracts that make them **incomplete**. Remember that defining and measuring outcomes is an important source of contractual incompleteness.

The team might be a group of professionals sharing a practice (common among doctors and lawyers) or employees of a restaurant who all contribute to putting the meal on the table. In recognition of the team nature of production many firms base a portion of compensation to employees on some measure of team output. A major airline, for example, paid the ground staff who manage departures and arrivals of planes a team bonus based on the fraction of on-time departures.

For concreteness think about a group of *n* software engineers working together to write code for new applications. Their production has the following characteristics, similar to that in Chapter 11:

- *Disutility of Effort*: The engineers devote effort (which they consider to be onerous, that is, a bad because it results in disutility) to producing an output.
- *Effort produces a good:* They sell the output and with the resulting sales revenue they purchase goods that they consume for utility.

But we have the following conflict:

- *Total team output is verifiable*: Information on the total output the team produces is observable and verifiable.
- *Individual effort is unverifiable*: Effort levels of individual engineers and the quality of their contributions to the team's output are not verifiable.

Because of the conflict between the verifiability of total output versus the unverifiability of individual effort, the team confronts a problem: how to design a compensation system that will motivate the members to work hard and well even when their effort is unverifiable. The problem they face is different from the public goods problem explored above, for two reasons.

- *A key input is not verifiable:* In the public goods example we assumed that the mechanism designer could *observe* the levels of contribution to the public good and subsidize each member's contribution. The social planner could motivate the desired level of total effort contributed by fines or subsidies based on the amount worked. But because information on the level of effort of each team member is not verifiable, team members *cannot* enforce this kind of contract – either a work level or subsidies and fines – on themselves.
- *The output of the team is rival*: When the team sells its output, the revenues it receives will be divided among the members somehow. Each dollar more that a team member gets is one dollar less another team member gets. The team's output is *not* a public good.

Checkpoint 16.7: Team production in your experience.

Describe a team production process in which you or someone you know have engaged. Be clear about how this process fits the definition of team production.

How team production works

EXAMPLE Think about some team sport, football (soccer) for example. Each player on a team contributes to the team's success in winning games. Suppose you were asked to determine the contribution of each player to this objective. Can you think of a way to do this? You could, for example, substitute other players in games.

Each team member contributes some amount of effort to the team's production process, e^i, with $i = 1 \dots n$ for the n members of the team. The team produces an output, q that is just the sum of all of the contributions of the team members multiplied by a positive constant γ.

$$\text{Production:} \qquad q \quad = \quad \gamma(e^1 + e^2 + \dots e^n) \tag{16.11}$$

From the production function you can see that γ is the average productivity of effort and because this is a constant, it is also the marginal productivity of effort. To simplify we assume the team has no costs other than paying its members an amount y^i, member i's income from the team. Then the utility of team member i is:

$$\text{Utility:} \qquad u^i \quad = \quad u(y^i, e^i) \tag{16.12}$$

Which says that an individual team member's utility is derived from income (a good) and the amount and quality of effort contributed to the project (a bad). Recall the following:

- having more of the good results increases the individual's utility, so the marginal utility of income is positive, or $u_y > 0$
- more of the bad reduces utility (it is results in a *disutility*), so the marginal utility of effort is negative, or $u_e < 0$.

We will use an explicit utility function that expressees the above relationships:

$$u^i = y^i - \frac{1}{2}e^{i2} \tag{16.13}$$

To discover the marginal utilities of both income and effort, we differentiate this utility function, finding that

- the marginal utility of income u_y is equal to 1; and
- the marginal utility of effort u_e is $-e^i$ or what is the same thing, the marginal *dis*utility of effort $-u_e$ is e^i itself.

The second bullet makes sense: if you are not working hard, the disutility of working harder is not very great; but if you are providing a lot of effort, the marginal disutility to you is substantial.

The socially optimal effort by a team member

Suppose the members of the team ask a mechanism designer to advise them on the the compensation system they should adopt, determining the income of each member. They would like the compensation system to provide each member incentives to contribute to the team an amount of effort that will maximize the total utility of its members.

Just as in the case of the public good that you studied earlier in this chapter, the designer would proceed in two steps:

- First, she would determine what is the amount of work effort that if implemented by each team member would maximize the total welfare (sum of individual utilities) of the team; and then

- second, she would find (if one exists) a set of incentives (a “mechanism”) that would motivate of the team members – each maximizing their own utility – to implement the socialy optimal amount of effort.

M-Note 16.7 shows how the mechanism designer would determine the socially optimal level of work effort by each team member. In the case there are just two of them, A and B, and the marginal private cost (disutility) to each of providing effort is the level of effort itself.

The level of effort that the mechanism designer determines to maximize the sum of their utilities is that

Marginal disutility of effort $= e^A = e^B = \gamma =$ Marginal productivity of effort

This is the same thing as requiring that the marginal private cost of effort for each of the two team members is equal to the marginal social benefit (including the external benefits that each one working confers on the other.)

M-Note 16.7: The socially optimal effort of a team member

Supose $n = 2$ so the team is composed of just A and B. The mechanism designer selects e^A and e^B to maximize the sum of their two utilities:

$$W = (e^A + e^B)\frac{\gamma}{2} - \frac{1}{2}e^{A2} + (e^A + e^B)\frac{\gamma}{2} - \frac{1}{2}e^{B2} \tag{16.14}$$

Differentiating W with respect to e^A and e^B, and setting the result equal to zero we have:

$$\begin{aligned} \frac{\delta W}{\delta e^A} = \frac{\gamma}{2} - e^A + \frac{\gamma}{2} &= 0 \\ \frac{\delta W}{\delta e^B} = \frac{\gamma}{2} - e^B + \frac{\gamma}{2} &= 0 \\ \text{marg. priv. benefit - marg. priv. cost + marg. ext. benefit} &= 0 \end{aligned}$$

So, rearranging the above equations we find that $e^A = e^B = \gamma$.

The $\frac{1}{n}$ problem

Having figured out the optimal level of effort –each member should provide a level of effort equal to γ – is a start. But as we have already seen, the team members cannot simply agree to all work at that level. The reason is that information about individuals' effort levels is not verifiable and hence cannot be used in any enforceable agreement or contract.

REMINDER Return to Chapter 2 and think about the experimental games that people played. What kind of game is team production with equal sharing of the team's revenue like? What were the predictions about the ways in which self-interested players would play? What happened in reality when people played the game and what kinds of preferences might people have as a result?

They first explore the idea that the income of each team member is the total output of the team divided by the number of team members. So member i receives

$$y^i = \frac{q}{n} \tag{16.15}$$

$$= \frac{\gamma(e^1 + e^2 + \ldots + e^i + \ldots + e^n)}{n} \tag{16.16}$$

Notice that if he works harder not only does he receive more income (because his effort contributes to q): so does everyone else on the team. In fact they each get the same share, namely, $\frac{1}{n}$ of the contribution of his greater effort to increased team output. So we have the following results:

- The *marginal private benefit* of working harder is $\frac{\gamma}{n}$ ...
- the *marginal external benefit* of working harder (the increase income enjoyed by $n-1$ other members) is $(n-1)\frac{\gamma}{n}$ ...
- and the *marginal social benefit* is the sum of the two, or just γ itself.

How hard will the team member work under the q/n compensation system? Here the following rules give the answer:

- If the marginal private benefit of working harder ($\frac{\gamma}{n}$) exceeds the marginal cost (e), she will work harder, and
- ... if the marginal private benefit of working harder is less than the marginal cost, she will work less hard, ...
- so, she will work at the level of effort that equates the marginal private benefits and costs of working harder

She will set $e = \frac{\gamma}{n}$. Each other team member will choose the same level so in the Nash equilibrium of this game, we will have $e^N = \frac{\gamma}{n}$. We clarify the result mathematically in M-Note 16.8.

The Nash equilibrium level of effort is one n^{th} the socially optimal level of effort, which is γ. The fact that she is not compensated for the external benefits that her work confers on other team members means that she (and the other team members) will not put in as much effort as the they would if they could agree on how much each would work. Every worker could be better off if each worked harder.

