PRINCIPLES OF MICROECONOMICS

LibreTexts™

Principles of Microeconomics

The LibreTexts libraries are Powered by NICE CXOne and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

This text was compiled on 04/16/2025

TABLE OF CONTENTS

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing***.

CHAPTER OVERVIEW

1: Welcome to Economics!

This leads us to the topic of this chapter, an introduction to the world of making decisions, processing information, and understanding behavior in markets —the world of economics. Each chapter in this book will start with a discussion about current (or sometimes past) events and revisit it at chapter's end—to "bring home" the concepts in play.

1.1: What Economics Is and Why It's Important

Learning Objectives

- Discuss the importance of studying economics
- Explain the relationship between production and division of labor
- Evaluate the significance of scarcity

Economics is the study of how humans make decisions in the face of scarcity. These can be individual decisions, family decisions, business decisions or societal decisions. If you look around carefully, you will see that scarcity is a fact of life. Scarcity means that human wants for goods, services and resources exceed what is available. Resources, such as labor, tools, land, and raw materials are necessary to produce the goods and services we want but they exist in limited supply. Of course, the ultimate scarce resource is time- everyone, rich or poor, has just 24 hours in the day to try to acquire the goods they want. At any point in time, there is only a finite amount of resources available.

Think about it this way: In 2015 the labor force in the United States contained over 158.6 million workers, according to the U.S. Bureau of Labor Statistics. Similarly, the total area of the United States is $3,794,101$ square miles. These are large numbers for such crucial resources, however, they are limited. Because these resources are limited, so are the numbers of goods and services we produce with them. Combine this with the fact that human wants seem to be virtually infinite, and you can see why scarcity is a problem.

Figure 1.1.1: Homeless people are a stark reminder that scarcity of resources is real. (Credit: “daveynin”/Flickr Creative Commons)

If you still do not believe that scarcity is a problem, consider the following: Does everyone need food to eat? Does everyone need a decent place to live? Does everyone have access to healthcare? In every country in the world, there are people who are hungry, homeless (for example, those who call park benches their beds, as shown in Figure $1.1.1$), and in need of healthcare, just to focus on a few critical goods and services. Why is this the case? It is because of scarcity. Let’s delve into the concept of scarcity a little deeper, because it is crucial to understanding economics.

The Problem of Scarcity

Think about all the things you consume: food, shelter, clothing, transportation, healthcare, and entertainment. How do you acquire those items? You do not produce them yourself. You buy them. How do you afford the things you buy? You work for pay. Or if you do not, someone else does on your behalf. Yet most of us never have enough to buy all the things we want. This is because of scarcity. So how do we solve it?

Every society, at every level, must make choices about how to use its resources. Families must decide whether to spend their money on a new car or a fancy vacation. Towns must choose whether to put more of the budget into police and fire protection or into the school system. Nations must decide whether to devote more funds to national defense or to protecting the environment. In most cases, there just isn’t enough money in the budget to do everything. So why do we not each just produce all of the things we consume? The simple answer is most of us do not know how, but that is not the main reason. (When you study economics, you will discover that the obvious choice is not always the right answer—or at least the complete answer. Studying economics teaches you to think in a different of way.) Think back to pioneer days, when individuals knew how to do so much more than we do today, from building their homes, to growing their crops, to hunting for food, to repairing their equipment. Most of us do not know how to do all—or any—of those things. It is not because we could not learn. Rather, we do not have to. The reason why is something called the division and specialization of labor, a production innovation first put forth by Adam Smith, Figure 1.1.2, in his book, The Wealth of Nations.

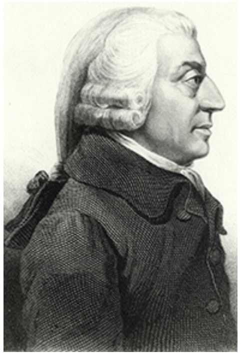


Figure 1.1.2: Adam Smith introduced the idea of dividing labor into discrete tasks. (Credit: Wikimedia Commons)

The Division of and Specialization of Labor

The formal study of economics began when Adam Smith (1723–1790) published his famous book The Wealth of Nations in 1776. Many authors had written on economics in the centuries before Smith, but he was the first to address the subject in a comprehensive way. In the first chapter, Smith introduces the division of labor, which means that the way a good or service is produced is divided into a number of tasks that are performed by different workers, instead of all the tasks being done by the same person.

To illustrate the division of labor, Smith counted how many tasks went into making a pin: drawing out a piece of wire, cutting it to the right length, straightening it, putting a head on one end and a point on the other, and packaging pins for sale, to name just a few. Smith counted 18 distinct tasks that were often done by different people—all for a pin, believe it or not!

Modern businesses divide tasks as well. Even a relatively simple business like a restaurant divides up the task of serving meals into a range of jobs like top chef, sous chefs, less-skilled kitchen help, servers to wait on the tables, a greeter at the door, janitors to clean up, and a business manager to handle paychecks and bills—not to mention the economic connections a restaurant has with suppliers of food, furniture, kitchen equipment, and the building where it is located. A complex business like a large manufacturing factory, such as the shoe factory shown in Figure 1.1.3, or a hospital can have hundreds of job classifications.

Figure 1.1.3: Workers on an assembly line are an example of the divisions of labor. (Credit: Nina Hale/Flickr Creative Commons)

Why the Division of Labor Increases Production

When the tasks involved with producing a good or service are divided and subdivided, workers and businesses can produce a greater quantity of output. In his observations of pin factories, Smith observed that one worker alone might make 20 pins in a day, but that a small business of 10 workers (some of whom would need to do two or three of the 18 tasks involved with pin-making), could make 48,000 pins in a day. How can a group of workers, each specializing in certain tasks, produce so much more than the same number of workers who try to produce the entire good or service by themselves? Smith offered three reasons.

First, specialization in a particular small job allows workers to focus on the parts of the production process where they have an advantage. (In later chapters, we will develop this idea by discussing comparative advantage.) People have different skills, talents, and interests, so they will be better at some jobs than at others. The particular advantages may be based on educational choices, which are in turn shaped by interests and talents. Only those with medical degrees qualify to become doctors, for instance. For some goods, specialization will be affected by geography—it is easier to be a wheat farmer in North Dakota than in Florida, but easier to run a tourist hotel in Florida than in North Dakota. If you live in or near a big city, it is easier to attract enough customers to operate a successful dry cleaning business or movie theater than if you live in a sparsely populated rural area. Whatever the reason, if people specialize in the production of what they do best, they will be more productive than if they produce a combination of things, some of which they are good at and some of which they are not.

Second, workers who specialize in certain tasks often learn to produce more quickly and with higher quality. This pattern holds true for many workers, including assembly line laborers who build cars, stylists who cut hair, and doctors who perform heart surgery. In fact, specialized workers often know their jobs well enough to suggest innovative ways to do their work faster and better.

A similar pattern often operates within businesses. In many cases, a business that focuses on one or a few products (sometimes called its "core competency") is more successful than firms that try to make a wide range of products.

Third, specialization allows businesses to take advantage of economies of scale, which means that for many goods, as the level of production increases, the average cost of producing each individual unit declines. For example, if a factory produces only 100 cars per year, each car will be quite expensive to make on average. However, if a factory produces $50,000$ cars each year, then it can set up an assembly line with huge machines and workers performing specialized tasks, and the average cost of production per car will be lower. The ultimate result of workers who can focus on their preferences and talents, learn to do their specialized jobs better, and work in larger organizations is that society as a whole can produce and consume far more than if each person tried to produce all of their own goods and services. The division and specialization of labor has been a force against the problem of scarcity.

Trade and Markets

Specialization only makes sense, though, if workers can use the pay they receive for doing their jobs to purchase the other goods and services that they need. In short, specialization requires trade.

You do not have to know anything about electronics or sound systems to play music—you just buy an iPod or MP3 player, download the music and listen. You do not have to know anything about artificial fibers or the construction of sewing machines if you need a jacket—you just buy the jacket and wear it. You do not need to know anything about internal combustion engines to operate a car—you just get in and drive. Instead of trying to acquire all the knowledge and skills involved in producing all of the goods and services that you wish to consume, the market allows you to learn a specialized set of skills and then use the pay you receive to buy the goods and services you need or want. This is how our modern society has evolved into a strong economy.

Why Study Economics?

Now that we have gotten an overview on what economics studies, let's quickly discuss why you are right to study it. Economics is not primarily a collection of facts to be memorized, though there are plenty of important concepts to be learned. Instead, economics is better thought of as a collection of questions to be answered or puzzles to be worked out. Most important, economics provides the tools to work out those puzzles. If you have yet to be been bitten by the economics "bug," there are other reasons why you should study economics.

- Virtually every major problem facing the world today, from global warming, to world poverty, to the conflicts in Syria, Afghanistan, and Somalia, has an economic dimension. If you are going to be part of solving those problems, you need to be able to understand them. Economics is crucial.
- It is hard to overstate the importance of economics to good citizenship. You need to be able to vote intelligently on budgets, regulations, and laws in general. When the U.S. government came close to a standstill at the end of 2012 due to the "fiscal cliff," what were the issues involved? Did you know?
- A basic understanding of economics makes you a well-rounded thinker. When you read articles about economic issues, you will understand and be able to evaluate the writer's argument. When you hear classmates, co-workers, or political candidates talking about economics, you will be able to distinguish between common sense and nonsense. You will find new ways of thinking about current events and about personal and business decisions, as well as current events and politics.

The study of economics does not dictate the answers, but it can illuminate the different choices.

Key Concepts and Summary

Economics seeks to solve the problem of scarcity, which is when human wants for goods and services exceed the available supply. A modern economy displays a division of labor, in which people earn income by specializing in what they produce and then use that income to purchase the products they need or want. The division of labor allows individuals and firms to specialize and to produce more for several reasons:

a. It allows the agents to focus on areas of advantage due to natural factors and skill levels
b. It encourages the agents to learn and invent
c. It allows agents to take advantage of economies of scale.

Division and specialization of labor only work when individuals can purchase what they do not produce in markets. Learning about economics helps you understand the major problems facing the world today, prepares you to be a good citizen, and helps you become a well-rounded thinker.

References

Bureau of Labor Statistics, U.S. Department of Labor. 2015. "The Employment Situation—February 2015." Accessed March 27, 2015. http://www.bls.gov/news.release/pdf/empsit.pdf.

Williamson, Lisa. “US Labor Market in 2012.” Bureau of Labor Statistics. Accessed December 1, 2013. http://www.bls.gov/opub/mlr/2013/03/art1full.pdf.

Glossary

division of labor
: the way in which the work required to produce a good or service is divided into tasks performed by different workers

economics
: the study of how humans make choices under conditions of scarcity

economies of scale
: when the average cost of producing each individual unit declines as total output increases

scarcity
: when human wants for goods and services exceed the available supply

specialization
: when workers or firms focus on particular tasks for which they are well-suited within the overall production process

Contributor

- Template:ContribOpenSTAXMicroEcon

1.2: How Economists Use Theories and Models to Understand Economic Issues

Learning Objectives

- Explain the Economists Toolkit, which includes abstraction, models and theories
- Interpret a simple model known as "The Circular Flow" diagram
- Describe goods and services markets and labor markets
- Apply these concepts to the American economy

John Maynard Keynes (1883–1946), one of the greatest economists of the twentieth century, pointed out that economics is not just a subject area but also a way of thinking. Keynes, shown in Figure 1.2.1, famously wrote in the introduction to a fellow economist's book: "[Economics] is a method rather than a doctrine, an apparatus of the mind, a technique of thinking, which helps its possessor to draw correct conclusions." In other words, economics teaches you how to think, not what to think.

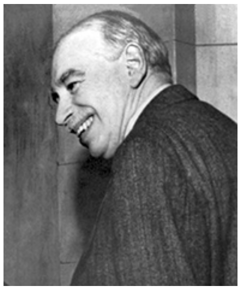


Figure 1.2.1: One of the most influential economists in modern times was John Maynard Keynes. (Credit: Wikimedia Commons)

The Economists Tool Kit

Economics is a social science which is different from the natural and physical sciences because it studies human behavior. Of course, humans behave in differing ways and with many more complications at times than the natural environment. Therefore, economists and other social scientists use several tools. They study history to better understand current and future events and institutions. They use mathematic reasoning to help explain events and phenomenon that they observe. They also use statistics to make inferences about data that they collect.

They also use a technique known as abstraction to simplify their tasks. Abstraction means to oversimplify a problem so that a researcher only focuses on the primary or important details in order to ignore the unnecessary ones.

Abstraction helps economists create economic models that can more easily represent and explain the complex world that they are studying. In more academic language, an economic model is a theoretical construct representing ***economic*** processes by a set of variables and a set of logical and/or quantitative relationships between them. This allows economists to derive hypothesis, or testable questions. Once these questions are tested and retested numerous times, they become adopted as the general set of principles in the field or the theories.

For example, many economists wondered what would happen if the government raised the minimum wage. This is a testable question or hypothesis. Economists set up many tests of this hypotheses. They tested the question for teens, restaurant workers, individuals who worked in retail, those in urban areas, in more 100 studies, The Federal Reserve Bank of San Francisco summarizes the findings and concludes that there is very little, if any, unemployment caused by raising the minimum wage. Therefore, the theory most widely accepted by most economists is that raising the minimum wage does not cause much, if any, minimum wage.

The Federal Reserve has summarized this research in their article entitled **"The Effects of Minimum Wage on Employment"**

Many people note, however, that there is an increase in some unemployment. Why is this occurring? Raising the minimum wage can create an environment where people who were not in the labor force wish to enter at the new higher wage (I won't work for $10/hr but I will for $15/hr). However, since they haven't been in the labor force, their employability is lower than those already in the labor force. This may take them a little time to find work as a result. In other words, the labor demand curve is highly inelastic but the labor supply curve is elastic (elasticity will be discussed in an upcoming chapter).

Correlation vs Causation

It is important to note that just because two events are correlated, that does not mean they are caused by one another. Moreover, the direction of the causation is important to note. For example, when it rains there are more accidents on the freeway. One could conclude that the accidents caused the rain. This would be a fallacious causal relationship. Therefore, it is important to consider all of the factors when attempting to determine correlation and causation.

Sometimes economists use the term **model** instead of **theory**. Strictly speaking, *a theory is a more abstract representation, while a model is more applied or empirical representation.* Models are used to test theories, but for this course we will use the terms interchangeably.

For example, an architect who is planning a major office building will often build a physical model that sits on a tabletop to show how the entire city block will look after the new building is constructed. Companies often build models of their new products, which are more rough and unfinished than the final product will be, but can still demonstrate how the new product will work.

The Circular Flow of Goods and Services

A good model to start with in economics is the circular flow diagram, which is shown in Figure 1.2.2. It pictures the economy as consisting of two groups—households and firms—that interact in two markets: the goods and services market in which firms sell and households buy and the labor market in which households sell labor to business firms or other employees.

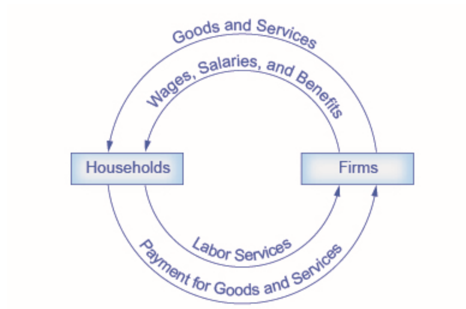


Figure 1.2.2: *The circular flow diagram shows how households and firms interact in the goods and services market, and in the labor market. The direction of the arrows shows that in the goods and services market, households receive goods and services and pay firms for them. In the labor market, households provide labor and receive payment from firms through wages, salaries, and benefits.*

Of course, in the real world, there are many different markets for goods and services and markets for many different types of labor. The circular flow diagram simplifies this to make the picture easier to grasp. In the diagram, firms produce goods and services, which they sell to households in return for revenues. This is shown in the outer circle, and represents the two sides of the product market (for example, the market for goods and services) in which households demand and firms supply. Households sell their labor as workers to firms in return for wages, salaries and benefits. This is shown in the inner circle and represents the two sides of the labor market in which households supply and firms demand. The labor market is an important market because households earn a majority of their income from this market and also because it comprises the largest cost for most firms. However, firms also pay rent for land and interest payments for capital.

This version of the circular flow model is stripped down to the essentials, but it has enough features to explain how the product and labor markets work in the economy. We could easily add details to this basic model if we wanted to introduce more real-world elements, like financial markets, governments, and interactions with the rest of the globe (imports and exports).

Economists carry a set of theories in their heads like a carpenter carries around a toolkit. When they see an economic issue or problem, they go through the theories they know to see if they can find one that fits. Then they use the theory to derive insights about the issue or problem. In economics, theories are expressed as diagrams, graphs, or even as mathematical equations. (Do not worry. In this course, we will mostly use graphs.) Economists do not figure out the answer to the problem first and then draw the graph to illustrate. Rather, they use the graph of the theory to help them figure out the answer. Although at the introductory level, you can sometimes figure out the right answer without applying a model, if you keep studying economics, before too long you will run into issues and problems that you will need to graph to solve. Both micro and macroeconomics are explained in terms of theories and models. The most well-known theories are probably those of supply and demand, but you will learn a number of others.

The American Economy

Current Size of the US Economy and US Population

Currently the US Economy produces the most goods and services in the world with the third largest population, according to the World Population Review.

Before the Industrial Revolution most workers in the U.S. were employed as subsistence agricultural producers, who worked on their own farms, producing for their own consumption. Currently, only 1.5% of the labor force is employed in agriculture. Since the 1970s, we have also seen a steady decline in the manufacturing sector with only a little more than 10% of the labor force employed in the manufacturing sector and the remaining 80%+ employed in the service sector, in such fields as finance and banking.

GDP Growth Over Time

The following graph published by the Federal Reserve Bank of St. Louis illustrates the growth rate of GDP over time from 1947 to the present.

It illustrates a few facts:

1. Growth has been uneven over time. However, fiscal and monetary policy seem to have lessened the severity of recessions and the high growth spurts since about the mid-1960s.
2. After women entered into the labor force in the 1960s, the growth rate of GDP increased.
3. Wars increase the GDP growth rate but every time the war ends, the transition to civilian goods tends to reduce the growth rate of GDP and the economy typically experiences a recession or slowing growth.

US Trade and Connection to World Economies

The US is considered a relatively **closed** economy. In other words, the US trades a small fraction of its goods or services annually. The OECD tracks the percentage of GDP that is imported and exported annually major nations. According to the data, the US exports 12% of GDP and imports 15% of GDP in 2016.

Which other nations export significantly more of their GDP? These nations are considered relatively more "open".

Key Concepts and Summary

Economists analyze problems differently than do other disciplinary experts. The main tools economists use are economic theories or models. A theory is not an illustration of the answer to a problem. Rather, a theory is a tool for determining the answer.

Glossary

circular flow diagram
: a diagram that views the economy as consisting of households and firms interacting in a goods and services market and a labor market

goods and services market
: a market in which firms are sellers of what they produce and households are buyers

labor market
: the market in which households sell their labor as workers to business firms or other employers

model
: see theory

theory
: a representation of an object or situation that is simplified while including enough of the key features to help us understand the object or situation

Contributor

- Template:ContribOpenSTAXMicroEcon

1.3: How Economies Can Be Organized- An Overview of Economic Systems

Skills to Develop

- Contrast traditional economies, command economies, and market economies
- Explain gross domestic product (GDP)
- Assess the importance and effects of globalization

Think about what a complex system a modern economy is. It includes all production of goods and services, all buying and selling, all employment. The economic life of every individual is interrelated, at least to a small extent, with the economic lives of thousands or even millions of other individuals. Who organizes and coordinates this system? Who insures that, for example, the number of televisions a society provides is the same as the amount it needs and wants? Who insures that the right number of employees work in the electronics industry? Who insures that televisions are produced in the best way possible? How does it all get done?

There are at least three ways societies have found to organize an economy. The first is the **traditional economy**, which is the oldest economic system and can be found in parts of Asia, Africa, and South America. Traditional economies organize their economic affairs the way they have always done (i.e., tradition). Occupations stay in the family. Most families are farmers who grow the crops they have always grown using traditional methods. What you produce is what you get to consume. Because things are driven by tradition, there is little economic progress or development.

Figure $\mathbf{1.3.1}$***:*** *Ancient Egypt was an example of a command economy. (Credit: Jay Bergesen/Flickr Creative Commons)*

Command economies are very different. In a command economy, economic effort is devoted to goals passed down from a ruler or ruling class. Ancient Egypt was a good example: a large part of economic life was devoted to building pyramids, like those shown in Figure $1.3.1$, for the pharaohs. Medieval manor life is another example: the lord provided the land for growing crops and protection in the event of war. In return, vassals provided labor and soldiers to do the lord’s bidding. In the last century, communism emphasized command economies.

In a command economy, the government decides what goods and services will be produced and what prices will be charged for them. The government decides what methods of production will be used and how much workers will be paid. Many necessities like healthcare and education are provided for free. Currently, Cuba and North Korea have command economies. Command economies have a very centralized structure for economic decisions.

Figure $\mathbf{1.3.2}$***:*** *Nothing says “market” more than The New York Stock Exchange. (Credit: Erik Drost/Flickr Creative Commons)*

Market economies have a very decentralized structure. A market is an institution that brings together buyers and sellers of goods or services, who may be either individuals or businesses. The New York Stock Exchange, shown in Figure $1.3.2$, is a prime example of market in which buyers and sellers are brought together. In a market economy, decision-making is decentralized. Market economies are based on private enterprise: the means of production (resources and businesses) are **owned and operated by**

private individuals or groups of private individuals. Businesses supply goods and services based on demand. (In a command economy, by contrast, resources and businesses are owned by the government.) What goods and services are supplied depends on what is demanded. A person's income is based on his or her ability to convert resources (especially labor) into something that society values. The more society values the person's output, the higher the income (think Lady Gaga or LeBron James). In this scenario, economic decisions are determined by market forces, not governments.

Socialism is another form of economic system that you will learn about in greater detail in *The 1st Chapter*. It is a theory and system of social and economic organization that advocates **public ownership** and control of the means of production and distribution, of capital, land. This can be complicated with so many voices and opinions and that is why socialism must also be a social organization that requires rules and norms about how to collectively operate.

Most economies in the real world are **mixed economies**; they combine elements of private ownership of the **means of production** and private ownership of the means of production (or socialism and market-based systems). The U.S. economy is positioned toward the market-oriented end of the spectrum. Many countries in Europe and Latin America, while primarily market-oriented, have a greater degree of government involvement in economic decisions than does the U.S. economy. China and Russia, while they are closer to having a market-oriented system now than several decades ago, remain closer to the command economy end of the spectrum. There are other nations that have market-based systems that also utilize socialism to promote social welfare such as Sweden and Scandinavian countries. A rich resource of information about countries and their economies can be found on the Heritage Foundation's website, as the following Clear It Up feature discusses.

Example 1.3.1: What countries are considered economically free?

Who is in control of economic decisions? Are people free to do what they want and to work where they want? Are businesses free to produce when they want and what they choose, and to hire and fire as they wish? Are banks free to choose who will receive loans? Or does the government control these kinds of choices? Each year, researchers at the Heritage Foundation and the Wall Street Journal look at 50 different categories of economic freedom for countries around the world. They give each nation a score based on the extent of economic freedom in each category.

The 2015 Heritage Foundation's Index of Economic Freedom report ranked 178 countries around the world: some examples of the most free and the least free countries are listed in Table 1.3.1. Several countries were not ranked because of extreme instability that made judgments about economic freedom impossible. These countries include Afghanistan, Iraq, Syria, and Somalia.

The assigned rankings are inevitably based on estimates, yet even these rough measures can be useful for discerning trends. In 2015, 101 of the 178 included countries shifted toward greater economic freedom, although 77 of the countries shifted toward less economic freedom. In recent decades, the overall trend has been a higher level of economic freedom around the world. *The chart below indicates that the U.S. is highly free economically. In other words, there is very little ownership, regulation, or coordination of resources by the government. Some might argue this is also a reason for the growing inequality in this nation.*

***Table 1.3.1:** Economic Freedoms, 2015(Source: The Heritage Foundation, 2015 Index of Economic Freedom, Country Rankings, http://www.heritage.org/index/ranking)*

Most Economic Freedom	Least Economic Freedom
1. Hong Kong	167. Timor-Leste
2. Singapore	168. Democratic Republic of Congo
3. New Zealand	169. Argentina
4. Australia	170. Republic of Congo
5. Switzerland	171. Iran
6. Canada	172. Turkamenistan
7. Chile	173. Equatorial Guinea

8. Estonia	174. Eritrea
9. Ireland	175. Zimbabwe
10. Mauritius	176. Venezuela
11. Denmark	177. Cuba
12. United States	178. North Korea

Regulations: The Rules of the Game

Markets and government regulations are always entangled. There is no such thing as an absolutely free market. Regulations always define the “rules of the game” in the economy. Economies that are primarily market-oriented have fewer regulations—ideally just enough to maintain an even playing field for participants. At a minimum, these laws govern matters like safeguarding private property against theft, protecting people from violence, enforcing legal contracts, preventing fraud, and collecting taxes. Conversely, even the most command-oriented economies operate using markets. How else would buying and selling occur? But the decisions of what will be produced and what prices will be charged are heavily regulated. Heavily regulated economies often have **underground economies**, which are markets where the buyers and sellers make transactions without the government’s approval.

The question of how to organize economic institutions is typically not a black-or-white choice between all market or all government, but instead involves a balancing act over the appropriate combination of market freedom and government rules.

Figure* 1.3.3*: *Cargo ships are one mode of transportation for shipping goods in the global economy. (Credit: Raul Valdez/Flickr Creative Commons)*

The Rise of Globalization

Recent decades have seen a trend toward **globalization**, which is the expanding cultural, political, and economic connections between people around the world. One measure of this is the increased buying and selling of goods, services, and assets across national borders—in other words, international trade and financial capital flows.

Globalization has occurred for a number of reasons. Improvements in shipping, as illustrated by the container ship shown in Figure 1.3.3, and air cargo have driven down transportation costs. Innovations in computing and telecommunications have made it easier and cheaper to manage long-distance economic connections of production and sales. Many valuable products and services in the modern economy can take the form of information—for example: computer software; financial advice; travel planning; music, books and movies; and blueprints for designing a building. These products and many others can be transported over telephones and computer networks at ever-lower costs. Finally, international agreements and treaties between countries have encouraged greater trade.

Table 1.3.2 presents one measure of globalization. It shows the percentage of domestic economic production that was exported for a selection of countries from 2010 to 2013, according to an entity known as The World Bank. **Exports** are the goods and services that are produced domestically and sold abroad. **Imports** are the goods and services that are produced abroad and then sold domestically. The size of total production in an economy is measured by the **gross domestic product (GDP)**. Thus, the ratio of exports divided by GDP measures what share of a country’s total economic production is sold in other countries.

***Table** 1.3.2: The Extent of Globalization (exports/GDP)(Source: http://databank.worldbank.org/data/)*

Country	2010	2011	2012	2013
Higher Income countries Measured in GDP/capita				
United States	12.4	13.6	13.6	13.5
Belgium	76.2	81.4	82.2	82.8
Canada	29.1	30.7	30.0	30.1
France	26.0	27.8	28.1	28.3
Middle Income Countries				
Brazil	10.9	11.9	12.6	12.6
Mexico	29.9	31.2	32.6	31.7
South Korea	49.4	55.7	56.3	53.9
Lower Income Countries				
Chad	36.8	38.9	36.9	32.2
China	29.4	28.5	27.3	26.4
India	22.0	23.9	24.0	24.8
Nigeria	25.3	31.3	31.4	18.0

In recent decades, the export/GDP ratio has generally risen, both worldwide and for the U.S. economy. *Although the US is a world leader in GDP per person or per capita (compared to India or China who trade much larger shares), the share of U.S. exports in proportion to the U.S. economy is well below the global average*, in part because large economies like the United States can contain more of the division of labor inside their national borders. However, smaller economies like Belgium, Korea, and Canada need to trade across their borders with other countries to take full advantage of division of labor, specialization, and economies of scale. In this sense, the enormous U.S. economy is less affected by globalization than most other countries. Currently, the U.S. exports roughly 14-15% of their GDP. This implies that the United States is considered a **relatively closed economy**, in other words, the amount traded with other nations is a small fraction of GDP, unlike most other industrial societies that are ***relatively open economies*** that trade larger portions of their GDP.

Table \(\PageIndex{2}\) also shows that many medium and low income countries around the world, like Mexico and China, have also experienced a surge of globalization in recent decades. If an astronaut in orbit could put on special glasses that make all economic transactions visible as brightly colored lines and look down at Earth, the astronaut would see the planet covered with connections.

So, hopefully, you now have an idea of what economics is about. It is essential that you learn more about how to read and use models in economics.

Example 1.3.1: Decisions ... Decisions in the Social Media Age

The world we live in today provides nearly instant access to a wealth of information. Consider that as recently as the late 1970s, the Farmer's Almanac, along with the Weather Bureau of the U.S. Department of Agriculture, were the primary sources American farmers used to determine when to plant and harvest their crops. Today, farmers are more likely to access, online, weather forecasts from the National Oceanic and Atmospheric Administration or watch the Weather Channel. After all, knowing the upcoming forecast could drive when to harvest crops. Consequently, knowing the upcoming weather could change the amount of crop harvested.

Some relatively new information forums, such as Facebook, are rapidly changing how information is distributed; hence, influencing decision making. In 2014, the Pew Research Center reported that 71% of online adults use Facebook. Facebook post topics range from the National Basketball Association, to celebrity singers and performers, to farmers.

Information helps us make decisions. Decisions as simple as what to wear today to how many reporters should be sent to cover a crash. Each of these decisions is an economic decision. After all, resources are scarce. If ten reporters are sent to cover an accident, they are not available to cover other stories or complete other tasks. Information provides the knowledge needed to make the best possible decisions on how to utilize scarce resources. Welcome to the world of economics!

Key Concepts and Summary

Societies can be organized as traditional, command, socialist or market-oriented economies. Most societies are a mix. The last few decades have seen globalization evolve as a result of growth in commercial and financial networks that cross national borders, making businesses and workers from different economies increasingly interdependent.

References

The Heritage Foundation. 2015. "2015 Index of Economic Freedom." Accessed March 11, 2015. http://www.heritage.org/index/ranking.

Garling, Caleb. “S.F. plane crash: Reporting, emotions on social media,” The San Francisco Chronicle. July 7, 2013. http://www.sfgate.com/news/article/S...al-4651639.php.

Irvine, Jessica. “Social Networking Sites are Factories of Modern Ideas.” The Sydney Morning Herald. November 25, 2011.http://www.smh.com.au/federal-politi...#ixzz2YZhPYeME.

Pew Research Center. 2015. "Social Networking Fact Sheet." Accessed March 11, 2015. http://www.pewinternet.org/fact-shee...ng-fact-sheet/.

The World Bank Group. 2015. "World Data Bank." Accessed March 30, 2014. http://databank.worldbank.org/data/.

Glossary

command economy

an economy where economic decisions are passed down from government authority and where resources are owned by the government

factors or means of production (resources)

the labor or human effort, the land or natural resources, and the capital good inputs needed to produce outputs or finished goods and services.

globalization

the trend in which buying and selling in markets have increasingly crossed national borders

gross domestic product (GDP)

measure of the size of total production in an economy

imports

products (goods and services) made abroad and then sold domestically

market

interaction between potential buyers and sellers; a combination of demand and supply

market economy

an economy where economic decisions are decentralized, resources are owned by private individuals, and businesses supply goods and services based on demand

private enterprise

system where the means of production (resources and businesses) are owned and operated by private individuals or groups of private individuals

traditional economy

typically an agricultural economy where things are done the same as they have always been done

underground economy

a market where the buyers and sellers do so without the government's approval

closed economy

when exports account for a small fraction of GDP

open economy

when exports account for a large fraction of GDP

mixed economy

an economic systems that contains both private and pubic ownership of the means of production.

socialism

an economic and social organization based on the idea of private ownership and control of the means of production

exports

products (goods and services) made domestically and sold abroad

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

2: Choice in a World of Scarcity

CHAPTER OVERVIEW

2.1: Welcome to Economics!

This leads us to the topic of this chapter, an introduction to the world of making decisions, processing information, and understanding behavior in markets —the world of economics. Each chapter in this book will start with a discussion about current (or sometimes past) events and revisit it at chapter's end—to "bring home" the concepts in play.

2.2: Prelude to Choice in a World of Scarcity

Learning Objectives

- How Individuals Make Choices Based on Their Budget Constraint
- The Production Possibilities Frontier and Social Choices
- Confronting Objections to the Economic Approach

Choices and Tradeoffs

Figure 2.2.1: *In general, the higher the degree, the higher the salary. So why aren't more people pursuing higher degrees? The short answer: choices and tradeoffs. (Credit: modification of work by "Jim, the Photographer"/Flickr Creative Commons)*

Choices ... To What Degree?

In 2015, the median income for workers who hold master's degrees varies from males to females. The average of the two is $\$2,951$ weekly. Multiply this average by 52 weeks, and you get an average salary of $\$153,452$ Compare that to the median weekly earnings for a full-time worker over 25 with no higher than a bachelor's degree: $\$1,224$ weekly and $\$63,648$ a year. What about those with no higher than a high school diploma in 2015? They earn just $\$664$ weekly and $\$34,528$ over 12 months. In other words, says the Bureau of Labor Statistics (BLS), earning a bachelor's degree boosted salaries 54% over what you would have earned if you had stopped your education after high school. A bachelor's degree yields a salary almost double that of a high school diploma.

Given these statistics, we might expect a lot of people to choose to go to college and at least earn a bachelor's degree. Assuming that people want to improve their material well-being, it seems like they would make those choices that give them the greatest opportunity to consume goods and services. As it turns out, the analysis is not nearly as simple as this. In fact, in 2014, the BLS reported that while almost 88% of the population in the United States had a high school diploma, only 33.6% of $25-65$ year olds had bachelor's degrees, and only 7.4% of $25-65$ year olds in 2014 had earned a master's.

This brings us to the subject of this chapter: why people make the choices they make and how economists go about explaining those choices.

You will learn quickly when you examine the relationship between economics and scarcity that choices involve tradeoffs. Every choice has a cost.

In 1968, the Rolling Stones recorded "*You Can't Always Get What You Want.*" Economists chuckled, because they had been singing a similar tune for decades. English economist Lionel Robbins (1898–1984), in his Essay on the Nature and Significance of Economic Science in 1932, described not always getting what you want in this way:

The time at our disposal is limited. There are only twenty-four hours in the day. We have to choose between the different uses to which they may be put. ... Everywhere we turn, if we choose one thing we must relinquish others which, in different circumstances, we would wish not to have relinquished. Scarcity of means to satisfy given ends is an almost ubiquitous condition of human nature.

Because people live in a world of scarcity, they cannot have all the time, money, possessions, and experiences they wish. Neither can society.

This chapter will continue our discussion of scarcity and the economic way of thinking by first introducing three critical concepts: opportunity cost, marginal decision making, and diminishing returns. Later, it will consider whether the economic way of thinking accurately describes either how choices are made or how they should be made.

Contributor

- Template:ContribOpenSTAXMicroEcon

2.3: How Individuals Make Choices Based on Their Budget Constraint

Learning Objectives

- Calculate and graph budgets constraints
- Explain opportunity sets and opportunity costs
- Evaluate the law of diminishing marginal utility
- Explain how marginal analysis and utility influence choices

Consider the typical consumer's budget problem. Consumers have a limited amount of income to spend on the things they need and want. Suppose Alphonso has $\$10$ in spending money each week that he can allocate between bus tickets for getting to work and the burgers that he eats for lunch. Burgers cost $\$2$ each, and bus tickets are 50 cents each. Figure $2.3.1$ shows Alphonso's budget constraint, that is, the outer boundary of his opportunity set. The opportunity set identifies all the opportunities for spending within his budget. The budget constraint indicates all the combinations of burgers and bus tickets Alphonso can afford when he exhausts his budget, given the prices of the two goods. (There are actually many different kinds of budget constraints. You will learn more about them in the chapter on Consumer Choices.)

The Budget Constraint: Alphonso's Consumption Choice Opportunity Frontier

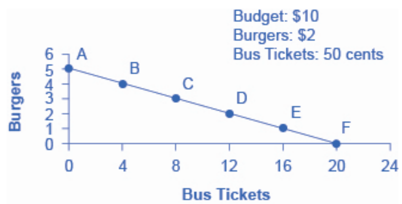


Figure $2.3.1$: Each point on the budget constraint represents a combination of burgers and bus tickets whose total cost adds up to Alphonso's budget of $10. The slope of the budget constraint is determined by the relative price of burgers and bus tickets. All along the budget set, giving up one burger means gaining four bus tickets.

The vertical axis in the figure shows burger purchases and the horizontal axis shows bus ticket purchases. If Alphonso spends all his money on burgers, he can afford five per week. ($\$10$ per week/$\$2$ per burger = 5 burgers per week.) But if he does this, he will not be able to afford any bus tickets. This choice (zero bus tickets and five burgers) is shown by point A in the figure. Alternatively, if Alphonso spends all his money on bus tickets, he can afford 20 per week. ($\$10$ per week/$\$0.50$ per bus ticket = 20 bus tickets per week.) Then, however, he will not be able to afford any burgers. This alternative choice (20 bus tickets and zero burgers) is shown by point F.

If Alphonso is like most people, he will choose some combination that includes both bus tickets and burgers. That is, he will choose some combination on the budget constraint that connects points A and F. Every point on (or inside) the constraint shows a combination of burgers and bus tickets that Alphonso can afford. Any point outside the constraint is not affordable, because it would cost more money than Alphonso has in his budget.

The budget constraint clearly shows the tradeoff Alphonso faces in choosing between burgers and bus tickets. Suppose he is currently at point D, where he can afford 12 bus tickets and two burgers. What would it cost Alphonso for one more burger? It would be natural to answer $\$2$, but that's not the way economists think. Instead they ask, how many bus tickets would Alphonso have to give up to get one more burger, while staying within his budget? The answer is four bus tickets. That is the true cost to Alphonso of one more burger.

The Concept of Opportunity Cost

Economists use the term opportunity cost to indicate what must be given up to obtain something that is desired. The idea behind opportunity cost is that the cost of one item is the lost opportunity to do or consume something else; in short, opportunity cost is the value of the next best alternative. For Alphonso, the opportunity cost of a burger is the four bus tickets he would have to give up. He would decide whether or not to choose the burger depending on whether the value of the burger exceeds the value of the forgone alternative—in this case, bus tickets. Since people must choose, they inevitably face tradeoffs in which they have to give up things they desire to get other things they desire more.

A fundamental principle of economics is that every choice has an opportunity cost. If you sleep through your economics class (not recommended, by the way), the opportunity cost is the learning you miss from not attending class. If you spend your income on video games, you cannot spend it on movies. If you choose to marry one person, you give up the opportunity to marry anyone else. In short, opportunity cost is all around us and part of human existence.

The following Work It Out feature shows a step-by-step analysis of a budget constraint calculation. Read through it to understand another important concept—slope—that is further explained in the appendix ***The Use of Mathematics in Principles of Economics***.

Example 2.3.1: Understanding Budget Constraints

Budget constraints are easy to understand if you apply a little math. The appendix ***The Use of Mathematics in Principles of Economics*** explains all the math you are likely to need in this book. So if math is not your strength, you might want to take a look at the appendix.

Step 1: The equation for any budget constraint is:

$$Budget = P1 \times Q1 + P2 \times Q2 \tag{2.3.1}$$

where P and Q are the price and quantity of items purchased and Budget is the amount of income one has to spend.

Step 2: Apply the budget constraint equation to the scenario. In Alphonso's case, this works out to be:

$$Budget = P1 \times Q1 + P2 \times Q2 \tag{2.3.2}$$

$$\$\text{budget} = \$2 \text{ per burger } \times \text{ quantity of burgers } + \$0.50 \text{ per bus ticket } \times \text{ quantity of bus tickets} \tag{2.3.3}$$

$$\$10 = \$2 \times Q_{burgers} \ + \$0.50 \times Q_{bustickets} \tag{2.3.4}$$

Step 3: Using a little algebra, we can turn this into the familiar equation of a line:

$$y = b + mx \tag{2.3.5}$$

For Alphonso, this is:

$$\$10 = \$2 \times Q_{burgers} \ + \$0.50 \times Q_{bustickets} \tag{2.3.6}$$

Step 4: Simplify the equation. Begin by multiplying both sides of the equation by 2:

$$2 \times 10 = 2 \times 2 \times Q_{burgers} \ + 2 \times 0.50 \times Q_{bustickets} \tag{2.3.7}$$

$$20 = 4 \times Q_{burgers} \ + 1 \times Q_{bustickets} \tag{2.3.8}$$

Step 5: Subtract one bus ticket from both sides:

$$20 - Q_{bustickets} \ = 4 \times Q_{burgers} \tag{2.3.9}$$

Divide each side by 4 to yield the answer:

$$5 - 0.25 \times Q_{bustickets} \ = Q_{burgers} \tag{2.3.10}$$

or

$$Q_{burgers} \ = 5 - 0.25 \times Q_{bustickets} \tag{2.3.11}$$

Step 6: Notice that this equation fits the budget constraint in Figure $2.3.1$. The vertical intercept is 5 and the slope is -0.25, just as the equation says. If you plug 20 bus tickets into the equation, you get 0 burgers. If you plug other numbers of bus tickets into the equation, you get the results shown in Table $2.3.1$, which are the points on Alphonso's budget constraint.

Table $2.3.1$: Budget results

Point	Quantity of Burgers (at $2)	Quantity of Bus Tickets (at 50 cents)
A	5	0
B	4	4
C	3	8

Point	Quantity of Burgers (at $2)	Quantity of Bus Tickets (at 50 cents)
D	2	12
E	1	16
F	0	20

Step 7: Notice that the slope of a budget constraint always shows the opportunity cost of the good which is on the horizontal axis. For Alphonso, the slope is -0.25, indicating that for every four bus tickets he buys, Alphonso must give up 1 burger.

There are two important observations here. First, the algebraic sign of the slope is negative, which means that the only way to get more of one good is to give up some of the other. Second, the slope is defined as the price of bus tickets (whatever is on the horizontal axis in the graph) divided by the price of burgers (whatever is on the vertical axis), in this case $\$0.50/\$2 = 0.25$ So if you want to determine the opportunity cost quickly, just divide the two prices.

Identifying Opportunity Cost

In many cases, it is reasonable to refer to the opportunity cost as the **price**. If your cousin buys a new bicycle for $\$300$, then $\$300$ measures the amount of "other consumption" that he has given up. For practical purposes, there may be no special need to identify the specific alternative product or products that could have been bought with that $\$300$, but sometimes the price as measured in dollars may not accurately capture the true opportunity cost. This problem can loom especially large when costs of time are involved.

For example, consider a boss who decides that all employees will attend a two-day retreat to "build team spirit." The out-of-pocket monetary cost of the event may involve hiring an outside consulting firm to run the retreat, as well as room and board for all participants. But an opportunity cost exists as well: during the two days of the retreat, none of the employees are doing any other work.

Attending college is another case where the opportunity cost exceeds the monetary cost. The out-of-pocket costs of attending college include tuition, books, room and board, and other expenses. But in addition, during the hours that you are attending class and studying, it is impossible to work at a paying job. Thus, college imposes both an out-of-pocket cost and an opportunity cost of lost earnings.

✓ Example 2.3.2: What is the opportunity cost associated with increased airport security measures?

After the terrorist plane hijackings on September 11, 2001, many steps were proposed to improve air travel safety. For example, the federal government could provide armed "sky marshals" who would travel inconspicuously with the rest of the passengers. The cost of having a sky marshal on every flight would be roughly $\$3$ billion per year. Retrofitting all U.S. planes with reinforced cockpit doors to make it harder for terrorists to take over the plane would have a price tag of $\$450$ million. Buying more sophisticated security equipment for airports, like three-dimensional baggage scanners and cameras linked to face recognition software, could cost another $\$2$ billion.

But the single biggest cost of greater airline security does not involve spending money. It is the opportunity cost of additional waiting time at the airport. According to the United States Department of Transportation (DOT), more than 800 million passengers took plane trips in the United States in 2012. Since the $9/11$ hijackings, security screening has become more intensive, and consequently, the procedure takes longer than in the past. Say that, on average, each air passenger spends an extra 30 minutes in the airport per trip. Economists commonly place a value on time to convert an opportunity cost in time into a monetary figure. Because many air travelers are relatively high-paid business people, conservative estimates set the average price of time for air travelers at $20 per hour. By these back-of-the-envelope calculations, the opportunity cost of delays in airports could be as much as $800 \text{ million} \times 0.5 \text{ hours} \times \$20/\text{hour}$, or $\$8 \text{ billion}$ per year. Clearly, the opportunity costs of waiting time can be just as important as costs that involve direct spending.

In some cases, realizing the opportunity cost can alter behavior. Imagine, for example, that you spend $\$8$ on lunch every day at work. You may know perfectly well that bringing a lunch from home would cost only $\$3$ a day, so the opportunity cost of buying lunch at the restaurant is $\$5$ each day (that is, the $\$8$ buying lunch costs minus the $\$3$ your lunch from home would cost). $\$5$ each day does not seem to be that much. However, if you project what that adds up to in a year—250 days a year × $\$5$ per day equals

$\$1,250$, the cost, perhaps, of a decent vacation. If the opportunity cost is described as "a nice vacation" instead of "$\$5$ a day," you might make different choices.