The team production problem is another coordination failure. The external effect of working harder and how it results in effort being too low is similar to the external effect of the citizen's contribution to the public good and how it means the public good will be under-provided.

M-Note 16.8: The individual chooses an effort level

Each of the two team members maximize their utility function, which for A is:

$$\text{Maximize} \quad u^A = y^i - \frac{1}{2}e^{A^2}$$
$$= (e^A + e^B)\frac{\gamma}{2} - \frac{1}{2}e^{A^2}.$$

To find the level of effort that maximizes utility we differentiate this function

$$\frac{\partial u^A}{\partial e^A} = \frac{\gamma}{2} - e^A \qquad \text{set the result equal to zero}$$
$$\frac{\gamma}{2} - e^A = 0$$

The same reasoning is true for team member B. Therefore we have

$$e^A = e^B = \frac{\gamma}{2}.$$

or the marginal private cost is equal to the marginal private benefit. This is the Nash equilibrium level of effort, or

$$e^N = \frac{\gamma}{n}.$$

Checkpoint 16.8: Team size and the extent of the coordination failure

- Use the equation that gives the socially optimal level of an individual's contribution to explain why there is no coordination failure if the team has a single member.
- Explain in words why there is no coordination if the team has only one member.
- Explain why the extent of the coordination failure – the difference between the socially optimal amount and the amount an individual will contribute under the q/n compensation system – is proportional to the number of team members, n.

Internalizing external effects by mechanism design

We have shown how paying team members a fraction of the output proportional to the total output (one n^{th}) results in a Pareto inefficient outcome. We therefore want to ask: is there a payment system can motivate individual team members to implement the socially optimal level of effort? Surprisingly, there is: pay each member the entire value of the output of the team, minus a constant sum.

So now each team member would be paid:

$$\begin{aligned} \text{Payment} &= \text{Total output} - \text{Constant} \\ y^i &= \gamma(e^1 + e^2 + \ldots + e^n) - k \end{aligned} \quad (16.17)$$

Subtracting a constant sum (k) in Equation 16.17 is necessary to balance the team's budget or even to turn a profit. Otherwise, they would pay out *n* times the team's total revenue.

This seemingly bizarre mechanism ensures the following:

- *Compensation:* any contribution by a member to the output of the team will be exactly compensated.
- *Incentives:* The compensation gives each team member the same incentives as an isolated individual who owns the entire fruits of his labor (what we previously called an "owner-operator").

M-Note 16.9: A socially optimal contract

To see how this mechanism works, we now rewrite the utility function of team member A as follows, substituting in $y = \gamma(e^A + e^B) - k$:

$$u^A = \gamma(e^A + e^B) - k - \frac{1}{2}e^{A2}, \quad (16.18)$$

To find the amount that will maximize the utility of member A under this compensation system we differentiate the utility function

$$\frac{\delta u^A}{\delta e^A} = \gamma - e^A$$

and set the result equal to zero, finding that $e^A = \gamma$. The contract is optimal: the private solution is the same than the solution that the mechanism designer aims.

M-CHECK To see that under the optimal contract the first order condition for choosing the team member's effort is exactly the same as Equation **??** just differentiate Equation 16.18 with respect to the team member's effort level, and set the result equal to zero .

As in the case of the optimal subsidy for the contribution to the public good, the compensation system proposed by the mechanism designer succeeds in internalizing the positive external effects of each member's work effort. It does this because the individual acting alone now is treating the external benefits that their effort confers on others as if it was their own income, *because it is*!

Checkpoint 16.9: Team production and public goods

Go back to Figure 16.2 which contrasts the socially optimal level of contribution to the public good with that which the citizen will provide in the absence of a subsidy.

- Draw a similar figure with team member's effort on the horizontal axis and dollars measuring marginal private and social benefits, and private costs of effort on the vertical axis.

- Draw in lines for the marginal cost of effort, the marginal private benefit of effort (when team members receive q/n as their income) and marginal social benefits.
- Use the figure to explain why paying the each team member q/n results in a coordination failure and why paying them $q - k$ eliminates the coordination failure.
- The social benefit of the public good per person ($na^N\gamma$) increases as the size of the population increases (with each citizen contributing an unchanged amount, a^N) but this is not the case in the team production case where the total output per person is just $e^N\gamma$ and does not depend on the size of the team. Why does the public goods game differ from the team production game in this way?

Risk and credit constraints: Mechanism design in an imperfect world

How clever! You might think. But you have probably never heard of such a compensation system in practice. To see why we do not see this kind of contract, we can introduce some real-world risk and credit market constraints on borrowing by the team members.

Suppose the team's production depends not only on the sum of the team members' efforts but also on chance events affecting production but not controlled by team members. Introducing risk changes the problem in unexpected ways. It means that in any period of time – say, week or month – the actual output of the firm – called its realized team income – may be either higher or lower than the average income of the firm over a longer period, say a year – called its expected income.

REMINDER Shirking is the behavior of not exerting effort or not exerting *enough* effort. Someone who shirks is called a shirker.

Call these positive or negative chance events 'shocks' and suppose, realistically, that like each member's effort, the shocks are not observable (or at least not verifiable) so the members cannot determine whether the team's unexpectedly low output in some year comes from bad luck or from workers who shirk.

Checkpoint 16.10: How team projects are (optimally) graded

Consider a team project in which 5 students conduct some research and write up a single paper. The students all receive the same grade on the paper depending on the teacher's assessment of its quality. Explain how this system of 'compensation' is similar to or differs from the optimal compensation for the team production members described above.

A positive shock may also raise the level of total output significantly. So, for teams of any significant size, each member's realized income (total output minus some large constant sufficient to balance the budget and to satisfy members' participation constraints) in any period could be many times larger than the workers' next-best alternative. For example, if a positive shock increases γ by 5 percent and there are 100 members of the team, then the income of the worker in a period with a positive shock could be many times larger than his average income. But the potentially large realized income is not the problem. A negative shock could mean that that the worker's realized income could also be much less than what she would have received in some alternative employment. It could, in fact, be a large negative number, meaning that the member would have to pay the team a substantial sum rather than the team paying the member.

M CHECK Here is an illustration. The team has one hundred members and the expected value of output of the firm is 100. So each team member is paid the realized value of output minus k=99. With this plan, on average the entire revenue is paid to the team members. But if a negative shock reduces team output to 95, then team members would owe the team 4 each, that is four times their expected income!

The problem arises because, for this mechanism to work, the pay of each member has to be tied to the entire team's realized output. But, both negative and positive shocks to total output would realistically dwarf any individual's average compensation in the long run. A contract under which, in some periods, a team member would not be paid and instead would be required to pay the team a substantial multiple of her expected salary is not likely attract many workers even if they were only modestly risk averse. No contract of this type would be voluntarily accepted by the team members.

If the team members could borrow an unlimited amount of money in the bad periods they might be okay with such a contract. But we know from Chapter 12 that this is impossible given that people of modest wealth are either credit market excluded or limited in how much they can borrow. So, implementing the optimal contract for the team members just displaces the contractual challenge to the analogous problem in the credit market: the incompleteness of contracts.

REMINDER A second-best outcome is an outcome that is Pareto-inferior to another ("first best") outcome that while technically feasible, cannot be implemented.

This is a reminder that economists, policy-makers and citizens concerned about economic policy all work in a *second-best* world, one in which ideal models are important in teaching basic ideas, but differ importantly from the best that the policy maker can do.

16.7 The limits of incentives: Crowding out and crowding in.

As you have seen from this chapter, mechanism design is all about incentives. And we have seen that well designed incentives can motivate citizens to contribute to a public good, to smoke less, and even to make decisions about where to live that can sustain ethnically integrated residential communities. But we have also seen that there are limits to what incentives can accomplish:

- In the double auction set up there are no incentives that can make truth telling a dominant strategy, and as a result of buyers and sellers misrepresenting their true values, some mutually beneficial exchanges will not occur.

- In the team production case there does exist a set of incentives that would motivate all members to contribute a Pareto efficient level of effort to the teams production; but the optimal contract that would introduce these incentives would expose team members to extraordinary risks and would never be accepted.
- And recall from Chapter **??** that imposing a fine on parents arriving late to pick up their kids at Haifa daycare centers led to a doubling of lateness, not a reduction; the introduction of incentives apparently crowded out social preferences.

HISTORY In honoring Myerson (along with Eric Maskin and Leo Hurwicz) the Prize Committee of the Royal Swedish Academy of Sciences summarized their findings: "no incentive compatible mechanism which satisfies the participation constraint can produce Pareto-efficient outcomes... In a large class of models Pareto efficiency is incompatible with voluntary participation, even if there are no public goods."

Crowding out

The most plausible explanation of crowding out in the Haifa case is that the fine changed the way the parents thought about lateness. Before the fine parents may have considered arriving on time to be something that is the "right thing to do" out of kindness to the staff of the daycare center. The fine may have suggested that lateness is simply something that is ok to purchase as long as you pay the price. In other words the message of the fine was "picking up your kids on time is like shopping, not like how you treat family or neighbors." The incentive provided a frame for the decision suggesting appropriate behavior.The result is what psychologists call *moral disengagement.*

MORAL DISENGAGEMENT is a process by which in some particular situations people come to feel that ethical considerations need not be applied to their actions.

But there are other reasons by crowding out occurs. Incentives have a purpose, and because the purpose is often evident the target of the incentives may infer information about the person who designed the incentive, about his beliefs concerning her (the target), and about the nature of the task to be done. As the social psychologist Mark Lepper and his coauthors wrote, incentives may affect preferences: because they indicate "the presumed motives of the person administering the reward."

By implementing an incentive, say, as an employer, you reveal information about your intentions (own-payoff-maximizing versus fair-minded, for example) as well as beliefs about the target (hardworking or not, for example) and the targeted behavior (how onerous it is, for example). This information may then affect the target's motivation to undertake the task.

The Boston fire commissioner's threat to dock the pay of firemen accumulating more than fifteen sick days (also from Chapter **??**) conveyed the information that he did not trust that the firemen were doing their very best to come to work, especially not on Mondays and Fridays. For the firemen, the new situation–working for a hostile boss–seems to have altered their motivation. In other words, the threatened reduction in pay conveyed some bad news: "the commissioner does not trust you." You will recall that the firemen responded with a pronounced spike in sick call-ins.

This "bad news" effect commonly occurs in relationships between a principal, who designs incentives (a wage rate, a schedule of penalties for late delivery of a promised service, and so forth), and an agent, who is being induced to behave more in the principal's interest than the agent otherwise would. To do this, the principal must know (or guess) how the agent will respond to each of the possible incentives he could deploy, for example the employees best (effort) response to each wage the employer may offer. The agent knows this, of course, and hence can ordinarily figure out what the principal was thinking when he chose one particular incentive over other possible ways of affecting the agent's behavior.

Here is an example of how this sometimes does not work out well in practice. In this experiment – called a Trust Game – German students in the role of 'investor,' the principal, were given the opportunity to transfer some amount to the agent, called the 'trustee.' The experimenter then tripled this amount. Then the trustee, knowing the investor's choice, could in turn return some (or all or none) of this tripled amount, returning a benefit to the investor.

When the investor transferred money to the trustee, he or she also specified a desired level of back-transfer. In addition, the experimenters implemented an incentive treatment: in some of the experimental sessions, the investor had the option of declaring that he would impose a fine if the trustee's back-transfer were less than the desired amount.

In this "fine treatment," the investor had a further option to decline to impose the fine, and this choice (forgoing the opportunity to fine a nonperforming trustee) was known to the trustee and taken before the trustee's decision about the amount to back-transfer. There was also the standard "trust" condition, in which no such incentives were available to the investor. Figure YY summarizes the results.

In the trust condition, trustees reciprocated generous initial transfers by investors with greater back-transfers. But stating the intent to fine a non-compliant trustee actually reduced return transfers for given levels of the investors transfer. The use of the fine appears to have diminished the trustees feelings of reciprocity toward the investor. Even more interesting is that renouncing use of the fine when it was available increased back-transfers (given the amount transferred by the investor).

Only one-third of the investors renounced the fine when it was available; their payoffs were 50 percent greater than those of investors who used the fines. The bad-news interpretation suggested by the authors of the experiment is that both in the trust condition and when the investor renounced the fine, a large initial transfer signaled that the investor trusted the trustee. The threat of the fine, however, conveyed a different message and diminished the trustee's reciprocity.

There are lessons here for the design of institutions and organizations.

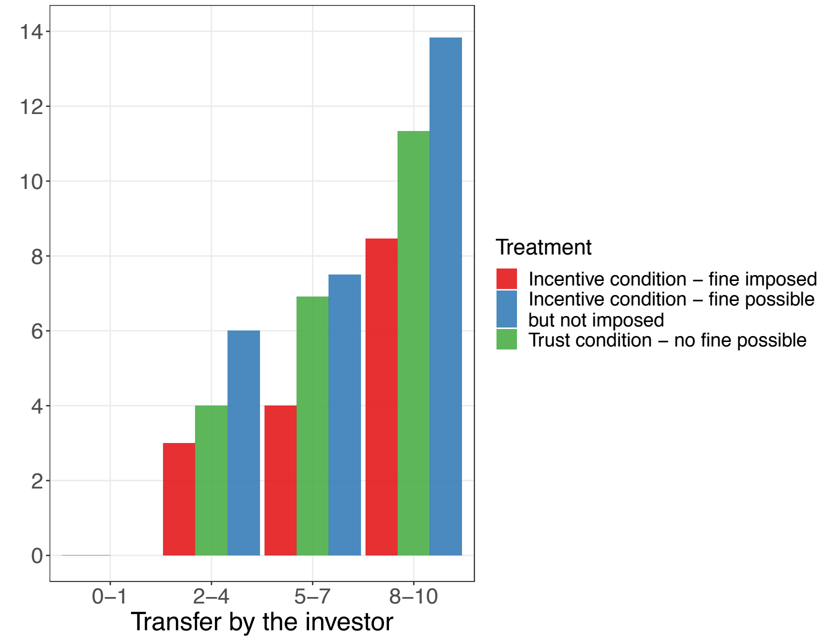


Figure 16.8: **Reciprocity in the trust game**. The height of each bar shows how much the second mover returned given how much the first mover transferred to the second mover. The figure shows that the second mover reciprocated the first mover's trust and or generosity: larger initial transfers were associated with larger back transfers. It also shows that threatening the use of a fine reduced reciprocity and foregoing the use of the fine when it was available increased reciprocity.

Crowding out as a result of the bad-news effect may be prevalent in principal-agent settings but can be averted where the principal has a means of signaling fairness or trust in the agent. The Trust Game experiment even gives us a glimpse of how incentives could crowd in social preferences: the non-use of an available fine resulted in greater reciprocity by the trustees than occurred when fines were not in the picture.

A third reason for crowding out is that people value their own autonomy and may feel that incentives are designed to limit their freedom of action, provoking a negative response. This is called "control aversion" which, analogous to risk aversion, is a preference for self-determination and a negative valuation of any attempt to control the individual.

FACT CHECK In a famous experiment in social psychology, children who did paintings expecting a monetary reward later chose to paint (as opposed to some other kind of activity) half as often as children in a control group in which there had been no experience of a reward. This is an example of control aversion.

To summarize: The problem of crowding out may arise

- when the information that an incentive conveys is off-putting about the person imposing the incentive ("bad news") or . . .
- when the presence of the incentive frames the problem as one in which self-interested motives are acceptable or even called for ("moral disengagement") or . . .
- when the incentive appears to be an attempt to control the individual who responds by acting contrary to the incentive in order to affirm their self determination ("control aversion").

The economics and psychology of getting and being

A combination of game theory and social psychology may help us better understand how incentives work and sometimes why they do not, that is crowding out.