Marginal Decision-Making and Diminishing Marginal Utility

The budget constraint framework helps to emphasize that most choices in the real world are not about getting all of one thing or all of another; that is, they are not about choosing either the point at one end of the budget constraint or else the point all the way at the other end. Instead, most choices involve marginal analysis, which means comparing the benefits and costs of choosing a little more or a little less of a good.

People desire goods and services for the satisfaction or utility those goods and services provide. Utility, as we will see in the chapter on Consumer Choices, is subjective but that does not make it less real. Economists typically assume that the more of some good one consumes (for example, slices of pizza), the more utility one obtains. At the same time, the utility a person receives from consuming the first unit of a good is typically more than the utility received from consuming the fifth or the tenth unit of that same good. When Alphonso chooses between burgers and bus tickets, for example, the first few bus rides that he chooses might provide him with a great deal of utility—perhaps they help him get to a job interview or a doctor's appointment. But later bus rides might provide much less utility—they may only serve to kill time on a rainy day. Similarly, the first burger that Alphonso chooses to buy may be on a day when he missed breakfast and is ravenously hungry. However, if Alphonso has a burger every single day, the last few burgers may taste pretty boring. The general pattern that consumption of the first few units of any good tends to bring a higher level of utility to a person than consumption of later units is a common pattern. Economists refer to this pattern as the law of diminishing marginal utility, which means that as a person receives more of a good, the additional (or marginal) utility from each additional unit of the good declines. In other words, the first slice of pizza brings more satisfaction than the sixth.

The law of diminishing marginal utility explains why people and societies rarely make all-or-nothing choices. You would not say, "*My favorite food is ice cream, so I will eat nothing but ice cream from now on.*" Instead, even if you get a very high level of utility from your favorite food, if you ate it exclusively, the additional or marginal utility from those last few servings would not be very high. Similarly, most workers do not say: "*I enjoy leisure, so I'll never work.*" Instead, workers recognize that even though some leisure is very nice, a combination of all leisure and no income is not so attractive. The budget constraint framework suggests that when people make choices in a world of scarcity, they will use marginal analysis and think about whether they would prefer a little more or a little less.

Sunk Costs

In the budget constraint framework, all decisions involve what will happen next: that is, what quantities of goods will you consume, how many hours will you work, or how much will you save. These decisions do not look back to past choices. Thus, the budget constraint framework assumes that **sunk costs**, which are costs that were incurred in the past and cannot be recovered, should not affect the current decision.

Consider the case of Selena, who pays $\$8$ to see a movie, but after watching the film for 30 minutes, she knows that it is truly terrible. Should she stay and watch the rest of the movie because she paid for the ticket, or should she leave? The money she spent is a sunk cost, and unless the theater manager is feeling kindly, Selena will not get a refund. But staying in the movie still means paying an opportunity cost in time. Her choice is whether to spend the next 90 minutes suffering through a cinematic disaster or to do something—anything—else. The lesson of sunk costs is to forget about the money and time that is irretrievably gone and instead to focus on the marginal costs and benefits of current and future options.

For people and firms alike, dealing with sunk costs can be frustrating. It often means admitting an earlier error in judgment. Many firms, for example, find it hard to give up on a new product that is doing poorly because they spent so much money in creating and launching the product. But the lesson of sunk costs is to ignore them and make decisions based on what will happen in the future.

From a Model with Two Goods to One of Many Goods

The budget constraint diagram containing just two goods, like most models used in this book, is not realistic. After all, in a modern economy people choose from thousands of goods. However, thinking about a model with many goods is a straightforward extension of what we discussed here. Instead of drawing just one budget constraint, showing the tradeoff between two goods, you can draw multiple budget constraints, showing the possible tradeoffs between many different pairs of goods. Or in more advanced classes in economics, you would use mathematical equations that include many possible goods and services that can be purchased, together with their quantities and prices, and show how the total spending on all goods and services is limited to the overall budget available. The graph with two goods that was presented here clearly illustrates that every choice has an opportunity cost, which is the point that does carry over to the real world.

Key Concepts and Summary

Economists see the real world as one of scarcity: that is, a world in which people's desires exceed what is possible. As a result, economic behavior involves tradeoffs in which individuals, firms, and society must give up something that they desire to obtain things that they desire more. Individuals face the tradeoff of what quantities of goods and services to consume. The budget constraint, which is the frontier of the opportunity set, illustrates the range of choices available. The slope of the budget constraint is determined by the relative price of the choices. Choices beyond the budget constraint are not affordable.

Opportunity cost measures cost by what is given up in exchange. Sometimes opportunity cost can be measured in money, but it is often useful to consider time as well, or to measure it in terms of the actual resources that must be given up.

Most economic decisions and tradeoffs are not all-or-nothing. Instead, they involve marginal analysis, which means they are about decisions on the margin, involving a little more or a little less. The law of diminishing marginal utility points out that as a person receives more of something—whether it is a specific good or another resource—the additional marginal gains tend to become smaller. Because sunk costs occurred in the past and cannot be recovered, they should be disregarded in making current decisions.

References

Bureau of Labor Statistics, U.S. Department of Labor. 2015. "Median Weekly Earnings by Educational Attainment in 2014." Accessed March 27, 2015. www.bls.gov/opub/ted/2015/med...ty-in-2014.htm.

Robbins, Lionel. An Essay on the Nature and Significance of Economic Science. London: Macmillan. 1932.

United States Department of Transportation. "Total Passengers on U.S Airlines and Foreign Airlines U.S. Flights Increased 1.3% in 2012 from 2011." Accessed October 2013. www.rita.dot.gov/bts/press_releases/bts016_13

Glossary

budget constraint

all possible consumption combinations of goods that someone can afford, given the prices of goods, when all income is spent; the boundary of the opportunity set

law of diminishing marginal utility

as we consume more of a good or service, the utility we get from additional units of the good or service tend to become smaller than what we received from earlier units

marginal analysis

examination of decisions on the margin, meaning a little more or a little less from the status quo

opportunity cost

measures cost by what is given up in exchange; opportunity cost measures the value of the forgone alternative

opportunity set

all possible combinations of consumption that someone can afford given the prices of goods and the individual's income

sunk costs

costs that are made in the past and cannot be recovered

utility

satisfaction, usefulness, or value one obtains from consuming goods and services

Contributor

- Template:ContribOpenSTAXMicroEcon

2.4: The Production Possibilities Frontier and Social Choices

Learning Objectives

- Interpret production possibilities frontier graphs
- Contrast a budget constraint and a production possibilities frontier
- Explain the relationship between a production possibilities frontier and the law of diminishing returns
- Contrast productive efficiency and allocative efficiency
- Define comparative advantage

Just as individuals cannot have everything they want and must instead make choices, society as a whole cannot have everything it might want, either. This section of the chapter will explain the constraints faced by society, using a model called the **production possibilities frontier (PPF)**. There are more similarities than differences between individual choice and social choice. As you read this section, focus on the similarities.

Because society has limited resources (e.g., labor, land, capital, raw materials) at any point in time, there is a limit to the quantities of goods and services it can produce. Suppose a society desires two products, healthcare and education. This situation is illustrated by the production possibilities frontier in Figure $2.4.1$.

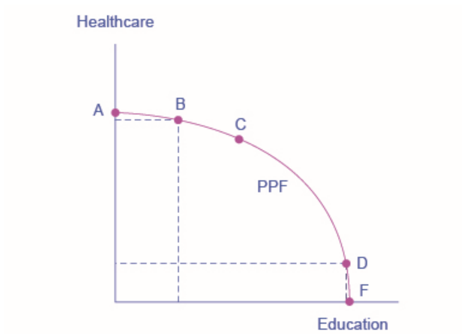


Figure $2.4.1$: This production possibilities frontier shows a tradeoff between devoting social resources to healthcare and devoting them to education. At A all resources go to healthcare and at B, most go to healthcare. At D most resources go to education, and at F, all go to education.

In Figure $2.4.1$, healthcare is shown on the vertical axis and education is shown on the horizontal axis. If the society were to allocate all of its resources to healthcare, it could produce at point A. But it would not have any resources to produce education. If it were to allocate all of its resources to education, it could produce at point F. Alternatively, the society could choose to produce any combination of healthcare and education shown on the production possibilities frontier. In effect, the production possibilities frontier plays the same role for society as the budget constraint plays for Alphonso. Society can choose any combination of the two goods on or inside the PPF. But it does not have enough resources to produce outside the PPF.

Most important, the production possibilities frontier clearly shows the tradeoff between healthcare and education. Suppose society has chosen to operate at point B, and it is considering producing more education. Because the PPF is downward sloping from left to right, the only way society can obtain more education is by giving up some healthcare. That is the tradeoff society faces. Suppose it considers moving from point B to point C. What would the opportunity cost be for the additional education? The opportunity cost would be the healthcare society has to give up. Just as with Alphonso's budget constraint, the opportunity cost is shown by the **slope** of the production possibilities frontier. By now you might be saying, "*Hey, this PPF is sounding like the budget constraint.*" If so, read the following Clear It Up feature.

Example $2.4.1$: What's the difference between a budget constraint and a PPF?

There are two major differences between a budget constraint and a production possibilities frontier. The first is the fact that the budget constraint is a straight line. This is because its slope is given by the relative prices of the two goods. In contrast, the PPF has a curved shape because of the law of the diminishing returns. The second is the absence of specific numbers on the axes of the PPF. There are no specific numbers because we do not know the exact amount of resources this imaginary economy has, nor do we know how many resources it takes to produce healthcare and how many resources it takes to produce education. If this were a real world example, that data would be available. An additional reason for the lack of numbers is that there is no single way to measure levels of education and healthcare. However, when you think of improvements in education, you can think of accomplishments like more years of school completed, fewer high-school dropouts, and higher scores on standardized tests. When you think of improvements in healthcare, you can think of longer life expectancies, lower levels of infant mortality, and fewer outbreaks of disease.

Whether or not we have specific numbers, conceptually we can measure the opportunity cost of additional education as society moves from point B to point C on the PPF. The additional education is measured by the horizontal distance between B and C. The foregone healthcare is given by the vertical distance between B and C. The slope of the PPF between B and C is (approximately) the vertical distance (the "rise") over the horizontal distance (the "run"). This is the opportunity cost of the additional education.

The Shape of the PPF and the Law of Diminishing Returns

The budget constraints presented earlier in this chapter, showing individual choices about what quantities of goods to consume, were all straight lines. The reason for these straight lines was that the slope of the budget constraint was determined by relative prices of the two goods in the **consumption budget constraint**. However, the production possibilities frontier for healthcare and education was drawn as a curved line. Why does the PPF have a different shape?

To understand why the PPF is curved, start by considering point A at the top left-hand side of the PPF. At point A, all available resources are devoted to healthcare and none are left for education. This situation would be extreme and even ridiculous. For example, children are seeing a doctor every day, whether they are sick or not, but not attending school. People are having cosmetic surgery on every part of their bodies, but no high school or college education exists. Now imagine that some of these resources are diverted from healthcare to education, so that the economy is at point B instead of point A. Diverting some resources away from A to B causes relatively little reduction in health because the last few marginal dollars going into healthcare services are not producing much additional gain in health. However, putting those marginal dollars into education, which is completely without resources at point A, can produce relatively large gains. For this reason, the shape of the PPF from A to B is relatively flat, representing a relatively small drop-off in health and a relatively large gain in education.

Now consider the other end, at the lower right, of the production possibilities frontier. Imagine that society starts at choice D, which is devoting nearly all resources to education and very few to healthcare, and moves to point F, which is devoting all spending to education and none to healthcare. For the sake of concreteness, you can imagine that in the movement from D to F, the last few doctors must become high school science teachers, the last few nurses must become school librarians rather than dispensers of vaccinations, and the last few emergency rooms are turned into kindergartens. The gains to education from adding these last few resources to education are very small. However, the opportunity cost lost to health will be fairly large, and thus the slope of the PPF between D and F is steep, showing a large drop in health for only a small gain in education.

The lesson is not that society is likely to make an extreme choice like devoting no resources to education at point A or no resources to health at point F. Instead, the lesson is that the gains from committing additional marginal resources to education depend on how much is already being spent. If on the one hand, very few resources are currently committed to education, then an increase in resources used can bring relatively large gains. On the other hand, if a large number of resources are already committed to education, then committing additional resources will bring relatively smaller gains.

This pattern is common enough that it has been given a name: the **law of diminishing returns**, which holds that as additional increments of resources are added to a certain purpose, the marginal benefit from those additional increments will decline. When government spends a certain amount more on reducing crime, for example, the original gains in reducing crime could be relatively large. But additional increases typically cause relatively smaller reductions in crime, and paying for enough police and security to reduce crime to nothing at all would be tremendously expensive.

The curvature of the production possibilities frontier shows that as additional resources are added to education, moving from left to right along the horizontal axis, the original gains are fairly large, but gradually diminish. Similarly, as additional resources are added to healthcare, moving from bottom to top on the vertical axis, the original gains are fairly large, but again gradually diminish. In this way, the law of diminishing returns produces the outward-bending shape of the production possibilities frontier.

Productive Efficiency and Allocative Efficiency

The study of economics does not presume to tell a society what choice it should make along its production possibilities frontier. In a market-oriented economy with a democratic government, the choice will involve a mixture of decisions by individuals, firms, and government. However, economics can point out that some choices are unambiguously better than others. This observation is based on the concept of efficiency. In everyday usage, efficiency refers to lack of waste. An inefficient machine operates at high **cost**, while an efficient machine operates at lower cost, because it is not wasting energy or materials. An inefficient organization operates with long delays and high costs, while an efficient organization meets schedules, is focused, and performs within budget.

The production possibilities frontier can illustrate two kinds of efficiency: productive efficiency and allocative efficiency. Figure $2.4.2$ illustrates these ideas using a production possibilities frontier between healthcare and education.

Productive and Allocative Efficiency

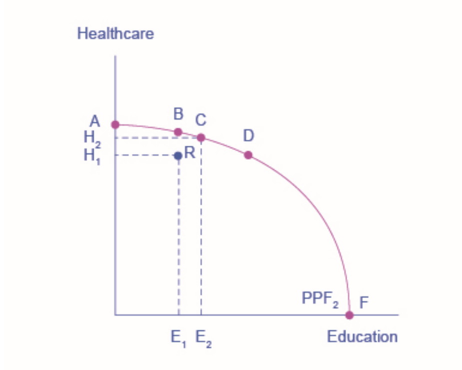


Figure $2.4.2$: Productive efficiency means it is impossible to produce more of one good without decreasing the quantity that is produced of another good. Thus, all choices along a given PPF like B, C, and D display productive efficiency, but R does not. Allocative efficiency means that the particular mix of goods being produced—that is, the specific choice along the production possibilities frontier—represents the allocation that society most desires.

Productive efficiency means that, given the available inputs and technology, it is impossible to produce more of one good without decreasing the quantity that is produced of another good. All choices on the PPF in Figure $2.4.2$, including A, B, C, D, and F, display productive efficiency. As a firm moves from any one of these choices to any other, either healthcare increases and education decreases or vice versa. However, any choice inside the production possibilities frontier is productively inefficient and wasteful because it is possible to produce more of one good, the other good, or some combination of both goods.

For example, point R is productively inefficient because it is possible at choice C to have more of both goods: education on the horizontal axis is higher at point C than point R (E_2 is greater than E_1), and healthcare on the vertical axis is also higher at point C than point R (H_2 is great than H_1).

The particular mix of goods and services being produced—that is, the specific combination of healthcare and education chosen along the production possibilities frontier—can be shown as a ray (line) from the origin to a specific point on the PPF. Output mixes that had more healthcare (and less education) would have a steeper ray, while those with more education (and less healthcare) would have a flatter ray.

Allocative efficiency means that the particular mix of goods a society produces represents the combination that society most desires. How to determine what a society desires can be a controversial question, and is usually discussed in political science, sociology, and philosophy classes as well as in economics. At its most basic, allocative efficiency means producers supply the quantity of each product that consumers demand. Only one of the productively efficient choices will be the allocatively efficient choice for society as a whole.

Why Society Must Choose

Every economy faces two situations in which it may be able to expand consumption of all goods. In the first case, a society may discover that it has been using its resources inefficiently, in which case by improving efficiency and producing on the production possibilities frontier, it can have more of all goods (or at least more of some and less of none). In the second case, as resources grow over a period of years (e.g., more labor and more capital), the economy grows. As it does, the production possibilities frontier for a society will tend to shift outward and society will be able to afford more of all goods.

But improvements in productive efficiency take time to discover and implement, and economic growth happens only gradually. So, a society must choose between tradeoffs in the present. For government, this process often involves trying to identify where additional spending could do the most good and where reductions in spending would do the least harm. At the individual and **firm** level, the **market economy** coordinates a process in which firms seek to produce goods and services in the quantity, quality, and price that people want. But for both the government and the market economy in the short term, increases in production of one good typically mean offsetting decreases somewhere else in the economy.

The PPF and Comparative Advantage

While every society must choose how much of each good it should produce, it does not need to produce every single good it consumes. Often how much of a good a country decides to produce depends on how expensive it is to produce it versus buying it from a different country. As we saw earlier, the curvature of a country's PPF gives us information about the tradeoff between devoting resources to producing one good versus another. In particular, its slope gives the opportunity cost of producing one more unit of the good in the x-axis in terms of the other good (in the y-axis). Countries tend to have different opportunity costs of producing a specific good, either because of different climates, geography, technology or skills.

Suppose two countries, the US and Brazil, need to decide how much they will produce of two crops: sugar cane and wheat. Due to its climatic conditions, Brazil can produce a lot of sugar cane per acre but not much wheat. Conversely, the U.S. can produce a lot of wheat per acre, but not much sugar cane. Clearly, Brazil has a lower opportunity cost of producing sugar cane (in terms of wheat) than the U.S. The reverse is also true; the U.S. has a lower opportunity cost of producing wheat than Brazil. This can be illustrated by the PPFs of the two countries in Figure $2.4.3$:

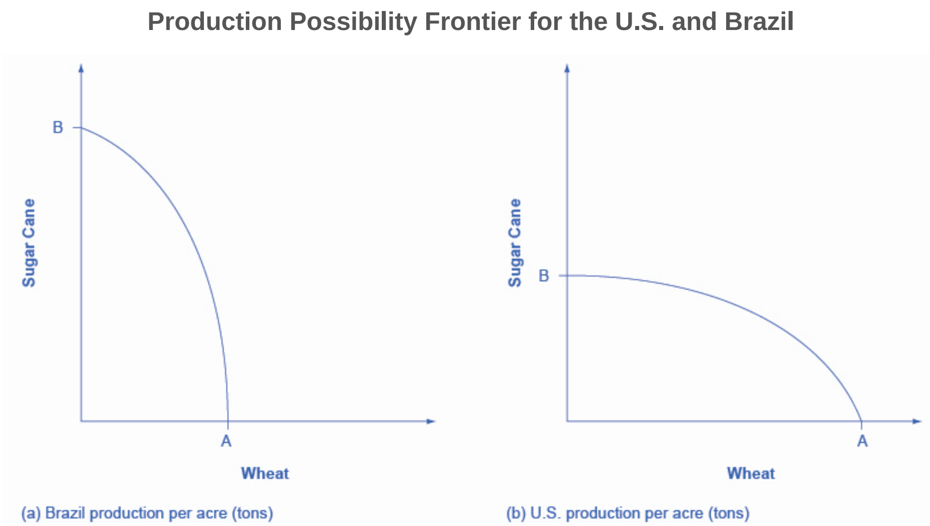


Figure $2.4.3$: The U.S. PPF is flatter than the Brazil PPF implying that the opportunity cost of wheat in term of sugar cane is lower in the U.S. than in Brazil. Conversely, the opportunity cost of sugar cane is lower in Brazil. The U.S. has comparative advantage in wheat and Brazil has comparative advantage in sugar cane.

When a country can produce a good at a lower opportunity cost than another country, we say that this country has a **comparative advantage** in that good. In our example, Brazil has a comparative advantage in sugar cane and the U.S. has a comparative advantage in wheat. One can easily see this with a simple observation of the extreme production points in the PPFs of the two countries. If Brazil devoted all of its resources to producing wheat, it would be producing at point A. If however it had devoted all of its resources to producing sugar cane instead, it would be producing a much larger amount, at point B. By moving from point A to point B Brazil would give up a relatively small quantity in wheat production to obtain a large production in sugar cane. The opposite is true for the U.S. If the U.S. moved from point A to B and produced only sugar cane, this would result in a large opportunity cost in terms of foregone wheat production.

The slope of the PPF gives the opportunity cost of producing an additional unit of wheat. While the slope is not constant throughout the PPFs, it is quite apparent that the PPF in Brazil is much steeper than in the U.S., and therefore the opportunity cost of wheat generally higher in Brazil. In the chapter on International Trade you will learn that countries' differences in comparative advantage determine which goods they will choose to produce and trade. When countries engage in trade, they specialize in the production of the goods that they have comparative advantage in, and trade part of that production for goods they do not have comparative advantage in. With trade, goods are produced where the opportunity cost is lowest, so total production increases, benefiting both trading parties.

Key Concepts and Summary

A production possibilities frontier defines the set of choices society faces for the combinations of goods and services it can produce given the resources available. The shape of the PPF is typically curved outward, rather than straight. Choices outside the PPF are unattainable and choices inside the PPF are wasteful. Over time, a growing economy will tend to shift the PPF outwards.

The law of diminishing returns holds that as increments of additional resources are devoted to producing something, the marginal increase in output will become smaller and smaller. All choices along a production possibilities frontier display productive efficiency; that is, it is impossible to use society's resources to produce more of one good without decreasing production of the other good. The specific choice along a production possibilities frontier that reflects the mix of goods society prefers is the choice with allocative efficiency. The curvature of the PPF is likely to differ by country, which results in different countries having comparative advantage in different goods. Total production can increase if countries specialize in the goods they have comparative advantage in and trade some of their production for the remaining goods.

Glossary

allocative efficiency
: when the mix of goods being produced represents the mix that society most desires

comparative advantage
: when a country can produce a good at a lower cost in terms of other goods; or, when a country has a lower opportunity cost of production

law of diminishing returns
: as additional increments of resources are added to producing a good or service, the marginal benefit from those additional increments will decline

production possibilities frontier (PPF)
: a diagram that shows the productively efficient combinations of two products that an economy can produce given the resources it has available.

productive efficiency
: when it is impossible to produce more of one good (or service) without decreasing the quantity produced of another good (or service)

Contributor

- Template:ContribOpenSTAXMicroEcon

2.5: Confronting Objections to the Economic Approach

Learning Objectives

- Analyze arguments against economic approaches to decision-making
- Interpret a tradeoff diagram
- Contrast normative statements and positive statements

It is one thing to understand the economic approach to decision-making and another thing to feel comfortable applying it. The sources of discomfort typically fall into two categories: that people do not act in the way that fits the economic way of thinking, and that even if people did act that way, they should try not to. Let's consider these arguments in turn.

First Objection: People, Firms, and Society Do Not Act Like This

The economic approach to decision-making seems to require more information than most individuals possess and more careful decision-making than most individuals actually display. After all, do you or any of your friends draw a budget constraint and mutter to yourself about maximizing utility before you head to the shopping mall? Do members of the U.S. Congress contemplate production possibilities frontiers before they vote on the annual budget? The messy ways in which people and societies operate somehow doesn't look much like neat budget constraints or smoothly curving production possibilities frontiers.

However, the economics approach can be a useful way to analyze and understand the tradeoffs of economic decisions even so. To appreciate this point, imagine for a moment that you are playing basketball, dribbling to the right, and throwing a bounce-pass to the left to a teammate who is running toward the basket. A physicist or engineer could work out the correct speed and trajectory for the pass, given the different movements involved and the weight and bounciness of the ball. But when you are playing basketball, you do not perform any of these calculations. You just pass the ball, and if you are a good player, you will do so with high accuracy.

Someone might argue: "The scientist's formula of the bounce-pass requires a far greater knowledge of physics and far more specific information about speeds of movement and weights than the basketball player actually has, so it must be an unrealistic description of how basketball passes are actually made." This reaction would be wrongheaded. The fact that a good player can throw the ball accurately because of practice and skill, without making a physics calculation, does not mean that the physics calculation is wrong.

Similarly, from an economic point of view, someone who goes shopping for groceries every week has a great deal of practice with how to purchase the combination of goods that will provide that person with utility, even if the shopper does not phrase decisions in terms of a budget constraint. Government institutions may work imperfectly and slowly, but in general, a democratic form of government feels pressure from voters and social institutions to make the choices that are most widely preferred by people in that society. So, when thinking about the economic actions of groups of people, firms, and society, it is reasonable, as a first approximation, to analyze them with the tools of economic analysis. For more on this, read about behavioral economics in the chapter on Consumer Choices.

Second Objection: People, Firms, and Society Should Not Act This Way

The economics approach portrays people as self-interested. For some critics of this approach, even if self-interest is an accurate description of how people behave, these behaviors are not moral. Instead, the critics argue that people should be taught to care more deeply about others. Economists offer several answers to these concerns.

First, economics is not a form of moral instruction. Rather, it seeks to describe economic behavior as it actually exists. Philosophers draw a distinction between **positive statements**, which describe the world as it is, and **normative statements**, which describe how the world should be. For example, an economist could analyze a proposed subway system in a certain city. If the expected benefits exceed the costs, he concludes that the project is worth doing—an example of positive analysis. Another economist argues for extended unemployment compensation during the Great Depression because a rich country like the United States should take care of its less fortunate citizens—an example of normative analysis.

Even if the line between positive and normative statements is not always crystal clear, economic analysis does try to remain rooted in the study of the actual people who inhabit the actual economy. Fortunately however, the assumption that individuals are purely self-interested is a simplification about human nature. In fact, we need to look no further than to **Adam Smith**, the very father of modern economics to find evidence of this. The opening sentence of his book, The Theory of Moral Sentiments, puts it very clearly: “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.” Clearly, individuals are both self-interested and altruistic.

Second, self-interested behavior and profit-seeking can be labeled with other names, such as personal choice and freedom. The ability to make personal choices about buying, working, and saving is an important personal freedom. Some people may choose high-pressure, high-paying jobs so that they can earn and spend a lot of money on themselves. Others may earn a lot of money and give it to charity or spend it on their friends and family. Others may devote themselves to a career that can require a great deal of time, energy, and expertise but does not offer high financial rewards, like being an elementary school teacher or a social worker. Still others may choose a job that does not take lots of their time or provide a high level of income, but still leaves time for family, friends, and contemplation. Some people may prefer to work for a large company; others might want to start their own business. People’s freedom to make their own economic choices has a moral value worth respecting.

Example $2.5.1$: Is a diagram by any other name the same?

When you study economics, you may feel buried under an avalanche of diagrams: diagrams in the text, diagrams in the lectures, diagrams in the problems, and diagrams on exams. Your goal should be to recognize the common underlying logic and pattern of the diagrams, not to memorize each of the individual diagrams.

This chapter uses only one basic diagram, although it is presented with different sets of labels. The consumption budget constraint and the production possibilities frontier for society, as a whole, are the same basic diagram. Figure $2.5.1$ shows an individual budget constraint and a production possibilities frontier for two goods, Good 1 and Good 2. The tradeoff diagram always illustrates three basic themes: scarcity, tradeoffs, and economic efficiency.

The first theme is **scarcity**. It is not feasible to have unlimited amounts of both goods. But even if the budget constraint or a PPF shifts, scarcity remains—just at a different level. The second theme is **tradeoffs**. As depicted in the budget constraint or the production possibilities frontier, it is necessary to give up some of one good to gain more of the other good. The details of this tradeoff vary. In a budget constraint, the tradeoff is determined by the relative prices of the goods: that is, the relative price of two goods in the consumption choice budget constraint. These tradeoffs appear as a straight line. However, the tradeoffs in many production possibilities frontiers are represented by a curved line because the law of diminishing returns holds that as resources are added to an area, the marginal gains tend to diminish. Regardless of the specific shape, tradeoffs remain.

The third theme is **economic efficiency**, or getting the most benefit from scarce resources. All choices on the production possibilities frontier show productive efficiency because in such cases, there is no way to increase the quantity of one good without decreasing the quantity of the other. Similarly, when an individual makes a choice along a budget constraint, there is no way to increase the quantity of one good without decreasing the quantity of the other. The choice on a production possibilities set that is socially preferred, or the choice on an individual’s budget constraint that is personally preferred, will display allocative efficiency.

The basic budget constraint/production possibilities frontier diagram will recur throughout this book. Some examples include using these tradeoff diagrams to analyze trade, labor supply versus leisure, saving versus consumption, environmental protection and economic output, equality of incomes and economic output, and the macroeconomic tradeoff between consumption and investment. Do not be confused by the different labels. The budget constraint/production possibilities frontier diagram is always just a tool for thinking carefully about scarcity, tradeoffs, and efficiency in a particular situation.

The Tradeoff Diagram

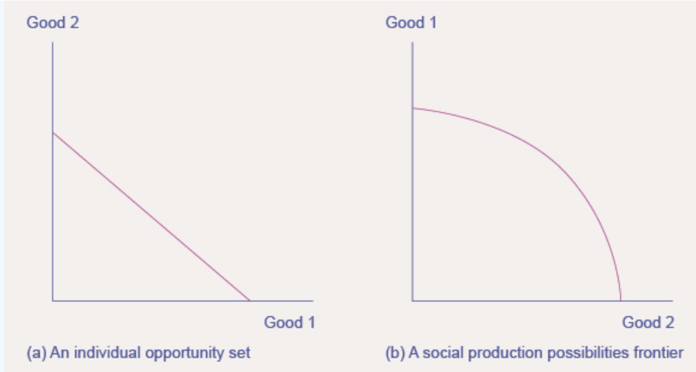


Figure 2.5.1: Both the individual opportunity set (or budget constraint) and the social production possibilities frontier show the constraints under which individual consumers and society as a whole operate. Both diagrams show the tradeoff in choosing more of one good at the cost of less of the other.

Third, self-interested behavior can lead to positive social results. For example, when people work hard to make a living, they create economic output. Consumers who are looking for the best deals will encourage businesses to offer goods and services that meet their needs. Adam Smith, writing in The Wealth of Nations, christened this property the **invisible hand**. In describing how consumers and producers interact in a market economy, Smith wrote:

> *"Every individual…generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it. By preferring the support of domestic to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, 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…By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it."*

The metaphor of the invisible hand suggests the remarkable possibility that broader social good can emerge from selfish individual actions.

Fourth, even people who focus on their own self-interest in the economic part of their life often set aside their own narrow self-interest in other parts of life. For example, you might focus on your own self-interest when asking your employer for a raise or negotiating to buy a car. But then you might turn around and focus on other people when you volunteer to read stories at the local library, help a friend move to a new apartment, or donate money to a charity. Self-interest is a reasonable starting point for analyzing many economic decisions, without needing to imply that people never do anything that is not in their own immediate self-interest.

✔ Example 2.5.2: Choices ... To What Degree?

What have we learned? We know that scarcity impacts all the choices we make. So, an economist might argue that people do not go on to get bachelor's degrees or master's degrees because they do not have the resources to make those choices or because their incomes are too low and/or the price of these degrees is too high. A bachelor's degree or a master's degree may not be available in their opportunity set.

The price of these degrees may be too high not only because the actual price, college tuition (and perhaps room and board), is too high. An economist might also say that for many people, the full opportunity cost of a bachelor's degree or a master's degree is too high. For these people, they are unwilling or unable to make the tradeoff of giving up years of working, and earning an income, to earn a degree.

Finally, the statistics introduced at the start of the chapter reveal information about **intertemporal choices**. An economist might say that people choose not to get a college degree because they may have to borrow money to go to college, and the interest they have to pay on that loan in the future will affect their decisions today. Also, it could be that some people have a preference for current consumption over future consumption, so they choose to work now at a lower salary and consume now, rather than putting that consumption off until after they graduate college.

Key Concepts and Summary

The economic way of thinking provides a useful approach to understanding human behavior. Economists make the careful distinction between positive statements, which describe the world as it is, and normative statements, which describe how the world should be. Even when economics analyzes the gains and losses from various events or policies, and thus draws normative conclusions about how the world should be, the analysis of economics is rooted in a positive analysis of how people, firms, and governments actually behave, not how they should behave.

References

Smith, Adam. “Of Restraints upon the Importation from Foreign Countries.” In The Wealth of Nations. London: Methuen & Co., 1904, first pub 1776), I.V. 2.9.

Smith, Adam. “Of the Propriety of Action.” In The Theory of Moral Sentiments. London: A. Millar, 1759, 1.

Glossary

invisible hand
idea that self-interested behavior by individuals can lead to positive social outcomes

normative statement
statement which describes how the world should be

positive statement
statement which describes the world as it is

Contributor

- Template:ContribOpenSTAXMicroEcon

2.E: Choice in a World of Scarcity (Exercises)

2.1: How Individuals Make Choices Based on Their Budget Constraint

Self-Check Questions

Q1

Suppose Alphonso's town raised the price of bus tickets to $\$1$ per trip (while the price of burgers stayed at $\$2$ and his budget remained $\$10$ per week.) Draw Alphonso's new budget constraint. What happens to the opportunity cost of bus tickets?

Review Questions

Q2

Explain why scarcity leads to tradeoffs.

Q3

Explain why individuals make choices that are directly on the budget constraint, rather than inside the budget constraint or outside it.

Critical Thinking Questions

Q4

Suppose Alphonso's town raises the price of bus tickets from $\$0.50$ to $\$1$ and the price of burgers rises from $\$2$ to $\$4$. Why is the opportunity cost of bus tickets unchanged? Suppose Alphonso's weekly spending money increases from $\$10$ to $\$20$. How is his budget constraint affected from all three changes? Explain.

Problems

Q5

Use this information to answer the following 4 questions: Marie has a weekly budget of $\$24$, which she likes to spend on magazines and pies.

a. If the price of a magazine is $\$4$ each, what is the maximum number of magazines she could buy in a week?
b. If the price of a pie is $12, what is the maximum number of pies she could buy in a week?
c. Draw Marie's budget constraint with pies on the horizontal axis and magazines on the vertical axis. What is the slope of the budget constraint?
d. What is Marie's opportunity cost of purchasing a pie?

Solution

S1

The opportunity cost of bus tickets is the number of burgers that must be given up to obtain one more bus ticket. Originally, when the price of bus tickets was 50 cents per trip, this opportunity cost was $0.50/2 = 0.25$ burgers. The reason for this is that at the original prices, one burger ($\$2$) costs the same as four bus tickets ($\$0.50$), so the opportunity cost of a burger is four bus tickets, and the opportunity cost of a bus ticket is 0.25 (the inverse of the opportunity cost of a burger). With the new, higher price of bus tickets, the opportunity cost rises to $\$1/\$2$ or 0.50. You can see this graphically since the slope of the new budget constraint is flatter than the original one. If Alphonso spends all of his budget on burgers, the higher price of bus tickets has no impact so the horizontal intercept of the budget constraint is the same. If he spends all of his budget on bus tickets, he can now afford only half as many, so the vertical intercept is half as much. In short, the budget constraint rotates clockwise around the horizontal intercept, flattening as it goes and the opportunity cost of bus tickets increases.

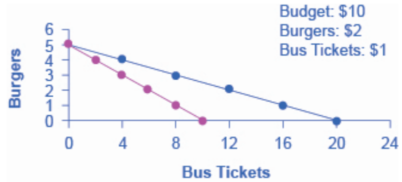


Figure 2.E.1: Alphonso's budget

2.2: The Production Possibilities Frontier and Social Choices

Self-Check Questions

Q1

Return to the example in Figure 2.2.2. Suppose there is an improvement in medical technology that enables more healthcare to be provided with the same amount of resources. How would this affect the production possibilities curve and, in particular, how would it affect the opportunity cost of education?

Q2

Could a nation be producing in a way that is allocatively efficient, but productively inefficient?

Q3

What are the similarities between a consumer's budget constraint and society's production possibilities frontier, not just graphically but analytically?

Review Questions

Q4

What is comparative advantage?

Q5

What does a production possibilities frontier illustrate?

Q6

Why is a production possibilities frontier typically drawn as a curve, rather than a straight line?

Q7

Explain why societies cannot make a choice above their production possibilities frontier and should not make a choice below it.

Q8

What are diminishing marginal returns?

Q9

What is productive efficiency? Allocative efficiency?

Critical Thinking Questions

Q10

During the Second World War, Germany's factories were decimated. It also suffered many human casualties, both soldiers and civilians. How did the war affect Germany's production possibilities curve?

Q11

It is clear that productive inefficiency is a waste since resources are being used in a way that produces less goods and services than a nation is capable of. Why is allocative inefficiency also wasteful?

Solution

S1

Because of the improvement in technology, the vertical intercept of the PPF would be at a higher level of healthcare. In other words, the PPF would rotate clockwise around the horizontal intercept. This would make the PPF steeper, corresponding to an increase in the opportunity cost of education, since resources devoted to education would now mean forgoing a greater quantity of healthcare.

S2

No. Allocative efficiency requires productive efficiency, because it pertains to choices along the production possibilities frontier.

S3

Both the budget constraint and the PPF show the constraint that each operates under. Both show a tradeoff between having more of one good but less of the other. Both show the opportunity cost graphically as the slope of the constraint (budget or PPF).

2.3: Confronting Objections to the Economic Approach

Self-Check Questions

Q1

Individuals may not act in the rational, calculating way described by the economic model of decision making, measuring utility and costs at the margin, but can you make a case that they behave approximately that way?

Q2

Would an op-ed piece in a newspaper urging the adoption of a particular economic policy be considered a positive or normative statement?

Q3

Would a research study on the effects of soft drink consumption on children's cognitive development be considered a positive or normative statement?

Review Questions

Q4

What is the difference between a positive and a normative statement?

Q5

Is the economic model of decision-making intended as a literal description of how individuals, firms, and the governments actually make decisions?

Q6

What are four responses to the claim that people should not behave in the way described in this chapter?

Critical Thinking Questions

Q7

What assumptions about the economy must be true for the invisible hand to work? To what extent are those assumptions valid in the real world?

Q8

Do economists have any particular expertise at making normative arguments? In other words, they have expertise at making positive statements (i.e., what will happen) about some economic policy, for example, but do they have special expertise to judge whether or not the policy should be undertaken?

Solution

S1

When individuals compare cost per unit in the grocery store, or characteristics of one product versus another, they are behaving approximately like the model describes.

S2

Since an op-ed makes a case for what should be, it is considered normative.

S3

Assuming that the study is not taking an explicit position about whether soft drink consumption is good or bad, but just reporting the science, it would be considered positive.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

3: Demand and Supply

3.1: Prelude to Demand and Supply

Skills to Develop

- Demand, Supply, and Equilibrium in Markets for Goods and Services
- Shifts in Demand and Supply for Goods and Services
- Changes in Equilibrium Price and Quantity: The Four-Step Process
- Price Ceilings and Price Floors

Farmer's Market

***Figure 3.1.1**: Organic vegetables and fruits that are grown and sold within a specific geographical region should, in theory, cost less than conventional produce because the transportation costs are less. That is not, however, usually the case. (Credit: modification of work by Natalie Maynor/Flickr Creative Commons)*

Why Can We Not Get Enough of Organic?

Organic food is increasingly popular, not just in the United States, but worldwide. At one time, consumers had to go to specialty stores or farmer's markets to find organic produce. Now it is available in most grocery stores. In short, organic is part of the mainstream.

Ever wonder why organic food costs more than conventional food? Why, say, does an organic Fuji apple cost $1.99 a pound, while its conventional counterpart costs $1.49 a pound? The same price relationship is true for just about every organic product on the market. If many organic foods are locally grown, would they not take less time to get to market and therefore be cheaper? What are the forces that keep those prices from coming down? Turns out those forces have a lot to do with this chapter's topic: demand and supply.

An auction bidder pays thousands of dollars for a dress Whitney Houston wore. A collector spends a small fortune for a few drawings by John Lennon. People usually react to purchases like these in two ways: their jaw drops because they think these are high prices to pay for such goods or they think these are rare, desirable items and the amount paid seems right.

When economists talk about prices, they are less interested in making judgments than in gaining a practical understanding of what determines prices and why prices change. Consider a price most of us contend with weekly: that of a gallon of gas. Why was the average price of gasoline in the United States $3.71 per gallon in June 2014? Why did the price for gasoline fall sharply to $2.07 per gallon by January 2015? To explain these price movements, economists focus on the determinants of what gasoline buyers are willing to pay and what gasoline sellers are willing to accept.

As it turns out, the price of gasoline in June of any given year is nearly always higher than the price in January of that same year; over recent decades, gasoline prices in midsummer have averaged about 10 cents per gallon more than their midwinter low. The likely reason is that people drive more in the summer, and are also willing to pay more for gas, but that does not explain how steeply gas prices fell. Other factors were at work during those six months, such as increases in supply and decreases in the demand for crude oil.

This chapter introduces the economic model of demand and supply—one of the most powerful models in all of economics. The discussion here begins by examining how demand and supply determine the price and the quantity sold in markets for goods and services, and how changes in demand and supply lead to changes in prices and quantities.

Contributor

- Template:ContribOpenSTAXMicroEcon

3.2: Demand, Supply, and Equilibrium in Markets for Goods and Services

> **Learning Objectives**
>
> - Explain demand, quantity demanded, and the law of demand
> - Identify a demand curve and a supply curve
> - Explain supply, quantity supply, and the law of supply
> - Explain equilibrium, equilibrium price, and equilibrium quantity

First let's first focus on what economists mean by demand, what they mean by supply, and then how demand and supply interact in a market.

Demand for Goods and Services

Economists use the term **demand** to refer to the amount of some good or service consumers are willing and able to purchase at each price. Demand is based on needs and wants—a consumer may be able to differentiate between a need and a want, but from an economist's perspective they are the same thing. Demand is also based on ability to pay. If you cannot pay for it, you have no effective demand.

What a buyer pays for a unit of the specific good or service is called **price**. The total number of units purchased at that price is called the quantity demanded. A rise in price of a good or service almost always decreases the **quantity demanded** of that good or service. Conversely, a fall in price will increase the quantity demanded. When the price of a gallon of gasoline goes up, for example, people look for ways to reduce their consumption by combining several errands, commuting by carpool or mass transit, or taking weekend or vacation trips closer to home. Economists call this inverse relationship between price and quantity demanded the law of demand. The **law of demand** assumes that all other variables that affect demand (to be explained in the next module) are held constant.

An example from the market for gasoline can be shown in the form of a table or a graph. A table that shows the quantity demanded at each price, such as Table $3.2.1$, is called a **demand schedule**. Price in this case is measured in dollars per gallon of gasoline. The quantity demanded is measured in millions of gallons over some time period (for example, per day or per year) and over some geographic area (like a state or a country). A **demand curve** shows the relationship between price and quantity demanded on a graph like Figure $3.2.1$, with quantity on the horizontal axis and the price per gallon on the vertical axis. (Note that this is an exception to the normal rule in mathematics that the independent variable (x) goes on the horizontal axis and the dependent variable (y) goes on the vertical. Economics is not math.)

The demand schedule shown by Table $3.2.1$ and the demand curve shown by the graph in Figure $3.2.1$are two ways of describing the same relationship between price and quantity demanded.

A Demand Curve for Gasoline

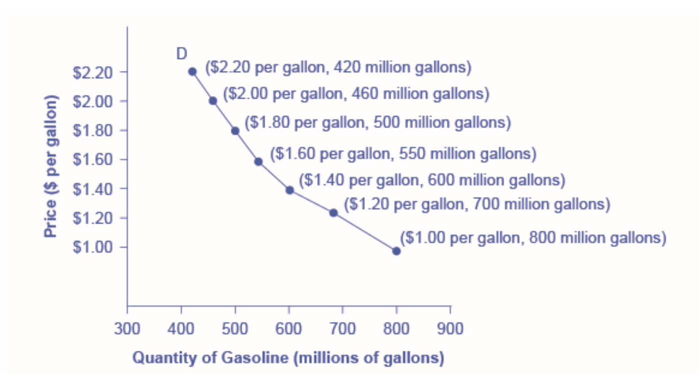


Figure $3.2.1$: The demand schedule shows that as price rises, quantity demanded decreases, and vice versa. These points are then graphed, and the line connecting them is the demand curve (D). The downward slope of the demand curve again illustrates the law of demand—the inverse relationship between prices and quantity demanded.

Table $3.2.1$: Price and Quantity Demanded of Gasoline

Price (per gallon)	Quantity Demanded (millions of gallon)

$1.00	800
$1.20	700
$1.40	600
$1.60	550
$1.80	500
$2.00	460
$2.20	420

Demand curves will appear somewhat different for each product. They may appear relatively steep or flat, or they may be straight or curved. Nearly all demand curves share the fundamental similarity that they slope down from left to right. So demand curves embody the law of demand: As the price increases, the quantity demanded decreases, and conversely, as the price decreases, the quantity demanded increases.

Confused about these different types of demand? Read the next Clear It Up feature.

Is demand the same as quantity demanded?

In economic terminology, demand is not the same as quantity demanded. When economists talk about demand, they mean the relationship between a range of prices and the quantities demanded at those prices, as illustrated by a demand curve or a demand schedule. When economists talk about quantity demanded, they mean only a certain point on the demand curve, or one quantity on the demand schedule. In short, demand refers to the curve and quantity demanded refers to the (specific) point on the curve.

Supply of Goods and Services

When economists talk about **supply**, they mean the amount of some **good** or **service** a producer is willing to supply at each price. Price is what the producer receives for selling one unit of a good or service. A rise in price almost always leads to an increase in the quantity supplied of that good or service, while a fall in price will decrease the **quantity supplied**. When the price of gasoline rises, for example, it encourages profit-seeking firms to take several actions: expand exploration for oil reserves; drill for more oil; invest in more pipelines and oil tankers to bring the oil to plants where it can be refined into gasoline; build new oil refineries; purchase additional pipelines and trucks to ship the gasoline to gas stations; and open more gas stations or keep existing gas stations open longer hours. Economists call this positive relationship between price and quantity supplied—that a higher price leads to a higher quantity supplied and a lower price leads to a lower quantity supplied—the **law of supply**. The law of supply assumes that all other variables that affect supply (to be explained in the next module) are held constant.

Still unsure about the different types of supply? See the following Clear It Up feature.

Is supply the same as quantity supplied?

In economic terminology, supply is not the same as quantity supplied. When economists refer to supply, they mean the relationship between a range of prices and the quantities supplied at those prices, a relationship that can be illustrated with a supply curve or a supply schedule. When economists refer to quantity supplied, they mean only a certain point on the supply curve, or one quantity on the supply schedule. In short, supply refers to the curve and quantity supplied refers to the (specific) point on the curve.

Figure $3.2.2$ illustrates the law of supply, again using the market for gasoline as an example. Like demand, supply can be illustrated using a table or a graph. A supply schedule is a table, like Table $3.2.2$, that shows the quantity supplied at a range of different prices. Again, price is measured in dollars per gallon of gasoline and quantity supplied is measured in millions of gallons. A supply curve is a graphic illustration of the relationship between price, shown on the vertical axis, and quantity, shown on the horizontal axis. The supply schedule and the supply curve are just two different ways of showing the same information. Notice that the horizontal and vertical axes on the graph for the supply curve are the same as for the demand curve.

A Supply Curve for Gasoline

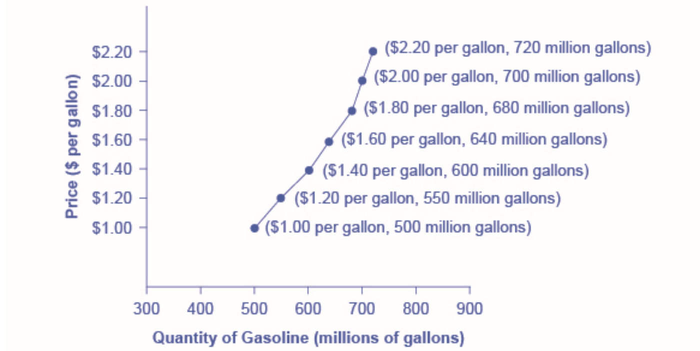


Figure 3.2.2: The supply schedule is the table that shows quantity supplied of gasoline at each price. As price rises, quantity supplied also increases, and vice versa. The supply curve (S) is created by graphing the points from the supply schedule and then connecting them. The upward slope of the supply curve illustrates the law of supply—that a higher price leads to a higher quantity supplied, and vice versa.

Table 3.2.2: Price and Supply of Gasoline

Price (per gallon)	Quantity Supplied (millions of gallon)
$1.00	500
$1.20	550
$1.40	600
$1.60	640
$1.80	680
$2.00	700
$2.20	720

The shape of supply curves will vary somewhat according to the product: steeper, flatter, straighter, or curved. Nearly all supply curves, however, share a basic similarity: they slope up from left to right and illustrate the law of supply: as the price rises, say, from $\$1.00$ per gallon to $\$2.20$ per gallon, the quantity supplied increases from 500 gallons to 720 gallons. Conversely, as the price falls, the quantity supplied decreases.

Equilibrium—Where Demand and Supply Intersect

Because the graphs for demand and supply curves both have price on the vertical axis and quantity on the horizontal axis, the demand curve and supply curve for a particular good or service can appear on the same graph. Together, demand and supply determine the price and the quantity that will be bought and sold in a market.

Figure $3.2.3$ illustrates the interaction of demand and supply in the market for gasoline. The demand curve (D) is identical to Figure $3.2.1$. The supply curve (S) is identical to Figure $3.2.1$. Table $3.2.3$ contains the same information in tabular form.

Demand and Supply for Gasoline

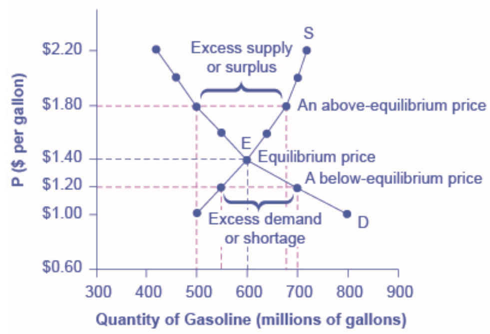


Figure 3.2.3: The demand curve (D) and the supply curve (S) intersect at the equilibrium point E, with a price of $1.40 and a quantity of 600. The equilibrium is the only price where quantity demanded is equal to quantity supplied. At a price above equilibrium like $1.80, quantity supplied exceeds the quantity demanded, so there is excess supply. At a price below equilibrium such as $1.20, quantity demanded exceeds quantity supplied, so there is excess demand.

Table 3.2.3: Price, *Quantity Demanded* and Quantity Supplied of Gasoline

Price (per gallon)	Quantity Demanded (millions of gallon)	Quantity Supplied (millions of gallon)
$1.00	800	500
$1.20	700	550
$1.40	**600**	**600**
$1.60	550	640
$1.80	500	680
$2.00	460	700
$2.20	420	720

Remember this: When two lines on a diagram cross, this intersection usually means something. The point where the supply curve (S) and the demand curve (D) cross, designated by point E in Figure $3.2.3$, is called the **equilibrium**. The **equilibrium price** is the only price where the plans of consumers and the plans of producers agree—that is, where the amount of the product consumers want to buy (quantity demanded) is equal to the amount producers want to sell (quantity supplied). This common quantity is called the **equilibrium quantity**. At any other price, the quantity demanded does not equal the quantity supplied, so the market is not in equilibrium at that price.

In Figure $3.2.3$, the equilibrium price is $\$1.40$ per gallon of gasoline and the equilibrium quantity is 600 million gallons. If you had only the demand and supply schedules, and not the graph, you could find the equilibrium by looking for the price level on the tables where the quantity demanded and the quantity supplied are equal.

The word "equilibrium" means "balance." If a market is at its equilibrium price and quantity, then it has no reason to move away from that point. However, if a market is not at equilibrium, then economic pressures arise to move the market toward the equilibrium price and the equilibrium quantity.

Imagine, for example, that the price of a gallon of gasoline was above the equilibrium price—that is, instead of $\$1.40$ per gallon, the price is $\$1.80$ per gallon. This above-equilibrium price is illustrated by the dashed horizontal line at the price of $\$1.80$ in Figure $3.2.3$. At this higher price, the quantity demanded drops from 600 to 500. This decline in quantity reflects how consumers react to the higher price by finding ways to use less gasoline.

Moreover, at this higher price of $\$1.80$, the quantity of gasoline supplied rises from the 600 to 680, as the higher price makes it more profitable for gasoline producers to expand their output. Now, consider how quantity demanded and quantity supplied are related at this above-equilibrium price. Quantity demanded has fallen to 500 gallons, while quantity supplied has risen to 680 gallons. In fact, at any above-equilibrium price, the quantity supplied exceeds the quantity demanded. We call this an **excess supply** or a **surplus**.

With a surplus, gasoline accumulates at gas stations, in tanker trucks, in pipelines, and at oil refineries. This accumulation puts pressure on gasoline sellers. If a surplus remains unsold, those firms involved in making and selling gasoline are not receiving enough cash to pay their workers and to cover their expenses. In this situation, some producers and sellers will want to cut prices, because it is better to sell at a lower price than not to sell at all. Once some sellers start cutting prices, others will follow to avoid losing sales. These price reductions in turn will stimulate a higher quantity demanded. So, if the price is above the equilibrium level, incentives built into the structure of demand and supply will create pressures for the price to fall toward the equilibrium.

Now suppose that the price is below its equilibrium level at $\$1.20$ per gallon, as the dashed horizontal line at this price in Figure $3.2.3$ shows. At this lower price, the quantity demanded increases from 600 to 700 as drivers take longer trips, spend more minutes warming up the car in the driveway in wintertime, stop sharing rides to work, and buy larger cars that get fewer miles to the gallon. However, the below-equilibrium price reduces gasoline producers' incentives to produce and sell gasoline, and the quantity supplied falls from 600 to 550.

When the price is below equilibrium, there is **excess demand**, or a **shortage**—that is, at the given price the quantity demanded, which has been stimulated by the lower price, now exceeds the quantity supplied, which had been depressed by the lower price. In this situation, eager gasoline buyers mob the gas stations, only to find many stations running short of fuel. Oil companies and gas stations recognize that they have an opportunity to make higher profits by selling what gasoline they have at a higher price. As a result, the price rises toward the equilibrium level. Read Demand, Supply, and Efficiency for more discussion on the importance of the demand and supply model.

Key Concepts and Summary

A demand schedule is a table that shows the quantity demanded at different prices in the market. A demand curve shows the relationship between quantity demanded and price in a given market on a graph. The law of demand states that a higher price typically leads to a lower quantity demanded.

A supply schedule is a table that shows the quantity supplied at different prices in the market. A supply curve shows the relationship between quantity supplied and price on a graph. The law of supply says that a higher price typically leads to a higher quantity supplied.

The equilibrium price and equilibrium quantity occur where the supply and demand curves cross. The equilibrium occurs where the quantity demanded is equal to the quantity supplied. If the price is below the equilibrium level, then the quantity demanded will exceed the quantity supplied. Excess demand or a shortage will exist. If the price is above the equilibrium level, then the quantity supplied will exceed the quantity demanded. Excess supply or a surplus will exist. In either case, economic pressures will push the price toward the equilibrium level.

References

Costanza, Robert, and Lisa Wainger. "No Accounting For Nature: How Conventional Economics Distorts the Value of Things." The Washington Post. September 2, 1990.

European Commission: Agriculture and Rural Development. 2013. "Overview of the CAP Reform: 2014-2024." Accessed April 13, 205. http://ec.europa.eu/agriculture/cap-post-2013/.

Radford, R. A. "The Economic Organisation of a P.O.W. Camp." Economica. no. 48 (1945): 189-201. www.jstor.org/stable/2550133.

Glossary

demand curve

a graphic representation of the relationship between price and quantity demanded of a certain good or service, with quantity on the horizontal axis and the price on the vertical axis

demand schedule

a table that shows a range of prices for a certain good or service and the quantity demanded at each price

demand
: the relationship between price and the quantity demanded of a certain good or service

equilibrium price
: the price where quantity demanded is equal to quantity supplied

equilibrium quantity
: the quantity at which quantity demanded and quantity supplied are equal for a certain price level

equilibrium
: the situation where quantity demanded is equal to the quantity supplied; the combination of price and quantity where there is no economic pressure from surpluses or shortages that would cause price or quantity to change

excess demand
: at the existing price, the quantity demanded exceeds the quantity supplied; also called a shortage

excess supply
: at the existing price, quantity supplied exceeds the quantity demanded; also called a surplus

law of demand
: the common relationship that a higher price leads to a lower quantity demanded of a certain good or service and a lower price leads to a higher quantity demanded, while all other variables are held constant

law of supply
: the common relationship that a higher price leads to a greater quantity supplied and a lower price leads to a lower quantity supplied, while all other variables are held constant

price
: what a buyer pays for a unit of the specific good or service

quantity demanded
: the total number of units of a good or service consumers are willing to purchase at a given price

quantity supplied
: the total number of units of a good or service producers are willing to sell at a given price

shortage
: at the existing price, the quantity demanded exceeds the quantity supplied; also called excess demand

supply curve
: a line that shows the relationship between price and quantity supplied on a graph, with quantity supplied on the horizontal axis and price on the vertical axis

supply schedule
: a table that shows a range of prices for a good or service and the quantity supplied at each price

supply
: the relationship between price and the quantity supplied of a certain good or service

surplus
: at the existing price, quantity supplied exceeds the quantity demanded; also called excess supply

Contributor

- Template:ContribOpenSTAXMicroEcon

3.3: Shifts in Demand and Supply for Goods and Services

Learning Objectives

- Identify factors that affect demand
- Graph demand curves and demand shifts
- Identify factors that affect supply
- Graph supply curves and supply shifts

The previous module explored how **price** affects the quantity demanded and the quantity supplied. The result was the demand curve and the supply curve. Price, however, is not the only thing that influences demand. Nor is it the only thing that influences supply. For example, how is demand for vegetarian food affected if, say, health concerns cause more consumers to avoid eating meat? Or how is the supply of diamonds affected if diamond producers discover several new diamond mines? What are the major factors, in addition to the price, that influence demand or supply?

What Factors Affect Demand?

We defined demand as the amount of some product a consumer is willing and able to purchase at each price. That suggests at least two factors in addition to price that affect demand. Willingness to purchase suggests a desire, based on what economists call tastes and preferences. If you neither need nor want something, you will not buy it. Ability to purchase suggests that income is important. Professors are usually able to afford better housing and transportation than students, because they have more income. Prices of substitute and complementary goods can affect demand also. If you need a new car, the price of a Honda may affect your demand for a Ford. Finally, the size or composition of the population can affect demand. The more children a family has, the greater their demand for clothing. The more driving-age children a family has, the greater their demand for car insurance, and the less for diapers and baby formula.

These factors matter both for demand by an individual and demand by the market as a whole. Exactly how do these various factors affect demand, and how do we show the effects graphically? To answer those questions, we need the *ceteris paribus* assumption.

The *Ceteris Paribus* Assumption

A **demand curve** or a **supply curve** is a relationship between two, and only two, variables: quantity on the horizontal axis and price on the vertical axis. The assumption behind a demand curve or a supply curve is that no relevant economic factors, other than the product's price, are changing. Economists call this assumption **ceteris paribus**, a Latin phrase meaning "other things being equal." Any given demand or supply curve is based on the ceteris paribus assumption that all else is held equal. A demand curve or a supply curve is a relationship between two, and only two, variables when all other variables are kept constant. If all else is not held equal, then the laws of supply and demand will not necessarily hold, as the following Clear It Up feature shows.

Example $3.3.1$: When does ceteris paribus apply?

Ceteris paribus is typically applied when we look at how changes in price affect demand or supply, but ceteris paribus can be applied more generally. In the real world, demand and supply depend on more factors than just price. For example, a consumer's demand depends on income and a producer's supply depends on the cost of producing the product. How can we analyze the effect on demand or supply if multiple factors are changing at the same time—say price rises and income falls? The answer is that we examine the changes one at a time, assuming the other factors are held constant.

For example, we can say that an increase in the price reduces the amount consumers will buy (assuming income, and anything else that affects demand, is unchanged). Additionally, a decrease in income reduces the amount consumers can afford to buy (assuming price, and anything else that affects demand, is unchanged). This is what the *ceteris paribus* assumption really means. In this particular case, after we analyze each factor separately, we can combine the results. The amount consumers buy falls for two reasons: first because of the higher price and second because of the lower income.

How Does Income Affect Demand?

Let's use income as an example of how factors other than price affect demand. Figure $3.3.1$ shows the initial demand for automobiles as D_0. At point Q, for example, if the price is $\$20,000$ per car, the quantity of cars demanded is 18 million. D_0 also shows how the quantity of cars demanded would change as a result of a higher or lower price. For example, if the price of a car rose to $\$22,000$, the quantity demanded would decrease to 17 million, at point R.

The original demand curve D_0, like every demand curve, is based on the ceteris paribus assumption that no other economically relevant factors change. Now imagine that the economy expands in a way that raises the incomes of many people, making cars more affordable. How will this affect demand? How can we show this graphically?

Return to Figure $3.3.1$. The price of cars is still $\$20,000$, but with higher incomes, the quantity demanded has now increased to 20 million cars, shown at point S. As a result of the higher income levels, the demand curve shifts to the right to the new demand curve D_1, indicating an increase in demand. Table $3.3.1$ shows clearly that this increased demand would occur at every price, not just the original one.

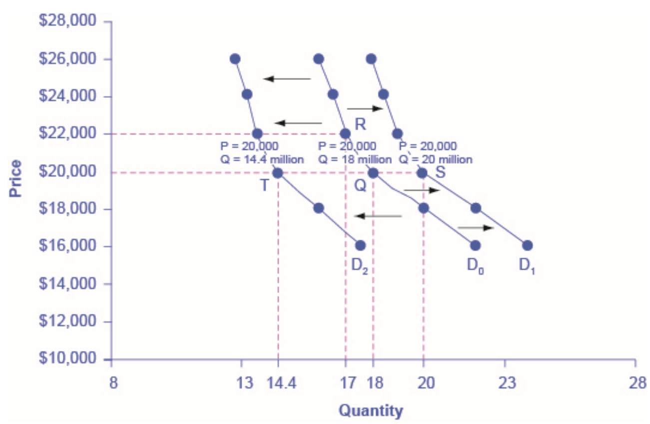


Figure $3.3.1$: *Increased demand means that at every given price, the quantity demanded is higher, so that the demand curve shifts to the right from D_0 to D_1. Decreased demand means that at every given price, the quantity demanded is lower, so that the demand curve shifts to the left from D_0 to D_2.*

Table $3.3.1$: *Price and Demand Shifts: A Car Example*

Price	Decrease to D_2	Original Quantity Demanded D_0	Increase to D_1
$16,000	17.6 million	22.0 million	24.0 million
$18,000	16.0 million	20.0 million	22.0 million
$20,000	14.4 million	18.0 million	20.0 million
$22,000	13.6 million	17.0 million	19.0 million
$24,000	13.2 million	16.5 million	18.5 million
$26,000	12.8 million	16.0 million	18.0 million

Now, imagine that the economy slows down so that many people lose their jobs or work fewer hours, reducing their incomes. In this case, the decrease in income would lead to a lower quantity of cars demanded at every given price, and the original demand curve D_0 would shift left to D_2. The shift from D_0 to D_2 represents such a decrease in demand: At any given price level, the quantity demanded is now lower. In this example, a price of $\$20,000$ means 18 million cars sold along the original demand curve, but only 14.4 million sold after demand fell.

When a demand curve shifts, it does not mean that the quantity demanded by every individual buyer changes by the same amount. In this example, not everyone would have higher or lower income and not everyone would buy or not buy an additional car. Instead, a shift in a demand curve captures an pattern for the market as a whole.

In the previous section, we argued that higher income causes greater demand at every price. This is true for most goods and services. For some—luxury cars, vacations in Europe, and fine jewelry—the effect of a rise in income can be especially pronounced. A product whose demand rises when income rises, and vice versa, is called a **normal good**. A few exceptions to this pattern do exist. As incomes rise, many people will buy fewer generic brand groceries and more name brand groceries. They are less likely to buy used cars and more likely to buy new cars. They will be less likely to rent an apartment and more likely to own a home, and so on. A product whose demand falls when income rises, and vice versa, is called an **inferior good**. In other words, when income increases, the demand curve shifts to the left.

Other Factors That Shift Demand Curves

Income is not the only factor that causes a shift in demand. Other things that change demand include changes to prices of substitutes and complementary goods, tastes and preferences, the composition or size of the population, and even expectations. A change in any one of the underlying factors that determine what quantity people are willing to buy at a given price will cause a shift in demand. Graphically, the new demand curve lies either to the right (an increase) or to the left (a decrease) of the original demand curve. Let's look at these factors.

Changes in the Prices of Substitute and Complementary Goods

The demand for a product can also be affected by changes in the prices of goods that are as substitutes or complements. A **substitute** is a good or service that can be used in place of another good or service. As electronic books, like this one, become more available, you would expect to see a decrease in demand for traditional printed books. A lower price for a substitute decreases demand for the other product. For example, in recent years as the price of tablet computers has fallen, the quantity demanded has increased (because of the law of demand). Since people are purchasing tablets, there has been a decrease in demand for laptops, which can be shown graphically as a leftward shift in the demand curve for laptops. A higher price for a substitute good has the reverse effect.

Other goods are **complements** for each other, meaning that the goods are often used together, because consumption of one good tends to enhance consumption of the other. Examples include breakfast cereal and milk; notebooks and pens or pencils, golf balls and golf clubs; gasoline and sport utility vehicles; and the five-way combination of bacon, lettuce, tomato, mayonnaise, and bread. If the price of golf clubs rises, since the quantity demanded of golf clubs falls (because of the law of demand), demand for a complement good like golf balls decreases, too. Similarly, a higher price for skis would shift the demand curve for a complement good like ski resort trips to the left, while a lower price for a complement has the reverse effect.

Changing Tastes or Preferences

From 1980 to 2014, the per-person consumption of chicken by Americans rose from 48 pounds per year to 85 pounds per year, and consumption of beef fell from 77 pounds per year to 54 pounds per year, according to the U.S. Department of Agriculture (USDA). Changes like these are largely due to movements in taste, which change the quantity of a good demanded at every price: that is, they shift the demand curve for that good, rightward for chicken and leftward for beef.

Changes in the Composition of the Population

The proportion of elderly citizens in the United States population is rising. It rose from 9.8% in 1970 to 12.6% in 2000, and will be a projected (by the **U.S. Census Bureau**) 20% of the population by 2030. A society with relatively more children, like the United States in the 1960s, will have greater demand for goods and services like tricycles and day care facilities. A society with relatively more elderly persons, as the United States is projected to have by 2030, has a higher demand for nursing homes and hearing aids. Similarly, changes in the size of the population can affect the demand for housing and many other goods. Each of these changes in demand will be shown as a shift in the demand curve.

Changes in Expectations about Future Prices or Other Factors that Affect Demand

While it is clear that the price of a good affects the quantity demanded, it is also true that expectations about the future price (or expectations about tastes and preferences, income, and so on) can affect demand. For example, if people hear that a hurricane is coming, they may rush to the store to buy flashlight batteries and bottled water. If people learn that the price of a good like coffee is likely to rise in the future, they may head for the store to stock up on coffee now. These changes in demand are shown as shifts in the curve. Therefore, a shift in demand happens when a change in some economic factor (other than price) causes a different quantity to be demanded at every price. The following Work It Out feature shows how this happens.

Example $3.3.1$: Shift in Demand

A shift in demand means that at any price (and at every price), the quantity demanded will be different than it was before. Following is an example of a shift in demand due to an income increase.

Step 1: Draw the graph of a demand curve for a normal good like pizza. Pick a price (like P_0). Identify the corresponding Q_0. An example is shown in Figure $3.3.2$.

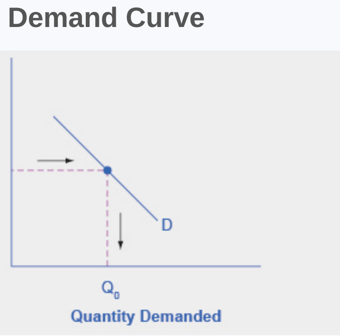


Figure $\mathbf{3.3.2}$***:*** *The demand curve can be used to identify how much consumers would buy at any given price.*

Step 2: Suppose income increases. As a result of the change, are consumers going to buy more or less pizza? The answer is more. Draw a dotted horizontal line from the chosen price, through the original quantity demanded, to the new point with the new Q_1. Draw a dotted vertical line down to the horizontal axis and label the new Q_1. An example is provided in Figure $3.3.3$.

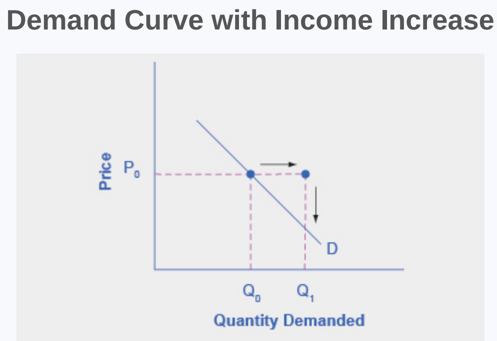


Figure $\mathbf{3.3.3}$***:*** *With an increase in income, consumers will purchase larger quantities, pushing demand to the right.*

Step 3: Now, shift the curve through the new point. You will see that an increase in income causes an upward (or rightward) shift in the demand curve, so that at any price the quantities demanded will be higher, as shown in Figure $3.3.4$.

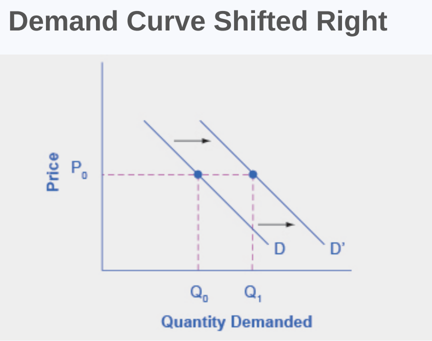


Figure $\mathbf{3.3.4}$***:*** *With an increase in income, consumers will purchase larger quantities, pushing demand to the right, and causing the demand curve to shift right.*

Summing Up Factors That Change Demand

Six factors that can shift demand curves are summarized in Figure $3.3.5$. The direction of the arrows indicates whether the demand curve shifts represent an increase in demand or a decrease in demand. Notice that a change in the price of the good or service itself is not listed among the factors that can shift a demand curve. A change in the price of a good or service causes a movement along a specific demand curve, and it typically leads to some change in the quantity demanded, but it does not shift the demand curve.

Factors That Shift Demand Curves

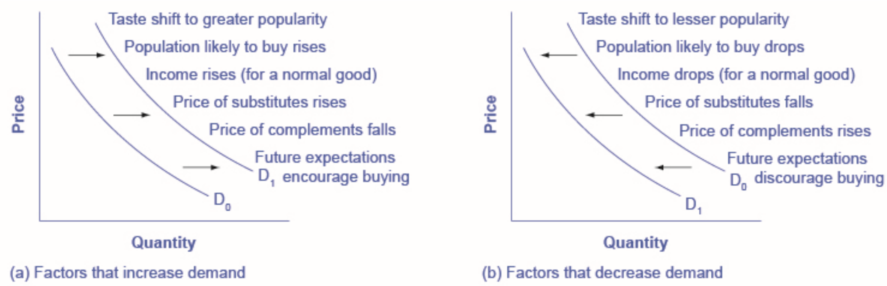


Figure $3.3.5$: *(a) A list of factors that can cause an increase in demand from D_0 to D_1. (b) The same factors, if their direction is reversed, can cause a decrease in demand from D_0 to D_1.*

When a demand curve shifts, it will then intersect with a given supply curve at a different equilibrium price and quantity. We are, however, getting ahead of our story. Before discussing how changes in demand can affect equilibrium price and quantity, we first need to discuss shifts in supply curves.

How Production Costs Affect Supply

A supply curve shows how quantity supplied will change as the price rises and falls, assuming *ceteris paribus* so that no other economically relevant factors are changing. If other factors relevant to supply do change, then the entire supply curve will shift. Just as a shift in demand is represented by a change in the quantity demanded at every price, a **shift in supply** means a change in the quantity supplied at every price.

Prices of Inputs

In thinking about the factors that affect supply, remember what motivates firms: profits, which are the difference between revenues and costs. Goods and services are produced using combinations of labor, materials, and machinery, or what we call **inputs** or **factors of production**. If a firm faces lower costs of production, while the prices for the good or service the firm produces remain unchanged, a firm's profits go up. When a firm's profits increase, it is more motivated to produce output, since the more it produces the more profit it will earn. So, when costs of production fall, a firm will tend to supply a larger quantity at any given price for its output. This can be shown by the supply curve shifting to the right.

Take, for example, a messenger company that delivers packages around a city. The company may find that buying gasoline is one of its main costs. If the price of gasoline falls, then the company will find it can deliver messages more cheaply than before. Since lower costs correspond to higher profits, the messenger company may now supply more of its services at any given price. For example, given the lower gasoline prices, the company can now serve a greater area, and increase its supply.

Conversely, if a firm faces higher costs of production, then it will earn lower profits at any given selling price for its products. As a result, a higher cost of production typically causes a firm to supply a smaller quantity at any given price. In this case, the supply curve shifts to the left.

Consider the supply for cars, shown by curve S_0 in Figure $3.3.6$. Point J indicates that if the price is $\$20,000$, the quantity supplied will be 18 million cars. If the price rises to $\$22,000$ per car, ceteris paribus, the quantity supplied will rise to 20 million cars, as point K on the S_0 curve shows. The same information can be shown in table form, as in Table $3.3.2$.

Shifts in Supply: A Car Example

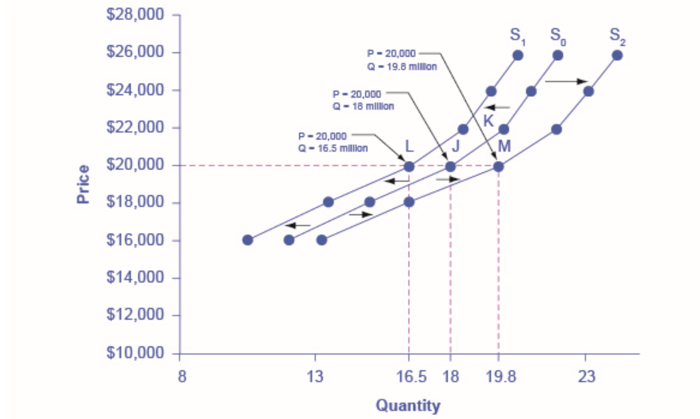


***Figure 3.3.6**: Decreased supply means that at every given price, the quantity supplied is lower, so that the supply curve shifts to the left, from S_0 to S_1. Increased supply means that at every given price, the quantity supplied is higher, so that the supply curve shifts to the right, from S_0 to S_2.*

***Table 3.3.2**: Price and Shifts in Supply: A Car Example*

Price	Decrease to S_1	Original Quantity Supplied S_0	Increase to S_2
$16,000	10.5 million	12.0 million	13.2 million
$18,000	13.5 million	15.0 million	16.5 million
$20,000	16.5 million	18.0 million	19.8 million
$22,000	18.5 million	20.0 million	22.0 million
$24,000	19.5 million	21.0 million	23.1 million
$26,000	20.5 million	22.0 million	24.2 million

Now, imagine that the price of steel, an important ingredient in manufacturing cars, rises, so that producing a car has become more expensive. At any given price for selling cars, car manufacturers will react by supplying a lower quantity. This can be shown graphically as a leftward shift of supply, from S_0 to S_1, which indicates that at any given price, the quantity supplied decreases. In this example, at a price of $\$20,000$, the quantity supplied decreases from 18 million on the original supply curve (S_0) to 16.5 million on the supply curve S_1, which is labeled as point L.

Conversely, if the price of steel decreases, producing a car becomes less expensive. At any given price for selling cars, car manufacturers can now expect to earn higher profits, so they will supply a higher quantity. The shift of supply to the right, from S_0 to S_2, means that at all prices, the quantity supplied has increased. In this example, at a price of $\$20,000$, the quantity supplied increases from 18 million on the original supply curve (S_0) to 19.8 million on the supply curve S_2, which is labeled M.

Other Factors That Affect Supply

In the example above, we saw that changes in the prices of inputs in the production process will affect the cost of production and thus the supply. Several other things affect the cost of production, too, such as changes in weather or other natural conditions, new technologies for production, and some government policies.

While the prices of inputs are utmost on the minds of firms, especially labor, since this is the largest costs for most firms, agriculture, experiences a great deal of uncertainty in their production costs in other ways. The cost of production for many agricultural products will be affected by changes in natural conditions. For example, in 2014 the Manchurian Plain in Northeastern

China, which produces most of the country's wheat, corn, and soybeans, experienced its most severe drought in 50 years. A drought decreases the supply of agricultural products, which means that at any given price, a lower quantity will be supplied; conversely, especially good weather would shift the supply curve to the right.

Price of Related Outputs

Firm's can also produce several goods that are related but not exactly similar. For example, yogurt manufacturers have experienced a change in their sales with the greek yogurt craze. Standard yogurt sales are no longer as popular as they once were and thus, are not as profitable. Therefore, firms may decide to reduce their supply of standard yogurt in favor of producing more greek-style yogurt.

Technology

When a **firm** discovers a new technology that allows the firm to produce at a lower cost, the supply curve will shift to the right, as well. For instance, in the 1960s a major scientific effort nicknamed the Green Revolution focused on breeding improved seeds for basic crops like wheat and rice. By the early 1990s, more than two-thirds of the wheat and rice in low-income countries around the world was grown with these Green Revolution seeds—and the harvest was twice as high per acre. A technological improvement that reduces costs of production will shift supply to the right, so that a greater quantity will be produced at any given price.

Size of the Industry

The size of the industry also plays an important role in determining the output of a market. The larger and more competitive an industry, the more production we are likely to see. Monopolists, an industry of one seller, typically ration, or limit their supply, to increase the price consumers must pay. We will learn more about this in future chapters but, the larger the size usually means the larger the supply of the product.

Government Policies (Taxes and Subsidies)

Government policies can affect the cost of production and the supply curve through taxes, regulations, and subsidies. For example, the U.S. government imposes a tax on alcoholic beverages that collects about $8 billion per year from producers. Taxes are treated as costs by businesses. Higher costs decrease supply for the reasons discussed above. Other examples of policy that can affect cost are the wide array of government regulations that require firms to spend money to provide a cleaner environment or a safer workplace; complying with regulations increases costs.

A government subsidy, on the other hand, is the opposite of a tax. A subsidy occurs when the government pays a firm directly or reduces the firm’s taxes if the firm carries out certain actions. From the firm’s perspective, taxes or regulations are an additional cost of production that shifts supply to the left, leading the firm to produce a lower quantity at every given price. Government subsidies reduce the cost of production and increase supply at every given price, shifting supply to the right. The following Work It Out feature shows how this shift happens.

Example 3.3.1: Shift in Supply

We know that a supply curve shows the minimum price a firm will accept to produce a given quantity of output. What happens to the supply curve when the cost of production goes up? Following is an example of a shift in supply due to a production cost increase.

Step 1: Draw a graph of a supply curve for pizza. Pick a quantity (like Q_0). If you draw a vertical line up from Q_0 to the supply curve, you will see the price the firm chooses. An example is shown in Figure 3.3.7.

Suppy Curve

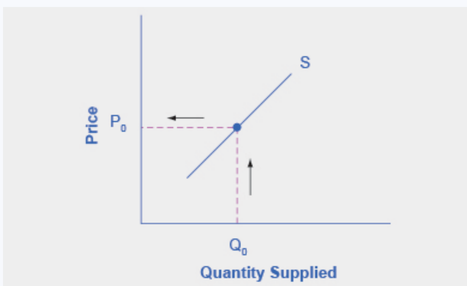


Figure 3.3.7: The supply curve can be used to show the minimum price a firm will accept to produce a given quantity of output.

Step 2: Why did the firm choose that price and not some other? One way to think about this is that the price is composed of two parts. The first part is the cost of producing pizzas at the margin; in this case, the cost of producing the pizza, including cost of ingredients (dough, sauce, cheese, pepperoni, and so on), the cost of the pizza oven, the rent on the shop, and the wages of the workers. The second part is the firm's desired profit, which is determined, among other factors, by the profit margins in that particular business. If you add these two parts together, you get the price the firm wishes to charge. The quantity Q_0 and associated price P_0 give you one point on the firm's supply curve, as shown in Figure 3.3.8.

Setting Prices

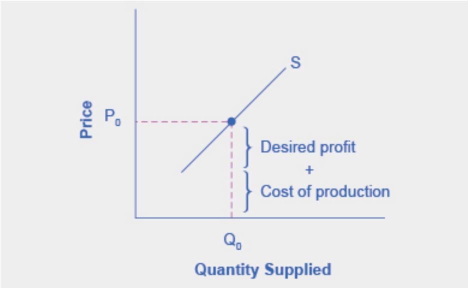


Figure 3.3.8: The cost of production and the desired profit equal the price a firm will set for a product.

Step 3: Now, suppose that the cost of production goes up. Perhaps cheese has become more expensive by $0.75 per pizza. If that is true, the firm will want to raise its price by the amount of the increase in cost ($0.75). Draw this point on the supply curve directly above the initial point on the curve, but $0.75 higher, as shown in Figure 3.3.9.

Increasing Costs Leads to Increasing Price

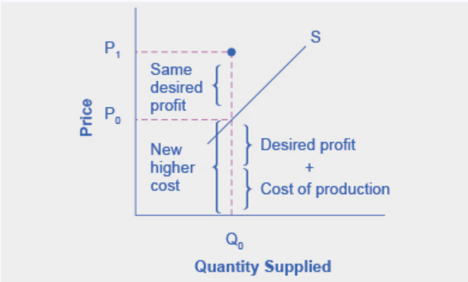


Figure 3.3.9: Because the cost of production and the desired profit equal the price a firm will set for a product, if the cost of production increases, the price for the product will also need to increase.

Step 4: Shift the supply curve through this point. You will see that an increase in cost causes an upward (or a leftward) shift of the supply curve so that at any price, the quantities supplied will be smaller, as shown in Figure 3.3.10.

Supply Curve Shifts

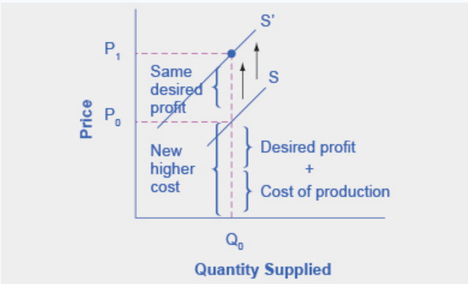


Figure 3.3.10: When the cost of production increases, the supply curve shifts upwardly to a new price level.

Summing Up Factors That Change Supply

Changes in the cost of inputs (including natural disasters), the price of related goods that firms could produce instead, new technologies, size of the industry, and the impact of government decisions (taxes and subsidies) all affect the cost of production. In turn, these factors affect how much firms are willing to supply at any given price.

Figure 3.3.11 summarizes factors that change the supply of goods and services. Notice that a change in the price of the product itself is not among the factors that shift the supply curve. Although a change in price of a good or service typically causes a change in quantity supplied or a movement along the supply curve for that specific good or service, it does not cause the supply curve itself to shift.

Factors That Shift Supply Curves

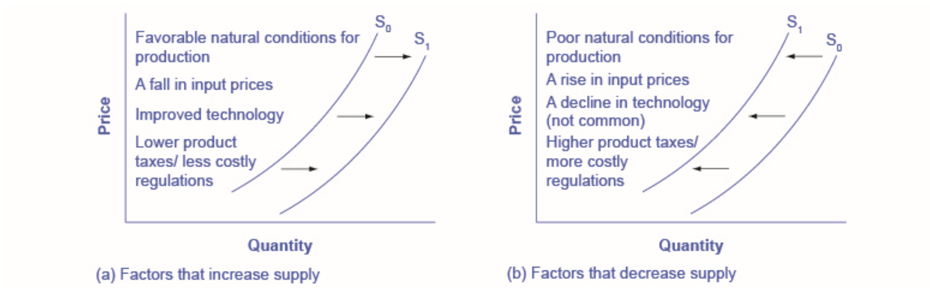


Figure 3.3.11: *(a) A list of factors that can cause an increase in supply from S_0 to S_1. (b) The same factors, if their direction is reversed, can cause a decrease in supply from S_0 to S_1.*

Because demand and supply curves appear on a two-dimensional diagram with only price and quantity on the axes, an unwary visitor to the land of economics might be fooled into believing that economics is about only four topics: demand, supply, price, and quantity. However, demand and supply are really "umbrella" concepts: demand covers all the factors that affect demand, and supply covers all the factors that affect supply. Factors other than price that affect demand and supply are included by using shifts in the demand or the supply curve. In this way, the two-dimensional demand and supply model becomes a powerful tool for analyzing a wide range of economic circumstances.

Key Concepts and Summary

Economists often use the *ceteris paribus* or "other things being equal" assumption: while examining the economic impact of one event, all other factors remain unchanged for the purpose of the analysis. Factors that can shift the demand curve for goods and services, causing a different quantity to be demanded at any given price, include changes in tastes, population, income, prices of substitute or complement goods, and expectations about future conditions and prices. Factors that can shift the supply curve for goods and services, causing a different quantity to be supplied at any given price, include input prices, natural conditions, changes in technology, and government taxes, regulations, or subsidies.

References

Landsburg, Steven E. The Armchair Economist: Economics and Everyday Life. New York: The Free Press. 2012. specifically Section IV: How Markets Work.

National Chicken Council. 2015. "Per Capita Consumption of Poultry and Livestock, 1965 to Estimated 2015, in Pounds." Accessed April 13, 2015. http://www.nationalchickencouncil.or...012-in-pounds/.

Wessel, David. "Saudi Arabia Fears $40-a-Barrel Oil, Too." The Wall Street Journal. May 27, 2004, p. 42. http://online.wsj.com/news/articles/...61000087822300.

Glossary

ceteris paribus

other things being equal

complements

goods that are often used together so that consumption of one good tends to enhance consumption of the other

factors of production

the combination of labor, materials, and machinery that is used to produce goods and services; also called inputs

inferior good

a good in which the quantity demanded falls as income rises, and in which quantity demanded rises and income falls

inputs

the combination of labor, materials, and machinery that is used to produce goods and services; also called factors of production

normal good

a good in which the quantity demanded rises as income rises, and in which quantity demanded falls as income falls

shift in demand

when a change in some economic factor (other than price) causes a different quantity to be demanded at every price

shift in supply

when a change in some economic factor (other than price) causes a different quantity to be supplied at every price

substitute

a good that can replace another to some extent, so that greater consumption of one good can mean less of the other

Contributor

- Template:ContribOpenSTAXMicroEcon

3.4: Changes in Equilibrium Price and Quantity- The Four-Step Process

Learning Objectives

- Identify equilibrium price and quantity through the four-step process
- Graph equilibrium price and quantity
- Contrast shifts of demand or supply and movements along a demand or supply curve
- Graph demand and supply curves, including equilibrium price and quantity, based on real-world examples

Let's begin this discussion with a single economic event. It might be an event that affects demand, like a change in income, population, tastes, prices of substitutes or complements, or expectations about future prices. It might be an event that affects supply, like a change in natural conditions, input prices, or technology, or government policies that affect production. How does this economic event affect equilibrium price and quantity? We will analyze this question using a four-step process.

Step 1: Draw a demand and supply model before the economic change took place. To establish the model requires four standard pieces of information: The law of demand, which tells us the slope of the demand curve; the law of supply, which gives us the slope of the supply curve; the shift variables for demand; and the shift variables for supply. From this model, find the initial equilibrium values for price and quantity.

Step 2: Decide whether the economic change being analyzed affects demand or supply. In other words, does the event refer to something in the list of demand factors or supply factors?

Step 3: Decide whether the effect on demand or supply causes the curve to shift to the right or to the left, and sketch the new demand or supply curve on the diagram. In other words, does the event increase or decrease the amount consumers want to buy or producers want to sell?

Step 4: Identify the new equilibrium and then compare the original equilibrium price and quantity to the new equilibrium price and quantity.

Let's consider one example that involves a shift in supply and one that involves a shift in demand. Then we will consider an example where both supply and demand shift.

Good Weather for Salmon Fishing

In the summer of 2000, weather conditions were excellent for commercial salmon fishing off the California coast. Heavy rains meant higher than normal levels of water in the rivers, which helps the salmon to breed. Slightly cooler ocean temperatures stimulated the growth of plankton, the microscopic organisms at the bottom of the ocean food chain, providing everything in the ocean with a hearty food supply. The ocean stayed calm during fishing season, so commercial fishing operations did not lose many days to bad weather. How did these climate conditions affect the quantity and price of salmon? Figure $3.4.1$ illustrates the four-step approach, which is explained below, to work through this problem. Table $3.4.1$ provides the information to work the problem as well.

Good Weather for Salmon Fishing: The Four-Step Process

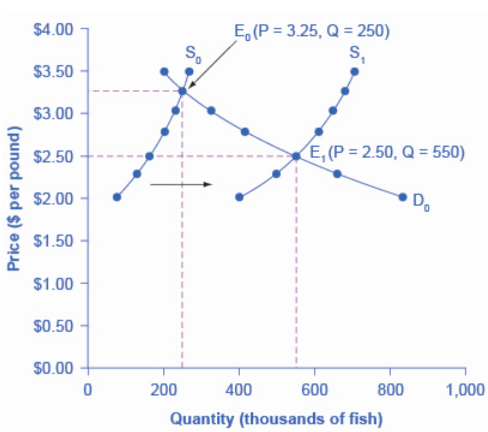


Figure $3.4.1$***:*** *Unusually good weather leads to changes in the price and quantity of salmon.*

Table 3.4.1: Salmon Fishing

Price per Pound	Quantity Supplied in 1999	Quantity Supplied in 2000	Quantity Demanded
$2.00	80	400	840
$2.25	120	480	680
$2.50	160	550	550
$2.75	200	600	450
$3.00	230	640	350
$3.25	250	670	250
$3.50	270	700	200

Step 1: Draw a demand and supply model to illustrate the market for salmon in the year before the good weather conditions began. The demand curve D_0 and the supply curve S_0 show that the original equilibrium price is $\$3.25$ per pound and the original equilibrium quantity is $250,000$ fish. (This price per pound is what commercial buyers pay at the fishing docks; what consumers pay at the grocery is higher.)