Recall that in a sequential Prisoner's Dilemma, the second mover most often mimics the first mover, reciprocating cooperation or defection depending on what the first mover did. The fact that second movers reciprocate cooperation means that in comparison with what he would have received had the second mover defected they place a positive subjective value either on jointly cooperating per se or on the payoffs that will be received by the cooperative and trusting first mover, should the second mover cooperate. This value is sufficient to offset the higher payoff the second mover could receive by defecting on the cooperator, which is why they cooperate.

When second movers defect on defecting first movers, they are taking the action that maximizes payoffs under the circumstances, but evidently are not doing so from exclusively acquisitive motives. The same individuals would have forgone payoffs that could have been gained by defecting, in order to cooperate with a cooperator, as we have seen. But cooperating with a defector has a different meaning, identifying the second mover as a "loser," someone easily taken advantage of. Thus, part of the motivation behind the "mimic the first move" pattern that we observe is what the second mover wants to say about herself: "I am the kind of person who rewards those who cooperate and stands up to defectors who would exploit the cooperation of others."

This kind of motivation goes way beyond how we play games.

When people engage in trade, produce goods and services, save and invest, vote and advocate policies, they are attempting not only to get things, but also to be someone, both in their own eyes and in the eyes of others. Sometimes constitutive and acquisitive motives are closely aligned. For example, the second mover in the sequential Prisoner's Dilemma game who defects on a defecting first mover is both making a statement about who she is and maximizing her payoffs.

There are lessons here too about the use of incentives. Acquisitive and constitutive reasons for actions may sometimes clash. And we know from experiments and from observing ourselves and others that being good sometimes is more important to us than doing well in money terms. Responding to an incentive in the manner intended (that is, as a payoff maximizer) may make the responder a victim. But not always. A self-interested response to an incentive may constitute the actor as a good citizen or an intelligent shopper, indicating that constitutive and acquisitive motives were closely aligned. The same reasoning, we will see, suggests how we can make incentives and social preferences synergistic.

How acquisitive ends interact with constitutive motives may explain why incentives sometimes work exactly as economists predict on the basis of unmitigated self-interest–and sometimes don't.

Crowding in

In the trust game described above, the 'investor's' threatened imposition of a fine if the 'trustee' did not return enough money to the investor reduced rather than increased the amount the trustee sent to the investor, consistent with the "bad news" interpretation of crowding out. But on closer scrutiny the incentive itself — the threatened fine — seems not to have been the problem. When we look at the data from the experiment to see who among the trustees responded negatively to the incentive, it appears that crowding out was almost exclusively a reaction not to the incentive per se, but to the apparent greed of the investor. Crowding out occurred when the back-transfer demanded of trustee would have given most of the joint surplus (total payoffs for the two) to the investor. There was no backlash against the fines threatened by investors who asked for back-transfers that allocated both the investor and the trustee substantial shares of the surplus.

The key difference was the message sent by the fine. Where the stipulated back-transfer would have captured most of the surplus for the investor, the fine conveyed greed. Where it would have split the surplus more equally, the fine conveyed a commitment to fairness, and perhaps the investor's desire not to be exploited by the trustee. The use of the fine to enforce a seemingly unfair demand provided an acquisitive motive to comply, but to the trustee, it also may have transformed the meaning of complying with the investor's stipulated back-transfer. Going along with the investor's demands no longer made the trustee a cooperative and ethical person, as it would have had the investor's demands been modest, but instead possibly a person easily manipulated, or a victim or a loser.

It therefore appears that it was the relationship between the investor and the trustee, not the threatened fine alone, that was the source of crowding out. That suspicion is reinforced by a diametrically opposite reaction to fines in a Public Goods with Punishment experiment that you saw in 2. The imposition of fines by peers who have to pay to levy them, when they had nothing to gain personally from doing so, appear to have crowded in social preferences.

Why is the fine counterproductive when imposed by an overreaching investor in the Trust game, but in many subject pools so effective when imposed by peers in the Public Goods game? A plausible explanation is that when punished by a peer who had nothing to gain by doing so, players saw the fine as a signal of public-spirited social disapproval by fellow group members. If this were the case, targeted free riders would feel shame, which they would re-

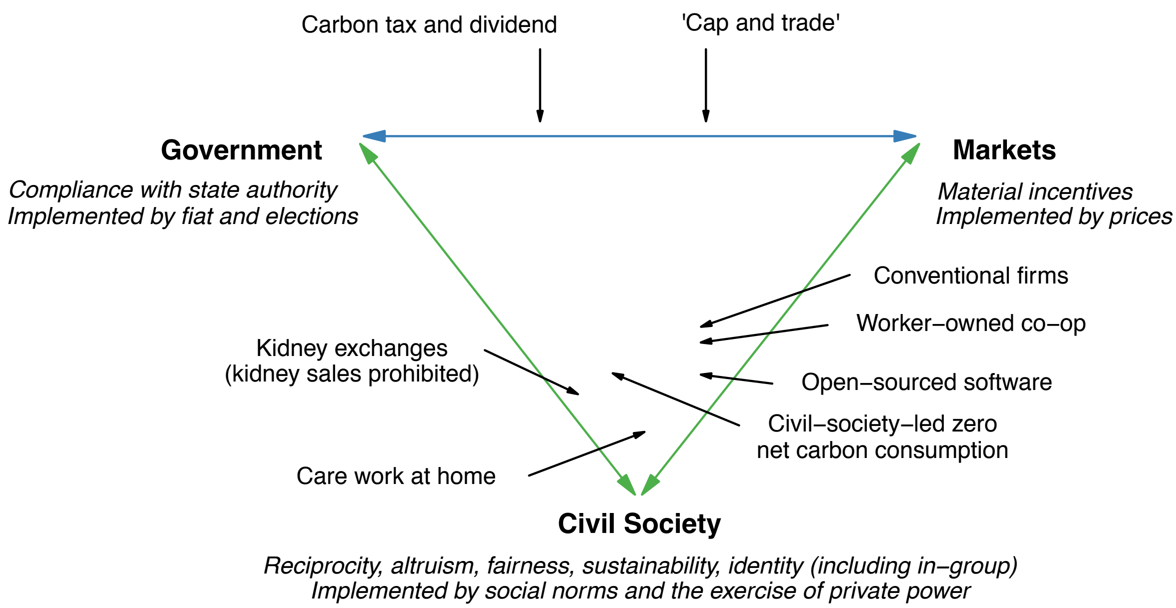


Figure 16.9: **Caption** caption

dress by contributing more. If so, the incentive (the prospect of peer-imposed fines) has crowded in social preferences.

Checkpoint 16.11: Crowding out: Reasons and responses

- Can you think of examples from your own or others' experiences that illustrate the three reasons for crowding out (above)?
- In these examples is there some way that the incentive could have been explained or framed that would have been more likely to produce the intended response?

16.8 Beyond market versus government: Expanding the space for policies and institutions

16.9 Application: A worker-owned cooperative.

When considering the already-existing relationships between owners, managers and workers, one of the puzzles for economists to solve has been the puzzle of why so many firms have the structure that owners of firms own or

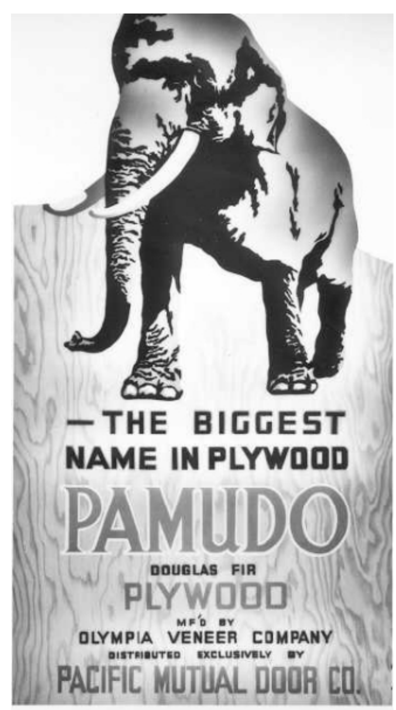


Figure 16.10: **An advertisement for Olympia Veneer, the first cooperatively owned plywood company in the U.S. Pacific Northwest**

rent capital and hire workers to work with that capital (machines, computers, and so on). Why don't we see more examples of the relationship being the other way around: workers hiring managers to monitor their work and renting capital to produce goods so that they, as workers, own the full profits from their effort? This is exactly the problem that Paul Samuelson — one of the path-breakers in modern mathematical economics — meant when he said "Remember that in a perfectly competitive market, it really does not matter who hires whom; so have labor hire capital." We don't see labor hiring capital happen that often. But, in studying the few examples where it has happened — and continues to happen — we can learn some important lessons about capitalism.