Step 2: Did the economic event affect supply or demand? Good weather is an example of a natural condition that affects supply.

Step 3: Was the effect on supply an increase or a decrease? Good weather is a change in natural conditions that increases the quantity supplied at any given price. The supply curve shifts to the right, moving from the original supply curve S_0 to the new supply curve S_1, which is shown in both the table and the figure.

Step 4: Compare the new equilibrium price and quantity to the original equilibrium. At the new equilibrium E_1, the equilibrium price falls from $\$3.25$ to $\$2.50$, but the equilibrium quantity increases from $250,000$ to $550,000$ salmon. Notice that the equilibrium quantity demanded increased, even though the demand curve did not move.

In short, good weather conditions increased supply of the California commercial salmon. The result was a higher equilibrium quantity of salmon bought and sold in the market at a lower price.

Newspapers and the Internet

According to the **Pew Research Center for People and the Press**, more and more people, especially younger people, are getting their news from online and digital sources. The majority of U.S. adults now own smartphones or tablets, and most of those Americans say they use them in part to get the news. From 2004 to 2012, the share of Americans who reported getting their news from digital sources increased from 24% to 39%. How has this affected consumption of print news media, and radio and television news? Figure $3.4.2$ and the text below illustrates using the four-step analysis to answer this question.

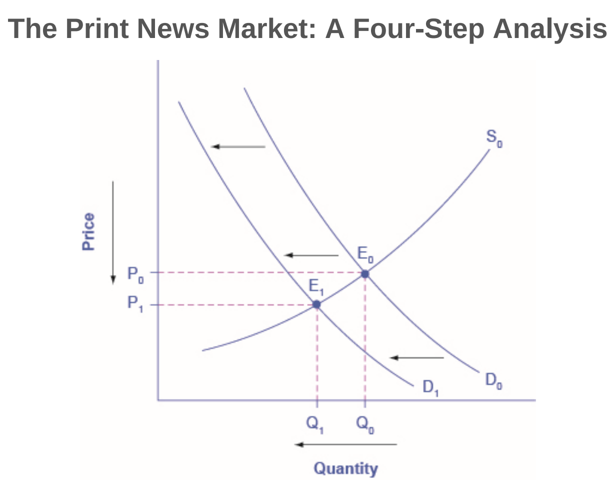


***Figure 3.4.2:** A change in tastes from print news sources to digital sources results in a leftward shift in demand for the former. The result is a decrease in both equilibrium price and quantity.*

Step 1: Develop a demand and supply model to think about what the market looked like before the event. The demand curve D_0 and the supply curve S_0 show the original relationships. In this case, the analysis is performed without specific numbers on the price and quantity axis.

Step 2: Did the change described affect supply or demand? A change in tastes, from traditional news sources (print, radio, and television) to digital sources, caused a change in demand for the former.

Step 3: Was the effect on demand positive or negative? A shift to digital news sources will tend to mean a lower quantity demanded of traditional news sources at every given price, causing the demand curve for print and other traditional news sources to shift to the left, from D_0 to D_1.

Step 4: Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_1) occurs at a lower quantity and a lower price than the original equilibrium (E_0).

The decline in print news reading predates 2004. Print newspaper circulation peaked in 1973 and has declined since then due to competition from television and radio news. In 1991, 55% of Americans indicated they got their news from print sources, while only 29% did so in 2012. Radio news has followed a similar path in recent decades, with the share of Americans getting their news from radio declining from 54% in 1991 to 33% in 2012. Television news has held its own over the last 15 years, with a market share staying in the mid to upper fifties. What does this suggest for the future, given that two-thirds of Americans under 30 years old say they do not get their news from television at all?

The Interconnections and Speed of Adjustment in Real Markets

In the real world, many factors that affect demand and supply can change all at once. For example, the demand for cars might increase because of rising incomes and population, and it might decrease because of rising gasoline prices (a complementary good). Likewise, the supply of cars might increase because of innovative new technologies that reduce the cost of car production, and it might decrease as a result of new government regulations requiring the installation of costly pollution-control technology.

Moreover, rising incomes and population or changes in gasoline prices will affect many markets, not just cars. How can an economist sort out all these interconnected events? The answer lies in the ***ceteris paribus*** assumption. Look at how each economic event affects each market, one event at a time, holding all else constant. Then combine the analyses to see the net effect.

A Combined Example

The U.S. Postal Service is facing difficult challenges. Compensation for postal workers tends to increase most years due to cost-of-living increases. At the same time, more and more people are using email, text, and other digital message forms such as Facebook and Twitter to communicate with friends and others. What does this suggest about the continued viability of the Postal Service? Figure $3.4.3$ and the text below illustrates using the four-step analysis to answer this question.

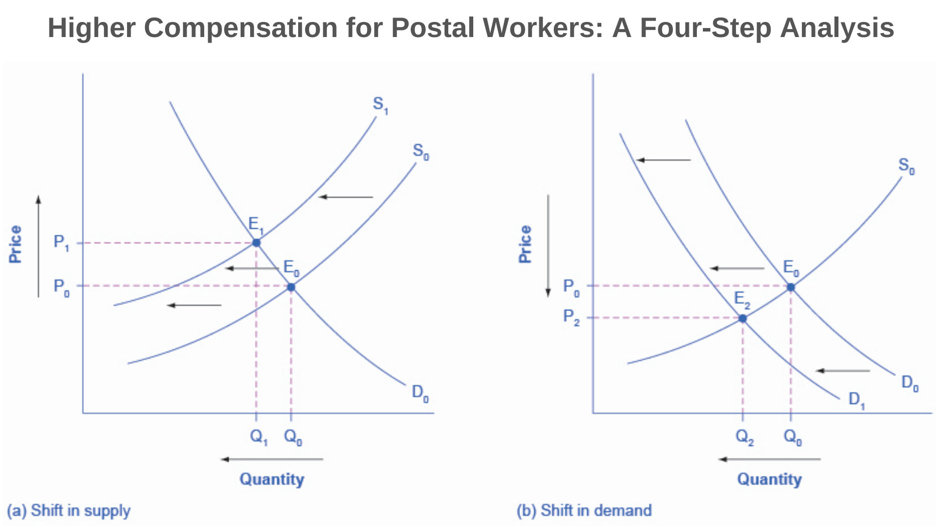


Figure $3.4.3$: *(a) Higher labor compensation causes a leftward shift in the supply curve, a decrease in the equilibrium quantity, and an increase in the equilibrium price. (b) A change in tastes away from Postal Services causes a leftward shift in the demand curve, a decrease in the equilibrium quantity, and a decrease in the equilibrium price.*

Since this problem involves two disturbances, we need two four-step analyses, the first to analyze the effects of higher compensation for postal workers, the second to analyze the effects of many people switching from "snailmail" to email and other digital messages.

Figure 3.4.3 (a) shows the shift in supply discussed in the following steps.

Step 1: Draw a demand and supply model to illustrate what the market for the U.S. Postal Service looked like before this scenario starts. The demand curve D_0 and the supply curve S_0 show the original relationships.

Step 2: Did the change described affect supply or demand? Labor compensation is a cost of production. A change in production costs caused a change in supply for the Postal Service.

Step 3: Was the effect on supply positive or negative? Higher labor compensation leads to a lower quantity supplied of postal services at every given price, causing the supply curve for postal services to shift to the left, from S_0 to S_1.

Step 4: Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_1) occurs at a lower quantity and a higher price than the original equilibrium (E_0).

Figure 3.4.3 (b) shows the shift in demand discussed in the following steps.

Step 1: Draw a demand and supply model to illustrate what the market for U.S. Postal Services looked like before this scenario starts. The demand curve D_0 and the supply curve S_0 show the original relationships. Note that this diagram is independent from the diagram in panel (a).

Step 2: Did the change described affect supply or demand? A change in tastes away from snailmail toward digital messages will cause a change in demand for the Postal Service.

Step 3: Was the effect on demand positive or negative? A change in tastes away from snailmail toward digital messages causes lower quantity demanded of postal services at every given price, causing the demand curve for postal services to shift to the left, from D_0 to D_1.

Step 4: Compare the new equilibrium price and quantity to the original equilibrium price. The new equilibrium (E_2) occurs at a lower quantity and a lower price than the original equilibrium (E_0).

The final step in a scenario where both supply and demand shift is to combine the two individual analyses to determine what happens to the equilibrium quantity and price. Graphically, we superimpose the previous two diagrams one on top of the other, as in Figure 3.4.4.

Combined Effect of Decreased Demand and Decreased Supply

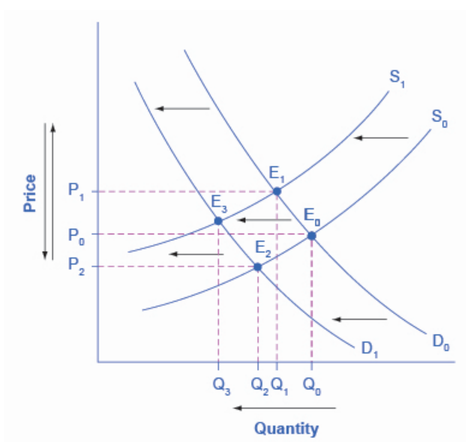


Figure 3.4.4: *Supply and demand shifts cause changes in equilibrium price and quantity.*

Following are the results:

Effect on Quantity: The effect of higher labor compensation on Postal Services because it raises the cost of production is to decrease the equilibrium quantity. The effect of a change in tastes away from snailmail is to decrease the equilibrium quantity. Since both shifts are to the left, the overall impact is a decrease in the equilibrium quantity of Postal Services (Q_3). This is easy to see graphically, since Q_3 is to the left of Q_0.

Effect on Price: The overall effect on price is more complicated. The effect of higher labor compensation on Postal Services, because it raises the cost of production, is to increase the equilibrium price. The effect of a change in tastes away from snailmail is to decrease the equilibrium price. Since the two effects are in opposite directions, unless we know the magnitudes of the two effects, the overall effect is unclear. This is not unusual. When both curves shift, typically we can determine the overall effect on price or on quantity, but not on both. In this case, we determined the overall effect on the equilibrium quantity, but not on the equilibrium price. In other cases, it might be the opposite.

The next Clear It Up feature focuses on the difference between shifts of supply or demand and movements along a curve.

Example $3.4.1$: What is the difference between shifts of demand or supply versus movements along a demand or supply curve?

One common mistake in applying the demand and supply framework is to confuse the shift of a demand or a supply curve with movement along a demand or supply curve. As an example, consider a problem that asks whether a drought will increase or decrease the equilibrium quantity and equilibrium price of wheat. Lee, a student in an introductory economics class, might reason:

> *"Well, it is clear that a drought reduces supply, so I will shift back the supply curve, as in the shift from the original supply curve S0 to S1 shown on the diagram (called Shift 1). So the equilibrium moves from E_0 to E_1, the equilibrium quantity is lower and the equilibrium price is higher. Then, a higher price makes farmers more likely to supply the good, so the supply curve shifts right, as shown by the shift from S_1 to S_2, on the diagram (shown as Shift 2), so that the equilibrium now moves from E_1 to E_2. The higher price, however, also reduces demand and so causes demand to shift back, like the shift from the original demand curve, D_0 to D_1 on the diagram (labeled Shift 3), and the equilibrium moves from E_2 to E_3."*

Shifts of Demand or Supply versus Movements along a Demand or Supply Curve

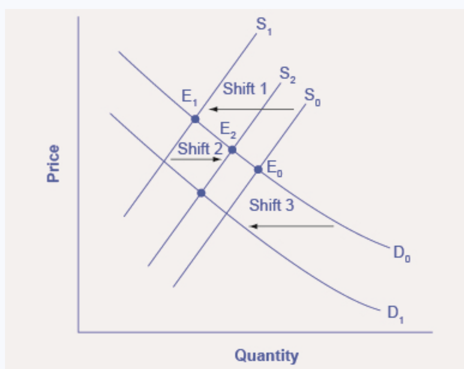


***Figure $3.4.5$:** A shift in one curve never causes a shift in the other curve. Rather, a shift in one curve causes a movement along the second curve.*

At about this point, Lee suspects that this answer is headed down the wrong path. Think about what might be wrong with Lee's logic, and then read the answer that follows.

Answer: Lee's first step is correct: that is, a drought shifts back the supply curve of wheat and leads to a prediction of a lower equilibrium quantity and a higher equilibrium price. This corresponds to a movement along the original demand curve (D_0), from E_0 to E_1. The rest of Lee's argument is wrong, because it mixes up shifts in supply with quantity supplied, and shifts in demand with quantity demanded. A higher or lower price never shifts the supply curve, as suggested by the shift in supply from S_1 to S_2. Instead, a price change leads to a movement along a given supply curve. Similarly, a higher or lower price never shifts a demand curve, as suggested in the shift from D_0 to D_1. Instead, a price change leads to a movement along a given demand curve. Remember, a change in the price of a good never causes the demand or supply curve for that good to shift.

Think carefully about the timeline of events: What happens first, what happens next? What is cause, what is effect? If you keep the order right, you are more likely to get the analysis correct.

In the four-step analysis of how economic events affect equilibrium price and quantity, the movement from the old to the new equilibrium seems immediate. As a practical matter, however, prices and quantities often do not zoom straight to equilibrium. More realistically, when an economic event causes demand or supply to shift, prices and quantities set off in the general direction of equilibrium. Indeed, even as they are moving toward one new equilibrium, prices are often then pushed by another change in demand or supply toward another equilibrium.

Key Concepts and Summary

When using the supply and demand framework to think about how an event will affect the equilibrium price and quantity, proceed through four steps:

1. sketch a supply and demand diagram to think about what the market looked like before the event
2. decide whether the event will affect supply or demand
3. decide whether the effect on supply or demand is negative or positive, and draw the appropriate shifted supply or demand curve
4. compare the new equilibrium price and quantity to the original ones.

References

Pew Research Center. "Pew Research: Center for the People & the Press." http://www.people-press.org/.

Contributor

- Template:ContribOpenSTAXMicroEcon

3.5: Price Ceilings and Price Floors

Learning Objectives

- Explain price controls, price ceilings, and price floors
- Analyze demand and supply as a social adjustment mechanism

Controversy sometimes surrounds the prices and quantities established by demand and supply, especially for products that are considered necessities. In some cases, discontent over prices turns into public pressure on politicians, who may then pass legislation to prevent a certain price from climbing "too high" or falling "too low."

The demand and supply model shows how people and firms will react to the incentives provided by these laws to control prices, in ways that will often lead to undesirable consequences. Alternative policy tools can often achieve the desired goals of price control laws, while avoiding at least some of their costs and tradeoffs.

Price Ceilings

Laws that government enacts to regulate prices are called **Price controls**. Price controls come in two flavors. A **price ceiling** keeps a price from rising above a certain level (the "ceiling"), while a **price floor** keeps a price from falling below a certain level (the "floor"). This section uses the demand and supply framework to analyze price ceilings. The next section discusses price floors.

In many markets for goods and services, demanders outnumber suppliers. Consumers, who are also potential voters, sometimes unite behind a political proposal to hold down a certain price. In some cities, such as Albany, renters have pressed political leaders to pass rent control laws, a price ceiling that usually works by stating that rents can be raised by only a certain maximum percentage each year.

Rent control becomes a politically hot topic when rents begin to rise rapidly. Everyone needs an affordable place to live. Perhaps a change in tastes makes a certain suburb or town a more popular place to live. Perhaps locally-based businesses expand, bringing higher incomes and more people into the area. Changes of this sort can cause a change in the demand for rental housing, as Figure $3.5.1$ illustrates. The original equilibrium (E_0) lies at the intersection of supply curve S_0 and demand curve D_0, corresponding to an equilibrium price of $\$500$ and an equilibrium quantity of $15,000$ units of rental housing. The effect of greater income or a change in tastes is to shift the demand curve for rental housing to the right, as shown by the data in Table $3.5.1$ and the shift from D_0 to D_1 on the graph. In this market, at the new equilibrium E_1, the price of a rental unit would rise to $\$600$ and the equilibrium quantity would increase to $17,000$ units.

A Price Ceiling Example—Rent Control

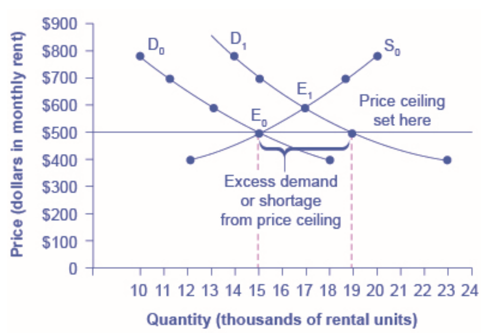


***Figure** $3.5.1$: The original intersection of demand and supply occurs at E_0. If demand shifts from D_0 to D_1, the new equilibrium would be at E_1—unless a price ceiling prevents the price from rising. If the price is not permitted to rise, the quantity supplied remains at 15,000. However, after the change in demand, the quantity demanded rises to 19,000, resulting in a shortage.*

***Table** $3.5.1$: Rent Control*

Price	Original Quantity Supplied	Original Quantity Demanded	New Quantity Demanded
$400	12,000	18,000	23,000
$500	15,000	15,000	19,000

Price	Original Quantity Supplied	Original Quantity Demanded	New Quantity Demanded
$600	17,000	13,000	17,000
$700	19,000	11,000	15,000
$800	20,000	10,000	14,000

Suppose that a rent control law is passed to keep the price at the original equilibrium of $\$500$ for a typical apartment. In Figure $3.5.1$, the horizontal line at the price of $\$500$ shows the legally fixed maximum price set by the rent control law. However, the underlying forces that shifted the demand curve to the right are still there. At that price ($\$500$), the quantity supplied remains at the same $15,000$ rental units, but the quantity demanded is $19,000$ rental units. In other words, the quantity demanded exceeds the quantity supplied, so there is a shortage of rental housing. One of the ironies of price ceilings is that while the price ceiling was intended to help renters, there are actually fewer apartments rented out under the price ceiling ($15,000$ rental units) than would be the case at the market rent of $\$600$ ($17,000$ rental units).

Price ceilings do not simply benefit renters at the expense of landlords. Rather, some renters (or potential renters) lose their housing as landlords convert apartments to co-ops and condos. Even when the housing remains in the rental market, landlords tend to spend less on maintenance and on essentials like heating, cooling, hot water, and lighting. The first rule of economics is you do not get something for nothing—everything has an opportunity cost. So if renters get "cheaper" housing than the market requires, they tend to also end up with lower quality housing.

Price ceilings have been proposed for other products. For example, price ceilings to limit what producers can charge have been proposed in recent years for prescription drugs, doctor and hospital fees, the charges made by some automatic teller bank machines, and auto insurance rates. Price ceilings are enacted in an attempt to keep prices low for those who demand the product. But when the market price is not allowed to rise to the equilibrium level, quantity demanded exceeds quantity supplied, and thus a shortage occurs. Those who manage to purchase the product at the lower price given by the price ceiling will benefit, but sellers of the product will suffer, along with those who are not able to purchase the product at all. Quality is also likely to deteriorate.

Price Floors

A price floor is the lowest legal price that can be paid in markets for goods and services, labor, or financial capital. Perhaps the best-known example of a price floor is the minimum wage, which is based on the normative view that someone working full time ought to be able to afford a basic standard of living. The federal minimum wage at the end of 2014 was $\$7.25$ per hour, which yields an income for a single person slightly higher than the poverty line. As the cost of living rises over time, the Congress periodically raises the federal minimum wage.

Price floors are sometimes called "price supports," because they support a price by preventing it from falling below a certain level. Around the world, many countries have passed laws to create agricultural price supports. Farm prices and thus farm incomes fluctuate, sometimes widely. So even if, on average, farm incomes are adequate, some years they can be quite low. The purpose of price supports is to prevent these swings.

The most common way price supports work is that the government enters the market and buys up the product, adding to demand to keep prices higher than they otherwise would be. According to the Common Agricultural Policy reform passed in 2013, the **European Union (EU)** will spend about 60 billion euros per year, or 67 billion dollars per year, or roughly 38% of the EU budget, on price supports for Europe's farmers from 2014 to 2020.

Figure $3.5.2$ illustrates the effects of a government program that assures a price above the equilibrium by focusing on the market for wheat in Europe. In the absence of government intervention, the price would adjust so that the quantity supplied would equal the quantity demanded at the equilibrium point E_0, with price P_0 and quantity Q_0. However, policies to keep prices high for farmers keeps the price above what would have been the market equilibrium level—the price P_f shown by the dashed horizontal line in the diagram. The result is a quantity supplied in excess of the quantity demanded (Q_d). When quantity supplied exceeds quantity demanded, a surplus exists.

The high-income areas of the world, including the United States, Europe, and Japan, are estimated to spend roughly $\$1$ billion per day in supporting their farmers. If the government is willing to purchase the excess supply (or to provide payments for others to purchase it), then farmers will benefit from the price floor, but taxpayers and consumers of food will pay the costs. Numerous proposals have been offered for reducing farm subsidies. In many countries, however, political support for subsidies for farmers remains strong. Either because this is viewed by the population as supporting the traditional rural way of life or because of the lobbying power of the agro-business industry.

For more detail on the effects price ceilings and floors have on demand and supply, see the following Clear It Up feature.

European Wheat Prices: A Price Floor Example

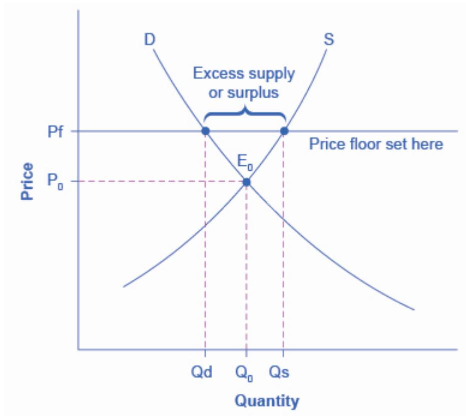


Figure* 3.5.2*: *The intersection of demand* (D) *and supply* (S) *would be at the equilibrium point* E_0*. However, a price floor set at* P_f *holds the price above* E_0 *and prevents it from falling. The result of the price floor is that the quantity supplied* Q_s *exceeds the quantity demanded* Q_d*. There is excess supply, also called a surplus.*

Do price ceilings and floors change demand or supply?

Neither price ceilings nor price floors cause demand or supply to change. They simply set a price that limits what can be legally charged in the market. Remember, changes in price do not cause demand or supply to change. Price ceilings and price floors can cause a different choice of quantity demanded along a demand curve, but they do not move the demand curve. Price controls can cause a different choice of quantity supplied along a supply curve, but they do not shift the supply curve.

Key Concepts and Summary

Price ceilings prevent a price from rising above a certain level. When a price ceiling is set below the equilibrium price, quantity demanded will exceed quantity supplied, and excess demand or shortages will result. Price floors prevent a price from falling below a certain level. When a price floor is set above the equilibrium price, quantity supplied will exceed quantity demanded, and excess supply or surpluses will result. Price floors and price ceilings often lead to unintended consequences.

Glossary

price ceiling

a legal maximum price

price control

government laws to regulate prices instead of letting market forces determine prices

price floor

a legal minimum price

total surplus

see social surplus

Contributor

- Template:ContribOpenSTAXMicroEcon

3.6: Demand, Supply, and Efficiency

Learning Objectives

- Contrast consumer surplus, producer surplus, and social surplus
- Explain why price floors and price ceilings can be inefficient
- Analyze demand and supply as a social adjustment mechanism

The familiar **demand and supply diagram** holds within it the concept of economic efficiency. One typical way that economists define **efficiency** is when it is impossible to improve the situation of one party without imposing a cost on another. Conversely, if a situation is inefficient, it becomes possible to benefit at least one party without imposing costs on others.

Efficiency in the demand and supply model has the same basic meaning: The economy is getting as much benefit as possible from its scarce resources and all the possible gains from trade have been achieved. In other words, the optimal amount of each good and service is being produced and consumed.

Consumer Surplus, Producer Surplus, Social Surplus

Consider a market for tablet computers, as shown in Figure $3.6.1$. The equilibrium price is $\$80$ and the equilibrium quantity is 28 million. To see the benefits to consumers, look at the segment of the demand curve above the **equilibrium** point and to the left. This portion of the demand curve shows that at least some demanders would have been willing to pay more than $\$80$ for a tablet.

For example, point J shows that if the price was $\$90$, 20 million tablets would be sold. Those consumers who would have been willing to pay $\$90$ for a tablet based on the utility they expect to receive from it, but who were able to pay the equilibrium price of $\$80$, clearly received a benefit beyond what they had to pay for. Remember, the demand curve traces consumers' willingness to pay for different quantities. The amount that individuals would have been willing to pay, minus the amount that they actually paid, is called **consumer surplus**. Consumer surplus is the area labeled F—that is, the area above the market price and below the demand curve.

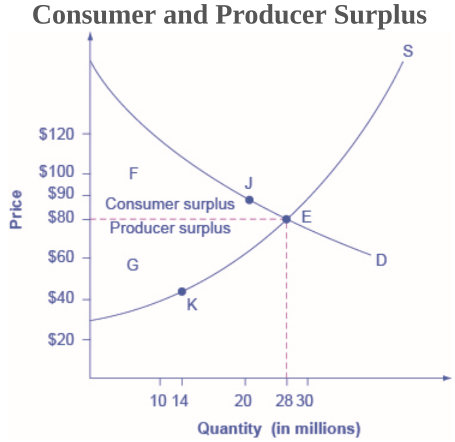


Figure $3.6.1$: The somewhat triangular area labeled by F shows the area of consumer surplus, which shows that the equilibrium price in the market was less than what many of the consumers were willing to pay. Point J on the demand curve shows that, even at the price of \$90, consumers would have been willing to purchase a quantity of 20 million. The somewhat triangular area labeled by G shows the area of producer surplus, which shows that the equilibrium price received in the market was more than what many of the producers were willing to accept for their products. For example, point K on the supply curve shows that at a price of \$45, firms would have been willing to supply a quantity of 14 million.

The supply curve shows the quantity that firms are willing to supply at each price. For example, point K in Figure $3.6.1$ illustrates that, at $\$45$, firms would still have been willing to supply a quantity of 14 million. Those producers who would have been willing to supply the tablets at $\$45$, but who were instead able to charge the equilibrium price of $\$80$, clearly received an extra benefit beyond what they required to supply the product. The amount that a seller is paid for a good minus the seller's actual cost is called producer surplus. In Figure $3.6.1$, producer surplus is the area labeled G—that is, the area between the market price and the segment of the supply curve below the equilibrium.

The sum of consumer surplus and producer surplus is social surplus, also referred to as economic surplus or total surplus. In Figure $3.6.1$, social surplus would be shown as the area $F+G$. Social surplus is larger at equilibrium quantity and price than it would be at any other quantity. This demonstrates the economic efficiency of the market equilibrium. In addition, at the efficient level of output, it is impossible to produce greater consumer surplus without reducing producer surplus, and it is impossible to produce greater producer surplus without reducing consumer surplus.

Inefficiency of Price Floors and Price Ceilings

The imposition of a price floor or a price ceiling will prevent a market from adjusting to its equilibrium price and quantity, and thus will create an inefficient outcome. But there is an additional twist here. Along with creating inefficiency, price floors and ceilings will also transfer some consumer surplus to producers, or some producer surplus to consumers.

Imagine that several firms develop a promising but expensive new drug for treating back pain. If this therapy is left to the market, the equilibrium price will be $\$600$ per month and $20,000$ people will use the drug, as shown in Figure $3.6.2$ (a). The original level of consumer surplus is $T+U$ and producer surplus is $V+W+X$. However, the government decides to impose a price ceiling of $\$400$ to make the drug more affordable. At this price ceiling, firms in the market now produce only $15,000$.

As a result, two changes occur. First, an inefficient outcome occurs and the total surplus of society is reduced. The loss in social surplus that occurs when the economy produces at an inefficient quantity is called deadweight loss. In a very real sense, it is like money thrown away that benefits no one. In Figure $3.6.2$ (a), the deadweight loss is the area $U+W$. When deadweight loss exists, it is possible for both consumer and producer surplus to be higher, in this case because the price control is blocking some suppliers and demanders from transactions they would both be willing to make.

A second change from the price ceiling is that some of the producer surplus is transferred to consumers. After the price ceiling is imposed, the new consumer surplus is $T+V$, while the new producer surplus is X. In other words, the price ceiling transfers the area of surplus (V) from producers to consumers. Note that the gain to consumers is less than the loss to producers, which is just another way of seeing the deadweight loss.

Efficiency and Price Floors and Ceilings

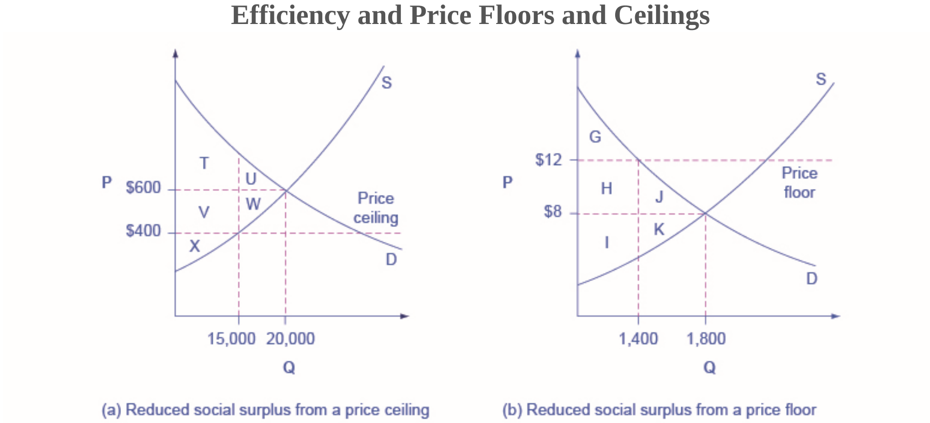


Figure $3.6.2$: (a) The original equilibrium price is \$600 with a quantity of 20,000. Consumer surplus is $T+U$, and producer surplus is $V+W+X$. A price ceiling is imposed at \$400, so firms in the market now produce only a quantity of 15,000. As a result, the new consumer surplus is $T+V$, while the new producer surplus is X. (b) The original equilibrium is \$8 at a quantity of 1,800. Consumer surplus is $G+H+J$, and producer surplus is $I+K$. A price floor is imposed at \$12, which means that quantity demanded falls to 1,400. As a result, the new consumer surplus is G, and the new producer surplus is $H+I$.

Figure $3.6.2$ (b) shows a price floor example using a string of struggling movie theaters, all in the same city. The current equilibrium is $\$8$ per movie ticket, with $1,800$ people attending movies. The original consumer surplus is $G+H+J$, and producer surplus is $I+K$. The city government is worried that movie theaters will go out of business, reducing the entertainment options available to citizens, so it decides to impose a price floor of $\$12$ per ticket. As a result, the quantity demanded of movie tickets falls to $1,400$. The new consumer surplus is G, and the new producer surplus is $H+I$. In effect, the **price floor** causes the area H to be transferred from consumer to producer surplus, but also causes a deadweight loss of $J+K$.

This analysis shows that a price ceiling, like a law establishing rent controls, will transfer some producer surplus to consumers—which helps to explain why consumers often favor them. Conversely, a price floor like a guarantee that farmers will receive a certain price for their crops will transfer some consumer surplus to producers, which explains why producers often favor them. However, both price floors and price ceilings block some transactions that buyers and sellers would have been willing to make, and creates deadweight loss. Removing such barriers, so that prices and quantities can adjust to their equilibrium level, will increase the economy's social surplus.

Demand and Supply as a Social Adjustment Mechanism

The demand and supply model emphasizes that prices are not set only by demand or only by supply, but by the interaction between the two. In 1890, the famous economist **Alfred Marshall** wrote that asking whether supply or demand determined a price was like arguing "whether it is the upper or the under blade of a pair of scissors that cuts a piece of paper." The answer is that both blades of the demand and supply scissors are always involved.

The adjustments of equilibrium price and quantity in a market-oriented economy often occur without much government direction or oversight. If the coffee crop in Brazil suffers a terrible frost, then the supply curve of coffee shifts to the left and the price of coffee rises. Some people—call them the coffee addicts—continue to drink coffee and pay the higher price. Others switch to tea or soft drinks. No government commission is needed to figure out how to adjust coffee prices, which companies will be allowed to process the remaining supply, which supermarkets in which cities will get how much coffee to sell, or which consumers will ultimately be allowed to drink the brew. Such adjustments in response to price changes happen all the time in a market economy, often so smoothly and rapidly that we barely notice them.

Think for a moment of all the seasonal foods that are available and inexpensive at certain times of the year, like fresh corn in midsummer, but more expensive at other times of the year. People alter their diets and restaurants alter their menus in response to these fluctuations in prices without fuss or fanfare. For both the U.S. economy and the world economy as a whole, markets—that is, demand and supply—are the primary social mechanism for answering the basic questions about what is produced, how it is produced, and for whom it is produced.

Why Can We Not Get Enough of Organic?

Organic food is grown without synthetic pesticides, chemical fertilizers or genetically modified seeds. In recent decades, the demand for organic products has increased dramatically. The Organic Trade Association reported sales increased from $\$1$ billion in 1990 to $\$35.1$ billion in 2013, more than 90% of which were sales of food products.

Why, then, are organic foods more expensive than their conventional counterparts? The answer is a clear application of the theories of supply and demand. As people have learned more about the harmful effects of chemical fertilizers, growth hormones, pesticides and the like from large-scale factory farming, our tastes and preferences for safer, organic foods have increased. This change in tastes has been reinforced by increases in income, which allow people to purchase pricier products, and has made organic foods more mainstream. This has led to an increased demand for organic foods. Graphically, the demand curve has shifted right, and we have moved up the supply curve as producers have responded to the higher prices by supplying a greater quantity.

In addition to the movement along the supply curve, we have also had an increase in the number of farmers converting to organic farming over time. This is represented by a shift to the right of the supply curve. Since both demand and supply have shifted to the right, the resulting equilibrium quantity of organic foods is definitely higher, but the price will only fall when the increase in supply is larger than the increase in demand. We may need more time before we see lower prices in organic foods. Since the production costs of these foods may remain higher than conventional farming, because organic fertilizers and pest management techniques are more expensive, they may never fully catch up with the lower prices of non-organic foods.

As a final, specific example: The Environmental Working Group's "Dirty Dozen" list of fruits and vegetables, which test high for pesticide residue even after washing, was released in April 2013. The inclusion of strawberries on the list has led to an increase in demand for organic strawberries, resulting in both a higher equilibrium price and quantity of sales.

Key Concepts and Summary

Consumer surplus is the gap between the price that consumers are willing to pay, based on their preferences, and the market equilibrium price. Producer surplus is the gap between the price for which producers are willing to sell a product, based on their costs, and the market equilibrium price. Social surplus is the sum of consumer surplus and producer surplus. Total surplus is larger at the equilibrium quantity and price than it will be at any other quantity and price. Deadweight loss is loss in total surplus that occurs when the economy produces at an inefficient quantity.

Glossary

consumer surplus
: the extra benefit consumers receive from buying a good or service, measured by what the individuals would have been willing to pay minus the amount that they actually paid

deadweight loss
: the loss in social surplus that occurs when a market produces an inefficient quantity

economic surplus
: see social surplus

producer surplus
: the extra benefit producers receive from selling a good or service, measured by the price the producer actually received minus the price the producer would have been willing to accept

social surplus
: the sum of consumer surplus and producer surplus

Contributor

- Template:ContribOpenSTAXMicroEcon

SECTION OVERVIEW

3.7: Labor and Financial Markets

Topic hierarchy

3.8: Demand and Supply at Work in Labor Markets

Learning Objectives

- Predict shifts in the demand and supply curves of the labor market
- Explain the impact of new technology on the demand and supply curves of the labor market
- Explain price floors in the labor market such as minimum wage or a living wage

Markets for labor have demand and supply curves, just like markets for goods. The law of demand applies in labor markets this way: A higher salary or wage—that is, a higher price in the labor market—leads to a decrease in the quantity of labor demanded by employers, while a lower salary or wage leads to an increase in the quantity of labor demanded. The law of supply functions in labor markets, too: A higher price for labor leads to a higher quantity of labor supplied; a lower price leads to a lower quantity supplied.

Equilibrium in the Labor Market

In 2013, about $34,000$ registered nurses worked in the Minneapolis-St. Paul-Bloomington, Minnesota-Wisconsin metropolitan area, according to the BLS. They worked for a variety of employers: hospitals, doctors' offices, schools, health clinics, and nursing homes. Figure $3.8.1$ illustrates how demand and supply determine equilibrium in this labor market. The demand and supply schedules in Table $3.8.1$ list the quantity supplied and quantity demanded of nurses at different salaries.

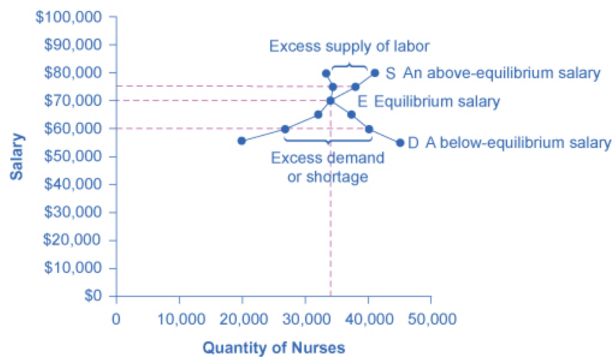


Figure $3.8.1$: The demand curve (D) of those employers who want to hire nurses intersects with the supply curve (S) of those who are qualified and willing to work as nurses at the equilibrium point (E). The equilibrium salary is $70,000 and the equilibrium quantity is 34,000 nurses. At an above-equilibrium salary of $75,000, quantity supplied increases to 38,000, but the quantity of nurses demanded at the higher pay declines to 33,000. At this above-equilibrium salary, an excess supply or surplus of nurses would exist. At a below-equilibrium salary of $60,000, quantity supplied declines to 27,000, while the quantity demanded at the lower wage increases to 40,000 nurses. At this below-equilibrium salary, excess demand or a shortage exists.

Table $3.8.1$: Demand and Supply of Nurses in Minneapolis-St. Paul-Bloomington

Annual Salary	Quantity Demanded	Quantity Supplied
$55,000	45,000	20,000
$60,000	40,000	27,000
$65,000	37,000	31,000
$70,000	34,000	34,000
$75,000	33,000	38,000
$80,000	32,000	41,000

The horizontal axis shows the quantity of nurses hired. In this example, labor is measured by number of workers, but another common way to measure the quantity of labor is by the number of hours worked. The vertical axis shows the price for nurses' labor —that is, how much they are paid. In the real world, this "price" would be total labor compensation: salary plus benefits. It is not obvious, but benefits are a significant part (as high as 30%) of labor compensation. In this example, the price of labor is measured by salary on an annual basis, although in other cases the price of labor could be measured by monthly or weekly pay, or even the wage paid per hour. As the salary for nurses rises, the quantity demanded will fall. Some hospitals and nursing homes may cut back on the number of nurses they hire, or they may lay off some of their existing nurses, rather than pay them higher salaries. Employers who face higher nurses' salaries may also try to replace some nursing functions by investing in physical equipment, like computer monitoring and diagnostic systems to monitor patients, or by using lower-paid health care aides to reduce the number of nurses they need.

As the salary for nurses rises, the quantity supplied will rise. If nurses' salaries in Minneapolis-St. Paul-Bloomington are higher than in other cities, more nurses will move to Minneapolis-St. Paul-Bloomington to find jobs, more people will be willing to train as nurses, and those currently trained as nurses will be more likely to pursue nursing as a full-time job. In other words, there will be more nurses looking for jobs in the area.

At **equilibrium**, the quantity supplied and the quantity demanded are equal. Thus, every employer who wants to hire a nurse at this equilibrium wage can find a willing worker, and every nurse who wants to work at this equilibrium salary can find a job. In Figure $3.8.1$, the supply curve (S) and demand curve (D) intersect at the equilibrium point (E). The equilibrium quantity of nurses in the Minneapolis-St. Paul-Bloomington area is $34,000$, and the equilibrium salary is $\$70,000$ per year. This example simplifies the nursing market by focusing on the "average" nurse. In reality, of course, the market for nurses is actually made up of many smaller markets, like markets for nurses with varying degrees of experience and credentials. Many markets contain closely related products that differ in quality; for instance, even a simple product like gasoline comes in regular, premium, and super-premium, each with a different price. Even in such cases, discussing the average price of gasoline, like the average salary for nurses, can still be useful because it reflects what is happening in most of the submarkets.

When the price of labor is not at the equilibrium, economic incentives tend to move salaries toward the equilibrium. For example, if salaries for nurses in Minneapolis-St. Paul-Bloomington were above the equilibrium at $\$75,000$ per year, then $38,000$ people want to work as nurses, but employers want to hire only $33,000$ nurses. At that above-equilibrium salary, excess supply or a surplus results. In a situation of excess supply in the **labor market**, with many applicants for every job opening, employers will have an incentive to offer lower wages than they otherwise would have. Nurses' salary will move down toward equilibrium.

In contrast, if the salary is below the equilibrium at, say, $\$60,000$ per year, then a situation of excess demand or a shortage arises. In this case, employers encouraged by the relatively lower wage want to hire $40,000$ nurses, but only $27,000$ individuals want to work as nurses at that salary in Minneapolis-St. Paul-Bloomington. In response to the shortage, some employers will offer higher pay to attract the nurses. Other employers will have to match the higher pay to keep their own employees. The higher salaries will encourage more nurses to train or work in Minneapolis-St. Paul-Bloomington. Again, price and quantity in the labor market will move toward equilibrium.

Shifts in Labor Demand

The demand curve for labor shows the quantity of labor employers wish to hire at any given salary or wage rate, under the ***ceteris paribus*** assumption. A change in the wage or salary will result in a change in the quantity demanded of labor. If the wage rate increases, employers will want to hire fewer employees. The quantity of labor demanded will decrease, and there will be a movement upward along the demand curve. If the wages and salaries decrease, employers are more likely to hire a greater number of workers. The quantity of labor demanded will increase, resulting in a downward movement along the demand curve.

Shifts in the demand curve for labor occur for many reasons. One key reason is that the demand for labor is based on the demand for the good or service that is being produced. For example, the more new automobiles consumers demand, the greater the number of workers automakers will need to hire. Therefore the demand for labor is called a "derived demand." Here are some examples of derived demand for labor:

- The demand for chefs is dependent on the demand for restaurant meals.
- The demand for pharmacists is dependent on the demand for prescription drugs.
- The demand for attorneys is dependent on the demand for legal services.

As the demand for the goods and services increases, the demand for labor will increase, or shift to the right, to meet employers' production requirements. As the demand for the goods and services decreases, the demand for labor will decrease, or shift to the left. Table $3.8.2$ shows that in addition to the derived demand for labor, demand can also increase or decrease (shift) in response to several factors.

Table $3.8.2$: Factors That Can Shift Demand

Factors	Results

Factors	Results
Demand for Output	When the demand for the good produced (output) increases, both the output price and profitability increase. As a result, producers demand more labor to ramp up production.
Education and Training	A well-trained and educated workforce causes an increase in the demand for that labor by employers. Increased levels of productivity within the workforce will cause the demand for labor to shift to the right. If the workforce is not well-trained or educated, employers will not hire from within that labor pool, since they will need to spend a significant amount of time and money training that workforce. Demand for such will shift to the left.
Technology	Technology changes can act as either substitutes for or complements to labor. When technology acts as a substitute, it replaces the need for the number of workers an employer needs to hire. For example, word processing decreased the number of typists needed in the workplace. This shifted the demand curve for typists left. An increase in the availability of certain technologies may increase the demand for labor. Technology that acts as a complement to labor will increase the demand for certain types of labor, resulting in a rightward shift of the demand curve. For example, the increased use of word processing and other software has increased the demand for information technology professionals who can resolve software and hardware issues related to a firm's network. More and better technology will increase demand for skilled workers who know how to use technology to enhance workplace productivity. Those workers who do not adapt to changes in technology will experience a decrease in demand.
Number of Companies	An increase in the number of companies producing a given product will increase the demand for labor resulting in a shift to the right. A decrease in the number of companies producing a given product will decrease the demand for labor resulting in a shift to the left.
Government Regulations	Complying with government regulations can increase or decrease the demand for labor at any given wage. In the healthcare industry, government rules may require that nurses be hired to carry out certain medical procedures. This will increase the demand for nurses. Less-trained healthcare workers would be prohibited from carrying out these procedures, and the demand for these workers will shift to the left.

Factors	Results
Price and Availability of Other Inputs	Labor is not the only input into the production process. For example, a salesperson at a call center needs a telephone and a computer terminal to enter data and record sales. The demand for salespersons at the call center will increase if the number of telephones and computer terminals available increases. This will cause a rightward shift of the demand curve. As the amount of inputs increases, the demand for labor will increase. If the terminal or the telephones malfunction, then the demand for that labor force will decrease. As the quantity of other inputs decreases, the demand for labor will decrease. Similarly, if prices of other inputs fall, production will become more profitable and suppliers will demand more labor to increase production. The opposite is also true. Higher input prices lower demand for labor

Shifts in Labor Supply

The supply of labor is upward-sloping and adheres to the law of supply: The higher the price, the greater the quantity supplied and the lower the price, the less quantity supplied. The supply curve models the tradeoff between supplying labor into the market or using time in leisure activities at every given price level. The higher the wage, the more labor is willing to work and forego leisure activities. Table $3.8.3$ lists some of the factors that will cause the supply to increase or decrease.