In 1921, a group of loggers, carpenters and mechanics in Olympia, Washington in the United States formed the Olympia Veneer plywood cooperative. In return for an investment of $1000, a cooperative member gained the right to work in the plywood plant and to share equally in any profit. Members wishing to leave had to sell their shares, and prospective members, if approved by the membership, were required to purchase shares, which by 1923 were selling for $2550 (Craig and Pencavel, 1992, 1995; Pencavel, 2002).

This was a generation before Samuelson, but these workers certainly would not have agreed that "it really does not matter who hires whom." For them it very much mattered, and they set out to do exactly what Samuelson would later whimsically suggest: "have labor hire capital."

Olympia Veneer was not a capitalist firm: members owned the buildings and equipment that made the plywood. They were their own employers.

The conventional and cooperative plywood firms exemplify differing assignments of the relevant rights. In cooperatives, both residual claimancy and control is assigned to the member-owners who supply labor, while in the conventional firms, the suppliers of capital and labor are distinct individuals, and residual claimancy and control is assigned to the capital suppliers (Putterman and Dow, 2000; Dow, 2002). We use the terms *residual claimancy* and *control rights* rather than the more general term *ownership* to allow for situations in which control rights (disposition of the use of an asset, including its sale and excluding others from its use) and residual claimancy on the income generated by an asset are assigned to different parties.

In 1939, 250 workers in nearby Anacortes invested $2000 each in a second cooperative plywood mill. Strong war-time demand for plywood boosted the value of their shares to $28,000 in 1951, and members were paying themselves at rates double the union wage in nearby conventionally organized (capitalist) plywood mills. Stimulated by the success of Olympia Veneer and Anacortes, between 1949 and 1956 twenty-one more co-ops entered the plywood industry in the states of Washington and Oregon, nine of them by buying out existing conventional firms.

Eventually, though, some co-ops either transformed themselves into *de facto* conventional firms, or sold out to conventional firms. For example, by mid-century the remaining handful of member-owners of Olympia Veneer were employing a thousand workers on conventional wage contracts, remaining a cooperative in name only. In 1954, they sold their shares to the U.S. Plywood Corporation. In the sale, twenty three early members realized a return averaging $652,000 (in 1954 dollars) on their initial investment.

Until the entire industry moved from the North-West to the South-East in the 1980s and 1990s, about half of the plywood firms were co-ops, the rest being conventional firms, some with unionized labor forces and some not. Though the co-ops and conventional firms used virtually identical machinery, the co-ops specialized in the more labor-intensive "sanded" plywood because, as one analyst of the co-ops commented, sanded plywood "puts a premium on worker effort" (Bellas, 1972:30).

The structure of the typical plywood co-op was both egalitarian and democratic. With few exceptions, worker-owners received equal pay, and jobs were often rotated. Management was elected by the body of worker-members. Some non-members were hired under conventional wage contracts, their numbers making up an average of a quarter of the total workforce. High levels of productivity were maintained through a strong work ethic among members, enforced by peer pressure and mutual monitoring. The resulting saving in supervision costs was substantial: when one conventional firm converted to a coop, the number of supervisors was reduced to a quarter of its previous level.

FACT CHECK James Andreoni and Laura Gee explore how mutual monitoring works experimentally in a public goods game where the team members are able to appoint one of the team members as a 'hired gun' with the ability to punish the lowest performing team member. The institution of the hired gun, like a manager in a cooperative, ensured that the experimental subjects cooperated in the public goods game.

Shares being relinquished by retiring or departing members were advertised in local newspapers. Average share prices ranged from the equivalent of a single year's annual earnings to three times that amount. While considerable, these shares were undervalued. The shares could be purchased only by a person wishing to work in the factory, and most of those job-seekers were not wealthy enough to put up a lot of money for a share. This limited the demand for shares and therefore lowered the share prices.

As a result, share values were substantially less than the long-term value of the difference between earnings in the co-ops and in the unionized mills: an individual who purchased a share and worked in a co-op for a number of years had a much higher long-term value of income than an individual who put the value of a share in a Portland savings bank and worked at union wages in a conventional firm.

The coexistence of cooperatives and conventional firms producing the same goods using virtually identical technologies over a period of three quarters of a century provides a remarkable opportunity for us to compare and contrast the success and failures of different institutional structures. Conventional firms and cooperatives alike were able to attract both labor and capital over this period. But the firms differed markedly in a number of ways.

TOTAL FACTOR PRODUCTIVITY is a measure of the share of output produced that *cannot* be explained only by the *amounts* of the factors of production, capital and labor. In the context of co-ops, then, the co-ops that had the same amount of labor and capital inputs as conventional firms had *more* output. This additional output suggests that either the techniques used by co-ops or their method of organization were better for producing output than the techniques or organization of capitalist firms.

Productivity **Total factor productivity** of the co-ops – a measure of the productivity of labor and capital goods combined – was substantially higher (from 6 percent higher to 45 percent higher, depending on the method of estimation).

Equality and security Cooperatives also adjusted to insufficient product demand in a very distinctive way: rather than laying off members, they reduced hours and pay of all workers, thereby spreading the impact of negative shocks among the membership.

In this particular case, contrary to Samuelson, it mattered very much "who hired whom."

Reasons why the co-operative firms were more productive than their conventional competitors include the superior work effort and reduced cost of monitoring of the co-op workers. This would occur because each co-op member shares in the income that they and their fellow workers produce, so as a result:

- worker-owners have a greater incentive to work hard and well; and
- to assist monitoring other workers; and also
- they may experience less disutility of effort because they are working under a system of discipline that they have devised and agreed to, not one imposed by an outsider.

16.10 The distributional impact of public policies: Rent control

RENT CONTROL is a policy regulating the rent that a landlord can charge, most commonly limiting the size of a rent increase that is permitted.

The final policy that we consider is not aimed at internalizing the external effects of the actions that people take so as to address market failures, but instead is motivated by concerns about the distribution of wealth or income.

A policy that is frequently advocated on grounds of fairness is rent control. Rent control is a legally binding limit on the rents that landlords can charge tenants. Landlords (owners of housing that is rented out) are typically much wealthier than the people they rent to. Rent control is advocated as a way to redistribute income from the landlords to the tenants. Adequate housing is also considered by many to be a merit good, one that on moral grounds should be available to all irrespective of their income. This is a second reason commonly proposed in support of rent control.

Rent control laws are common in some major U.S. cities, including Los Angeles, San Francisco, New York and Washington D.C. Rent control is typically bundled with restrictions on the conditions under which a landlord can evict a tenant. In other counties residential rents are also regulated by governments. In Germany, for example, there are legal restrictions on rent increases in excess of the rate of inflation (so that the real value of rents varies little over time).

Rent control has similar economic logic to the minimum wage: it seeks to improve the economic conditions of less well-off people (renters, low wage workers) by imposing a price (a lower rent, a higher wage) that favors their interests. As with the minimum wage example in Chapter 11, a policy change like rent control produces winners and losers.

In 1994 San Francisco voters approved rent control on rental housing built before 1980. Landlords responded by demolishing older rental structures and building new rental housing which was excempt from controls. Many offered their tenants large sums to vacate their apartments, which could then be converted to owner occupied condominiums, which were also not covered by the new law.

By comparing what happened in otherwise similar rental units built before and after this key date, researchers were able to assess the impact of the policy. The gains to tenants in the affected apartments were substantial, averaging between $2300 and $6600 per year, and totaling $214 million annually.

But the reduction in the supply of rental housing resulted in a city wide average increase in rent of 5 percent. So among the losers, in addition to landlords, were those who were renting units not covered by the new law.