Table $3.8.3$: Factors that Can Shift Supply

Factors	Results
Number of Workers	An increased number of workers will cause the supply curve to shift to the right. An increased number of workers can be due to several factors, such as immigration, increasing population, an aging population, and changing demographics. Policies that encourage immigration will increase the supply of labor, and vice versa. Population grows when birth rates exceed death rates; this eventually increases supply of labor when the former reach working age. An aging and therefore retiring population will decrease the supply of labor. Another example of changing demographics is more women working outside of the home, which increases the supply of labor.
Required Education	The more required education, the lower the supply. There is a lower supply of PhD mathematicians than of high school mathematics teachers; there is a lower supply of cardiologists than of primary care physicians; and there is a lower supply of physicians than of nurses.

Factors	Results
Government Policies	Government policies can also affect the supply of labor for jobs. On the one hand, the government may support rules that set high qualifications for certain jobs: academic training, certificates or licenses, or experience. When these qualifications are made tougher, the number of qualified workers will decrease at any given wage. On the other hand, the government may also subsidize training or even reduce the required level of qualifications. For example, government might offer subsidies for nursing schools or nursing students. Such provisions would shift the supply curve of nurses to the right. In addition, government policies that change the relative desirability of working versus not working also affect the labor supply. These include unemployment benefits, maternity leave, child care benefits and welfare policy. For example, child care benefits may increase the labor supply of working mothers. Long term unemployment benefits may discourage job searching for unemployed workers. All these policies must therefore be carefully designed to minimize any negative labor supply effects.

A change in salary will lead to a movement along labor demand or labor supply curves, but it will not shift those curves. However, other events like those outlined here will cause either the demand or the supply of labor to shift, and thus will move the labor market to a new equilibrium salary and quantity.

Technology and Wage Inequality: The Four-Step Process

Economic events can change the equilibrium salary (or wage) and quantity of labor. Consider how the wave of new information technologies, like computer and telecommunications networks, has affected low-skill and high-skill workers in the U.S. economy. From the perspective of employers who demand labor, these new technologies are often a substitute for low-skill laborers like file clerks who used to keep file cabinets full of paper records of transactions. However, the same new technologies are a complement to high-skill workers like managers, who benefit from the technological advances by being able to monitor more information, communicate more easily, and juggle a wider array of responsibilities. So, how will the new technologies affect the wages of high-skill and low-skill workers? For this question, the four-step process of analyzing how shifts in supply or demand affect a market (introduced in Demand and Supply) works in this way:

Step 1: What did the markets for low-skill labor and high-skill labor look like before the arrival of the new technologies? In Figure 3.8.2 (a) and Figure 3.8.2 (b), S_0 is the original supply curve for labor and D_0 is the original demand curve for labor in each market. In each graph, the original point of equilibrium, E_0, occurs at the price W_0 and the quantity Q_0.

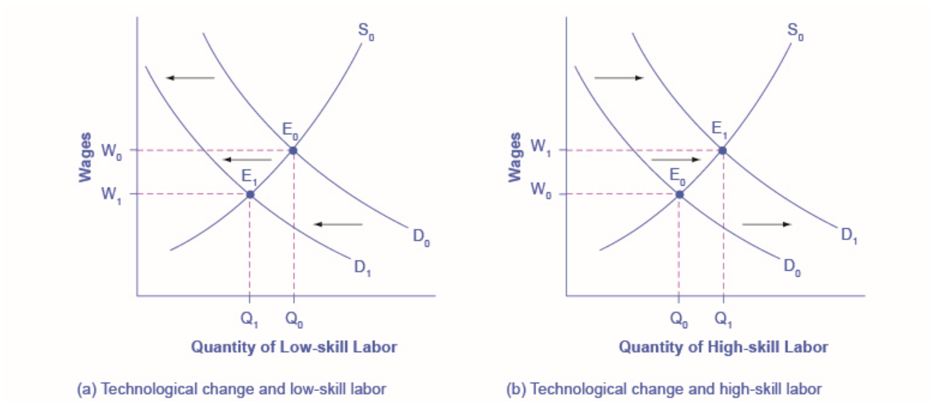


Figure 3.8.2: (a) The demand for low-skill labor shifts to the left when technology can do the job previously done by these workers. (b) New technologies can also increase the demand for high-skill labor in fields such as information technology and network administration.

Step 2: Does the new technology affect the supply of labor from households or the demand for labor from firms? The technology change described here affects demand for labor by firms that hire workers.

Step 3: Will the new technology increase or decrease demand? Based on the description earlier, as the substitute for low-skill labor becomes available, demand for low-skill labor will shift to the left, from D_0 to D_1. As the technology complement for high-skill labor becomes cheaper, demand for high-skill labor will shift to the right, from D_0 to D_1.

Step 4: The new equilibrium for low-skill labor, shown as point E_1 with price W_1 and quantity Q_1, has a lower wage and quantity hired than the original equilibrium, E_0. The new equilibrium for high-skill labor, shown as point E_1 with price W_1 and quantity Q_1, has a higher wage and quantity hired than the original equilibrium (E_0).

So, the demand and supply model predicts that the new computer and communications technologies will raise the pay of high-skill workers but reduce the pay of low-skill workers. Indeed, from the 1970s to the mid-2000s, the wage gap widened between high-skill and low-skill labor. According to the National Center for Education Statistics, in 1980, for example, a college graduate earned about 30% more than a high school graduate with comparable job experience, but by 2012, a college graduate earned about 60% more than an otherwise comparable high school graduate. Many economists believe that the trend toward greater wage inequality across the U.S. economy was primarily caused by the new technologies.

Price Floors in the Labor Market: Living Wages and Minimum Wages

In contrast to goods and services markets, price ceilings are rare in labor markets, because rules that prevent people from earning income are not politically popular. There is one exception: sometimes limits are proposed on the high incomes of top business executives.

The labor market, however, presents some prominent examples of price floors, which are often used as an attempt to increase the wages of low-paid workers. The U.S. government sets a minimum wage, a price floor that makes it illegal for an employer to pay employees less than a certain hourly rate. In mid-2009, the U.S. **minimum wage** was raised to $\$7.25$ per hour. Local political movements in a number of U.S. cities have pushed for a higher minimum wage, which they call a **living wage**. Promoters of living wage laws maintain that the minimum wage is too low to ensure a reasonable standard of living. They base this conclusion on the calculation that, if you work 40 hours a week at a minimum wage of $\$7.25$ per hour for 50 weeks a year, your annual income is $\$14,500$, which is less than the official U.S. government definition of what it means for a family to be in poverty. (A family with two adults earning minimum wage and two young children will find it more cost efficient for one parent to provide childcare while the other works for income. So the family income would be $\$14,500$, which is significantly lower than the federal poverty line for a family of four, which was $\$23,850$ in 2014.)

Supporters of the living wage argue that full-time workers should be assured a high enough wage so that they can afford the essentials of life: food, clothing, shelter, and healthcare. Since Baltimore passed the first living wage law in 1994, several dozen cities enacted similar laws in the late 1990s and the 2000s. The living wage ordinances do not apply to all employers, but they have specified that all employees of the city or employees of firms that are hired by the city be paid at least a certain wage that is usually a few dollars per hour above the U.S. minimum wage.

Figure $3.8.3$ illustrates the situation of a city considering a living wage law. For simplicity, we assume that there is no federal minimum wage. The wage appears on the vertical axis, because the wage is the price in the labor market. Before the passage of the living wage law, the equilibrium wage is $\$10$ per hour and the city hires $1,200$ workers at this wage. However, a group of concerned citizens persuades the city council to enact a living wage law requiring employers to pay no less than $\$12$ per hour. In response to the higher wage, $1,600$ workers look for jobs with the city. At this higher wage, the city, as an employer, is willing to hire only 700 workers. At the price floor, the quantity supplied exceeds the quantity demanded, and a surplus of labor exists in this market. For workers who continue to have a job at a higher salary, life has improved. For those who were willing to work at the old wage rate but lost their jobs with the wage increase, life has not improved. Table $3.8.4$ shows the differences in supply and demand at different wages.

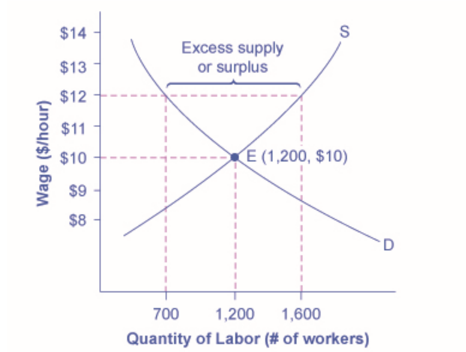


Figure 3.8.3: The original equilibrium in this labor market is a wage of $10/hour and a quantity of 1,200 workers, shown at point E. Imposing a wage floor at $12/hour leads to an excess supply of labor. At that wage, the quantity of labor supplied is 1,600 and the quantity of labor demanded is only 700.

Table 3.8.4: Living Wage - Example of a Price Floor

Wage	Quantity Labor Demanded	Quantity Labor Supplied
$8/hr	1,900	500
$9/hr	1,500	900
$10/hr	1,200	1,200
$11/hr	900	1,400
$12/hr	700	1,600
$13/hr	500	1,800
$14/hr	400	1,900

The Minimum Wage as an Example of a Price Floor

The U.S. minimum wage is a price floor that is set either very close to the equilibrium wage or even slightly below it. About 1% of American workers are actually paid the minimum wage. In other words, the vast majority of the U.S. labor force has its wages determined in the labor market, not as a result of the government price floor. But for workers with low skills and little experience, like those without a high school diploma or teenagers, the minimum wage is quite important. In many cities, the federal minimum wage is apparently below the market price for unskilled labor, because employers offer more than the minimum wage to checkout clerks and other low-skill workers without any government prodding.

Economists have attempted to estimate how much the minimum wage reduces the quantity demanded of low-skill labor. A typical result of such studies is that a 10% increase in the minimum wage would decrease the hiring of unskilled workers by 1% to 2%, which seems a relatively small reduction. In fact, some studies have even found no effect of a higher minimum wage on employment at certain times and places—although these studies are controversial.

Let's suppose that the minimum wage lies just slightly below the equilibrium wage level. Wages could fluctuate according to market forces above this price floor, but they would not be allowed to move beneath the floor. In this situation, the price floor minimum wage is said to be nonbinding —that is, the price floor is not determining the market outcome. Even if the minimum wage moves just a little higher, it will still have no effect on the quantity of employment in the economy, as long as it remains below the equilibrium wage. Even if the minimum wage is increased by enough so that it rises slightly above the equilibrium wage and becomes binding, there will be only a small excess supply gap between the quantity demanded and quantity supplied.

These insights help to explain why U.S. minimum wage laws have historically had only a small impact on employment. Since the minimum wage has typically been set close to the equilibrium wage for low-skill labor and sometimes even below it, it has not had a large effect in creating an excess supply of labor. However, if the minimum wage were increased dramatically—say, if it were doubled to match the living wages that some U.S. cities have considered—then its impact on reducing the quantity demanded of employment would be far greater. The following Clear It Up feature describes in greater detail some of the arguments for and against changes to minimum wage.

Example 3.8.1: What's the harm in raising the minimum wage?

Because of the law of demand, a higher required wage will reduce the amount of low-skill employment either in terms of employees or in terms of work hours. Although there is controversy over the numbers, let's say for the sake of the argument that a 10% rise in the minimum wage will reduce the employment of low-skill workers by 2%. Does this outcome mean that raising the minimum wage by 10% is bad public policy? Not necessarily.

If 98% of those receiving the minimum wage have a pay increase of 10%, but 2% of those receiving the minimum wage lose their jobs, are the gains for society as a whole greater than the losses? The answer is not clear, because job losses, even for a small group, may cause more pain than modest income gains for others. For one thing, we need to consider which minimum wage workers are losing their jobs. If the 2% of minimum wage workers who lose their jobs are struggling to support families, that is one thing. If those who lose their job are high school students picking up spending money over summer vacation, that is something else.

Another complexity is that many minimum wage workers do not work full-time for an entire year. Imagine a minimum wage worker who holds different part-time jobs for a few months at a time, with bouts of unemployment in between. The worker in this situation receives the 10% raise in the minimum wage when working, but also ends up working 2% fewer hours during the year because the higher minimum wage reduces how much employers want people to work. Overall, this worker's income would rise because the 10% pay raise would more than offset the 2% fewer hours worked.

Of course, these arguments do not prove that raising the minimum wage is necessarily a good idea either. There may well be other, better public policy options for helping low-wage workers. (The Poverty and Economic Inequality chapter discusses some possibilities.) The lesson from this maze of minimum wage arguments is that complex social problems rarely have simple answers. Even those who agree on how a proposed economic policy affects quantity demanded and quantity supplied may still disagree on whether the policy is a good idea.

Key Concepts and Summary

In the labor market, households are on the supply side of the market and firms are on the demand side. In the market for financial capital, households and firms can be on either side of the market: they are suppliers of financial capital when they save or make financial investments, and demanders of financial capital when they borrow or receive financial investments.

In the demand and supply analysis of labor markets, the price can be measured by the annual salary or hourly wage received. The quantity of labor can be measured in various ways, like number of workers or the number of hours worked.

Factors that can shift the demand curve for labor include: a change in the quantity demanded of the product that the labor produces; a change in the production process that uses more or less labor; and a change in government policy that affects the quantity of labor that firms wish to hire at a given wage. Demand can also increase or decrease (shift) in response to: workers' level of education and training, technology, the number of companies, and availability and price of other inputs.

The main factors that can shift the supply curve for labor are: how desirable a job appears to workers relative to the alternatives, government policy that either restricts or encourages the quantity of workers trained for the job, the number of workers in the economy, and required education.

References

American Community Survey. 2012. "School Enrollment and Work Status: 2011." Accessed April 13, 2015. www.census.gov/prod/2013pubs/acsbr11-14.pdf.

National Center for Educational Statistics. "Digest of Education Statistics." (2008 and 2010). Accessed December 11, 2013. nces.ed.gov.

Glossary

minimum wage

a price floor that makes it illegal for an employer to pay employees less than a certain hourly rate

Contributor

- Template:ContribOpenSTAXMicroEcon

3.9: Demand and Supply in Financial Markets

Learning Objectives

- Identify the demanders and suppliers in a financial market.
- Explain how interest rates can affect supply and demand
- Analyze the economic effects of U.S. debt in terms of domestic financial markets
- Explain the role of price ceilings and usury laws in the U.S.

United States' households, institutions, and domestic businesses saved almost $\$1.9$ trillion in 2013. Where did that savings go and what was it used for? Some of the savings ended up in banks, which in turn loaned the money to individuals or businesses that wanted to borrow money. Some was invested in private companies or loaned to government agencies that wanted to borrow money to raise funds for purposes like building roads or mass transit. Some firms reinvested their savings in their own businesses.

In this section, we will determine how the demand and supply model links those who wish to supply financial capital (i.e., savings) with those who demand financial capital (i.e., borrowing). Those who save money (or make financial investments, which is the same thing), whether individuals or businesses, are on the supply side of the financial market. Those who borrow money are on the demand side of the financial market. For a more detailed treatment of the different kinds of financial investments like bank accounts, stocks and bonds, see the Financial Markets chapter.

Who Demands and Who Supplies in Financial Markets?

In any market, the price is what suppliers receive and what demanders pay. In financial markets, those who supply financial capital through saving expect to receive a rate of return, while those who demand financial capital by receiving funds expect to pay a rate of return. This rate of return can come in a variety of forms, depending on the type of investment.

The simplest example of a rate of return is the **interest rate**. For example, when you supply money into a savings account at a bank, you receive interest on your deposit. The interest paid to you as a percent of your deposits is the interest rate. Similarly, if you demand a loan to buy a car or a computer, you will need to pay interest on the money you borrow.

Let's consider the market for borrowing money with credit cards. In 2014, almost 200 million Americans were cardholders. Credit cards allow you to borrow money from the card's issuer, and pay back the borrowed amount plus interest, though most allow you a period of time in which you can repay the loan without paying interest. A typical credit card interest rate ranges from 12% to 18% per year. In 2014, Americans had about $\$793$ billion outstanding in credit card debts. About half of U.S. families with credit cards report that they almost always pay the full balance on time, but one-quarter of U.S. families with credit cards say that they "hardly ever" pay off the card in full. In fact, in 2014, 56% of consumers carried an unpaid balance in the last 12 months. Let's say that, on average, the annual interest rate for credit card borrowing is 15% per year. So, Americans pay tens of billions of dollars every year in interest on their credit cards—plus basic fees for the credit card or fees for late payments.

Figure $3.9.1$ illustrates demand and supply in the financial market for credit cards. The horizontal axis of the financial market shows the quantity of money that is loaned or borrowed in this market. The vertical or price axis shows the rate of return, which in the case of credit card borrowing can be measured with an interest rate. Table $3.9.1$ shows the quantity of financial capital that consumers demand at various interest rates and the quantity that credit card firms (often banks) are willing to supply.

Demand and Supply for Borrowing Money with Credit Cards

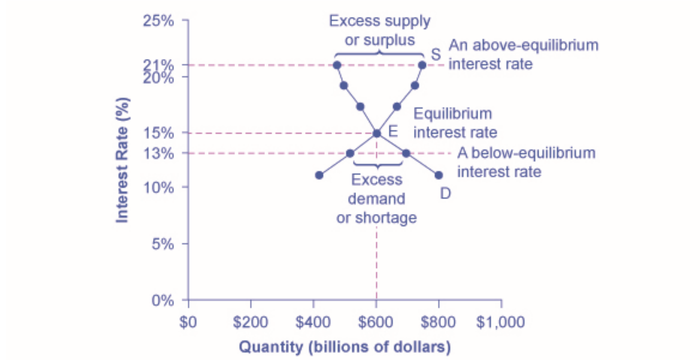


Figure 3.9.1: In this market for credit card borrowing, the demand curve (D) for borrowing financial capital intersects the supply curve (S) for lending financial capital at equilibrium €. At the equilibrium, the interest rate (the "price" in this market) is 15% and the quantity of financial capital being loaned and borrowed is $600 billion. The equilibrium price is where the quantity demanded and the quantity supplied are equal. At an above-equilibrium interest rate like 21%, the quantity of financial capital supplied would increase to $750 billion, but the quantity demanded would decrease to $480 billion. At a below-equilibrium interest rate like 13%, the quantity of financial capital demanded would increase to $700 billion, but the quantity of financial capital supplied would decrease to $510 billion.

Table 3.9.1: Demand and Supply for Borrowing Money with Credit Cards

Interest Rate (%)	Quantity of Financial Capital Demanded (Borrowing) ($ billions)	Quantity of Financial Capital Supplied (Lending) ($ billions)
11	$800	$420
13	$700	$510
15	$600	$600
17	$550	$660
19	$500	$720
21	$480	$750

The laws of demand and supply continue to apply in the financial markets. According to the **law of demand**, a higher rate of return (that is, a higher price) will decrease the quantity demanded. As the interest rate rises, consumers will reduce the quantity that they borrow. According to the law of supply, a higher price increases the quantity supplied. Consequently, as the interest rate paid on credit card borrowing rises, more firms will be eager to issue credit cards and to encourage customers to use them. Conversely, if the interest rate on credit cards falls, the quantity of financial capital supplied in the credit card market will decrease and the quantity demanded will fall.

Equilibrium in Financial Markets

In the financial market for credit cards shown in Figure $3.9.1$, the supply curve (S) and the demand curve (D) cross at the equilibrium point (E). The equilibrium occurs at an interest rate of 15%, where the quantity of funds demanded and the quantity supplied are equal at an equilibrium quantity of $\$600$ billion.

If the interest rate (remember, this measures the "price" in the financial market) is above the equilibrium level, then an excess supply, or a surplus, of financial capital will arise in this market. For example, at an interest rate of 21%, the quantity of funds supplied increases to $\$750$ billion, while the quantity demanded decreases to $\$480$ billion. At this above-equilibrium interest rate, firms are eager to supply loans to credit card borrowers, but relatively few people or businesses wish to borrow. As a result, some credit card firms will lower the interest rates (or other fees) they charge to attract more business. This strategy will push the interest rate down toward the equilibrium level.

If the interest rate is below the equilibrium, then excess demand or a shortage of funds occurs in this market. At an interest rate of 13%, the quantity of funds credit card borrowers demand increases to $\$700$ billion; but the quantity credit card firms are willing to supply is only $\$510$ billion. In this situation, credit card firms will perceive that they are overloaded with eager borrowers and conclude that they have an opportunity to raise interest rates or fees. The interest rate will face economic pressures to creep up toward the equilibrium level.

Shifts in Demand and Supply in Financial Markets

Those who supply financial capital face two broad decisions: how much to save, and how to divide up their savings among different forms of financial investments. We will discuss each of these in turn.

Participants in financial markets must decide when they prefer to consume goods: now or in the future. Economists call this **intertemporal decision making** because it involves decisions across time. Unlike a decision about what to buy from the grocery store, decisions about investment or saving are made across a period of time, sometimes a long period.

Most workers save for retirement because their income in the present is greater than their needs, while the opposite will be true once they retire. So they save today and supply financial markets. If their income increases, they save more. If their perceived situation in the future changes, they change the amount of their saving. For example, there is some evidence that Social Security, the program that workers pay into in order to qualify for government checks after retirement, has tended to reduce the quantity of financial capital that workers save. If this is true, Social Security has shifted the supply of financial capital at any interest rate to the left.

By contrast, many college students need money today when their income is low (or nonexistent) to pay their college expenses. As a result, they borrow today and demand from financial markets. Once they graduate and become employed, they will pay back the loans. Individuals borrow money to purchase homes or cars. A business seeks financial investment so that it has the funds to build a factory or invest in a research and development project that will not pay off for five years, ten years, or even more. So when consumers and businesses have greater confidence that they will be able to repay in the future, the quantity demanded of financial capital at any given interest rate will shift to the right.

For example, in the technology boom of the late 1990s, many businesses became extremely confident that investments in new technology would have a high rate of return, and their demand for financial capital shifted to the right. Conversely, during the Great Recession of 2008 and 2009, their demand for financial capital at any given interest rate shifted to the left.

To this point, we have been looking at saving in total. Now let us consider what affects saving in different types of financial investments. In deciding between different forms of financial investments, suppliers of financial capital will have to consider the rates of return and the risks involved. Rate of return is a positive attribute of investments, but risk is a negative. If Investment A becomes more risky, or the return diminishes, then savers will shift their funds to Investment B—and the supply curve of financial capital for Investment A will shift back to the left while the supply curve of capital for Investment B shifts to the right.

The United States as a Global Borrower

In the global economy, trillions of dollars of financial investment cross national borders every year. In the early 2000s, financial investors from foreign countries were investing several hundred billion dollars per year more in the U.S. economy than U.S. financial investors were investing abroad. The following Work It Out deals with one of the macroeconomic concerns for the U.S. economy in recent years.

Example $3.9.1$: The Effect of Growing U.S. Debt

Imagine that the U.S. economy became viewed as a less desirable place for foreign investors to put their money because of fears about the growth of the U.S. public debt. Using the four-step process for analyzing how changes in supply and demand affect equilibrium outcomes, how would increased U.S. public debt affect the equilibrium price and quantity for capital in U.S. financial markets?

Step 1: Draw a diagram showing demand and supply for financial capital that represents the original scenario in which foreign investors are pouring money into the U.S. economy. Figure $3.9.2$ shows a demand curve, D, and a supply curve, S, where the supply of capital includes the funds arriving from foreign investors. The original equilibrium E_0 occurs at interest rate R_0 and quantity of financial investment Q_0.

The United States as a Global Borrower Before U.S. Debt Uncertainty

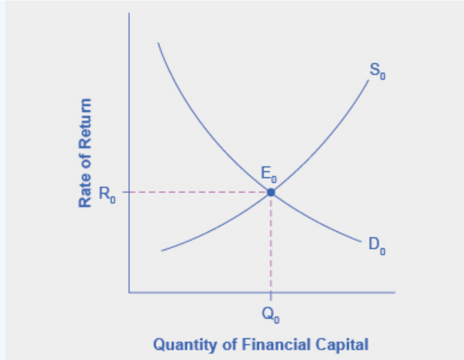


Figure $3.9.2$: The graph shows the demand for financial capital from and supply of financial capital into the U.S. financial markets by the foreign sector before the increase in uncertainty regarding U.S. public debt. The original equilibrium (E_0) occurs at an equilibrium rate of return (R_0) and the equilibrium quantity is at Q_0.

Step 2: Will the diminished confidence in the U.S. economy as a place to invest affect demand or supply of financial capital? Yes, it will affect supply. Many foreign investors look to the U.S. financial markets to store their money in safe financial vehicles with low risk and stable returns. As the U.S. debt increases, debt servicing will increase—that is, more current income will be used to pay the interest rate on past debt. Increasing U.S. debt also means that businesses may have to pay higher interest rates to borrow money, because business is now competing with the government for financial resources.

Step 3: Will supply increase or decrease? When the enthusiasm of foreign investors' for investing their money in the U.S. economy diminishes, the supply of financial capital shifts to the left. Figure $3.9.3$ shows the supply curve shift from S_0 to S_1.

The United States as a Global Borrower Before and After U.S. Debt Uncertainty

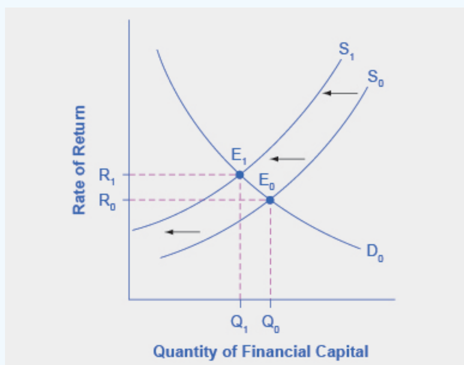


Figure $3.9.3$: The graph shows the demand for financial capital and supply of financial capital into the U.S. financial markets by the foreign sector before and after the increase in uncertainty regarding U.S. public debt. The original equilibrium (E_0) occurs at an equilibrium rate of return (R_0) and the equilibrium quantity is at Q_0.

Step 4: Thus, foreign investors' diminished enthusiasm leads to a new equilibrium, E_1, which occurs at the higher interest rate, R_1, and the lower quantity of financial investment, Q_1.

The economy has experienced an enormous inflow of foreign capital. According to the U.S. Bureau of Economic Analysis, by the third quarter of 2014, U.S. investors had accumulated $\$24.6$ trillion of foreign assets, but foreign investors owned a total of $\$30.8$ trillion of U.S. assets. If foreign investors were to pull their money out of the U.S. economy and invest elsewhere in the world, the result could be a significantly lower quantity of financial investment in the United States, available only at a higher interest rate. This reduced inflow of foreign financial investment could impose hardship on U.S. consumers and firms interested in borrowing.

In a modern, developed economy, financial capital often moves invisibly through electronic transfers between one bank account and another. Yet these flows of funds can be analyzed with the same tools of demand and supply as markets for goods or labor.

Price Ceilings in Financial Markets: Usury Laws

As we noted earlier, about 200 million Americans own credit cards, and their interest payments and fees total tens of billions of dollars each year. It is little wonder that political pressures sometimes arise for setting limits on the interest rates or fees that credit card companies charge. The firms that issue credit cards, including banks, oil companies, phone companies, and retail stores, respond that the higher interest rates are necessary to cover the losses created by those who borrow on their credit cards and who do not repay on time or at all. These companies also point out that cardholders can avoid paying interest if they pay their bills on time.

Consider the credit card market as illustrated in Figure $3.9.4$. In this financial market, the vertical axis shows the interest rate (which is the price in the financial market). Demanders in the credit card market are households and businesses; suppliers are the companies that issue credit cards. This figure does not use specific numbers, which would be hypothetical in any case, but instead focuses on the underlying economic relationships. Imagine a law imposes a price ceiling that holds the interest rate charged on credit cards at the rate R_c, which lies below the interest rate R_0 that would otherwise have prevailed in the market. The price ceiling is shown by the horizontal dashed line in Figure $3.9.4$. The demand and supply model predicts that at the lower price ceiling interest rate, the quantity demanded of credit card debt will increase from its original level of Q_0 to Q_d; however, the quantity supplied of credit card debt will decrease from the original Q_0 to Q_s. At the price ceiling (R_c), quantity demanded will exceed quantity supplied. Consequently, a number of people who want to have credit cards and are willing to pay the prevailing interest rate will find that companies are unwilling to issue cards to them. The result will be a credit shortage.

Credit Card Interest Rates: Another Price Ceiling Example

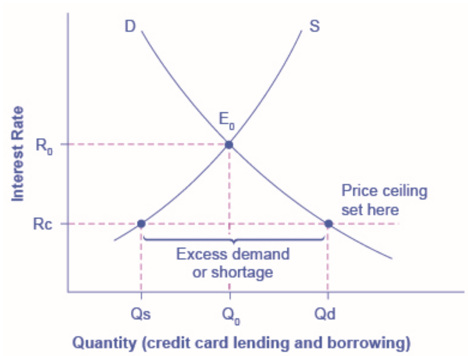


Figure $3.9.4$: The original intersection of demand D and supply S occurs at equilibrium E_0. However, a price ceiling is set at the interest rate R_c, below the equilibrium interest rate R_0, and so the interest rate cannot adjust upward to the equilibrium. At the price ceiling, the quantity demanded, Q_d, exceeds the quantity supplied, Q_s. There is excess demand, also called a shortage.

Many states do have **usury laws**, which impose an upper limit on the interest rate that lenders can charge. However, in many cases these upper limits are well above the market interest rate. For example, if the interest rate is not allowed to rise above 30% per year, it can still fluctuate below that level according to market forces. A price ceiling that is set at a relatively high level is nonbinding, and it will have no practical effect unless the equilibrium price soars high enough to exceed the price ceiling.

Key Concepts and Summary

In the demand and supply analysis of financial markets, the “price” is the rate of return or the interest rate received. The quantity is measured by the money that flows from those who supply financial capital to those who demand it.

Two factors can shift the supply of financial capital to a certain investment: if people want to alter their existing levels of consumption, and if the riskiness or return on one investment changes relative to other investments. Factors that can shift demand for capital include business confidence and consumer confidence in the future—since financial investments received in the present are typically repaid in the future.

References

CreditCards.com. 2013. www.creditcards.com/credit-ca...stics-1276.php.

Glossary

interest rate

the “price” of borrowing in the financial market; a rate of return on an investment

usury laws

laws that impose an upper limit on the interest rate that lenders can charge

Contributor

- Template:ContribOpenSTAXMicroEcon

3.E: Demand and Supply (Exercises)

3.1: Demand, Supply, and Equilibrium in Markets for Goods and Services

Self-Check Questions

Q1

Review Figure 3.1.3. Suppose the price of gasoline is $\$1.60$ per gallon. Is the quantity demanded higher or lower than at the equilibrium price of $\$1.40$ per gallon? And what about the quantity supplied? Is there a shortage or a surplus in the market? If so, of how much?

Review Questions

Q2

What determines the level of prices in a market?

Q3

What does a downward-sloping demand curve mean about how buyers in a market will react to a higher price?

Q4

Will demand curves have the same exact shape in all markets? If not, how will they differ?

Q5

Will supply curves have the same shape in all markets? If not, how will they differ?

Q6

What is the relationship between quantity demanded and quantity supplied at equilibrium? What is the relationship when there is a shortage? What is the relationship when there is a surplus?

Q7

How can you locate the equilibrium point on a demand and supply graph?

Q8

If the price is above the equilibrium level, would you predict a surplus or a shortage? If the price is below the equilibrium level, would you predict a surplus or a shortage? Why?

Q9

When the price is above the equilibrium, explain how market forces move the market price to equilibrium. Do the same when the price is below the equilibrium.

Q10

What is the difference between the demand and the quantity demanded of a product, say milk? Explain in words and show the difference on a graph with a demand curve for milk.

Q11

What is the difference between the supply and the quantity supplied of a product, say milk? Explain in words and show the difference on a graph with the supply curve for milk.

Critical Thinking Questions

Q12

Review Figure 3.1.3. Suppose the government decided that, since gasoline is a necessity, its price should be legally capped at $\$1.30$ per gallon. What do you anticipate would be the outcome in the gasoline market?

Q13

Explain why the following statement is false: "*In the goods market, no buyer would be willing to pay more than the equilibrium price.*"

Q14

Explain why the following statement is false: "*In the goods market, no seller would be willing to sell for less than the equilibrium price.*"

Problems

Q15

Review Figure 3.1.3 again. Suppose the price of gasoline is $\$1.00$. Will the quantity demanded be lower or higher than at the equilibrium price of $\$1.40$ per gallon? Will the quantity supplied be lower or higher? Is there a shortage or a surplus in the market? If so, of how much?

Solution

S1

Since $\$1.60$ per gallon is above the equilibrium price, the quantity demanded would be lower at 550 gallons and the quantity supplied would be higher at 640 gallons. (These results are due to the laws of demand and supply, respectively.) The outcome of lower Q_d and higher Q_s would be a surplus in the gasoline market of $640 - 550 = 90$ gallons.

3.2: Shifts in Demand and Supply for Goods and Services

Self-Check Questions

Q1

Why do economists use the *ceteris paribus* assumption?

Q2

In an analysis of the market for paint, an economist discovers the facts listed below. State whether each of these changes will affect supply or demand, and in what direction.

a. There have recently been some important cost-saving inventions in the technology for making paint.
b. Paint is lasting longer, so that property owners need not repaint as often.
c. Because of severe hailstorms, many people need to repaint now.
d. The hailstorms damaged several factories that make paint, forcing them to close down for several months.

Q3

Many changes are affecting the market for oil. Predict how each of the following events will affect the equilibrium price and quantity in the market for oil. In each case, state how the event will affect the supply and demand diagram. Create a sketch of the diagram if necessary.

a. Cars are becoming more fuel efficient, and therefore get more miles to the gallon.
b. The winter is exceptionally cold.
c. A major discovery of new oil is made off the coast of Norway.
d. The economies of some major oil-using nations, like Japan, slow down.
e. A war in the Middle East disrupts oil-pumping schedules.
f. Landlords install additional insulation in buildings.
g. The price of solar energy falls dramatically.
h. Chemical companies invent a new, popular kind of plastic made from oil.

Review Questions

Q4

When analyzing a market, how do economists deal with the problem that many factors that affect the market are changing at the same time?

Q5

Name some factors that can cause a shift in the demand curve in markets for goods and services.

Q6

Name some factors that can cause a shift in the supply curve in markets for goods and services.

Critical Thinking Questions

Q7

Consider the demand for hamburgers. If the price of a substitute good (for example, hot dogs) increases and the price of a complement good (for example, hamburger buns) increases, can you tell for sure what will happen to the demand for hamburgers? Why or why not? Illustrate your answer with a graph.

Q8

How do you suppose the demographics of an aging population of “Baby Boomers” in the United States will affect the demand for milk? Justify your answer.

Q9

We know that a change in the price of a product causes a movement along the demand curve. Suppose consumers believe that prices will be rising in the future. How will that affect demand for the product in the present? Can you show this graphically?

Q10

Suppose there is soda tax to curb obesity. What should a reduction in the soda tax do to the supply of sodas and to the equilibrium price and quantity? Can you show this graphically? *Hint*: assume that the soda tax is collected from the sellers

Problems

Q11

Table below shows information on the demand and supply for bicycles, where the quantities of bicycles are measured in thousands.

Price	Q_d	Q_s
$120	50	36
$150	40	40
$180	32	48
$210	28	56
$240	24	70

a. What is the quantity demanded and the quantity supplied at a price of $\$210$?
b. At what price is the quantity supplied equal to $48,000$?
c. Graph the demand and supply curve for bicycles. How can you determine the equilibrium price and quantity from the graph? How can you determine the equilibrium price and quantity from the table? What are the equilibrium price and equilibrium quantity?
d. If the price was $\$120$, what would the quantities demanded and supplied be? Would a shortage or surplus exist? If so, how large would the shortage or surplus be?

Q12

The computer market in recent years has seen many more computers sell at much lower prices. What shift in demand or supply is most likely to explain this outcome? Sketch a demand and supply diagram and explain your reasoning for each.

a. A rise in demand
b. A fall in demand
c. A rise in supply
d. A fall in supply

Solution

S1

To make it easier to analyze complex problems. *Ceteris paribus* allows you to look at the effect of one factor at a time on what it is you are trying to analyze. When you have analyzed all the factors individually, you add the results together to get the final answer.

S2

a. An improvement in technology that reduces the cost of production will cause an increase in supply. Alternatively, you can think of this as a reduction in price necessary for firms to supply any quantity. Either way, this can be shown as a rightward (or downward) shift in the supply curve.
b. An improvement in product quality is treated as an increase in tastes or preferences, meaning consumers demand more paint at any price level, so demand increases or shifts to the right. If this seems counterintuitive, note that demand in the future for the longer-lasting paint will fall, since consumers are essentially shifting demand from the future to the present.
c. An increase in need causes an increase in demand or a rightward shift in the demand curve.
d. Factory damage means that firms are unable to supply as much in the present. Technically, this is an increase in the cost of production. Either way you look at it, the supply curve shifts to the left.

S3

a. More fuel-efficient cars means there is less need for gasoline. This causes a leftward shift in the demand for gasoline and thus oil. Since the demand curve is shifting down the supply curve, the equilibrium price and quantity both fall.
b. Cold weather increases the need for heating oil. This causes a rightward shift in the demand for heating oil and thus oil. Since the demand curve is shifting up the supply curve, the equilibrium price and quantity both rise.
c. A discovery of new oil will make oil more abundant. This can be shown as a rightward shift in the supply curve, which will cause a decrease in the equilibrium price along with an increase in the equilibrium quantity. (The supply curve shifts down the demand curve so price and quantity follow the law of demand. If price goes down, then the quantity goes up.)
d. When an economy slows down, it produces less output and demands less input, including energy, which is used in the production of virtually everything. A decrease in demand for energy will be reflected as a decrease in the demand for oil, or a leftward shift in demand for oil. Since the demand curve is shifting down the supply curve, both the equilibrium price and quantity of oil will fall.
e. Disruption of oil pumping will reduce the supply of oil. This leftward shift in the supply curve will show a movement up the demand curve, resulting in an increase in the equilibrium price of oil and a decrease in the equilibrium quantity.
f. Increased insulation will decrease the demand for heating. This leftward shift in the demand for oil causes a movement down the supply curve, resulting in a decrease in the equilibrium price and quantity of oil.
g. Solar energy is a substitute for oil-based energy. So if solar energy becomes cheaper, the demand for oil will decrease as consumers switch from oil to solar. The decrease in demand for oil will be shown as a leftward shift in the demand curve. As the demand curve shifts down the supply curve, both equilibrium price and quantity for oil will fall.
h. A new, popular kind of plastic will increase the demand for oil. The increase in demand will be shown as a rightward shift in demand, raising the equilibrium price and quantity of oil.

3.3: Changes in Equilibrium Price and Quantity: The Four-Step Process

Self-Check Questions

Q1

Let’s think about the market for air travel. From August 2014 to January 2015, the price of jet fuel decreased roughly 47%. Using the four-step analysis, how do you think this fuel price decrease affected the equilibrium price and quantity of air travel?

Q2

A tariff is a tax on imported goods. Suppose the U.S. government cuts the tariff on imported flat screen televisions. Using the four-step analysis, how do you think the tariff reduction will affect the equilibrium price and quantity of flat screen TVs?

Review Questions

Q3

How does one analyze a market where both demand and supply shift?

Q4

What causes a movement along the demand curve? What causes a movement along the supply curve?

Critical Thinking Questions

Q5

Use the four-step process to analyze the impact of the advent of the iPod (or other portable digital music players) on the equilibrium price and quantity of the Sony Walkman (or other portable audio cassette players).

Q6

Use the four-step process to analyze the impact of a reduction in tariffs on imports of iPods on the equilibrium price and quantity of Sony Walkman-type products.

Q7

Suppose both of these events took place at the same time. Combine your analyses of the impacts of the iPod and the tariff reduction to determine the likely impact on the equilibrium price and quantity of Sony Walkman-type products. Show your answer graphically.

Problems

Q8

Demand and supply in the market for cheddar cheese is illustrated in Table below. Graph the data and find the equilibrium. Next, create a table showing the change in quantity demanded or quantity supplied, and a graph of the new equilibrium, in each of the following situations:

a. The price of milk, a key input for cheese production, rises, so that the supply decreases by 80 pounds at every price.
b. A new study says that eating cheese is good for your health, so that demand increases by 20% at every price.

Price per Pound	Q_d	Q_s
$3.00	750	540
$3.20	700	600
$3.40	650	650
$3.60	620	700
$3.80	600	720
$4.00	590	730

Q9

Supply and demand for movie tickets in a city are shown in Table below. Graph demand and supply and identify the equilibrium. Then calculate in a table and graph the effect of the following two changes.

a. Three new nightclubs open. They offer decent bands and have no cover charge, but make their money by selling food and drink. As a result, demand for movie tickets falls by six units at every price.
b. The city eliminates a tax that it had been placing on all local entertainment businesses. The result is that the quantity supplied of movies at any given price increases by 10%.

Price per Pound	Q_d	Q_s
$5.00	26	16
$6.00	24	18

Price per Pound	Q_d	Q_s
$7.00	22	20
$8.00	21	21
$9.00	20	22

Solution

S1

Step 1. Draw the graph with the initial supply and demand curves. Label the initial equilibrium price and quantity.

Step 2. Did the economic event affect supply or demand? Jet fuel is a cost of producing air travel, so an increase in jet fuel price affects supply.

Step 3. An increase in the price of jet fuel caused a decrease in the cost of air travel. We show this as a downward or rightward shift in supply.

Step 4. A rightward shift in supply causes a movement down the demand curve, lowering the equilibrium price of air travel and increasing the equilibrium quantity.

S2

Step 1. Draw the graph with the initial supply and demand curves. Label the initial equilibrium price and quantity.

Step 2. Did the economic event affect supply or demand? A tariff is treated like a cost of production, so this affects supply.

Step 3. A tariff reduction is equivalent to a decrease in the cost of production, which we can show as a rightward (or downward) shift in supply.

Step 4. A rightward shift in supply causes a movement down the demand curve, lowering the equilibrium price and raising the equilibrium quantity.

3.4: Price Ceilings and Price Floors

Self-Check Questions

Q1

What is the effect of a price ceiling on the quantity demanded of the product? What is the effect of a price ceiling on the quantity supplied? Why exactly does a price ceiling cause a shortage?

Q2

Does a price ceiling change the equilibrium price?

Q3

What would be the impact of imposing a price floor below the equilibrium price?

Review Questions

Q4

Does a price ceiling attempt to make a price higher or lower?

Q5

How does a price ceiling set below the equilibrium level affect quantity demanded and quantity supplied?

Q6

Does a price floor attempt to make a price higher or lower?

Q7

How does a price floor set above the equilibrium level affect quantity demanded and quantity supplied?

Critical Thinking Questions

Q8

Most government policy decisions have winners and losers. What are the effects of raising the minimum wage? It is more complex than simply producers lose and workers gain. Who are the winners and who are the losers, and what exactly do they win and lose? To what extent does the policy change achieve its goals?

Q9

Agricultural price supports result in governments holding large inventories of agricultural products. Why do you think the government cannot simply give the products away to poor people?

Q10

Can you propose a policy that would induce the market to supply more rental housing units?

Problems

Q11

A low-income country decides to set a price ceiling on bread so it can make sure that bread is affordable to the poor. The conditions of demand and supply are given in Table below. What are the equilibrium price and equilibrium quantity before the price ceiling? What will the excess demand or the shortage (that is, quantity demanded minus quantity supplied) be if the price ceiling is set at $2.40? At $2.00? At $3.60?

Price	Q_d	Q_s
$1.60	9,000	5,000
$2.00	8,500	5,500
$2.40	8,000	6,400
$2.80	7,500	7,500
$3.20	7,000	9,000
$3.60	6,500	11,000

Price	Q_d	Q_s
$4.00	6,000	15,000

Solution

S1

A price ceiling (which is below the equilibrium price) will cause the quantity demanded to rise and the quantity supplied to fall. This is why a price ceiling creates a shortage.

S2

A price ceiling is just a legal restriction. Equilibrium is an economic condition. People may or may not obey the price ceiling, so the actual price may be at or above the price ceiling, but the price ceiling does not change the equilibrium price.

S3

A price ceiling is a legal maximum price, but a price floor is a legal minimum price and, consequently, it would leave room for the price to rise to its equilibrium level. In other words, a price floor below equilibrium will not be binding and will have no effect.

3.5: Demand, Supply, and Efficiency

Self-Check Questions

Q1

Does a price ceiling increase or decrease the number of transactions in a market? Why? What about a price floor?

Q2

If a price floor benefits producers, why does a price floor reduce social surplus?

Review Questions

Q3

What is consumer surplus? How is it illustrated on a demand and supply diagram?

Q4

What is producer surplus? How is it illustrated on a demand and supply diagram?

Q5

What is total surplus? How is it illustrated on a demand and supply diagram?

Q6

What is the relationship between total surplus and economic efficiency?

Q7

What is deadweight loss?

Critical Thinking Questions

Q8

What term would an economist use to describe what happens when a shopper gets a “good deal” on a product?

Q9

Explain why voluntary transactions improve social welfare.

Q10

Why would a free market never operate at a quantity greater than the equilibrium quantity? Hint: What would be required for a transaction to occur at that quantity?

Solution

S1

Assuming that people obey the price ceiling, the market price will be below equilibrium, which means that Q_d will be more than Q_s. Buyers can only buy what is offered for sale, so the number of transactions will fall to Q_s. This is easy to see graphically. By analogous reasoning, with a price floor the market price will be above the equilibrium price, so Q_d will be less than Q_s. Since the limit on transactions here is demand, the number of transactions will fall to Q_d. Note that because both price floors and price ceilings reduce the number of transactions, social surplus is less.

S2

Because the losses to consumers are greater than the benefits to producers, so the net effect is negative. Since the lost consumer surplus is greater than the additional producer surplus, social surplus falls.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

4: Elasticity

Anyone who has studied economics knows the law of demand: a higher price will lead to a lower quantity demanded. What you may not know is how much lower the quantity demanded will be. Similarly, the law of supply shows that a higher price will lead to a higher quantity supplied. The question is: How much higher? This chapter will explain how to answer these questions and why they are critically important in the real world.

4.1: Prelude to Elasticity

Learning Objectives

- Price Elasticity of Demand and Price Elasticity of Supply
- Polar Cases of Elasticity and Constant Elasticity
- Elasticity and Pricing
- Elasticity in Areas Other Than Price

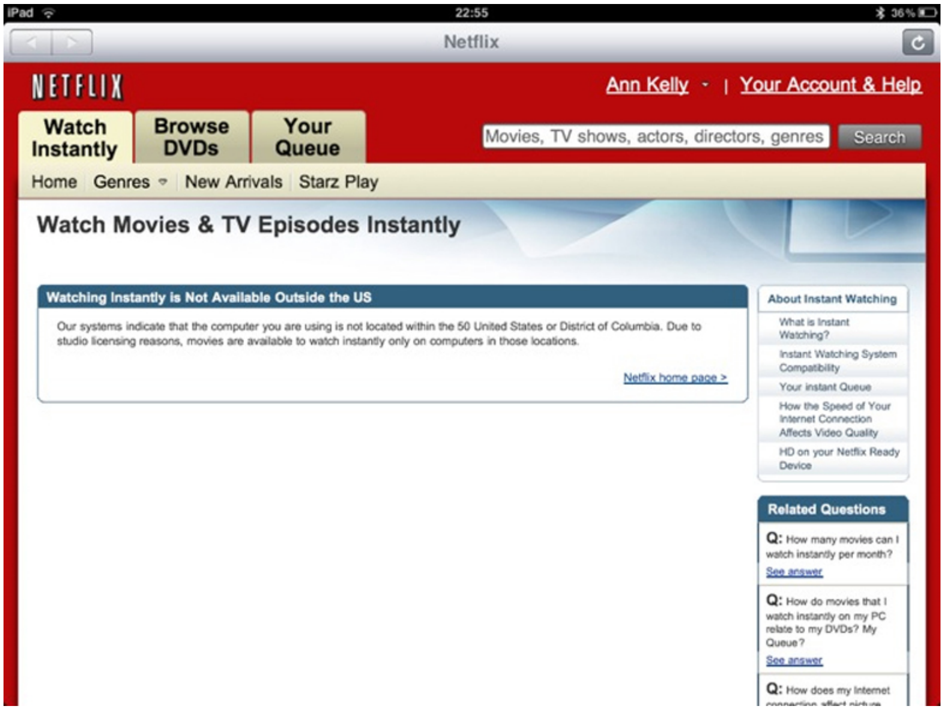


Figure 4.1.1: Netflix, Inc. is an American provider of on-demand Internet streaming media to many countries around the world, including the United States, and of flat rate DVD-by-mail in the United States. (Credit: modification of work by Traci Lawson/Flickr Creative Commons)

That Will Be How Much?

Imagine going to your favorite coffee shop and having the waiter inform you the pricing has changed. Instead of $3 for a cup of coffee, you will now be charged $2 for coffee, $1 for creamer, and $1 for your choice of sweetener. If you pay your usual $3 for a cup of coffee, you must choose between creamer and sweetener. If you want both, you now face an extra charge of $1. Sound absurd? Well, that is the situation Netflix customers found themselves in—a 60% price hike to retain the same service in 2011.

In early 2011, Netflix consumers paid about $10 a month for a package consisting of streaming video and DVD rentals. In July 2011, the company announced a packaging change. Customers wishing to retain both streaming video and DVD rental would be charged $15.98 per month, a price increase of about 60%. In 2014, Netflix also raised its streaming video subscription price from $7.99 to $8.99 per month for new U.S. customers. The company also changed its policy of 4K streaming content from $9.00 to $12.00 per month that year.

How would customers of the 18-year-old firm react? Would they abandon Netflix? Would the ease of access to other venues make a difference in how consumers responded to the Netflix price change? The answers to those questions will be explored in this chapter: the change in quantity with respect to a change in price, a concept economists call elasticity.

Anyone who has studied economics knows the law of demand: a higher price will lead to a lower quantity demanded. What you may not know is how much lower the quantity demanded will be. Similarly, the law of supply shows that a higher price will lead to a higher quantity supplied. The question is: How much higher? This chapter will explain how to answer these questions and why they are critically important in the real world.

To find answers to these questions, we need to understand the concept of elasticity. **Elasticity** is an economics concept that measures responsiveness of one variable to changes in another variable. Suppose you drop two items from a second-floor balcony. The first item is a tennis ball. The second item is a brick. Which will bounce higher? Obviously, the tennis ball. We would say that the tennis ball has greater elasticity.

Consider an economic example. Cigarette taxes are an example of a “sin tax,” a tax on something that is bad for you, like alcohol. Cigarettes are taxed at the state and national levels. State taxes range from a low of 17 cents per pack in Missouri to $\$4.35$ per pack in New York. The average state cigarette tax is $\$1.51$ per pack. The 2014 federal tax rate on cigarettes was $\$1.01$ per pack, but in 2015 the Obama Administration proposed raising the federal tax nearly a dollar to $\$1.95$ per pack. The key question is: How much would cigarette purchases decline?

Taxes on cigarettes serve two purposes: to raise tax revenue for government and to discourage consumption of cigarettes. However, if a higher cigarette tax discourages consumption by quite a lot, meaning a greatly reduced quantity of cigarettes is sold, then the cigarette tax on each pack will not raise much revenue for the government. Alternatively, a higher cigarette tax that does not discourage consumption by much will actually raise more tax revenue for the government. Thus, when a government agency tries to calculate the effects of altering its cigarette tax, it must analyze how much the tax affects the quantity of cigarettes consumed. This issue reaches beyond governments and taxes; every firm faces a similar issue. Every time a firm considers raising the price that it charges, it must consider how much a price increase will reduce the quantity demanded of what it sells. Conversely, when a firm puts its products on sale, it must expect (or hope) that the lower price will lead to a significantly higher quantity demanded.

Contributor

- Template:ContribOpenSTAXMicroEcon

4.2: Price Elasticity of Demand and Price Elasticity of Supply

Learning Objectives

- Calculate the price elasticity of demand
- Calculate the price elasticity of supply

Both the demand and supply curve show the relationship between price and the number of units demanded or supplied. **Price elasticity** is the ratio between the percentage change in the quantity demanded (Q_d) or supplied (Q_s) and the corresponding percent change in price. The **price elasticity of demand** is the percentage change in the quantity demanded of a good or service divided by the percentage change in the price. The **price elasticity of supply** is the percentage change in quantity supplied divided by the percentage change in price.

Elasticities can be usefully divided into three broad categories: elastic, inelastic, and unitary. An **elastic demand** or **elastic supply** is one in which the elasticity is greater than one, indicating a high responsiveness to changes in price. Elasticities that are less than one indicate low responsiveness to price changes and correspond to **inelastic demand** or **inelastic supply**. **Unitary elasticities** indicate proportional responsiveness of either demand or supply, as summarized in Table $4.2.1$.

Table $4.2.1$: Elastic, Inelastic, and Unitary - Three Cases of Elasticity

If . . .	Then . . .	And It Is Called . . .
% change in quantity > % change in price	$\frac{\%\text{ change in quantity}}{\%\text{ change in price}} > 1$	Elastic
% change in quantity = % change in price	$\frac{\%\text{ change in quantity}}{\%\text{ change in price}} = 1$	Unitary
% change in quantity < % change in price	$\frac{\%\text{ change in quantity}}{\%\text{ change in price}} < 1$	Inelastic

To calculate elasticity, instead of using simple percentage changes in quantity and price, economists use the average percent change in both quantity and price. This is called the Midpoint Method for Elasticity, and is represented in the following equations:

$$\%\text{ change in quantity} = \frac{Q_2 - Q_1}{(Q_2 + Q_1)/2} \times 100 \tag{4.2.1}$$

$$\%\text{ change in price} = \frac{P_2 - P_1}{(P_2 + P_1)/2} \times 100 \tag{4.2.2}$$

The advantage of the is **Midpoint Method** is that one obtains the same elasticity between two price points whether there is a price increase or decrease. This is because the formula uses the same base for both cases.

Calculating Price Elasticity of Demand

Let's calculate the elasticity between points A and B and between points G and H shown in Figure $4.2.1$.

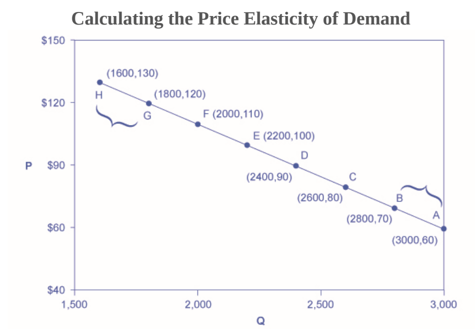


Figure $4.2.1$: The price elasticity of demand is calculated as the percentage change in quantity divided by the percentage change in price.

First, apply the formula to calculate the elasticity as price decreases from $\$70$ at point B to $\$60$ at point A:

$$\begin{aligned} \%\text{ change in quantity } &= \frac{3,000 - 2,800}{(3,000 + 2,800)/2} \times 100 \\ &= \frac{200}{2,900} \times 100 \\ &= 6.9 \end{aligned}$$

$$\begin{aligned} \%\text{ change in price } &= \frac{60 - 70}{(60 + 70)/2} \times 100 \\ &= \frac{-10}{65} \times 100 \\ &= -15.4 \end{aligned}$$

$$\begin{aligned} \text{Price Elasticity of Demand} &= \frac{6.9\%}{-15.4\%} \\ &= 0.45 \end{aligned}$$

Therefore, the elasticity of demand between these two points is $\frac{6.9\%}{-15.4\%}$ which is 0.45, an amount smaller than one, showing that the demand is inelastic in this interval. Price elasticities of demand are always negative since price and quantity demanded always move in oppos

This means that, along the demand curve between point B and A, if the price changes by 1%, the quantity demanded will change by 0.45%. A change in the price will result in a smaller percentage change in the quantity demanded. For example, a 10% increase in the price will result in only a 4.5% decrease in quantity demanded. A 10% decrease in the price will result in only a 4.5% increase in the quantity demanded. Price elasticities of demand are negative numbers indicating that the demand curve is downward sloping, but are read as absolute values. The following Work It Out feature will walk you through calculating the price elasticity of demand.

Example $4.2.1$: Finding the Price Elasticity of Demand

Calculate the price elasticity of demand using the data in Figure $4.2.1$ for an increase in price from G to H. Has the elasticity increased or decreased?

Solution

Step 1: We know that:

$$\text{Price Elasticity of Demand} = \frac{\%\text{ change in quantity}}{\%\text{ change in price}} \tag{4.2.3}$$

Step 2: From the **Midpoint Formula** we know that:

$$\%\text{ change in quantity} = \frac{Q_2 - Q_1}{(Q_2 + Q_1)/2} \times 100 \tag{4.2.4}$$

$$\%\text{ change in price} = \frac{P_2 - P_1}{(P_2 + P_1)/2} \times 100 \tag{4.2.5}$$

Step 3: So we can use the values provided in the figure in each equation:

$$\begin{aligned}\%\text{ change in quantity } &= \frac{1,600 - 1,800}{(1,600 + 1,800)/2} \times 100 \\ &= \frac{-200}{1,700} \times 100 \\ &= -11.76\end{aligned}$$

$$\begin{aligned}\%\text{ change in price } &= \frac{130 - 120}{(130 + 120)/2} \times 100 \\ &= \frac{10}{125} \times 100 \\ &= 8.0\end{aligned}$$

Step 4: Then, those values can be used to determine the price elasticity of demand:

$$\begin{aligned}\text{Price Elasticity of Demand} &= \frac{\%\text{ change in quantity}}{\%\text{ change in price}} \\ &= \frac{-11.76}{8.0} \\ &= 1.47\end{aligned}$$

Therefore, the elasticity of demand from G to H is 1.47. The magnitude of the elasticity has increased (in absolute value) as we moved up along the **demand curve** from points A to B. Recall that the elasticity between these two points was 0.45. Demand was inelastic between points A and B and elastic between points G and H. This shows us that price elasticity of demand changes at different points along a **straight-line demand curve**.

Calculating the Price Elasticity of Supply

Assume that an apartment rents for $\$650$ per month and at that price $10,000$ units are rented as shown in Figure $4.2.2$. When the price increases to $\$700$ per month, $13,000$ units are supplied into the market. By what percentage does apartment supply increase? What is the price sensitivity?

Price Elasticity of Supply

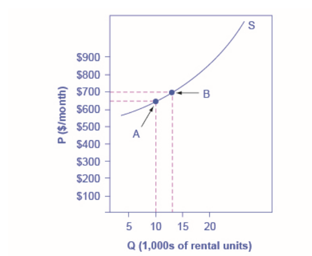


Figure $4.2.2$: The price elasticity of supply is calculated as the percentage change in quantity divided by the percentage change in price.

Using the **Midpoint Method**,

$$\begin{aligned}\%\text{ change in quantity } &= \frac{13,000 - 10,000}{(13,000 + 10,000)/2} \times 100 \\ &= \frac{3,000}{11,500} \times 100 \\ &= 26.1\end{aligned}$$

$$\begin{aligned}\%\text{ change in price } &= \frac{700 - 650}{(700 + 650)/2} \times 100 \\ &= \frac{50}{675} \times 100 \\ &= 7.4\end{aligned}$$

$$\begin{aligned}\text{Price Elasticity of Supply} &= \frac{\%\text{ change in quantity}}{\%\text{ change in price}} \\ &= \frac{26.1}{7.4} \\ &= 3.53\end{aligned}$$

Again, as with the elasticity of demand, the elasticity of supply is not followed by any units. Elasticity is a ratio of one percentage change to another percentage change—nothing more—and is read as an absolute value. In this case, a 1% rise in price causes an increase in quantity supplied of 3.5%. The greater than one elasticity of supply means that the percentage change in quantity supplied will be greater than a one percent price change. If you're starting to wonder if the concept of slope fits into this calculation, read the following Clear It Up box.

Is the elasticity the slope?

It is a common mistake to confuse the slope of either the supply or demand curve with its elasticity. The slope is the rate of change in units along the curve, or the rise/run (change in y over the change in x). For example, in Figure $4.2.1$, each point shown on the demand curve, price drops by $\$10$ and the number of units demanded increases by 200. So the slope is $-10/200$ along the entire demand curve and does not change. The price elasticity, however, changes along the curve. Elasticity between points A and B was 0.45 and increased to 1.47 between points G and H. Elasticity is the percentage change, which is a different calculation from the slope and has a different meaning.

When we are at the upper end of a demand curve, where price is high and the quantity demanded is low, a small change in the quantity demanded, even in, say, one unit, is pretty big in percentage terms. A change in price of, say, a dollar, is going to be much less important in percentage terms than it would have been at the bottom of the demand curve. Likewise, at the bottom of the demand curve, that one unit change when the quantity demanded is high will be small as a percentage.

So, at one end of the demand curve, where we have a large percentage change in quantity demanded over a small percentage change in price, the elasticity value would be high, or demand would be relatively elastic. Even with the same change in the price and the same change in the quantity demanded, at the other end of the demand curve the quantity is much higher, and the price is much lower, so the percentage change in quantity demanded is smaller and the percentage change in price is much higher. That means at the bottom of the curve we'd have a small numerator over a large denominator, so the elasticity measure would be much lower, or inelastic.

As we move along the demand curve, the values for quantity and price go up or down, depending on which way we are moving, so the percentages for, say, a $1 difference in price or a one unit difference in quantity, will change as well, which means the ratios of those percentages will change.

Key Concepts and Summary

Price elasticity measures the responsiveness of the quantity demanded or supplied of a good to a change in its price. It is computed as the percentage change in quantity demanded (or supplied) divided by the percentage change in price. Elasticity can be described as elastic (or very responsive), unit elastic, or inelastic (not very responsive). Elastic demand or supply curves indicate that quantity demanded or supplied respond to price changes in a greater than proportional manner. An inelastic demand or supply curve is one where a given percentage change in price will cause a smaller percentage change in quantity demanded or supplied. A unitary elasticity means that a given percentage change in price leads to an equal percentage change in quantity demanded or supplied.

Glossary

elastic demand
when the elasticity of demand is greater than one, indicating a high responsiveness of quantity demanded or supplied to changes in price

elastic supply
when the elasticity of either supply is greater than one, indicating a high responsiveness of quantity demanded or supplied to changes in price

elasticity
an economics concept that measures responsiveness of one variable to changes in another variable

inelastic demand
when the elasticity of demand is less than one, indicating that a 1 percent increase in price paid by the consumer leads to less than a 1 percent change in purchases (and vice versa); this indicates a low responsiveness by consumers to price changes

inelastic supply
when the elasticity of supply is less than one, indicating that a 1 percent increase in price paid to the firm will result in a less than 1 percent increase in production by the firm; this indicates a low responsiveness of the firm to price increases (and vice versa if prices drop)

price elasticity
the relationship between the percent change in price resulting in a corresponding percentage change in the quantity demanded or supplied

price elasticity of demand
percentage change in the quantity *demanded* of a good or service divided the percentage change in price

price elasticity of supply
percentage change in the quantity *supplied* divided by the percentage change in price

unitary elasticity
when the calculated elasticity is equal to one indicating that a change in the price of the good or service results in a proportional change in the quantity demanded or supplied

Contributor

- Template:ContribOpenSTAXMicroEcon

4.3: Polar Cases of Elasticity and Constant Elasticity

Learning Objectives

- Differentiate between infinite and zero elasticity
- Analyze graphs in order to classify elasticity as constant unitary, infinite, or zero

There are two extreme cases of elasticity: when elasticity equals zero and when it is infinite. A third case is that of constant unitary elasticity. We will describe each case. **Infinite elasticity** or **perfect elasticity** refers to the extreme case where either the quantity demanded (Q_d) or supplied (Q_s) changes by an infinite amount in response to any change in price at all. In both cases, the supply and the **demand curve** are horizontal as shown in Figure $4.3.1$. While perfectly elastic supply curves are unrealistic, goods with readily available inputs and whose production can be easily expanded will feature highly elastic supply curves. Examples include pizza, bread, books and pencils. Similarly, perfectly elastic demand is an extreme example. But luxury goods, goods that take a large share of individuals' income, and goods with many substitutes are likely to have highly elastic demand curves. Examples of such goods are Caribbean cruises and sports vehicles.

Infinite Elasticity

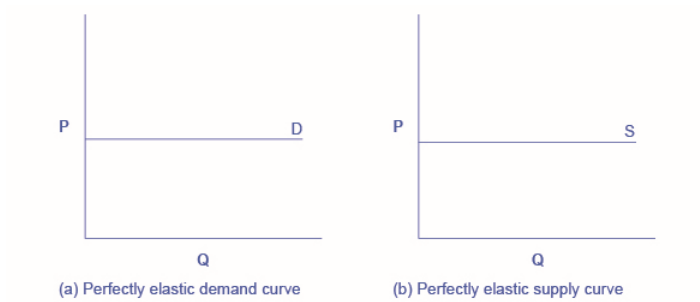


Figure $4.3.1$: The horizontal lines show that an infinite quantity will be demanded or supplied at a specific price. This illustrates the cases of a perfectly (or infinitely) elastic demand curve and supply curve. The quantity supplied or demanded is extremely responsive to price changes, moving from zero for prices close to P to infinite when price reach P.

Zero elasticity or perfect inelasticity, as depicted in Figure $4.3.2$ refers to the extreme case in which a percentage change in price, no matter how large, results in zero change in quantity. While a perfectly inelastic supply is an extreme example, goods with limited supply of inputs are likely to feature highly inelastic supply curves. Examples include diamond rings or housing in prime locations such as apartments facing Central Park in New York City. Similarly, while perfectly inelastic demand is an extreme case, necessities with no close substitutes are likely to have highly inelastic demand curves. This is the case of life-saving drugs and gasoline.

Zero Elasticity

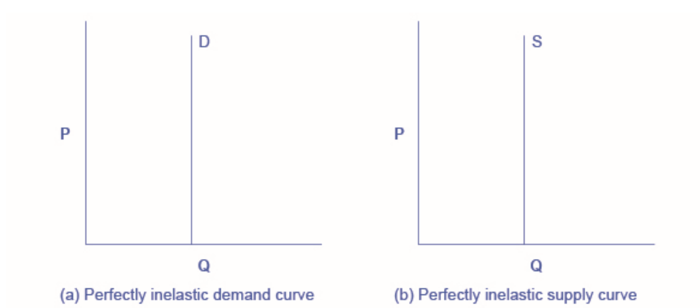


Figure $4.3.2$: The vertical supply curve and vertical demand curve show that there will be zero percentage change in quantity (a) demanded or (b) supplied, regardless of the price.

Constant unitary elasticity, in either a supply or demand curve, occurs when a price change of one percent results in a quantity change of one percent. Figure $4.3.3$ shows a demand curve with constant unit elasticity. As we move down the demand curve from A to B, the price falls by 33% and quantity demanded rises by 33%; as you move from B to C, the price falls by 25% and the quantity demanded rises by 25%; as you move from C to D, the price falls by 16% and the quantity rises by 16%. Notice that in absolute value, the declines in price, as you step down the demand curve, are not identical. Instead, the price falls by $\$3$ from A to

B, by a smaller amount of $\$1.50$ from B to C, and by a still smaller amount of $\$0.75$ from C to D. As a result, a demand curve with constant unitary elasticity moves from a steeper slope on the left and a flatter slope on the right—and a curved shape overall.

A Constant Unitary Elasticity Demand Curve

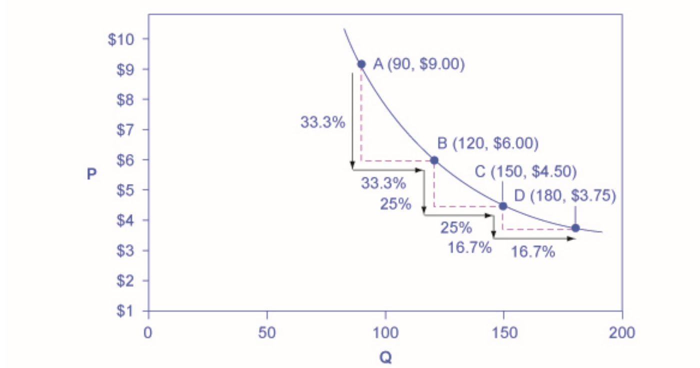


Figure $4.3.3$: A demand curve with constant unitary elasticity will be a curved line. Notice how price and quantity demanded change by an identical amount in each step down the demand curve.

Unlike the demand curve with unitary elasticity, the supply curve with unitary elasticity is represented by a straight line. In moving up the supply curve from left to right, each increase in quantity of 30, from 90 to 120 to 150 to 180, is equal in absolute value. However, in percentage value, the steps are decreasing, from 33.3% to 25% to 16.7%, because the original quantity points in each percentage calculation are getting larger and larger, which expands the denominator in the elasticity calculation.

Consider the price changes moving up the supply curve in Figure $4.3.4$. From points D to E to F and to G on the supply curve, each step of $\$1.50$ is the same in absolute value. However, if the price changes are measured in percentage change terms, they are also decreasing, from 33.3% to 25% to 16.7%, because the original price points in each percentage calculation are getting larger and larger in value. Along the constant unitary elasticity supply curve, the percentage quantity increases on the horizontal axis exactly match the percentage price increases on the vertical axis—so this supply curve has a constant unitary elasticity at all points.

A Constant Unitary Elasticity Supply Curve

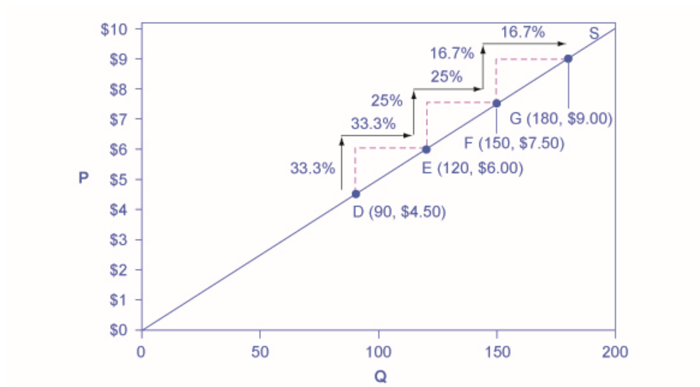


Figure $4.3.4$: A constant unitary elasticity supply curve is a straight line reaching up from the origin. Between each point, the percentage increase in quantity supplied is the same as the percentage increase in price.

Key Concepts and Summary

Infinite or perfect elasticity refers to the extreme case where either the quantity demanded or supplied changes by an infinite amount in response to any change in price at all. Zero elasticity refers to the extreme case in which a percentage change in price, no matter how large, results in zero change in quantity. Constant unitary elasticity in either a supply or demand curve refers to a situation where a price change of one percent results in a quantity change of one percent.

Glossary

constant unitary elasticity

when a given percent price change in price leads to an equal percentage change in quantity demanded or supplied

infinite elasticity

the extremely elastic situation of demand or supply where quantity changes by an infinite amount in response to any change in price; horizontal in appearance

perfect elasticity

see infinite elasticity

zero inelasticity

the highly inelastic case of demand or supply in which a percentage change in price, no matter how large, results in zero change in the quantity; vertical in appearance

perfect inelasticity

see zero elasticity

Contributor

- Template:ContribOpenSTAXMicroEcon

4.4: Elasticity and Pricing

Learning Objectives

- Analyze how price elasticities impact revenue
- Evaluate how elasticity can cause shifts in demand and supply
- Predict how the long-run and short-run impacts of elasticity affect equilibrium
- Explain how the elasticity of demand and supply determine the incidence of a tax on buyers and sellers

Studying elasticities is useful for a number of reasons, pricing being most important. Let's explore how elasticity relates to revenue and pricing, both in the long run and short run. But first, let's look at the elasticities of some common goods and services.

Table $4.4.1$ shows a selection of demand elasticities for different goods and services drawn from a variety of different studies by economists, listed in order of increasing elasticity.

Table $4.4.1$: Some Selected Elasticities of Demand

Goods and Services	Elasticity of Price
Housing	0.12
Transatlantic air travel (economy class)	0.12
Rail transit (rush hour)	0.15
Electricity	0.20
Taxi cabs	0.22
Gasoline	0.35
Transatlantic air travel (first class)	0.40
Wine	0.55
Beef	0.59
Transatlantic air travel (business class)	0.62
Kitchen and household appliances	0.63
Cable TV (basic rural)	0.69
Chicken	0.64
Soft drinks	0.70
Beer	0.80
New vehicle	0.87
Rail transit (off-peak)	1.00
Computer	1.44
Cable TV (basic urban)	1.51
Cable TV (premium)	1.77
Restaurant meals	2.27

Note that necessities such as housing and electricity are inelastic, while items that are not necessities such as restaurant meals are more price-sensitive. If the price of the restaurant meal increases by 10%, the quantity demanded will decrease by 22.7%. A 10% increase in the price of housing will cause a slight decrease of 1.2% in the quantity of housing demanded.

Does Raising Price Bring in More Revenue?

Imagine that a band on tour is playing in an indoor arena with $15,000$ seats. To keep this example simple, assume that the band keeps all the money from ticket sales. Assume further that the band pays the costs for its appearance, but that these costs, like travel, setting up the stage, and so on, are the same regardless of how many people are in the audience. Finally, assume that all the tickets have the same price. (The same insights apply if ticket prices are more expensive for some seats than for others, but the calculations become more complicated.) The band knows that it faces a downward-sloping demand curve; that is, if the band raises the price of tickets, it will sell fewer tickets. How should the band set the price for tickets to bring in the most total revenue, which in this example, because costs are fixed, will also mean the highest profits for the band? Should the band sell more tickets at a lower price or fewer tickets at a higher price?

The key concept in thinking about collecting the most revenue is the price elasticity of demand. Total revenue is price times the quantity of tickets sold. Imagine that the band starts off thinking about a certain price, which will result in the sale of a certain quantity of tickets. The three possibilities are laid out in Table $4.4.2$. If demand is elastic at that price level, then the band should cut the price, because the percentage drop in price will result in an even larger percentage increase in the quantity sold—thus raising total revenue. However, if demand is inelastic at that original quantity level, then the band should raise the price of tickets, because a certain percentage increase in price will result in a smaller percentage decrease in the quantity sold—and total revenue will rise. If demand has a unitary elasticity at that quantity, then a moderate percentage change in the price will be offset by an equal percentage change in quantity—so the band will earn the same revenue whether it (moderately) increases or decreases the price of tickets.

Table $4.4.2$: Will the Band Earn More Revenue by Changing Ticket Prices?

If Demand Is . . .	Then . . .	Therefore . . .
Elastic	% change in Q_d > % change in P	A given % rise in P will be more than offset by a larger % fall in Q so that total revenue ($P \times Q$) falls.
Unitary	% change in Q_d = % change in P	A given % rise in P will be exactly offset by an equal % fall in Q so that total revenue ($P \times Q$) is unchanged.
Inelastic	% change in Q_d < % change in P	A given % rise in P will cause a smaller % fall in Q so that total revenue ($P \times Q$) rises.

What if the band keeps cutting price, because demand is elastic, until it reaches a level where all $15,000$ seats in the available arena are sold? If demand remains elastic at that quantity, the band might try to move to a bigger arena, so that it could cut ticket prices further and see a larger percentage increase in the quantity of tickets sold. Of course, if the $15,000$-seat arena is all that is available or if a larger arena would add substantially to costs, then this option may not work.

Conversely, a few bands are so famous, or have such fanatical followings, that demand for tickets may be inelastic right up to the point where the arena is full. These bands can, if they wish, keep raising the price of tickets. Ironically, some of the most popular bands could make more revenue by setting prices so high that the arena is not filled—but those who buy the tickets would have to pay very high prices. However, bands sometimes choose to sell tickets for less than the absolute maximum they might be able to charge, often in the hope that fans will feel happier and spend more on recordings, T-shirts, and other paraphernalia.

Can Costs Be Passed on to Consumers?

Most businesses face a day-to-day struggle to figure out ways to produce at a lower cost, as one pathway to their goal of earning higher profits. However, in some cases, the price of a key input over which the firm has no control may rise. For example, many chemical companies use petroleum as a **key input**, but they have no control over the world market price for crude oil. Coffee shops use coffee as a key input, but they have no control over the world market price of coffee. If the cost of a key input rises, can the firm pass those higher costs along to consumers in the form of higher prices? Conversely, if new and less expensive ways of producing are invented, can the firm keep the benefits in the form of higher profits, or will the market pressure them to pass the gains along to consumers in the form of lower prices? The price elasticity of demand plays a key role in answering these questions.

Imagine that as a consumer of legal pharmaceutical products, you read a newspaper story that a technological breakthrough in the production of aspirin has occurred, so that every aspirin factory can now make aspirin more cheaply than it did before. What does this discovery mean to you? Figure $4.4.1$ illustrates two possibilities. In Figure $4.4.1$ (a), the demand curve is drawn as highly inelastic. In this case, a technological breakthrough that shifts supply to the right, from S0 to S1, so that the equilibrium shifts from

E_0 to E_1, creates a substantially lower price for the product with relatively little impact on the quantity sold. In Figure 4.4.1(b), the demand curve is drawn as highly elastic. In this case, the technological breakthrough leads to a much greater quantity being sold in the market at very close to the original price. Consumers benefit more, in general, when the demand curve is more inelastic because the shift in the supply results in a much lower price for consumers.

Passing along Cost Savings to Consumers

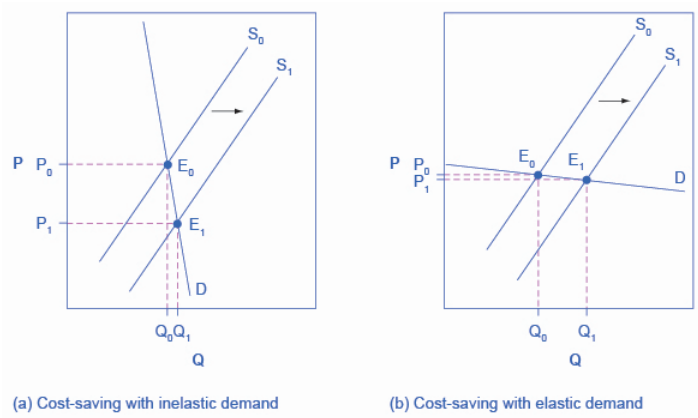


Figure 4.4.1: Cost-saving gains cause supply to shift out to the right from S_0 to S_1; that is, at any given price, firms will be willing to supply a greater quantity. If demand is inelastic, as in (a), the result of this cost-saving technological improvement will be substantially lower prices. If demand is elastic, as in (b), the result will be only slightly lower prices. Consumers benefit in either case, from a greater quantity at a lower price, but the benefit is greater when demand is inelastic, as in (a).

Producers of aspirin may find themselves in a nasty bind here. The situation shown in Figure 4.4.1, with extremely inelastic demand, means that a new invention may cause the price to drop dramatically while quantity changes little. As a result, the new production technology can lead to a drop in the revenue that firms earn from sales of aspirin. However, if strong competition exists between producers of aspirin, each producer may have little choice but to search for and implement any breakthrough that allows it to reduce production costs. After all, if one firm decides not to implement such a cost-saving technology, it can be driven out of business by other firms that do.

Since demand for food is generally inelastic, farmers may often face the situation in Figure 4.4.1(a). That is, a surge in production leads to a severe drop in price that can actually decrease the total revenue received by farmers. Conversely, poor weather or other conditions that cause a terrible year for farm production can sharply raise prices so that the total revenue received increases. The Clear It Up box discusses how these issues relate to coffee.

How do coffee prices fluctuate?

Coffee is an international crop. The top five coffee-exporting nations are Brazil, Vietnam, Colombia, Indonesia, and Ethiopia. In these nations and others, 20 million families depend on selling coffee beans as their main source of income. These families are exposed to enormous risk, because the world price of coffee bounces up and down. For example, in 1993, the world price of coffee was about 50 cents per pound; in 1995 it was four times as high, at $2 per pound. By 1997 it had fallen by half to $1.00 per pound. In 1998 it leaped back up to $2 per pound. By 2001 it had fallen back to 46 cents a pound; by early 2011 it went back up to about $2.31 per pound. By the end of 2012, the price had fallen back to about $1.31 per pound.

The reason for these price bounces lies in a combination of inelastic demand and shifts in supply. The elasticity of coffee demand is only about 0.3; that is, a 10% rise in the price of coffee leads to a decline of about 3% in the quantity of coffee consumed. When a major frost hit the Brazilian coffee crop in 1994, coffee supply shifted to the left with an inelastic demand curve, leading to much higher prices. Conversely, when Vietnam entered the world coffee market as a major producer in the late 1990s, the supply curve shifted out to the right. With a highly inelastic demand curve, coffee prices fell dramatically. This situation is shown in Figure 4.4.1(a).

Elasticity also reveals whether firms can pass higher costs that they incur on to consumers. Addictive substances tend to fall into this category. For example, the demand for cigarettes is relatively inelastic among regular smokers who are somewhat addicted; economic research suggests that increasing the price of cigarettes by 10% leads to about a 3% reduction in the quantity of cigarettes smoked by adults, so the elasticity of demand for cigarettes is 0.3. If society increases taxes on companies that make cigarettes, the result will be, as in Figure 4.4.2 (a), that the supply curve shifts from S_0 to S_1. However, as the equilibrium moves from E_0 to E_1, these taxes are mainly passed along to consumers in the form of higher prices. These higher taxes on cigarettes will raise tax revenue for the government, but they will not much affect the quantity of smoking.

If the goal is to reduce the quantity of cigarettes demanded, it must be achieved by shifting this inelastic demand back to the left, perhaps with public programs to discourage the use of cigarettes or to help people to quit. For example, anti-smoking advertising campaigns have shown some ability to reduce smoking. However, if demand for cigarettes was more elastic, as in Figure 4.4.2 (b), then an increase in taxes that shifts supply from S_0 to S_1 and equilibrium from E_0 to E_1 would reduce the quantity of cigarettes smoked substantially. Youth smoking seems to be more elastic than adult smoking—that is, the quantity of youth smoking will fall by a greater percentage than the quantity of adult smoking in response to a given percentage increase in price.

Passing along Higher Costs to Consumers

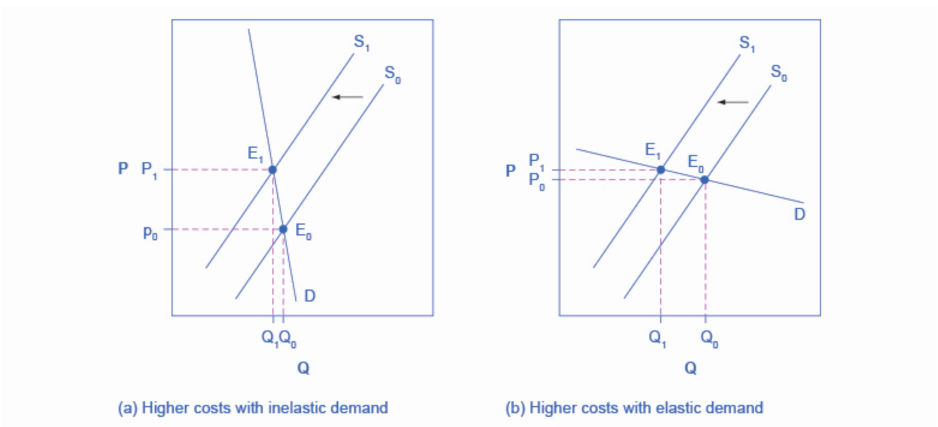


Figure 4.4.2: Higher costs, like a higher tax on cigarette companies for the example given in the text, lead supply to shift to the left. This shift is identical in (a) and (b). However, in (a), where demand is inelastic, the cost increase can largely be passed along to consumers in the form of higher prices, without much of a decline in equilibrium quantity. In (b), demand is elastic, so the shift in supply results primarily in a lower equilibrium quantity. Consumers suffer in either case, but in (a), they suffer from paying a higher price for the same quantity, while in (b), they suffer from buying a lower quantity (and presumably needing to shift their consumption elsewhere).

Elasticity and Tax Incidence

The example of cigarette taxes showed that because demand is inelastic, taxes are not effective at reducing the equilibrium quantity of smoking, and they are mainly passed along to consumers in the form of higher prices. The analysis, or manner, of how the burden of a tax is divided between consumers and producers is called **tax incidence**. Typically, the incidence, or burden, of a tax falls both on the consumers and producers of the taxed good. But if one wants to predict which group will bear most of the burden, all one needs to do is examine the elasticity of demand and supply. In the tobacco example, the tax burden falls on the most inelastic side of the market.

If demand is more inelastic than supply, consumers bear most of the tax burden, and if supply is more inelastic than demand, sellers bear most of the tax burden.

The intuition for this is simple. When the demand is inelastic, consumers are not very responsive to price changes, and the quantity demanded remains relatively constant when the tax is introduced. In the case of smoking, the demand is inelastic because consumers are addicted to the product. The government can then pass the tax burden along to consumers in the form of higher prices, without much of a decline in the equilibrium quantity.

Similarly, when a tax is introduced in a market with an inelastic supply, such as, for example, beachfront hotels, and sellers have no alternative than to accept lower prices for their business, taxes do not greatly affect the equilibrium quantity. The tax burden is now passed on to the sellers. If the supply was elastic and sellers had the possibility of reorganizing their businesses to avoid supplying the taxed good, the tax burden on the sellers would be much smaller. The tax would result in a much lower quantity sold instead of lower prices received. Figure 4.4.3 illustrates this relationship between the tax incidence and elasticity of demand and supply.

Elasticity and Tax Incidence

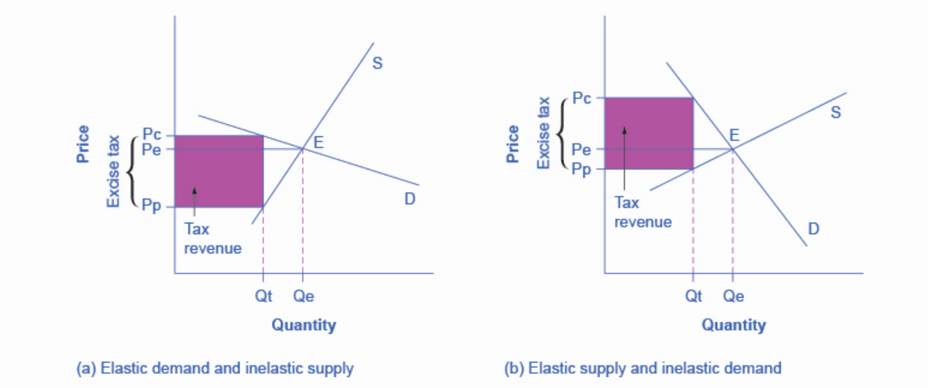


Figure $4.4.3$: An excise tax introduces a wedge between the price paid by consumers (P_c) and the price received by producers (P_p). (a) When the demand is more elastic than supply, the tax incidence on consumers $P_c - P_e$ is lower than the tax incidence on producers $P_e - P_p$. (b) When the supply is more elastic than demand, the tax incidence on consumers $P_c - P_e$ is larger than the tax incidence on producers $P_e - P_p$. The more elastic the demand and supply curves are, the lower the tax revenue.

In Figure $4.4.3$ (a), the supply is inelastic and the demand is elastic, such as in the example of beachfront hotels. While consumers may have other vacation choices, sellers can't easily move their businesses. By introducing a tax, the government essentially creates a wedge between the price paid by consumers P_c and the price received by producers P_p. In other words, of the total price paid by consumers, part is retained by the sellers and part is paid to the government in the form of a tax. The distance between P_c and P_p is the tax rate. The new market price is P_c, but sellers receive only P_p per unit sold, as they pay $P_c - P_p$ to the government. Since a tax can be viewed as raising the costs of production, this could also be represented by a leftward shift of the supply curve, where the new supply curve would intercept the demand at the new quantity Q_t. For simplicity, Figure $4.4.3$ omits the shift in the supply curve.

The tax revenue is given by the shaded area, which is obtained by multiplying the tax per unit by the total quantity sold Q_t. The tax incidence on the consumers is given by the difference between the price paid P_c and the initial equilibrium price P_e. The tax incidence on the sellers is given by the difference between the initial equilibrium price P_e and the price they receive after the tax is introduced P_p. In Figure $4.4.3$ (a), the tax burden falls disproportionately on the sellers, and a larger proportion of the tax revenue (the shaded area) is due to the resulting lower price received by the sellers than by the resulting higher prices paid by the buyers. The example of the tobacco excise tax could be described by Figure $4.4.3$ (b) where the supply is more elastic than demand. The tax incidence now falls disproportionately on consumers, as shown by the large difference between the price they pay, P_c, and the initial equilibrium price, P_e. Sellers receive a lower price than before the tax, but this difference is much smaller than the change in consumers' price. From this analysis one can also predict whether a tax is likely to create a large revenue or not. The more elastic the demand curve, the easier it is for consumers to reduce quantity instead of paying higher prices. The more elastic the supply curve, the easier it is for sellers to reduce the quantity sold, instead of taking lower prices. In a market where both the demand and supply are very elastic, the imposition of an excise tax generates low revenue.

Excise taxes tend to be thought to hurt mainly the specific industries they target. For example, the medical device excise tax, in effect since 2013, has been controversial for it can delay industry profitability and therefore hamper start-ups and medical innovation. But ultimately, whether the tax burden falls mostly on the medical device industry or on the patients depends simply on the elasticity of demand and supply.

Long-Run vs. Short-Run Impact

Elasticities are often lower in the short run than in the long run. On the demand side of the market, it can sometimes be difficult to change Q_d in the short run, but easier in the long run. Consumption of energy is a clear example. In the short run, it is not easy for a person to make substantial changes in the energy consumption. Maybe you can carpool to work sometimes or adjust your home thermostat by a few degrees if the cost of energy rises, but that is about all. However, in the long-run you can purchase a car that gets more miles to the gallon, choose a job that is closer to where you live, buy more energy-efficient home appliances, or install more insulation in your home. As a result, the elasticity of demand for energy is somewhat inelastic in the short run, but much more elastic in the long run.