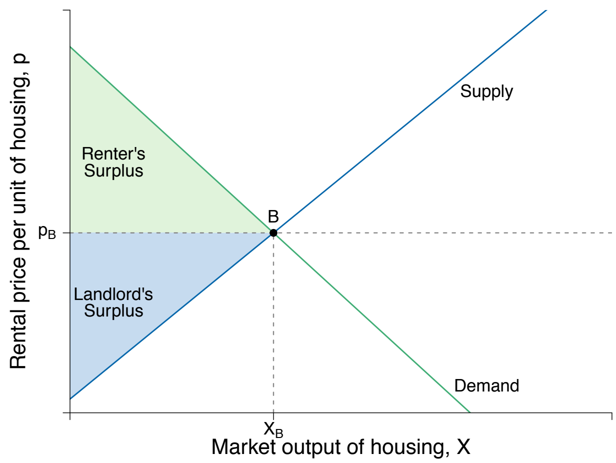
The rental housing market: Renters' surplus and landlords' surplus

We can use the model of supply and demand to study the impact of rent control. Recall that in Chapter 7 the supply and demand model allowed us to identify two components of the gains from trade:

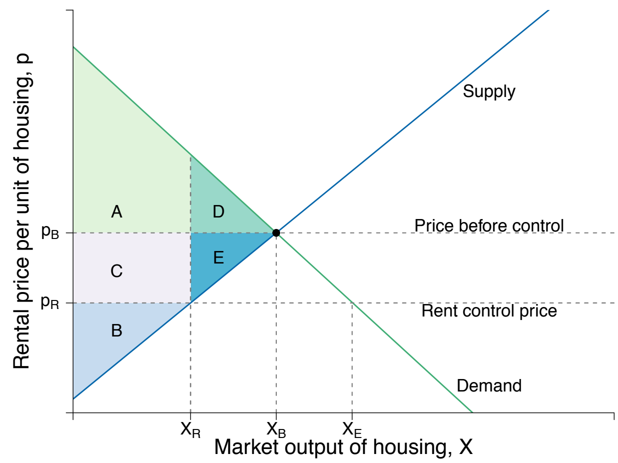
- *consumers' surplus* because for most buyers their willingness to pay exceeded the price they paid, and
- *producers' surplus* because the price at which a good is sold exceeds the marginal cost of its production, contributing to the profits of the owners of the firms producing the good.

Here we adapt those concepts to the rental market, so that we now have

- *renter surplus* arising from the fact that for most renters their willingness to pay for their apartment exceeds the rent they actually pay (similar to consumers' surplus) and



(a) Housing Market clears before rent control



(b) Housing Market with rent-control

Figure 16.11: **A model of supply and demand for housing with rent controls.** A price control lowers the price of housing from P_B to P_R and results in excess demand $(X_E - X_R)$ relative to the original quantity demanded and supplied X_B. Those occupants who have homes obtain an economic rent (consumer surplus) that is greater than they previously would have had at a higher price. Landlords have a lower surplus than before the rent control. As a consequence of the rent control there is lost consumer and producer surplus, called deadweight loss. The deadweight loss is similar to the deadweight loss of taxes that we considered in Chapter 7 and the deadweight loss of the price-making power of firms in Chapter 9.

- *landlord surplus*, which exists because the rent that most landlords receive exceeds the marginal cost of providing a unit of housing to the market (similar to producers' surplus).

The left panel of Figure 16.11 illustrates these concepts at the equilibrium of a hypothetical rental market. To make sense of the model assume that there are two classes of people in a city: landlords and renters. Renters considerably outnumber landlords, which in a democracy gives them the possibility of passing legislation limiting the rents that landlords can charge.

The horizontal axis is the number of units of housing. To simplify we assume all housing units are identical in quality and that landlords are unable to charge different rents to different people so there will be just a single rental price, which is measured on the vertical axis. The supply curve tells you, for any given rent, how many units of housing will be offered. A higher price will bring more units onto the market even in the short run as landlords find ways of converting unused space into apartments. In the long run higher rentals will raise the profitability of owning rental apartments and stimulate new construction.

The demand curve provides the answer to the question: if the rental price is p, how many units of housing will be demanded? At a lower rent, more units are demanded, as more people choose to live in the city, or not to live with their parents or with roommates.

In Figure 16.11 you can see that prior to the introduction of rent control, the rent was p_0 and the number of units rented was X_0, the renters' surplus and the landlords' surplus are as shown by the shaded areas.

Rent control reduces the total surplus and rearranges who gets it

The introduction of the rent control reduces the rental price to p_R and the landlords respond by supplying fewer units, reducing the number available to X_R. With fewer units being rented, notice, the willingness to pay of the 'least willing' renter (the height of the demand curve at X_R), exceeds the marginal cost of putting additional units on the market. Therefore, there are people who would have been willing to rent units beyond the X_R, being offered at a price exceeding the marginal cost. So the demand for rent controlled housing exceeds the supply of rent controlled housing. Rent control has two effects:

- *Redistribution to renters*. A portion of what was before the landlords' surplus, is now part of the renters' surplus. This was the intended effect of the policy.
- *Reduction of the total surplus due to deadweight loss* The deadweight loss (foregone surplus) resulting from the reduced supply of rental housing under rent control is partly lost by renters (the top triangle of the deadweight loss space) and partly by landlords' (the bottom triangle).

The net effect of these two changes is that landlords definitely lost. Their surplus is less than before for two reasons: first, they experienced some of the deadweight loss and second, they transferred some of what was before their surplus to the renters.

The effect on the renters is more complicated to evaluate. Like the landlords, they experienced some deadweight loss, but they also gained some of what was previously landlords' surplus. Their net gain is the green rectangle minus the orange triangle.

In Figure the surplus gained at the expense of the landlords is greater than the deadweight loss experienced by the renters. So, the policy benefited them, as intended, even though it reduced the supply of housing. Rent control is a way of dividing up a smaller pie, with a larger slice going to the renters.

Because the pie is smaller as a result of the deadweight loss, the surplus lost by the landlords must be greater than the surplus gained by the renters. To see this:

$$
\begin{aligned}
\text{Landlord surplus lost} &= \text{Landlord surplus transferred to Renters (rectangle} p_B \text{CD} p_R) \\
&+ \text{ Landlords' share of deadweight loss (triangle CDG)} \\
\text{Renter surplus gained } &= \text{ Landlord surplus transferred to Renters (rectangle } p_B \text{CD} p_R) \\
&- \text{ Renters' share of deadweight loss (triangle BCG)}
\end{aligned}
$$

So:

	Landlord and Renter Surplus gained or lost	*Area in Figure 16.11*	*Amount* ($ '000s)
1	Previously Landlord surplus, now Renter surplus	C	175,000
2	Landlord's share of deadweight loss	E	37,500
3	Renters's share of deadweight loss	D	37,500
4	Landlord surplus lost	C + E	212,500
5	Renter net surplus gained	C - D	137,500

Table 16.2: Monthly gains and losses compared to the no rent control market equilibrium. (Entries in the table are based on Figure 12.16, with p_B = \$1,500, p_R = \$1,000, X_B = \$500,000, X_R = \$350,000), and the price at A = \$2,000.

$$
\begin{aligned}
\text{Landlords' surplus lost - Renter's surplus gained} &= \text{Landlords' share of deadweight loss (triangle CDG)} \\
&+ \text{Renters' share of deadweight loss (triangle BCG)} \\
&= \text{Total deadweight loss (triangle BDG)} > 0
\end{aligned}
$$

Is there a better way to help the less well off? Coasean bargaining

It is also the case that rent control could hurt the less well off, rather than helping them as it did in this example. The costs inflicted on the renters in the form of deadweight loss could have exceeded the gains they made by capturing some of what previously had been the landlords' surplus.

> **Checkpoint 16.12: Distribution of surplus under rent control**
>
> Using a diagram similar to Figure 16.11 sketch supply and demand curves such that the costs experienced by renters (deadweight loss due to a smaller total surplus) will exceed the gains made possible by the fact that they capture a larger share of the surplus.

Is there a better way to help the less well off? Suppose the renters have joined an organization to promote their interests, and their President sits down for a talk with the head of the Landlords' Association.

Because it was she who had asked for the meeting, the renter's representative starts off: "we are willing to vote to rescind the rent control if you and your landlord pals will simply transfer some money to us so that we are as well off or better than we are under rent control."

'How much would you need?' is his reply. She simply hands him Table 16.2.

From Table 16.2 you can see that if the landlords transferred \$137.5 million per month to the renters, the renters would be as well off as they would be under rent control, and the landlords would be much better off (paying \$137.5 million directly to the renters is better than losing a total of \$212.5 million in lower rents and deadweight losses). Buying off the renters by paying them directly (rather than enduring the losses imposed by rent control) would strike the landlord's representative as a bargain.

Of course, the President of the renters group would be quick to point out that were the landlords to transfer $212.5 million to the renters, then the landlords would be no worse off than they were under the rent control, and the renters much better off (getting a transfer of $212.5 million beats the net benefits to the renters from lower rents but on a reduced number of apartments rented).

The two might then bargain and agree on some intermediate amount, under which both landlords and renters would be better off. This illustrates the logic of what you have learned about Coasean bargaining (in Chapter 14).

This episode of course is fanciful. It is difficult of think of how the transfers from landlords to renters could take place in any practical way. But it underlines an important objective: if possible, policies to grant a larger slice of the pie to the less well-off should be designed to make the pie larger or at least not make it smaller.

The data in Figure 15.7 shows that many countries have found ways to give a larger slice to the less well off while also growing the pie. By comparison to the US, France, and Italy, for example, Germany, Norway and Finland, have enjoyed both more rapid growth in average incomes and a larger share of income going to the less well off.

Checkpoint 16.13: Barriers to Coasean bargaining

Imagine that a city council and mayor of a city wanted to implement the kind of bargain we just described between the landlords association and the renters group. Make a list of the practical difficulties that writing such an agreement into an enforceable and unambiguous agreement (whether a private contract or a city ordinance) might face. Why might the landlords' association have greater political influence than the renters' group with a much larger membership, even in a democracy?

Weighing the gains and losses to different groups in society

What these countries have accomplished – granting a larger slice of a larger pie to the less well off – is impressive. But we do not conclude that policies that redistribute the pie, even while shrinking it, should be ruled out. The fact that in the case of rent control, the surplus lost to landlords must exceed the surplus gained by renters is not a reason to oppose the policy: remember, it was intended to help the renters, and it did.

Policies to redistribute income are often advocated on the grounds that providing additional income to one group (typically less well off) is more highly valued by the policy maker or the electorate than the incomes lost by some other (typically higher income). The basic idea here was introduced in Chapter

2? when we considered the difference between cardinal and ordinal utility.

M-Note 16.10: Gains and Losses of Rent Control

It is quite common among economists evaluating policies that affect the incomes of well off and less well off members of the population differently to place a greater value on gains or losses in income made by poorer people than the same gains or losses experienced by richer people.

But how much more?

One way to reason about this is to express these values as statements about the marginal utility of income to each of these groups. A higher value placed on income gained by the poor would the be expressed as saying that their marginal utility of income is higher.

If you treat the utilities of different people as comparable then taxing a person with a low marginal utility of income will inflict less harm on him than the benefit resulting from transferring the tax revenues to someone with a higher marginal utility of income.

If this is your way of addressing the problem, then you might specify some utility function such as

$$\text{Utility function:} \quad u(y) = \alpha \ln(y)$$

where a is some positive constant and y is the individual's income. Then (differentiating this utility function with respect to y and recalling that the derivative of lny with respect to y is just $1/y$) we see that

$$\text{Marginal Utility:} \quad u_y = \frac{\alpha}{y}$$

In this case the marginal utility in income of someone with 50,000 in income is four times as great as the person with 200,000 in income.

If you were to use this approach to evaluate the benefits and costs of the rent control using the numbers in Figure 16. Zz, and if you knew that landlords on average had incomes three times that of their tenants, then you would weight gains of the tenants dollar for dollar as three times as important the losses of the landlords.

Placing a greater weight on the gains of the renters than on the losses of the landlords can be done if the utility of people can be compared. In this case, the policy maker might conclude that the needs that will be more likely to be met by the renter family – more adequate housing, for example – are more important than the reduction in spending – perhaps on a vacation home – that the better off landlord will experience.

Going back to Table 16. .xx you can see that if we placed a value on the gains by the renters that is twice the value placed on the costs to the landlords then the former (2$\times$ 137.5 million) greatly outweigh the latter (212.5 million).

When we compare the gains and losses made to particular people or members of groups, we necessarily make value judgments, that is, we make judgments based on moral or ethical values. This is clear when we are evaluating the dollar gains of the renters as ethically 'more important' than the dollar losses of the landlords. But the point applies with equal force when we simply treat all of the dollar gains and losses as equivalent.

Checkpoint 16.14: Valuing the loses and benefits of rent control

Imagine you are a policymaker considering imposing the rent control whose distributional effects are shown in Table 16.xx. You place a higher value on the gains to the renters than on the losses to the landlords, because you wish to raise the living standards of the less well off (the renters) even at a cost to those who are better off (the landlords). You just saw that if your value on the gains to the renters is twice your value on the losses to the landlords then the benefits of the policy exceed the cost.

What is the smallest value placed on the gains of the renters that would make the benefits of the policy exceed the costs?

16.11 Egalitarian redistribution to address market failures

Redistributing wealth can help to address incentive problems in principal-agent relationships. Redistribution works when wealth disparities are sufficiently great so that a small reduction in the assets of the rich would not stop them from engaging in feasible contracts. Then granting the redistributed assets to poor people would open up opportunities for them that they would not have otherwise.

More egalitarian distributions are likely to be more efficient. Why? Because asset-poor people, not wealthy people, are the people prevented from engaging in efficient contacts. Competitive markets can efficiently allocate resources when potential borrowers are wealthy, ensuring that the wealthy borrower with better ideas and a higher quality project obtains funds for their idea or project.

But we cannot say the same for competitive markets allocating credit contracts to poor people. When an asset-poor person would be the most efficient owner of an asset – for example of a residence, or of farm land – then economic policy could enhance the contractual opportunities of asset poor people.

Sometimes, though, redistributing wealth may not be efficient. For example, concentrated wealth may allow the solution of collective action problems in the provision of public goods. We saw in Chapter 5 that if a single person owned the "lake" in which the fishers worked, then he would have no incentive to over fish, and the other fisher would not have any say in the matter. Similarly, if a firm is owned by a single very wealthy person, she will have strong incentives to monitor the managers of the firm. But if the firm is owned by a large number of owners, then each owner may be tempted to free ride on the monitoring done by the other owners.

A challenge to economic policy and institutional design is to devise ways of addressing the problem of misaligned incentives to perform non-contractual actions when wealth in concentrated.

The source of inefficiency in the principal-agent contracts that we have studied in this and the previous two chapters is that those people performing non-contractual actions – providing quality of a good produced, work effort, prudence in the conduct of a project, for example – are not the residual claimants on the consequences of their actions. Incentives to perform the action well are compromised as a result.

There is little reason to expect that loans will be assigned to those who can make best use of the asset, as we saw in the examples of poor borrowers being unable to fund high-quality projects. Where the allocation of effort to a task by an agent and the allocation of resources to monitoring effort by a principal are non-optimal (as in the model of Chapter 11), a reallocation of residual claimancy and control to the agent may improve the allocation. But this reassignment of rights is sometimes impeded by the agent's restricted access to credit markets. So re-assignment of property rights may fail to solve the effort regulation problem because of credit market failures.

16.12 Why governments sometimes fail: A caveat

This chapter has demonstrated how the concepts and models you have learned can be used to design better public policy. But before concluding, we need to walk back our too-good-to-be-true representation of the mechanism designer. The imaginary figure we introduced deployed the tools of economics to address market failures and unfairness in the economy. But you may have wondered: if the clever mechanisms that she designed are available, then why are market failures so common and why is unfairness such a feature of real economies?