Figure $4.4.4$ is an example, based roughly on historical experience, for the responsiveness of Q_d to price changes. In 1973, the price of crude oil was $\$12$ per barrel and total consumption in the U.S. economy was 17 million barrels per day. That year, the nations who were members of the Organization of Petroleum Exporting Countries (OPEC) cut off oil exports to the United States for six months because the Arab members of OPEC disagreed with the U.S. support for Israel. OPEC did not bring exports back to their earlier levels until 1975—a policy that can be interpreted as a shift of the supply curve to the left in the U.S. petroleum market. Figure 4.4.4 (a) and Figure 4.4.4 (b) show the same original equilibrium point and the same identical shift of a supply curve to the left from S_0 to S_1.

How a Shift in Supply Can Affect Price or Quantity

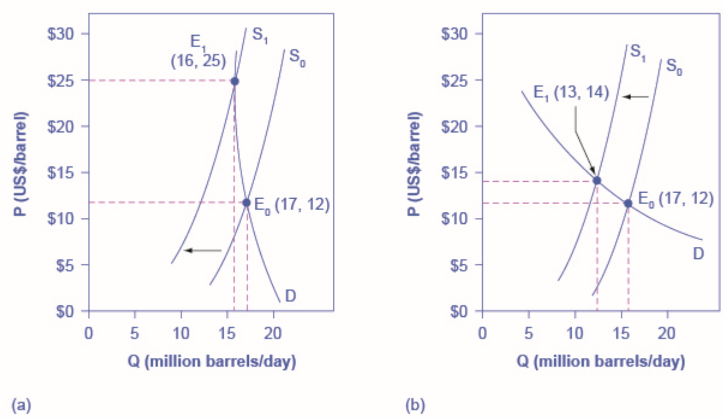


Figure 4.4.4: The intersection (E_0) between demand curve D and supply curve S_0 is the same in both (a) and (b). The shift of supply to the left from S_0 to S_1 is identical in both (a) and (b). The new equilibrium (E_1) has a higher price and a lower quantity than the original equilibrium (E_0) in both (a) and (b). However, the shape of the demand curve D is different in (a) and (b). As a result, the shift in supply can result either in a new equilibrium with a much higher price and an only slightly smaller quantity, as in (a), or in a new equilibrium with only a small increase in price and a relatively larger reduction in quantity, as in (b).

Figure 4.4.4 (a) shows inelastic demand for oil in the short run similar to that which existed for the United States in 1973. In Figure 4.4.4 (a), the new equilibrium (E_1) occurs at a price of $25 per barrel, roughly double the price before the OPEC shock, and an equilibrium quantity of 16 million barrels per day. Figure 4.4.4 (b) shows what the outcome would have been if the U.S. demand for oil had been more elastic, a result more likely over the long term. This alternative equilibrium (E_1) would have resulted in a smaller price increase to $14 per barrel and larger reduction in equilibrium quantity to 13 million barrels per day. In 1983, for example, U.S. petroleum consumption was 15.3 million barrels a day, which was lower than in 1973 or 1975. U.S. petroleum consumption was down even though the U.S. economy was about one-fourth larger in 1983 than it had been in 1973. The primary reason for the lower quantity was that higher energy prices spurred conservation efforts, and after a decade of home insulation, more fuel-efficient cars, more efficient appliances and machinery, and other fuel-conserving choices, the demand curve for energy had become more elastic.

On the supply side of markets, producers of goods and services typically find it easier to expand production in the long term of several years rather than in the short run of a few months. After all, in the short run it can be costly or difficult to build a new factory, hire many new workers, or open new stores. But over a few years, all of these are possible.

Indeed, in most markets for goods and services, prices bounce up and down more than quantities in the short run, but quantities often move more than prices in the long run. The underlying reason for this pattern is that supply and demand are often inelastic in the short run, so that shifts in either demand or supply can cause a relatively greater change in prices. But since supply and demand are more elastic in the long run, the long-run movements in prices are more muted, while quantity adjusts more easily in the long run.

Key Concepts and Summary

In the market for goods and services, quantity supplied and quantity demanded are often relatively slow to react to changes in price in the short run, but react more substantially in the long run. As a result, demand and supply often (but not always) tend to be relatively inelastic in the short run and relatively elastic in the long run. The tax incidence depends on the relative price elasticity of supply and demand. When supply is more elastic than demand, buyers bear most of the tax burden, and when demand is more elastic than supply, producers bear most of the cost of the tax. Tax revenue is larger the more inelastic the demand and supply are.

Glossary

tax incidence

manner in which the tax burden is divided between buyers and sellers

Contributors and Attributions

- Template:ContribOpenSTAXMicroEcon

4.5: Elasticity in Areas Other Than Price

Learning Objectives

- Calculate the income elasticity of demand and the cross-price elasticity of demand
- Calculate the elasticity in labor and financial capital markets through an understanding of the elasticity of labor supply and the elasticity of savings
- Apply concepts of price elasticity to real-world situations

The basic idea of elasticity—how a percentage change in one variable causes a percentage change in another variable—does not just apply to the responsiveness of supply and demand to changes in the price of a product. Recall that quantity demanded (Q_d) depends on income, tastes and preferences, the prices of related goods, and so on, as well as price. Similarly, quantity supplied (Q_s) depends on the cost of production, and so on, as well as price. Elasticity can be measured for any determinant of supply and demand, not just the price.

Income Elasticity of Demand

The **income elasticity of demand** is the percentage change in quantity demanded divided by the percentage change in income.

$$\text{Income elasticity of demand} = \frac{\%\ \text{change in quantity demanded}}{\%\ \text{change in income}} \tag{4.5.1}$$

For most products, most of the time, the income elasticity of demand is positive: that is, a rise in income will cause an increase in the quantity demanded. This pattern is common enough that these goods are referred to as **normal goods**. However, for a few goods, an increase in income means that one might purchase less of the good; for example, those with a higher income might buy fewer hamburgers, because they are buying more steak instead, or those with a higher income might buy less cheap wine and more imported beer. When the income elasticity of demand is negative, the good is called an **inferior good**.

The concepts of normal and inferior goods were introduced in Demand and Supply. A higher level of income for a normal good causes a demand curve to shift to the right for a normal good, which means that the income elasticity of demand is positive. How far the demand shifts depends on the income elasticity of demand. A higher income elasticity means a larger shift. However, for an inferior good, that is, when the income elasticity of demand is negative, a higher level of income would cause the demand curve for that good to shift to the left. Again, how much it shifts depends on how large the (negative) income elasticity is.

Cross-Price Elasticity of Demand

A change in the price of one good can shift the quantity demanded for another good. If the two goods are complements, like bread and peanut butter, then a drop in the price of one good will lead to an increase in the quantity demanded of the other good. However, if the two goods are substitutes, like plane tickets and train tickets, then a drop in the price of one good will cause people to substitute toward that good, and to reduce consumption of the other good. Cheaper plane tickets lead to fewer train tickets, and vice versa.

The **cross-price elasticity of demand** puts some meat on the bones of these ideas. The term “cross-price” refers to the idea that the price of one good is affecting the quantity demanded of a different good. Specifically, the cross-price elasticity of demand is the percentage change in the quantity of good A that is demanded as a result of a percentage change in the price of good B.

$$\text{Cross-price elasticity of demand} = \frac{\%\ \text{change in Qd of good A}}{\%\ \text{change in price of good B}} \tag{4.5.2}$$

Substitute goods have positive cross-price elasticities of demand: if good A is a substitute for good B, like coffee and tea, then a higher price for B will mean a greater quantity consumed of A. Complement goods have negative cross-price elasticities: if good A is a complement for good B, like coffee and sugar, then a higher price for B will mean a lower quantity consumed of A.

Elasticity in Labor and Financial Capital Markets

The concept of elasticity applies to any market, not just markets for goods and services. In the labor market, for example, the **wage elasticity of labor supply**—that is, the percentage change in hours worked divided by the percentage change in wages—will determine the shape of the labor supply curve. Specifically:

$$\text{Elasticity of labor supply} = \frac{\%\text{ change in quantity of labor supplied}}{\%\text{ change in wage}} \tag{4.5.3}$$

The wage elasticity of labor supply for teenage workers is generally thought to be fairly elastic: that is, a certain percentage change in wages will lead to a larger percentage change in the quantity of hours worked. Conversely, the wage elasticity of labor supply for adult workers in their thirties and forties is thought to be fairly inelastic. When wages move up or down by a certain percentage amount, the quantity of hours that adults in their prime earning years are willing to supply changes but by a lesser percentage amount.

In markets for financial capital, the **elasticity of savings**—that is, the percentage change in the quantity of savings divided by the percentage change in interest rates—will describe the shape of the supply curve for financial capital. That is:

$$\text{Elasticity of savings} = \frac{\%\text{ change in quantity of financial savings}}{\%\text{ change in interest rate}} \tag{4.5.4}$$

Sometimes laws are proposed that seek to increase the quantity of savings by offering tax breaks so that the return on savings is higher. Such a policy will increase the quantity if the supply curve for financial capital is elastic, because then a given percentage increase in the return to savings will cause a higher percentage increase in the quantity of savings. However, if the supply curve for financial capital is highly inelastic, then a percentage increase in the return to savings will cause only a small increase in the quantity of savings. The evidence on the supply curve of financial capital is controversial but, at least in the short run, the elasticity of savings with respect to the interest rate appears fairly inelastic.

Expanding the Concept of Elasticity

The elasticity concept does not even need to relate to a typical supply or demand curve at all. For example, imagine that you are studying whether the Internal Revenue Service should spend more money on auditing tax returns. The question can be framed in terms of the elasticity of tax collections with respect to spending on tax enforcement; that is, what is the percentage change in tax collections derived from a percentage change in spending on tax enforcement?

With all of the elasticity concepts that have just been described, some of which are listed in Table $4.5.1$, the possibility of confusion arises. When you hear the phrases "elasticity of demand" or "elasticity of supply," they refer to the elasticity with respect to price. Sometimes, either to be extremely clear or because a wide variety of elasticities are being discussed, the elasticity of demand or the demand elasticity will be called the price elasticity of demand or the "elasticity of demand with respect to price." Similarly, elasticity of supply or the supply elasticity is sometimes called, to avoid any possibility of confusion, the price elasticity of supply or "the elasticity of supply with respect to price." But in whatever context elasticity is invoked, the idea always refers to percentage change in one variable, almost always a price or money variable, and how it causes a percentage change in another variable, typically a quantity variable of some kind.

Table $4.5.1$: Formulas for Calculating Elasticity

$$\text{Income elasticity of demand} = \frac{\%\text{ change in Qd}}{\%\text{ change in income}} \tag{4.5.5}$$
$$\text{Cross-price elasticity of demand} = \frac{\%\text{ change in Qd of good A}}{\%\text{ change in price of good B}} \tag{4.5.6}$$
$$\text{Wage elasticity of labor supply} = \frac{\%\text{ change in quantity of labor supplied}}{\%\text{ change in wage}} \tag{4.5.7}$$
$$\text{Wage elasticity of labor demand} = \frac{\%\text{ change in quantity of labor demanded}}{\%\text{ change in wage}} \tag{4.5.8}$$
$$\text{Interest rate elasticity of savings} = \frac{\%\text{ change in quantity of savings}}{\%\text{ change in interest rate}} \tag{4.5.9}$$

$$\text{Interest rate elasticity of borrowing} = \frac{\%\text{ change in quantity of borrowing}}{\%\text{ change in interest rate}} \tag{4.5.10}$$

Example 4.5.1: That Will Be How Much?

How did the 60% price increase in 2011 end up for Netflix? It has been a very bumpy ride.

Before the price increase, there were about 24.6 million U.S. subscribers. After the price increase, $810,000$ infuriated U.S. consumers canceled their Netflix subscriptions, dropping the total number of subscribers to 23.79 million. Fast forward to June 2013, when there were 36 million streaming Netflix subscribers in the United States. This was an increase of 11.4 million subscribers since the price increase—an average per quarter growth of about 1.6 million. This growth is less than the 2 million per quarter increases Netflix experienced in the fourth quarter of 2010 and the first quarter of 2011.

During the first year after the price increase, the firm's stock price (a measure of future expectations for the firm) fell from about $\$300$ per share to just under $\$54$. In 2015, however, the stock price is at $\$448$ per share. Today, Netflix has 57 million subscribers in fifty countries.

What happened? Obviously, Netflix company officials understood the law of demand. Company officials reported, when announcing the price increase, this could result in the loss of about $600,000$ existing subscribers. Using the elasticity of demand formula, it is easy to see company officials expected an inelastic response:

$$\begin{aligned}
&= \frac{-600,000/[(24million + 24.6million)/2]}{\$6/[(\$10+\$16)/2]} \\
&= \frac{-600,000/24.3million}{\$6/\$13} \\
&= \frac{-0.025}{0.46} \\
&= -0.05
\end{aligned}$$

In addition, Netflix officials had anticipated the price increase would have little impact on attracting new customers. Netflix anticipated adding up to 1.29 million new subscribers in the third quarter of 2011. It is true this was slower growth than the firm had experienced—about 2 million per quarter.

Why was the estimate of customers leaving so far off? In the 18 years since Netflix had been founded, there was an increase in the number of close, but not perfect, substitutes. Consumers now had choices ranging from Vudu, Amazon Prime, Hulu, and Redbox, to retail stores. Jaime Weinman reported in Maclean's that Redbox kiosks are "a five-minute drive for less from 68 percent of Americans, and it seems that many people still find a five-minute drive more convenient than loading up a movie online." It seems that in 2012, many consumers still preferred a physical DVD disk over streaming video.

What missteps did the Netflix management make? In addition to misjudging the elasticity of demand, by failing to account for close substitutes, it seems they may have also misjudged customers' preferences and tastes. Yet, as the population increases, the preference for streaming video may overtake physical DVD disks. Netflix, the source of numerous late night talk show laughs and jabs in 2011, may yet have the last laugh.

Summary

Elasticity is a general term, referring to percentage change of one variable divided by percentage change of a related variable that can be applied to many economic connections. For instance, the income elasticity of demand is the percentage change in quantity demanded divided by the percentage change in income. The cross-price elasticity of demand is the percentage change in the quantity demanded of a good divided by the percentage change in the price of another good. Elasticity applies in labor markets and financial capital markets just as it does in markets for goods and services. The wage elasticity of labor supply is the percentage change in the quantity of hours supplied divided by the percentage change in the wage. The elasticity of savings with respect to interest rates is the percentage change in the quantity of savings divided by the percentage change in interest rates.

References

1. Abkowitz, A. "How Netflix got started: Netflix founder and CEO Reed Hastings tells Fortune how he got the idea for the DVD-by-mail service that now has more than eight million customers." CNN Money. Last Modified January 28, 2009. archive.fortune.com/2009/01/2...tune/index.htm.
2. Associated Press (a). "Analyst: Coinstar gains from Netflix pricing moves." Boston Globe Media Partners, LLC. Accessed June 24, 2013. www.boston.com/business/artic...pricing_moves/.
3. Associated Press (b). "Netflix loses 800,000 US subscribers in tough 3Q." ABC Inc. Accessed June 24, 2013. abclocal.go.com/wpvi/story?se...ess&id=8403368
4. Baumgardner, James. 2014. "Presentation on Raising the Excise Tax on Cigarettes: Effects on Health and the Federal Budget." Congressional Budget Office. Accessed March 27, 2015. www.cbo.gov/sites/default/fil...esentation.pdf.
5. Funding Universe. 2015. "Netflix, Inc. History." Accessed March 11, 2015. www.fundinguniverse.com/compa...x-inc-history/.
6. Laporte, Nicole. "A tale of two Netflix." Fast Company 177 (July 2013) 31-32. Accessed December 3 2013. www.fastcompany-digital.com/f...708?pg=33#pg33
7. Liedtke, Michael, The Associated Press. "Investors bash Netflix stock after slower growth forecast - fee hikes expected to take toll on subscribers most likely to shun costly bundled Net, DVD service." The Seattle Times. Accessed June 24, 2013 from NewsBank on-line database (Access World News).
8. Netflix, Inc. 2013. "A Quick Update On Our Streaming Plans And Prices." Netflix (blog). Accessed March 11, 2015. http://blog.netflix.com/2014/05/a-qu...ing-plans.html.
9. Organization for Economic Co-Operation and Development (OECC). n.d. "Average annual hours actually worked per worker." Accessed March 11, 2015. stats.oecd.org/Index.aspx?DataSetCode=ANHRS.
10. Savitz, Eric. "Netflix Warns DVD Subs Eroding; Q4 View Weak; Losses Ahead; Shrs Plunge." Forbes.com, 2011. Accessed December 3, 2013. www.forbes.com/sites/ericsavi...ak-q4-outlook/.
11. Statistica.com. 2014. "Coffee Export Volumes Worldwide in November 2014, by Leading Countries (in 60-kilo sacks)." Accessed March 27, 2015. www.statista.com/statistics/2...ing-countries/.
12. Stone, Marcie. "Netflix responds to customers angry with price hike; Netflix stock falls 9%." News & Politics Examiner, 2011. Clarity Digital Group. Accessed June 24, 2013. www.examiner.com/article/netf...-stock-falls-9.
13. Weinman, J. (2012). Die hard, hardly dying. Maclean's, 125(18), 44.
14. The World Bank Group. 2015. "Gross Savings (% of GDP)." Accessed March 11, 2015. http://data.worldbank.org/indicator/NY.GNS.ICTR.ZS.
15. Yahoo Finance. Retrieved from finance.yahoo.com/q?s=NFLX

Glossary

cross-price elasticity of demand
: the percentage change in the quantity of good A that is demanded as a result of a percentage change in good B

elasticity of savings
: the percentage change in the quantity of savings divided by the percentage change in interest rates

wage elasticity of labor supply
: the percentage change in hours worked divided by the percentage change in wages

Contributor

- Template:ContribOpenSTAXMicroEcon

4.E: Elasticity (Exercises)

5.1: Price Elasticity of Demand and Price Elasticity of Supply

Self-Check Questions

Q1

From the data shown in Table below about demand for smart phones, calculate the price elasticity of demand from: point B to point C, point D to point E, and point G to point H. Classify the elasticity at each point as elastic, inelastic, or unit elastic.

Points	P	Q
A	60	3,000
B	70	2,800
C	80	2,600
D	90	2,400
E	100	2,200
F	110	2,000
G	120	1,800
H	130	1,600

Q2

From the data shown in Table below about supply of alarm clocks, calculate the price elasticity of supply from: point J to point K, point L to point M, and point N to point P. Classify the elasticity at each point as elastic, inelastic, or unit elastic.

Point	Price	Quantity Supplied
J	$8	50
K	$9	70
L	$10	80
M	$11	88
N	$12	95
P	$13	100

Review Questions

Q3

What is the formula for calculating elasticity?

Q4

What is the price elasticity of demand? Can you explain it in your own words?

Q5

What is the price elasticity of supply? Can you explain it in your own words?

Critical Thinking Questions

Q6

Transatlantic air travel in business class has an estimated elasticity of demand of 0.40 less than transatlantic air travel in economy class, with an estimated price elasticity of 0.62. Why do you think this is the case?

Q7

What is the relationship between price elasticity and position on the demand curve? For example, as you move up the demand curve to higher prices and lower quantities, what happens to the measured elasticity? How would you explain that?

Problems

Q8

The equation for a demand curve is $P = 48 - 3Q$. What is the elasticity in moving from a quantity of 5 to a quantity of 6?

Q9

The equation for a demand curve is $P = 2/Q$. What is the elasticity of demand as price falls from 5 to 4? What is the elasticity of demand as the price falls from 9 to 8? Would you expect these answers to be the same?

Q10

The equation for a supply curve is $4P = Q$. What is the elasticity of supply as price rises from 3 to 4? What is the elasticity of supply as the price rises from 7 to 8? Would you expect these answers to be the same?

Q11

The equation for a supply curve is $P = 3Q - 8$. What is the elasticity in moving from a price of 4 to a price of 7?

Solution

S1

From point B to point C, price rises from $\$70$ to $\$80$, and Q_d decreases from $2,800$ to $2,600$. So:

$$\begin{aligned}\% \text{ change in quantity } &= \frac{2600-2800}{(2600+2800) \div 2} \times 100 \\ &= \frac{-200}{2700} \times 100 \\ &= -7.41\end{aligned}$$

$$\begin{aligned}\% \text{ change in price } &= \frac{80-70}{(80+70) \div 2} \times 100 \\ &= \frac{10}{75} \times 100 \\ &= 13.33\end{aligned}$$

$$\begin{aligned}\text{Elasticity of Demand} &= \frac{-7.41\%}{13.33\%}\\ &= 0.56\end{aligned}$$

The demand curve is inelastic in this area; that is, its elasticity value is less than one.

Answer from Point D to point E:

$$\begin{aligned}\%\text{ change in quantity } &= \frac{2200-2400}{(2200+2400)\div 2}\times 100\\ &= \frac{-200}{2300}\times 100\\ &= -8.7\end{aligned}$$

$$\begin{aligned}\%\text{ change in price } &= \frac{100-90}{(100+90)\div 2}\times 100\\ &= \frac{10}{95}\times 100\\ &= 10.53\end{aligned}$$

$$\begin{aligned}\text{Elasticity of Demand} &= \frac{-8.7\%}{10.53\%}\\ &= 0.83\end{aligned}$$

The demand curve is inelastic in this area; that is, its elasticity value is less than one.

Answer from Point G to point H:

$$\begin{aligned}\%\text{ change in quantity } &= \frac{1600-1800}{1700}\times 100\\ &= \frac{-200}{1700}\times 100\\ &= -11.76\end{aligned}$$

$$\begin{aligned}\%\text{ change in price } &= \frac{130-120}{125}\times 100\\ &= \frac{10}{125}\times 100\\ &= 8.00\end{aligned}$$

$$\begin{aligned}\text{Elasticity of Demand} &= \frac{-11.76\%}{8.00\%}\\ &= -1.47\end{aligned}$$

The demand curve is elastic in this interval.

S2

From point J to point K, price rises from $8 to $9, and quantity rises from 50 to 70. So:

$$\begin{aligned}\%\text{ change in quantity } &= \frac{70-50}{(70+50)\div 2}\times 100\\ &= \frac{20}{60}\times 100\\ &= 33.33\end{aligned}$$

$$\begin{aligned}\%\text{ change in price } &= \frac{\$9-\$8}{(\$9+\$8)\div 2}\times 100\\ &= \frac{1}{8.5}\times 100\\ &= 11.76\end{aligned}$$

$$\begin{aligned}\text{Elasticity of Supply} &= \frac{33.33\%}{11.76\%} \\ &= 2.83\end{aligned}$$

The supply curve is elastic in this area; that is, its elasticity value is greater than one.

From point L to point M, the price rises from $\$10$ to $\$11$, while the Q_s rises from 80 to 88:

$$\begin{aligned}\%\text{ change in quantity } &= \frac{88-80}{(88+80)\div 2} \times 100 \\ &= \frac{8}{84} \times 100 \\ &= 9.52\end{aligned}$$

$$\begin{aligned}\%\text{ change in price } &= \frac{\$11-\$10}{(\$11+\$10)\div 2} \times 100 \\ &= \frac{1}{10.5} \times 100 \\ &= 9.52\end{aligned}$$

$$\begin{aligned}\text{Elasticity of Demand} &= \frac{9.52\%}{9.52\%} \\ &= 1.0\end{aligned}$$

The supply curve has unitary elasticity in this area.

From point N to point P, the price rises from $\$12$ to $\$13$, and Q_s rises from 95 to 100:

$$\begin{aligned}\%\text{ change in quantity } &= \frac{100-95}{(100+95)\div 2} \times 100 \\ &= \frac{5}{97.5} \times 100 \\ &= 5.13\end{aligned}$$

$$\begin{aligned}\%\text{ change in price } &= \frac{\$13-\$12}{(\$13+\$12)\div 2} \times 100 \\ &= \frac{1}{12.5} \times 100 \\ &= 8.0\end{aligned}$$

$$\begin{aligned}\text{Elasticity of Supply} &= \frac{5.13\%}{8.0\%} \\ &= 0.64\end{aligned}$$

The supply curve is inelastic in this region of the supply curve.

5.2: Polar Cases of Elasticity and Constant Elasticity

Self-Check Questions

Q1

Why is the demand curve with constant unitary elasticity concave?

Q2

Why is the supply curve with constant unitary elasticity a straight line?

Review Questions

Q3

Describe the general appearance of a demand or a supply curve with zero elasticity.

Q4

Describe the general appearance of a demand or a supply curve with infinite elasticity.

Critical Thinking Questions

Q5

Can you think of an industry (or product) with near infinite elasticity of supply in the short term? That is, what is an industry that could increase Qs almost without limit in response to an increase in the price?

Problems

Q6

The supply of paintings by Leonardo Da Vinci, who painted the Mona Lisa and The Last Supper and died in 1519, is highly inelastic. Sketch a supply and demand diagram, paying attention to the appropriate elasticities, to illustrate that demand for these paintings will determine the price.

Q7

Say that a certain stadium for professional football has $70,000$ seats. What is the shape of the supply curve for tickets to football games at that stadium? Explain.

Q8

When someone's kidneys fail, the person needs to have medical treatment with a dialysis machine (unless or until they receive a kidney transplant) or they will die. Sketch a supply and demand diagram, paying attention to the appropriate elasticities, to illustrate that the supply of such dialysis machines will primarily determine the price.

Solution

S1

The demand curve with constant unitary elasticity is concave because at high prices, a one percent decrease in price results in more than a one percent increase in quantity. As we move down the demand curve, price drops and the one percent decrease in price causes less than a one percent increase in quantity.

S2

The constant unitary elasticity is a straight line because the curve slopes upward and both price and quantity are increasing proportionally.

Self-Check Questions

Q1

The federal government decides to require that automobile manufacturers install new anti-pollution equipment that costs $\$2,000$ per car. Under what conditions can carmakers pass almost all of this cost along to car buyers? Under what conditions can carmakers pass very little of this cost along to car buyers?

Q2

Suppose you are in charge of sales at a pharmaceutical company, and your firm has a new drug that causes bald men to grow hair. Assume that the company wants to earn as much revenue as possible from this drug. If the elasticity of demand for your company's product at the current price is 1.4, would you advise the company to raise the price, lower the price, or to keep the price the same? What if the elasticity were 0.6? What if it were 1? Explain your answer.

Review Questions

Q3

If demand is elastic, will shifts in supply have a larger effect on equilibrium quantity or on price?

Q4

If demand is inelastic, will shifts in supply have a larger effect on equilibrium price or on quantity?

Q5

If supply is elastic, will shifts in demand have a larger effect on equilibrium quantity or on price?

Q6

If supply is inelastic, will shifts in demand have a larger effect on equilibrium price or on quantity?

Q7

Would you usually expect elasticity of demand or supply to be higher in the short run or in the long run? Why?

Q8

Under which circumstances does the tax burden fall entirely on consumers?

Critical Thinking Questions

Q9

Would you expect supply to play a more significant role in determining the price of a basic necessity like food or a luxury like perfume? Explain. *Hint:* Think about how the price elasticity of demand will differ between necessities and luxuries.

Q10

A city has built a bridge over a river and it decides to charge a toll to everyone who crosses. For one year, the city charges a variety of different tolls and records information on how many drivers cross the bridge. The city thus gathers information about elasticity of demand. If the city wishes to raise as much revenue as possible from the tolls, where will the city decide to charge a toll: in the inelastic portion of the demand curve, the elastic portion of the demand curve, or the unit elastic portion? Explain.

Q11

In a market where the supply curve is perfectly inelastic, how does an excise tax affect the price paid by consumers and the quantity bought and sold?

Problems

Q12

Assume that the supply of low-skilled workers is fairly elastic, but the employers' demand for such workers is fairly inelastic. If the policy goal is to expand employment for low-skilled workers, is it better to focus on policy tools to shift the supply of unskilled labor or on tools to shift the demand for unskilled labor? What if the policy goal is to raise wages for this group? Explain your answers with supply and demand diagrams.

Solution

S1

Carmakers can pass this cost along to consumers if the demand for these cars is inelastic. If the demand for these cars is elastic, then the manufacturer must pay for the equipment.

S2

If the elasticity is 1.4 at current prices, you would advise the company to lower its price on the product, since a decrease in price will be offset by the increase in the amount of the drug sold. If the elasticity were 0.6, then you would advise the company to increase its price. Increases in price will offset the decrease in number of units sold, but increase your total revenue. If elasticity is 1, the total revenue is already maximized, and you would advise that the company maintain its current price level.

5.4: Elasticity in Areas Other Than Price

Self-Check Questions

Q1

What would the gasoline price elasticity of supply mean to UPS or FedEx?

Q2

The average annual income rises from $\$25,000$ to $\$38,000$, and the quantity of bread consumed in a year by the average person falls from 30 loaves to 22 loaves. What is the income elasticity of bread consumption? Is bread a normal or an inferior good?

Q3

Suppose the cross-price elasticity of apples with respect to the price of oranges is 0.4, and the price of oranges falls by 3%. What will happen to the demand for apples?

Review Questions

Q4

What is the formula for the income elasticity of demand?

Q5

What is the formula for the cross-price elasticity of demand?

Q6

What is the formula for the wage elasticity of labor supply?

Q7

What is the formula for elasticity of savings with respect to interest rates?

Critical Thinking Questions

Q8

Normal goods are defined as having a positive income elasticity. We can divide normal goods into two types: Those whose income elasticity is less than one and those whose income elasticity is greater than one. Think about products that would fall into each category. Can you come up with a name for each category?

Q9

Suppose you could buy shoes one at a time, rather than in pairs. What do you predict the cross-price elasticity for left shoes and right shoes would be?

Solution

S1

The percentage change in quantity supplied as a result of a given percentage change in the price of gasoline.

S2

$$\begin{aligned}\text{Percentage change in quantity demanded} &= [(\text{change in quantity})/(\text{original quantity})] \times 100 \\ &= [22-30]/[(22+30)/2] \times 100 \\ &= -8/26 \times 100 \\ &= -30.77\end{aligned}$$

$$\begin{aligned}\text{Percentage change in income} &= [(\text{change in income})/(\text{original income})] \times 100 \\ &= [38,000-25,000]/[(38,000+25,000)/2] \times 100 \\ &= 13/31.5 \times 100 \\ &= 41.27\end{aligned}$$

In this example, bread is an inferior good because its consumption falls as income rises.

S3

The formula for cross-price elasticity is $\% \text{ change in Qd for apples} / \% \text{ change in P of oranges}$. Multiplying both sides by $\% \text{ change in P of oranges}$ yields:

$$ \tag{4.E.1} $$

$$= 0.4 \times (-3\%) = -1.2\% \tag{4.E.2}$$

or a 1.2% decrease in demand for apples.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

5: Consumer Choices

The analysis in this chapter will build on the three budget constraints introduced in the Choice in a World of Scarcity chapter. These were the consumption choice budget constraint, the labor-leisure budget constraint, and the intertemporal budget constraint. This chapter will also illustrate how economic theory provides a tool to systematically look at the full range of possible consumption choices to predict how consumption responds to changes in prices or incomes. After reading this chapter, consult the appendix **Indifference Curves** to learn more about representing utility and choice through indifference curves.

5.1: Prelude to Consumer Choices

Learning Objectives

- Consumption Choices
- How Changes in Income and Prices Affect Consumption Choices
- Labor-Leisure Choices
- Intertemporal Choices in Financial Capital Markets

Investment Choices

Figure 5.1.1: Higher education is generally viewed as a good investment, if one can afford it, regardless of the state of the economy. (Credit: modification of work by Jason Bache/Flickr Creative Commons)

"Eeny, Meeny, Miney, Moe"—Making Choices

The Great Recession of 2008–2009 touched families around the globe. In too many countries, workers found themselves out of a job. In developed countries, unemployment compensation provided a safety net, but families still saw a marked decrease in disposable income and had to make tough spending decisions. Of course, non-essential, discretionary spending was the first to go.

Even so, there was one particular category that saw a universal increase in spending world-wide during that time—an 18% uptick in the United States, specifically. You might guess that consumers began eating more meals at home, increasing spending at the grocery store. But the Bureau of Labor Statistics' Consumer Expenditure Survey, which tracks U.S. food spending over time, showed "real total food spending by U.S. households declined five percent between 2006 and 2009." So, it was not groceries. Just what product would people around the world demand more of during tough economic times, and more importantly, why? (Find out at chapter's end.)

That question leads us to this chapter's topic—analyzing how consumers make choices. For most consumers, using "eeny, meeny, miney, moe" is not how they make decisions; their decision-making processes have been educated far beyond a children's rhyme.

Microeconomics seeks to understand the behavior of individual economic agents such as individuals and businesses. Economists believe that individuals' decisions, such as what goods and services to buy, can be analyzed as choices made within certain budget constraints. Generally, consumers are trying to get the most for their limited budget. In economic terms they are trying to maximize total utility, or satisfaction, given their budget constraint.

Everyone has their own personal tastes and preferences. The French say: Chacun à son goût, or "*Each to his own taste.*" An old Latin saying states, De gustibus non est disputandum or "*There's no disputing about taste.*" If people's decisions are based on their own tastes and personal preferences, however, then how can economists hope to analyze the choices consumers make?

An economic explanation for why people make different choices begins with accepting the proverbial wisdom that tastes are a matter of personal preference. But economists also believe that the choices people make are influenced by their incomes, by the prices of goods and services they consume, and by factors like where they live. This chapter introduces the economic theory of how consumers make choices about what to buy, how much to work, and how much to save.

The analysis in this chapter will build on the three budget constraints introduced in the Choice in a World of Scarcity chapter. These were the consumption choice budget constraint, the labor-leisure budget constraint, and the intertemporal budget constraint. This chapter will also illustrate how economic theory provides a tool to systematically look at the full range of possible consumption choices to predict how consumption responds to changes in prices or incomes. After reading this chapter, consult the appendix **Indifference Curves** to learn more about representing utility and choice through indifference curves.

Contributor

- Template:ContribOpenSTAXMicroEcon

5.2: Consumption Choices

Learning Objectives

- Calculate total utility
- Propose decisions that maximize utility
- Explain marginal utility and the significance of diminishing marginal utility

Information on the **consumption** choices of Americans is available from the Consumer Expenditure Survey carried out by the U.S. Bureau of Labor Statistics. Table $5.2.1$ shows spending patterns for the average U.S. household. The first row shows income and, after taxes and personal savings are subtracted, it shows that, in 2015, the average U.S. household spent $\$48,109$ on consumption. The table then breaks down consumption into various categories. The average U.S. household spent roughly one-third of its consumption on shelter and other housing expenses, another one-third on food and vehicle expenses, and the rest on a variety of items, as shown. Of course, these patterns will vary for specific households by differing levels of family income, by geography, and by preferences.

Table $5.2.1$: U.S. Consumption Choices in 2015(Source: http://www.bls.gov/cex/csxann13.pdf)

Average Household Income before Taxes	**$62,481**
Average Annual Expenditures	**$48.109**
Food at home	$3,264
Food away from home	$2,505
Housing	$16,557
Apparel and services	$1,700
Transportation	$7,677
Healthcare	$3,157
Entertainment	$2,504
Education	$1,074
Personal insurance and pensions	$5,357
All else: alcohol, tobacco, reading, personal care, cash contributions, miscellaneous	$3,356

Total Utility and Diminishing Marginal Utility

To understand how a household will make its choices, economists look at what consumers can afford, as shown in a **budget constraint line**, and the **total utility** or satisfaction derived from those choices. In a budget constraint line, the quantity of one good is measured on the horizontal axis and the quantity of the other good is measured on the vertical axis. The budget constraint line shows the various combinations of two goods that are affordable given consumer income. Consider the situation of José, shown in Figure $5.2.1$. José likes to collect T-shirts and watch movies.

In Figure $5.2.1$, the quantity of T-shirts is shown on the horizontal axis, while the quantity of movies is shown on the vertical axis. If José had unlimited income or goods were free, then he could consume without limit. But José, like all of us, faces a **budget constraint**. José has a total of $\$56$ to spend. The price of T-shirts is $\$14$ and the price of movies is $\$7$. Notice that the vertical intercept of the budget constraint line is at eight movies and zero T-shirts ($\$56/\$7 = 8$). The horizontal intercept of the budget constraint is four, where José spends of all of his money on T-shirts and no movies ($\$56/14 = 4$). The slope of the budget constraint line is rise/run or $-8/4 = -2$. The specific choices along the budget constraint line show the combinations of T-shirts and movies that are affordable.

A Choice between Consumption Goods

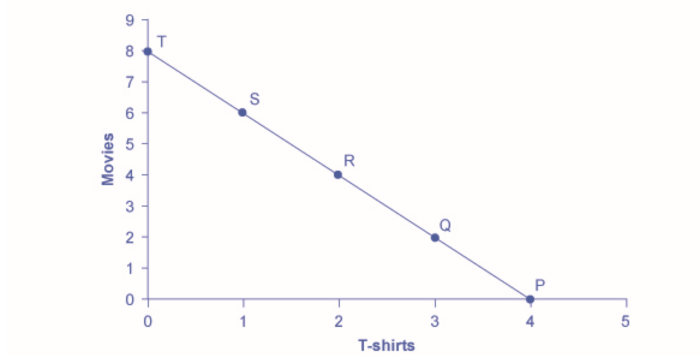


Figure 5.2.1: José has income of $56. Movies cost $7 and T-shirts cost $14. The points on the budget constraint line show the combinations of movies and T-shirts that are affordable.

José wishes to choose the combination that will provide him with the greatest **utility**, which is the term economists use to describe a person's level of satisfaction or happiness with his or her choices.

Let's begin with an assumption, which will be discussed in more detail later, that José can measure his own utility with something called utils. (It is important to note that you cannot make comparisons between the utils of individuals; if one person gets 20 utils from a cup of coffee and another gets 10 utils, this does not mean than the first person gets more enjoyment from the coffee than the other or that they enjoy the coffee twice as much.) Table 5.2.2 shows how José's utility is connected with his consumption of T-shirts or movies. The first column of the table shows the quantity of T-shirts consumed. The second column shows the total utility, or total amount of satisfaction, that José receives from consuming that number of T-shirts. The most common pattern of total utility, as shown here, is that consuming additional goods leads to greater total utility, but at a decreasing rate. The third column shows **marginal utility**, which is the additional utility provided by one additional unit of consumption. This equation for marginal utility is:

$$MU = \frac{\text{change in total utility}}{\text{change in quantity}} \tag{5.2.1}$$

Notice that marginal utility diminishes as additional units are consumed, which means that each subsequent unit of a good consumed provides less additional utility. For example, the first T-shirt José picks is his favorite and it gives him an addition of 22 utils. The fourth T-shirt is just to something to wear when all his other clothes are in the wash and yields only 18 additional utils. This is an example of the law of diminishing marginal utility, which holds that the additional utility decreases with each unit added.

The rest of Table 5.2.2 shows the quantity of movies that José attends, and his total and marginal utility from seeing each movie. Total utility follows the expected pattern: it increases as the number of movies seen rises. Marginal utility also follows the expected pattern: each additional movie brings a smaller gain in utility than the previous one. The first movie José attends is the one he wanted to see the most, and thus provides him with the highest level of utility or satisfaction. The fifth movie he attends is just to kill time. Notice that total utility is also the sum of the marginal utilities. Read the next Work It Out feature for instructions on how to calculate total utility.

Table 5.2.2: Total and Marginal Utility

T-Shirts (Quantity)	Total Utility	Marginal Utility	Movies (Quantity)	Total Utility	Marginal Utility
1	22	22	1	16	16
2	43	21	2	31	15
3	63	20	3	45	14
4	81	18	4	58	13
5	97	16	5	70	12
6	111	14	6	81	11
7	123	12	7	91	10
8	133	10	8	100	9

Table 5.2.3 looks at each point on the budget constraint in Figure 5.2.1, and adds up José's total utility for five possible combinations of T-

shirts and movies.

Table 5.2.3: Finding the Choice with the Highest Utility

Point	T-Shirts	Movies	Total Utility
P	4	0	81 + 0 = 81
Q	3	2	63 + 31 = 94
R	2	4	43 + 58 = 101
S	1	6	22 + 81 = 103
T	0	8	0 + 100 = 100

Example 5.2.1: Calculating Total Utility

Let's look at how José makes his decision in more detail.

- **Step 1:** Observe that, at point Q (for example), José consumes three T-shirts and two movies.
- **Step 2:** Look at Table 5.2.2. You can see from the fourth row/second column that three T-shirts are worth 63 utils. Similarly, the second row/fifth column shows that two movies are worth 31 utils.
- **Step 3:** From this information, you can calculate that point Q has a total utility of $94(63+31)$.
- **Step 4:** You can repeat the same calculations for each point on Table 5.2.3, in which the total utility numbers are shown in the last column.

For José, the highest total utility for all possible combinations of goods occurs at point S, with a total utility of 103 from consuming one T-shirt and six movies.

Choosing with Marginal Utility

Most people approach their utility-maximizing combination of choices in a step-by-step way. This step-by-step approach is based on looking at the tradeoffs, measured in terms of marginal utility, of consuming less of one good and more of another.

For example, say that José starts off thinking about spending all his money on T-shirts and choosing point P, which corresponds to four T-shirts and no movies, as illustrated in Figure 5.2.1. José chooses this starting point randomly; he has to start somewhere. Then he considers giving up the last T-shirt, the one that provides him the least marginal utility, and using the money he saves to buy two movies instead. Table 5.2.4 tracks the step-by-step series of decisions José needs to make (Key: T-shirts are \$14, movies are \$7, and income is \$56). The following Work It Out feature explains how marginal utility can effect decision making.

Table 5.2.4: A Step-by-Step Approach to Maximizing Utility

Try	Which Has	Total Utility	Marginal Gain and Loss of Utility, Compared with Previous Choice	Conclusion
Choice 1: P	4 T-shirts and 0 movies	81 from 4 T-shirts + 0 from 0 movies = 81	–	–
Choice 2: Q	3 T-shirts and 2 movies	63 from 3 T-shirts + 31 from 0 movies = 94	Loss of 18 from 1 less T-shirt, but gain of 31 from 2 more movies, for a net utility gain of 13	Q is preferred over P
Choice 3: R	2 T-shirts and 4 movies	43 from 2 T-shirts + 58 from 4 movies = 101	Loss of 20 from 1 less T-shirt, but gain of 27 from two more movies for a net utility gain of 7	R is preferred over Q
Choice 4: S	1 T-shirt and 6 movies	22 from 1 T-shirt + 81 from 6 movies = 103	Loss of 21 from 1 less T-shirt, but gain of 23 from two more movies, for a net utility gain of 2	S is preferred over R

Try	Which Has	Total Utility	Marginal Gain and Loss of Utility, Compared with Previous Choice	Conclusion
Choice 5: T	0 T-shirts and 8 movies	0 from 0 T-shirts + 100 from 8 movies = 100	Loss of 22 from 1 less T-shirt, but gain of 19 from two more movies, for a net utility loss of 3	S is preferred over T

Example 5.2.1: Decision Making by Comparing Marginal Utility

José could use the following thought process (if he thought in utils) to make his decision regarding how many T-shirts and movies to purchase:

- **Step 1:** From Table $5.2.2$, José can see that the marginal utility of the fourth T-shirt is 18. If José gives up the fourth T-shirt, then he loses 18 utils.
- **Step 2:** Giving up the fourth T-shirt, however, frees up $\$14$ (the price of a T-shirt), allowing José to buy the first two movies (at $\$7$ each).
- **Step 3:** José knows that the marginal utility of the first movie is 16 and the marginal utility of the second movie is 15. Thus, if José moves from point P to point Q, he gives up 18 utils (from the T-shirt), but gains 31 utils (from the movies).
- **Step 4:** Gaining 31 utils and losing 18 utils is a net gain of 13. This is just another way of saying that the total utility at Q (94 according to the last column in Table $5.2.3$) is 13 more than the total utility at P (81).
- **Step 5:** So, for José, it makes sense to give up the fourth T-shirt in order to buy two movies.

José clearly prefers point Q to point P. Now repeat this step-by-step process of decision making with marginal utilities. José thinks about giving up the third T-shirt and surrendering a marginal utility of 20, in exchange for purchasing two more movies that promise a combined marginal utility of 27. José prefers point R to point Q. What if José thinks about going beyond R to point S? Giving up the second T-shirt means a marginal utility loss of 21, and the marginal utility gain from the fifth and sixth movies would combine to make a marginal utility gain of 23, so José prefers point S to R.

However, if José seeks to go beyond point S to point T, he finds that the loss of marginal utility from giving up the first T-shirt is 22, while the marginal utility gain from the last two movies is only a total of 19. If José were to choose point T, his utility would fall to 100. Through these stages of thinking about marginal tradeoffs, José again concludes that S, with one T-shirt and six movies, is the choice that will provide him with the highest level of total utility. This step-by-step approach will reach the same conclusion regardless of José's starting point.