HISTORY In an 1896 essay the Swedish radical reformer and founder of neoclassical economics, Knut Wicksell challenged the common assumptions that policy makers are entirely public spirited "with no thought other than to promote the common weal...members of the representative body are ... precisely as interested in the general welfare as are their constituents, neither more nor less." A century later the libertarian James Buchanan in his Nobel prize lecture underlined this theme: "Wicksell's message was clear, . . . Economists should cease proffering policy advice as if they were employed by a benevolent despot . . . "

The short answer (which is all that we can provide here) is that the mechanism designer is not the government. A conceptual device for understanding ideal policies is not going to be an accurate representation of what governments actually do, any more than the 'auctioneer' describes how markets work, or the 'perfect competitor' gives an accurate picture of how firms really compete. Well designed government policies – things like social insurance and the stabilization of aggregate demand in the macroeconomy – have contributed substantially to the quality of people's lives. And economists can be proud of our contribution to these advances. But there remain significant instances of market failure – global climate change, for example – and perceived unfairness – income dispartities between men and women and the substantial increase in inequality of living standards since the late 20th century in many countries, for example. This is not a limitation of mechanism design but a government failure.

To understand the difference keep in mind:

- The mechanism designer is tasked with maximizing the utilities of citizens or at least avoiding outcomes that are Pareto inefficient when evaluated using the citizen's utility functions. Policy makers in governments have their own private objectives, and there is no reason to expect them be any more other regarding than the citizenry from which they were selected.

- The mechanisms in place to ensure that government policies are designed to promote the interests of citizens (rather the policy makers themselves) – in a democracy, for example, majority rule in fair elections – work imperfectly. This is true in part for intrinsic unavoidable reasons, just as the failure of the double auction market stemmed from an intrinsic problem: the impossibility of ensuring truth telling. But the limits of democratic mechanisms also arise in part because governing elites often succeed in subverting rules of the game intended to make them accountable.

- A successful mechanism implements its outcome by the use of incentives in ways that are consistent with the citizens' participation constraint (as well as their incentive compatibility constraints). The mechanism designer cannot propose policies that literally force citizens to contribute to the public good, or to stop smoking. By contrast, the distinctive characteristic of the government is that it can legitimately coerce people to do or not do particular actions: laws are enforced by the threat of incarceration, that is forced confinement. This ability to legitimately coerce is essential to the government doing its job well – requiring the payment of taxes or the purchase of insurance, or, when required, service in the military, for example. But it also means that the scope for harm done by a government is substantial.

HISTORY. In 1919, Max Weber one the founders of sociology, defined the state as the only entity that has a legitimate monopoly on the use of force within some given geographical territory. This is still the standard definition used in the social sciences.

Checkpoint 16.15: Government: part of the solution/part of the problem

"Any government with sufficient powers to address problems of coordination failures and unfairness will also be powerful enough to pursue the interests of a political elite against the interests of the vast majority."

Comment on this idea, and suggest ways of addressing the problem.

16.13 Conclusion

We began this course with the classical institutional challenge posed by Adam Smith, David Hume and the other great philosopher-economists of the 18th century: How should society be organized? We used modern economic language to update an important strand of the challenge they posed: "How can social interactions be structured so to avoid Pareto inefficient Nash equilibria resulting from people's free choice of their own actions?" Achieving a just distribution of the economy's burdens and bounties was a second strand of the challenge they raised.

Mechanism design carries on this tradition in economics (though substantially narrowed in its scope), seeking new rules of the game to better coordinate how we interact.

The outpouring of new ideas in the field of studies we now call economics in the 18th and 19th centuries was a response to a rapidly changing environment: new technologies introduced by the industrial revolution, new types of social interaction associated with the emergence of urban life and factory production, and new values of individual autonomy stemming from the Renaissance and the first halting steps towards political democracy. Today we face no less epoch-making changes in the world, changes that economics must, as it did then, seek to understand, and then to harness for the betterment of all.

Making connections

- The *Nash equilibrium* is a key concept for policy making: policies work by altering the environment in which people make their decisions, resulting in a Nash equilibrium that differs from the status quo.
- *Internalizing external effects* so as to prevent market failures and other coordination failures is a key idea in mechanism design.
- *Limited information* is both the fundamental reason for coordination failures (because it makes complete contracts impossible) and also a challenge facing the mechanism designer (or policy maker).
- *Fairness and Pareto efficiency* are widely shared values that people would like to see public policy advance, though people differ in the weights they would place on the two values. Some policies can advance the two together (government provided or tax financed subsidies for public goods) while in some cases the values may conflict, imposing a trade off (rent control).
- The *role of models and abstract thinking* in economics is illustrated by the imaginary figure of the public-spirited and well-informed mechanism designer; but the mechanism designer is not the government.
- The *classical institutional challenge* originated in the 17th and 18th century; but it is still with us as we seek to find a way that free people can coordinate their activities so as to achieve desired societal outcomes.

Mathematical Notation

Notation	Interpretation
u	utility function
γ	marginal private benefit of an individual contribution of a public good or effort
n	amount of people
a^i	contribution to a public good of person i
a^N	Nash equilibrium contribution to a public good of each citizen
ω	subsidy to public good
y	income
P, p	price of a good
p_0	price of a houses' rent without rent control
p_R	rental price with rent control
x^A, x^B	amount of cigarettes that A and B smoke
α, β	parameters utility function
x^{AN}, x^{AN}	Nash equilibrium level of smoking of A and B
X	market output
X_0	number of houses rented without rent control
X_R	supply of house to rent with rent control
m	social multiplier
q	output of the team production
e^i	level of effort of i
e^N	Nash equilibrium level of effort
W	aggregate utility function
k	constant sum subtracted to the income of team members
B	maximum willingness to pay of a buyer
b	buyer's offer price, $b < B$
S	seller's minimum price at which is willing to sell
s	seller's minimum announced price at which is willing to sell, $s > S$

Glossary

$\bar{p}$ is the person's *maximum willingness to pay* for good x. She won't pay more than $\bar{p}$ to get a unit of x.

section.12.1

12.0 Usury When a lender charges unreasonably, unethically, or illegally high interest rates, then they are said to be committing usury. Usury, or even charging any interest at all on a loan, is condemned in the scriptures of both Islam and Christianity. . In the United States, individual states have their own usury laws which put a cap on how much interest lenders can charge borrowers in order to protect consumers from usurious lending practices. 637

section.12.1

12.0 Wealth constraints Where borrowers face credit constraints due to their lack of wealth or low income, they are said to face **wealth constraints**. 637

section.12.2

12.0 Credit constraints A family or business is said to be credit constrained if they either

- are excluded from borrowing entirely or
- face limits on how much they can borrow or
- pay extraordinarily high rates of when they do succeed in getting a loan.

637

subsection*.219

12.0 Interest factor The interest factor is one plus the rate of interest. 644

equation.12.4.3

12.0 Wealth The stock of things (or the value of that stock) owned by a person or household that contribute to a flow of income or other benefits, is their wealth. 667

section.12.9

12.0 Income is the maximum a household or person can consume over a given period of time without reducing the value of the household's wealth (their stock of assets such as home or car, minus any outstanding debt). The borrower's income from the project is a part of their income. 667

subsection*.232

12.0 Equity is one's own wealth (rather than borrowed funds) invested in a project. There is a second entirely different use of the term, meaning the character of being fair, as in "an equitable division of the pie." 667

subsection*.232

12.0 Marginal borrower and excluded borrower A *marginal* borrower is a borrower with just enough wealth to be included in a credit contract with a lender. An *excluded* borrower is a borrower with insufficient wealth to obtain a credit contract with a lender and is therefore unable to obtain credit to fund a project. 671

subsection*.234

12.0 Aggregate demand is the sum of expenditures on goods and services produced in a country, including demand from the rest of the world. 679

subfigure.12.17.2

12.0 Monetary Policy Monetary policy implemented by a country's central bank (in the U.S. the Federal Reserve System) affects the rate of interest at which businesses and others can borrow and the amount of borrowing, thereby regulating aggregate demand to moderate the business cycle. 681