Another way to look at this is by focusing on satisfaction per dollar. **Marginal utility per dollar** is the amount of additional utility José receives given the price of the product. For José's T-shirts and movies, the marginal utility per dollar is shown in Table $5.2.5$.

$$\text{marginal utility per dollar} = \frac{\text{marginal utility}}{\text{price}} \tag{5.2.2}$$

José's first purchase will be a movie. Why? Because it gives him the highest marginal utility per dollar and it is affordable. José will continue to purchase the good which gives him the highest marginal utility per dollar until he exhausts the budget. José will keep purchasing movies because they give him a greater "bang or the buck" until the sixth movie is equivalent to a T-shirt purchase. José can afford to purchase that T-shirt. So José will choose to purchase six movies and one T-shirt.

Table $5.2.5$: Marginal Utility per Dollar

Quantity of T-Shirts	Total Utility	Marginal Utility	Marginal Utility per Dollar	Quantity of Movies	Total Utility	Marginal Utility	Marginal Utility per Dollar
1	22	22	22/$14=1.6	1	16	16	16/$7=2.3
2	43	21	21/$14=1.5	2	31	15	15/$7=2.14
3	63	20	20/$14=1.4	3	45	14	14/$7=2
4	81	18	18/$14=1.3	4	58	13	13/$7=1.9
5	97	16	16/$14=1.1	5	70	12	12/$7=1.7

Quantity of T-Shirts	Total Utility	Marginal Utility	Marginal Utility per Dollar	Quantity of Movies	Total Utility	Marginal Utility	Marginal Utility per Dollar
6	111	14	14/$14=1	6	81	11	11/$7=1.6
7	123	12	12/$14=1.2	7	91	10	10/$7=1.4

A Rule for Maximizing Utility

This process of decision making suggests a rule to follow when **maximizing utility**. Since the price of T-shirts is twice as high as the price of movies, to maximize utility the last T-shirt chosen needs to provide exactly twice the marginal utility (MU) of the last movie. If the last T-shirt provides less than twice the marginal utility of the last movie, then the T-shirt is providing less "bang for the buck" (i.e., marginal utility per dollar spent) than if the same money were spent on movies. If this is so, José should trade the T-shirt for more movies to increase his total utility. Marginal utility per dollar measures the additional utility that José will enjoy given what he has to pay for the good.

If the last T-shirt provides more than twice the marginal utility of the last movie, then the T-shirt is providing more "bang for the buck" or marginal utility per dollar, than if the money were spent on movies. As a result, José should buy more T-shirts. Notice that at José's optimal choice of point S, the marginal utility from the first T-shirt, of 22 is exactly twice the marginal utility of the sixth movie, which is 11. At this choice, the marginal utility per dollar is the same for both goods. This is a tell-tale signal that José has found the point with highest total utility.

This argument can be written as a general rule: the utility-maximizing choice between consumption goods occurs where the marginal utility per dollar is the same for both goods.

$$\frac{MU_1}{P_1} = \frac{MU_2}{P_2} \tag{5.2.3}$$

A sensible economizer will pay twice as much for something only if, in the marginal comparison, the item confers twice as much utility. Notice that the formula for the table above is:

$$\frac{M22}{\$14} = \frac{11}{\$7} \tag{5.2.4}$$

$$1.6 = 1.6 \tag{5.2.5}$$

The following Example provides step by step guidance for this concept of utility-maximizing choices.

✓ Example 5.2.1: Maximizing Utility

The general rule, $\frac{MU_1}{P_1} = \frac{MU_2}{P_2}$, means that the last dollar spent on each good provides exactly the same marginal utility. So:

- **Step 1:** If we traded a dollar more of movies for a dollar more of T-shirts, the marginal utility gained from T-shirts would exactly offset the marginal utility lost from fewer movies. In other words, the net gain would be zero.
- **Step 2:** Products, however, usually cost more than a dollar, so we cannot trade a dollar's worth of movies. The best we can do is trade two movies for another T-shirt, since in this example T-shirts cost twice what a movie does.
- **Step 3:** If we trade two movies for one T-shirt, we would end up at point R (two T-shirts and four movies).
- **Step 4:** Choice 4 in Table $5.2.4$ shows that if we move to point S, we would lose 21 utils from one less T-shirt, but gain 23 utils from two more movies, so we would end up with more total utility at point S.

In short, the general rule shows us the utility-maximizing choice.

There is another, equivalent way to think about this. The general rule can also be expressed as the ratio of the prices of the two goods should be equal to the ratio of the marginal utilities. When the price of good 1 is divided by the price of good 2, at the utility-maximizing point this will equal the marginal utility of good 1 divided by the marginal utility of good 2. This rule, known as the **consumer equilibrium**, can be written in algebraic form:

$$\frac{P_1}{P_2} = \frac{MU_1}{MU_2} \tag{5.2.6}$$

Along the budget constraint, the total price of the two goods remains the same, so the ratio of the prices does not change. However, the marginal utility of the two goods changes with the quantities consumed. At the optimal choice of one T-shirt and six movies, point S, the ratio of marginal utility to price for T-shirts $(22 : 14)$ matches the ratio of marginal utility to price for movies (of $11 : 7$).

Measuring Utility with Numbers

This discussion of utility started off with an assumption that it is possible to place numerical values on utility, an assumption that may seem questionable. You can buy a thermometer for measuring temperature at the hardware store, but what store sells an "utilimometer" for measuring utility? However, while measuring utility with numbers is a convenient assumption to clarify the explanation, the key assumption is not that utility can be measured by an outside party, but only that individuals can decide which of two alternatives they prefer.

To understand this point, think back to the step-by-step process of finding the choice with highest total utility by comparing the marginal utility that is gained and lost from different choices along the budget constraint. As José compares each choice along his budget constraint to the previous choice, what matters is not the specific numbers that he places on his utility—or whether he uses any numbers at all—but only that he personally can identify which choices he prefers.

In this way, the step-by-step process of choosing the highest level of utility resembles rather closely how many people make consumption decisions. We think about what will make us the happiest; we think about what things cost; we think about buying a little more of one item and giving up a little of something else; we choose what provides us with the greatest level of satisfaction. The vocabulary of comparing the points along a budget constraint and total and marginal utility is just a set of tools for discussing this everyday process in a clear and specific manner. It is welcome news that specific utility numbers are not central to the argument, since a good utilimometer is hard to find. Do not worry—while we cannot measure utils, by the end of the next module, we will have transformed our analysis into something we can measure—demand.

Key Concepts and Summary

Economic analysis of household behavior is based on the assumption that people seek the highest level of utility or satisfaction. Individuals are the only judge of their own utility. In general, greater consumption of a good brings higher total utility. However, the additional utility received from each unit of greater consumption tends to decline in a pattern of diminishing marginal utility.

The utility-maximizing choice on a consumption budget constraint can be found in several ways. You can add up total utility of each choice on the budget line and choose the highest total. You can choose a starting point at random and compare the marginal utility gains and losses of moving to neighboring points—and thus eventually seek out the preferred choice. Alternatively, you can compare the ratio of the marginal utility to price of good 1 with the marginal utility to price of good 2 and apply the rule that at the optimal choice, the two ratios should be equal:

$$\frac{MU_1}{P_1} = \frac{MU_2}{P_2} \tag{5.2.7}$$

References

U.S. Bureau of Labor Statistics. 2015. "Consumer Expenditures in 2013." Accessed March 11, 2015. http://www.bls.gov/cex/csxann13.pdf.

U.S. Bureau of Labor Statistics. 2015. "Employer Costs for Employee Compensation—December 2014." Accessed March 11, 2015. http://www.bls.gov/news.release/pdf/ecec.pdf.

U.S. Bureau of Labor Statistics. 2015. "Labor Force Statistics from the Current Population Survey." Accessed March 11, 2015. http://www.bls.gov/cps/cpsaat22.htm.

Glossary

budget constraint line

shows the possible combinations of two goods that are affordable given a consumer's limited income

consumer equilibrium

when the ratio of the prices of goods is equal to the ratio of the marginal utilities (point at which the consumer can get the most satisfaction)

diminishing marginal utility

the common pattern that each marginal unit of a good consumed provides less of an addition to utility than the previous unit

marginal utility

the additional utility provided by one additional unit of consumption

marginal utility per dollar

the additional satisfaction gained from purchasing a good given the price of the product; MU/Price

total utility

satisfaction derived from consumer choices

Contributor

- Template:ContribOpenSTAXMicroEcon

5.3: How Changes in Income and Prices Affect Consumption Choices

Learning Objectives

- Explain how income, prices, and preferences affect consumer choices
- Contrast the substitution effect and the income effect
- Utilize concepts of demand to analyze consumer choices
- Apply utility-maximizing choices to governments and businesses

Just as utility and marginal utility can be used to discuss making consumer choices along a budget constraint, these ideas can also be used to think about how consumer choices change when the budget constraint shifts in response to changes in income or price. Indeed, because the budget constraint framework can be used to analyze how quantities demanded change because of price movements, the budget constraint model can illustrate the underlying logic behind demand curves.

How Changes in Income Affect Consumer Choices

Let's begin with a concrete example illustrating how changes in income level affect consumer choices. Figure $5.3.1$ shows a budget constraint that represents Kimberly's choice between concert tickets at $\$50$ each and getting away overnight to a bed-and-breakfast for $\$200$ per night. Kimberly has $\$1,000$ per year to spend between these two choices. After thinking about her total utility and marginal utility and applying the decision rule that the ratio of the marginal utilities to the prices should be equal between the two products, Kimberly chooses point M, with eight concerts and three overnight getaways as her utility-maximizing choice.

How a Change in Income Affects Consumption Choices

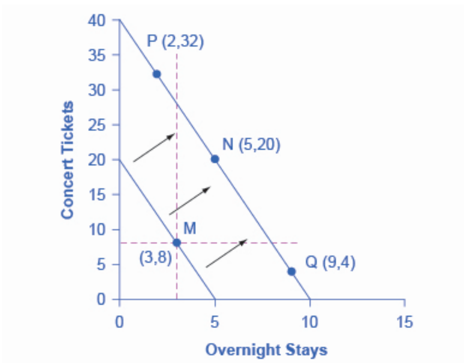


Figure $5.3.1$: The utility-maximizing choice on the original budget constraint is M. The dashed horizontal and vertical lines extending through point M allow you to see at a glance whether the quantity consumed of goods on the new budget constraint is higher or lower than on the original budget constraint. On the new budget constraint, a choice like N will be made if both goods are normal goods. If overnight stays is an inferior good, a choice like P will be made. If concert tickets are an inferior good, a choice like Q will be made.

Now, assume that the income Kimberly has to spend on these two items rises to $\$2,000$ per year, causing her budget constraint to shift out to the right. How does this rise in income alter her utility-maximizing choice? Kimberly will again consider the utility and marginal utility that she receives from concert tickets and overnight getaways and seek her **utility-maximizing choice** on the new budget line. But how will her new choice relate to her original choice?

The possible choices along the new budget constraint can be divided into three groups, which are divided up by the dashed horizontal and vertical lines that pass through the original choice M in the figure. All choices on the upper left of the new budget constraint that are to the left of the vertical dashed line, like choice P with two overnight stays and 32 concert tickets, involve less of the good on the horizontal axis but much more of the good on the vertical axis. All choices to the right of the vertical dashed line and above the horizontal dashed line—like choice N with five overnight getaways and 20 concert tickets—have more consumption of both goods. Finally, all choices that are to the right of the vertical dashed line but below the horizontal dashed line, like choice Q with four concerts and nine overnight getaways, involve less of the good on the vertical axis but much more of the good on the horizontal axis.

All of these choices are theoretically possible, depending on Kimberly's personal preferences as expressed through the total and marginal utility she would receive from consuming these two goods. When income rises, the most common reaction is to purchase more of both goods, like choice N, which is to the upper right relative to Kimberly's original choice M, although exactly how much more of each good will vary according to personal taste. Conversely, when income falls, the most typical reaction is to purchase less of both goods. As defined in the chapter on Demand and Supply and again in the chapter on Elasticity, goods and services are called normal goods when a rise in income leads to a rise in the quantity consumed of that good and a fall in income leads to a fall in quantity consumed.

However, depending on Kimberly's preferences, a rise in income could cause consumption of one good to increase while consumption of the other good declines. A choice like P means that a rise in income caused her quantity consumed of overnight stays to decline, while a choice like Q would mean that a rise in income caused her quantity of concerts to decline. Goods where demand declines as income rises (or conversely, where the demand rises as income falls) are called "inferior goods." An inferior good occurs when people trim back on a good as income rises, because they can now afford the more expensive choices that they prefer. For example, a higher-income household might eat fewer hamburgers or be less likely to buy a used car, and instead eat more steak and buy a new car.

How Price Changes Affect Consumer Choices

For analyzing the possible effect of a change in price on consumption, let's again use a concrete example. Figure $5.3.2$ represents the consumer choice of Sergei, who chooses between purchasing baseball bats and cameras. A price increase for baseball bats would have no effect on the ability to purchase cameras, but it would reduce the number of bats Sergei could afford to buy. Thus a price increase for baseball bats, the good on the horizontal axis, causes the budget constraint to rotate inward, as if on a hinge, from the vertical axis. As in the previous section, the point labeled M represents the originally preferred point on the original budget constraint, which Sergei has chosen after contemplating his total utility and marginal utility and the tradeoffs involved along the budget constraint. In this example, the units along the horizontal and vertical axes are not numbered, so the discussion must focus on whether more or less of certain goods will be consumed, not on numerical amounts.

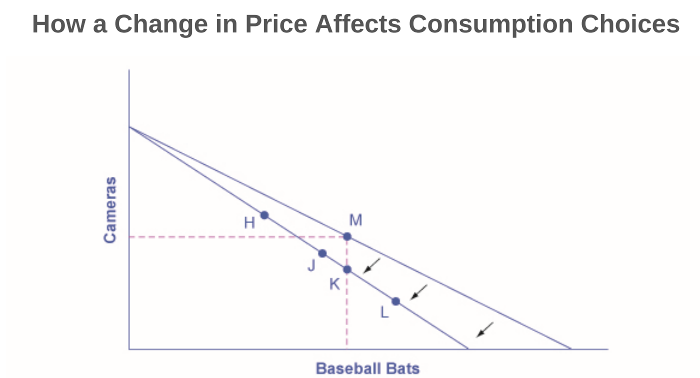


Figure $5.3.2$: The original utility-maximizing choice is M. When the price rises, the budget constraint shifts in to the left. The dashed lines make it possible to see at a glance whether the new consumption choice involves less of both goods, or less of one good and more of the other. The new possible choices would be fewer baseball bats and more cameras, like point H, or less of both goods, as at point J. Choice K would mean that the higher price of bats led to exactly the same quantity of bats being consumed, but fewer cameras. Choices like L are ruled out as theoretically possible but highly unlikely in the real world, because they would mean that a higher price for baseball bats means a greater quantity consumed of baseball bats.

After the price increase, Sergei will make a choice along the new budget constraint. Again, his choices can be divided into three segments by the dashed vertical and horizontal lines. In the upper left portion of the new budget constraint, at a choice like H, Sergei consumes more cameras and fewer bats. In the central portion of the new budget constraint, at a choice like J, he consumes less of both goods. At the right-hand end, at a choice like L, he consumes more bats but fewer cameras.

The typical response to higher prices is that a person chooses to consume less of the product with the higher price. This occurs for two reasons, and both effects can occur simultaneously. The **substitution effect** occurs when a price changes and consumers have an incentive to consume less of the good with a relatively higher price and more of the good with a relatively lower price. The **income effect** is that a higher price means, in effect, the buying power of income has been reduced (even though actual income has not changed), which leads to buying less of the good (when the good is normal). In this example, the higher price for baseball bats would cause Sergei to buy a fewer bats for both reasons. Exactly how much will a higher price for bats cause Sergei consumption of bats to fall? Figure $5.3.2$ suggests a range of possibilities. Sergei might react to a higher price for baseball bats by purchasing the same quantity of bats, but cutting his consumption of cameras. This choice is the point K on the new budget constraint, straight below the original choice M. Alternatively, Sergei might react by dramatically reducing his purchases of bats and instead buy more cameras.

The key is that it would be imprudent to assume that a change in the price of baseball bats will only or primarily affect the good whose price is changed, while the quantity consumed of other goods remains the same. Since Sergei purchases all his products out of the same budget, a change in the price of one good can also have a range of effects, either positive or negative, on the quantity consumed of other goods.

In short, a higher price typically causes reduced consumption of the good in question, but it can affect the consumption of other goods as well.

The Foundations of Demand Curves

Changes in the price of a good lead the budget constraint to shift. A shift in the budget constraint means that when individuals are seeking their highest utility, the quantity that is demanded of that good will change. In this way, the logical foundations of demand curves—which show a connection between prices and quantity demanded—are based on the underlying idea of individuals seeking utility. Figure $5.3.3$ (a) shows a budget constraint with a choice between housing and "everything else." (Putting "everything else" on the vertical axis can be a useful approach in some cases, especially when the focus of the analysis is on one particular good.) The preferred choice on the original budget constraint that provides the highest possible utility is labeled M_0. The other three budget constraints represent successively higher prices for housing of P_1, P_2, and P_3. As the budget constraint rotates in, and in, and in again, the utility-maximizing choices are labeled M_1, M_2, and M_3, and the quantity demanded of housing falls from Q_0 to Q_1 to Q_2 to Q_3.

The Foundations of a Demand Curve: An Example of Housing

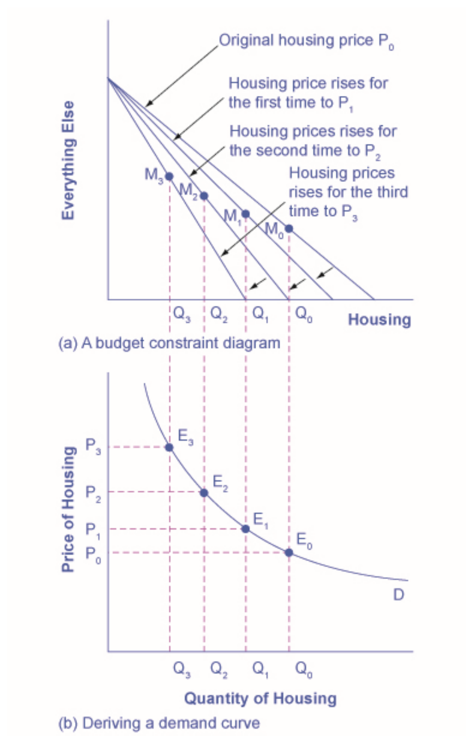


Figure $5.3.3$: (a) As the price increases from P_0 to P_1 to P_2 to P_3, the budget constraint on the upper part of the diagram shifts to the left. The utility-maximizing choice changes from M_0 to M_1 to M_2 to M_3. As a result, the quantity demanded of housing shifts from Q_0 to Q_1 to Q_2 to Q_3, ceteris paribus. (b) The demand curve graphs each combination of the price of housing and the quantity of housing demanded, ceteris paribus. Indeed, the quantities of housing are the same at the points on both (a) and (b). Thus, the original price of housing (P_0) and the original quantity of housing (Q_0) appear on the demand curve as point E_0. The higher price of housing (P_1) and the corresponding lower quantity demanded of housing (Q_1) appear on the demand curve as point E_1.

So, as the price of housing rises, the budget constraint shifts to the left, and the quantity consumed of housing falls, ceteris paribus (meaning, with all other things being the same). This relationship—the price of housing rising from P_0 to P_1 to P_2 to P_3, while the quantity of housing demanded falls from Q_0 to Q_1 to Q_2 to Q_3—is graphed on the demand curve in Figure $5.3.3$ (b). Indeed, the vertical dashed lines stretching between the top and bottom of Figure 5.3.3 show that the quantity of housing demanded at each point is the same in both (a) and (b). The shape of a demand curve is ultimately determined by the underlying choices about maximizing utility subject to a budget constraint. And while economists may not be able to measure “utils,” they can certainly measure price and quantity demanded.

Applications in Government and Business

The budget constraint framework for making utility-maximizing choices offers a reminder that people can react to a change in price or income in a range of different ways. For example, in the winter months of 2005, costs for heating homes increased significantly in many parts of the country as prices for natural gas and electricity soared, due in large part to the disruption caused by Hurricanes Katrina and Rita. Some people reacted by reducing the quantity demanded of energy; for example, by turning down the thermostats in their homes by a few degrees and wearing a heavier sweater inside. Even so, many home heating bills rose, so people adjusted their consumption in other ways, too. As you learned in the chapter on Elasticity, the short run demand for home heating is generally inelastic. Each household cut back on what it valued least on the margin; for some it might have been some dinners out, or a vacation, or postponing buying a new refrigerator or a new car. Indeed, sharply higher energy prices can have effects beyond the energy market, leading to a widespread reduction in purchasing throughout the rest of the economy.

A similar issue arises when the government imposes taxes on certain products, like it does on gasoline, cigarettes, and alcohol. Say that a tax on alcohol leads to a higher price at the liquor store, the higher price of alcohol causes the budget constraint to pivot left, and consumption of alcoholic beverages is likely to decrease. However, people may also react to the higher price of alcoholic beverages by cutting back on other purchases. For example, they might cut back on snacks at restaurants like chicken wings and nachos. It would be unwise to assume that the liquor industry is the only one affected by the tax on alcoholic beverages. Read the next Clear It Up to learn about how buying decisions are influenced by who controls the household income.

Does who controls household income make a difference?

In the mid-1970s, the United Kingdom made an interesting policy change in its “child allowance” policy. This program provides a fixed amount of money per child to every family, regardless of family income. Traditionally, the child allowance had been distributed to families by withholding less in taxes from the paycheck of the family wage earner—typically the father in this time period. The new policy instead provided the child allowance as a cash payment to the mother. As a result of this change, households have the same level of income and face the same prices in the market, but the money is more likely to be in the purse of the mother than in the wallet of the father.

Should this change in policy alter household consumption patterns? Basic models of consumption decisions, of the sort examined in this chapter, assume that it does not matter whether the mother or the father receives the money, because both parents seek to maximize the utility of the family as a whole. In effect, this model assumes that everyone in the family has the same preferences.

In reality, the share of income controlled by the father or the mother does affect what the household consumes. When the mother controls a larger share of family income a number of studies, in the United Kingdom and in a wide variety of other countries, have found that the family tends to spend more on restaurant meals, child care, and women’s clothing, and less on alcohol and tobacco. As the mother controls a larger share of household resources, children’s health improves, too. These findings suggest that when providing assistance to poor families, in high-income countries and low-income countries alike, the monetary amount of assistance is not all that matters: it also matters which member of the family actually receives the money.

The budget constraint framework serves as a constant reminder to think about the full range of effects that can arise from changes in income or price, not just effects on the one product that might seem most immediately affected.

Key Concepts and Summary

The budget constraint framework suggest that when income or price changes, a range of responses are possible. When income rises, households will demand a higher quantity of normal goods, but a lower quantity of inferior goods. When the price of a good rises, households will typically demand less of that good—but whether they will demand a much lower quantity or only a slightly lower quantity will depend on personal preferences. Also, a higher price for one good can lead to more or less of the other good being demanded.

Glossary

income effect

a higher price means that, in effect, the buying power of income has been reduced, even though actual income has not changed; always happens simultaneously with a substitution effect

substitution effect

when a price changes, consumers have an incentive to consume less of the good with a relatively higher price and more of the good with a relatively lower price; always happens simultaneously with an income effect

Contributor

- Template:ContribOpenSTAXMicroEcon

5.4: Labor-Leisure Choices

Learning Objectives

- Interpret labor-leisure budget constraint graphs
- Predict consumer choices based on wages and other compensation
- Explain the backward-bending supply curve of labor

People do not obtain utility just from products they purchase. They also obtain utility from leisure time. Leisure time is time not spent at work. The decision-making process of a utility-maximizing household applies to what quantity of hours to work in much the same way that it applies to purchases of goods and services. Choices made along the **labor-leisure budget constraint**, as wages shift, provide the logical underpinning for the labor supply curve. The discussion also offers some insights about the range of possible reactions when people receive higher wages, and specifically about the claim that if people are paid higher wages, they will work a greater quantity of hours—assuming that they have a say in the matter.

According to the Bureau of Labor Statistics, U.S. workers averaged 38.6 hours per week on the job in 2014. This average includes part-time workers; for full-time workers only, the average was 42.5 hours per week. Table 5.4.1 shows that more than half of all workers are on the job 35 to 48 hours per week, but significant proportions work more or less than this amount.

Table 5.4.2 breaks down the average hourly compensation received by private industry workers, including wages and benefits. Wages and salaries are about three-quarters of total compensation received by workers; the rest is in the form of health insurance, vacation pay, and other benefits. The compensation workers receive differs for many reasons, including experience, education, skill, talent, membership in a labor union, and the presence of discrimination against certain groups in the labor market. Issues surrounding the inequality of incomes in a market-oriented economy are explored in the chapters on Poverty and Economic Inequality and Issues in Labor Markets: Unions, Discrimination, Immigration.

Table 5.4.1: Persons at Work, by Average Hours Worked per Week in 2013 (Total number of workers: 137.7 million)(Source: http://www.bls.gov/news.release/empsit.t18.htm)

Hours Worked per Week	Number of Workers	Percentage of Workforce
1–14 hours	6.9 million	5.0%
15–34 hours	27.6 million	20.1%
35–40 hours	68.5 million	49.9%
41–48 hours	11.9 million	8.6%
49–59 hours	13.3 million	9.6%
60 hours and over	9.3 million	6.8%

Table 5.4.2: Hourly Compensation: Wages, Benefits, and Taxes in 2014(Source: http://www.bls.gov/news.release/pdf/ecec.pdf)

Compensation, Wage, Salary, and Benefits	**$30.92 per hour**
Wages and Salaries	**$20.92**
Benefits	
Vacation	$2.09
Supplemental Pay	$0.84
Insurance	$2.15
Health Benefits	$2.36
Retirement and Savings	$1.24
Defined Benefit	$0.57
Defined Contribution	$0.064
Legally Required	$2.46

The Labor-Leisure Budget Constraint

How do workers make decisions about the number of hours to work? Again, let's proceed with a concrete example. The economic logic is precisely the same as in the case of a **consumption choice budget constraint**, but the labels are different on a labor-leisure budget constraint.

Vivian has 70 hours per week that she could devote either to work or to leisure, and her wage is $\$10$/hour. The lower budget constraint in Figure $5.4.1$ shows Vivian's possible choices. The horizontal axis of this diagram measures both leisure and labor, by showing how Vivian's time is divided between leisure and labor. Hours of leisure are measured from left to right on the horizontal axis, while hours of labor are measured from right to left. Vivian will compare choices along this budget constraint, ranging from 70 hours of leisure and no income at point S to zero hours of leisure and $\$700$ of income at point L. She will choose the point that provides her with the highest total utility. For this example, let's assume that Vivian's utility-maximizing choice occurs at O, with 30 hours of leisure, 40 hours of work, and $\$400$ in weekly income.

How a Rise in Wages Alters the Utility-Maximizing Choice

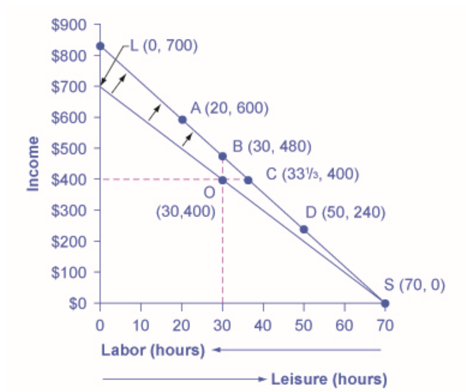


Figure $5.4.1$: Vivian's original choice is point O on the lower opportunity set. A rise in her wage causes her opportunity set to swing upward. In response to the increase in wages, Vivian can make a range of different choices available to her: a choice like D, which involves less work; and a choice like B, which involves the same amount of work but more income; or a choice like A, which involves more work and considerably more income. Vivian's personal preferences will determine which choice she makes.

For Vivian to discover the labor-leisure choice that will maximize her utility, she does not have to place numerical values on the total and marginal utility that she would receive from every level of income and leisure. All that really matters is that Vivian can compare, in her own mind, whether she would prefer more leisure or more income, given the tradeoffs she faces. If Vivian can say to herself: "*I'd really rather work a little less and have more leisure, even if it means less income,*" or "*I'd be willing to work more hours to make some extra income,*" then as she gradually moves in the direction of her preferences, she will seek out the utility-maximizing choice on her labor-leisure budget constraint.

Now imagine that Vivian's wage level increases to $\$12$/hour. A higher wage will mean a new budget constraint that tilts up more steeply; conversely, a lower wage would have led to a new budget constraint that was flatter. How will a change in the wage and the corresponding shift in the budget constraint affect Vivian's decisions about how many hours to work?

Vivian's choices of quantity of hours to work and income along her new budget constraint can be divided into several categories, using the dashed horizontal and vertical lines in Figure $5.4.1$ that go through her original choice (O). One set of choices in the upper-left portion of the new budget constraint involves more hours of work (that is, less leisure) and more income, at a point like A with 20 hours of leisure, 50 hours of work, and $\$600$ of income (that is, 50 hours of work multiplied by the new wage of $\$12$ per hour). A second choice would be to work exactly the same 40 hours, and to take the benefits of the higher wage in the form of income that would now be $\$480$, at choice B. A third choice would involve more leisure and the same income at point C (that is, $33-1/3$ hours of work multiplied by the new wage of $\$12$ per hour equals $\$400$ of total income). A fourth choice would involve less income and much more leisure at a point like D, with a choice like 50 hours of leisure, 20 hours of work, and $\$240$ in income.

In effect, Vivian can choose whether to receive the benefits of her wage increase in the form of more income, or more leisure, or some mixture of these two. With this range of possibilities, it would be unwise to assume that Vivian (or anyone else) will necessarily react to a wage increase by working substantially more hours. Maybe they will; maybe they will not.

Applications of Utility Maximizing with the Labor-Leisure Budget Constraint

The theoretical insight that higher wages will sometimes cause an increase in hours worked, sometimes cause hours worked not to change by much, and sometimes cause hours worked to decline, has led to labor supply curves that look like the one in Figure $5.4.2$. The bottom-left portion of the labor supply curve slopes upward, which reflects the situation of a person who reacts to a higher wage by supplying a greater quantity of labor. The middle, close-to-vertical portion of the labor supply curve reflects the situation of a person who reacts to a higher wage by supplying about the same quantity of labor. The very top portion of the labor supply curve is called a **backward-bending supply curve** for labor, which is the situation of high-wage people who can earn so much that they respond to a still-higher wage by working fewer hours. Read the following Clear It Up feature for more on the number of hours the average person works each year.

A Backward-Bending Supply Curve of Labor

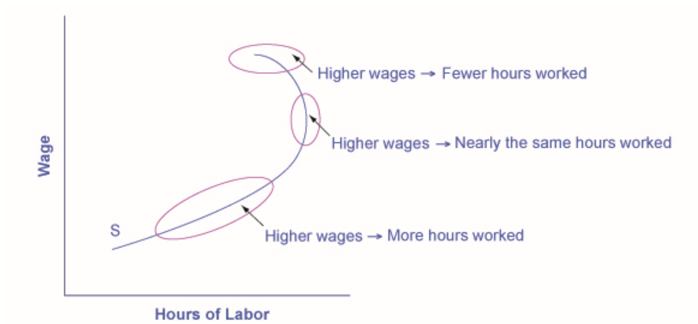


Figure $5.4.2$: The bottom upward-sloping portion of the labor supply curve shows that as wages increase over this range, the quantity of hours worked also increases. The middle, nearly vertical portion of the labor supply curve shows that as wages increase over this range, the quantity of hours worked changes very little. The backward-bending portion of the labor supply curve at the top shows that as wages increase over this range, the quantity of hours worked actually decreases. All three of these possibilities can be derived from how a change in wages causes movement in the labor-leisure budget constraint, and thus different choices by individuals.

✔ Example $5.4.1$: Is America a nation of workaholics?

Americans work a lot. Table $5.4.3$ shows average hours worked per year in the United States, Canada, Japan, and several European countries, with data from 2013. To get a perspective on these numbers, someone who works 40 hours per week for 50 weeks per year, with two weeks off, would work $2,000$ hours per year. The gap in hours worked is a little astonishing; the 250 to 300 hour gap between how much Americans work and how much Germans or the French work amounts to roughly six to seven weeks less of work per year. Economists who study these international patterns debate the extent to which average Americans and Japanese have a preference for working more than, say, Germans, or whether German workers and employers face particular kinds of taxes and regulations that lead to fewer hours worked. Many countries have laws that regulate the work week and dictate holidays and the standards of "normal" vacation time vary from country to country. It is also interesting to take the amount of time spent working in context; it is estimated that in the late nineteenth century in the United States, the average work week was over 60 hours per week—leaving little to no time for leisure.

Table $5.4.3$: Average Hours Worked Per Year in Select Countries(Source: stats.oecd.org/Index.aspx?DataSetCode=ANHRS)

Country	Average Annual Hours Actually Worked per Employed Person
United States	1,824
Spain	1,799
Japan	1,759
Canada	1,751
United Kingdom	1,669
Sweden	1,585
Germany	1,443
France	1,441

The different responses to a rise in wages—more hours worked, the same hours worked, or fewer hours worked—are patterns exhibited by different groups of workers in the U.S. economy. Many full-time workers have jobs where the number of hours is held relatively fixed, partly by their own choice and partly by their employer's practices. These workers do not much change their hours worked as wages rise or fall, so their supply curve of labor is inelastic. However, part-time workers and younger workers tend to be more flexible in their hours, and more ready to increase hours worked when wages are high or cut back when wages fall.

The backward-bending supply curve for labor, when workers react to higher wages by working fewer hours and having more income, is not observed often in the short run. However, some well-paid professionals, like dentists or accountants, may react to higher wages by choosing to limit the number of hours, perhaps by taking especially long vacations, or taking every other Friday off. Over a long-term perspective, the backward-bending supply curve for labor is common. Over the last century, Americans have reacted to gradually rising wages by working fewer hours; for example, the length of the average work-week has fallen from about 60 hours per week in 1900 to the present average of less than 40 hours per week.

Recognizing that workers have a range of possible reactions to a change in wages casts some fresh insight on a perennial political debate: the claim that a reduction in income taxes—which would, in effect, allow people to earn more per hour—will encourage people to work more. The leisure-income budget set points out that this connection will not hold true for all workers. Some people, especially part-timers, may react to higher wages by working more. Many will work the same number of hours. Some people, especially those whose incomes are already high, may react to the tax cut by working fewer hours. Of course, cutting taxes may be a good or a bad idea for a variety of reasons, not just because of its impact on work incentives, but the specific claim that tax cuts will lead people to work more hours is only likely to hold for specific groups of workers and will depend on how and for whom taxes are cut.

Key Concepts and Summary

When making a choice along the labor-leisure budget constraint, a household will choose the combination of labor, leisure, and income that provides the most utility. The result of a change in wage levels can be higher work hours, the same work hours, or lower work hours.

Glossary

backward-bending supply curve for labor

the situation when high-wage people can earn so much that they respond to a still-higher wage by working fewer hours

Contributor

- Template:ContribOpenSTAXMicroEcon

5.5: Intertemporal Choices in Financial Capital Markets

Learning Objectives

- Evaluate the reasons for making intertemporal choices
- Interpret an intertemporal budget constraint
- Analyze why people in America tend to save such a small percentage of their income

Rates of saving in America have never been especially high, but they seem to have dipped even lower in recent years, as the data from the Bureau of Economic Analysis in Figure $5.5.1$ shown below. A decision about how much to save can be represented using an **intertemporal budget constraint**. Household decisions about the quantity of financial savings show the same underlying pattern of logic as the consumption choice decision and the labor-leisure decision.

Personal Savings as a Percentage of Personal Income

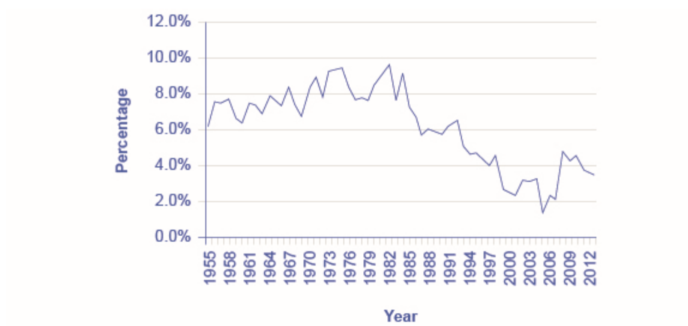


Figure $5.5.1$: Personal savings were about 7 to 11% of personal income for most of the years from the late 1950s up to the early 1990s. Since then, the rate of personal savings has fallen substantially, although it seems to have bounced back a bit since 2008. (Source: www.bea.gov/newsreleases/nati...ewsrelease.htm)

The discussion of financial saving here will not focus on the specific financial investment choices, like bank accounts, stocks, bonds, mutual funds, or owning a house or gold coins. The characteristics of these specific financial investments, along with the risks and tradeoffs they pose, are detailed in the Labor and Financial Markets chapter. Here, the focus is saving in total—that is, on how a household determines how much to consume in the present and how much to save, given the expected rate of return (or interest rate), and how the quantity of saving alters when the rate of return changes.

Using Marginal Utility to Make Intertemporal Choices

Savings behavior varies considerably across households. One factor is that households with higher incomes tend to save a larger percentage of their income. This pattern makes intuitive sense; a well-to-do family has the flexibility in its budget to save $20-25\%$ of income, while a poor family struggling to keep food on the table will find it harder to put money aside.

Another factor that causes personal saving to vary is personal preferences. Some people may prefer to consume more now, and let the future look after itself. Others may wish to enjoy a lavish retirement, complete with expensive vacations, or to pile up money that they can pass along to their grandchildren. There are savers and spendthrifts among the young, middle-aged, and old, and among those with high, middle, and low income levels.

Consider this example: Yelberton is a young man starting off at his first job. He thinks of the "present" as his working life and the "future" as after retirement. Yelberton's plan is to save money from ages 30 to 60, retire at age 60, and then live off his retirement money from ages 60 to 85. On average, therefore, he will be saving for 30 years. If the rate of return that he can receive is 6% per year, then $\$1$ saved in the present would build up to $\$5.74$ after 30 years (using the formula for compound interest, $\$1(1+0.06)30=\$5.74$). Say that Yelberton will earn $\$1,000,000$ over the 30 years from age 30 to age 60 (this amount is approximately an annual salary of $\$33,333$ multiplied by 30 years). The question for Yelberton is how much of those lifetime earnings to consume during his working life, and how much to put aside until after retirement. This example is obviously built on simplifying assumptions, but it does convey the basic life-cycle choice of saving during working life for future consumption after retirement.

Figure $5.5.2$ and Table $5.5.1$ show Yelberton's **intertemporal budget constraint**. Yelberton's choice involves comparing the utility of present consumption during his working life and future consumption after retirement. The rate of return that determines the slope of the intertemporal budget line between present consumption and future consumption in this example is the annual interest rate that he would earn on his savings, compounded over the 30 years of his working life. (For simplicity, we are assuming that any savings from current income will compound for 30 years.) Thus, in the lower budget constraint line on the figure, future consumption grows by increments of $\$574,000$ because each time $\$100,000$ is saved in the present, it compounds to $\$574,000$ after 30 years at a 6% interest rate. If some of the numbers on the future consumption axis look bizarrely large, remember that this occurs because of the power of compound interest over substantial periods of time, and because the figure is grouping together all of Yelberton's saving for retirement over his lifetime.

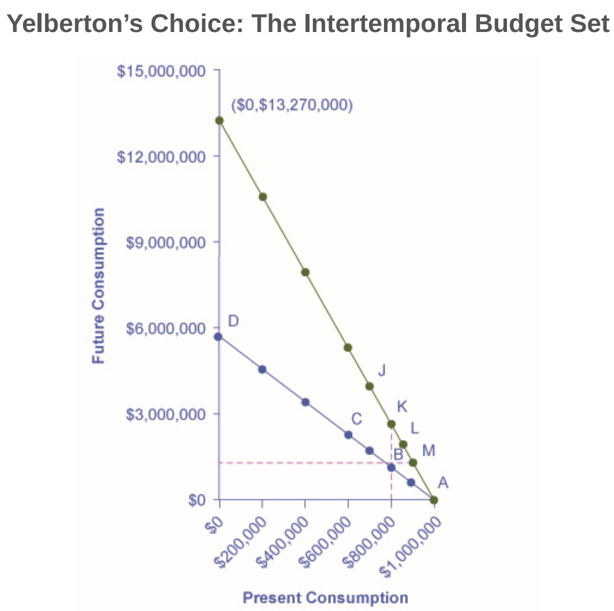


Figure $5.5.2$: Yelberton will make a choice between present and future consumption. With an annual rate of return of 6%, he decides that his utility will be highest at point B, which represents a choice of $800,000 in present consumption and $1,148,000 in future consumption. When the annual rate of return rises to 9%, the intertemporal budget constraint pivots up. Yelberton could choose to take the gains from this higher rate of return in several forms: more present saving and much higher future consumption (J), the same present saving and higher future consumption (K), more present consumption and more future consumption (L), or more present consumption and the same future consumption (M).

Table $5.5.1$: Yelburton's Intertemporal Budget Constraint

Present Consumption	Present Savings	Future Consumption (6% annual return)	Future Consumption (9% annual return)
$1,000,000	0	0	0
$900,000	$100,000	$574,000	$1,327,000
$800,000	$200,000	$1,148,000	$2,654,000
$700,000	$300,000	$1,722,000	$3,981,000
$600,000	$400,000	$2,296,000	$5,308,000
$400,000	$600,000	$3,444,000	$7,962,000
$200,000	$800,000	$4,592,000	$10,616,000
0	$1,000,000	$5,740,000	$13,270,000

Yelberton will compare the different choices along the budget constraint and choose the one that provides him with the highest utility. For example, he will compare the utility he would receive from a choice like point A, with consumption of $\$1$ million in the present, zero savings, and zero future consumption; point B, with present consumption of $\$800,000$, savings of $\$200,000$, and future consumption of $\$1,148,000$, point C, with present consumption of $\$600,000$, savings of $\$400,000$, and future consumption of $\$2,296,000$, or even choice D, with present consumption of zero, savings of $\$1,000,000$, and future consumption of $\$5,740,000$. Yelberton will also ask himself questions like these: "*Would I prefer to consume a little less in the present, save more, and have more future consumption?" or "Would I prefer to consume a little more in the present, save less, and have less* *future consumption?*" By considering marginal changes toward more or less consumption, he can seek out the choice that will provide him with the highest level of utility.

Let us say that Yelberton's preferred choice is B. Imagine that Yelberton's annual rate of return raises from 6% to 9%. In this case, each time he saves $\$100,000$ in the present, it will be worth $\$1,327,000$ in 30 years from now (using the formula for compound interest that $\$100,000(1+0.09)^{30}=\$1,327,000$). A change in rate of return alters the slope of the intertemporal budget constraint: a higher rate of return or interest rate will cause the budget line to pivot upward, while a lower rate of return will cause it to pivot downward. If Yelberton were to consume nothing in the present and save all $\$1,000,000$ with a 9% rate of return, his future consumption would be $\$13,270,000$ as shown on Figure $5.5.2$.

As the rate of return rises, Yelberton considers a range of choices on the new intertemporal budget constraint. The dashed vertical and horizontal lines running through the original choice B help to illustrate his range of options. One choice is to reduce present consumption (that is, to save more) and to have considerably higher future consumption at a point like J above and to the left of his original choice B. A second choice would be to keep the level of present consumption and savings the same, and to receive the benefits of the higher rate of return entirely in the form of higher future consumption, which would be choice K.

As a third choice Yelberton could have both more present consumption—that is, less savings—but still have higher future consumption because of the higher interest rate, which would be choice like L, above and to the right of his original choice B. Thus, the higher rate of return might cause Yelberton to save more, or less, or the same amount, depending on his own preferences. A fourth choice would be that Yelberton could react to the higher rate of return by increasing his current consumption and leaving his future consumption unchanged, as at point M directly to the right of his original choice B. The actual choice of what quantity to save and how saving will respond to changes in the rate of return will vary from person to person, according to the choice that will maximize each person's utility.

Applications of the Model of Intertemporal Choice

The theoretical model of the intertemporal budget constraint suggests that when the rate of return rises, the quantity of saving may rise, fall, or remain the same, depending on the preferences of individuals. For the U.S. economy as a whole, the most common pattern seems to be that the quantity of savings does not adjust much to changes in the rate of return. As a practical matter, many households either save at a fairly steady pace, by putting regular contributions into a retirement account or by making regular payments as they buy a house, or they do not save much at all. Of course, some people will have preferences that cause them to react to a higher rate of return by increasing their quantity of saving; others will react to a higher rate of return by noticing that with a higher rate of return, they can save less in the present and still have higher future consumption.

One prominent example in which a higher rate of return leads to a lower savings rate occurs when firms save money because they have promised to pay workers a certain fixed level of pension benefits after retirement. When rates of return rise, those companies can save less money in the present in their pension fund and still have enough to pay the promised retirement benefits in the future.

This insight suggests some skepticism about political proposals to encourage higher savings by providing savers with a higher rate of return. For example, **Individual Retirement Accounts (IRAs)** and **401(k)** accounts are special savings accounts where the money going into the account is not taxed until it is taken out many years later, after retirement. The main difference between these accounts is that an IRA is usually set up by an individual, while a 401(k) needs to be set up through an employer. By not taxing savings in the present, the effect of an IRA or a 401(k) is to increase the return to saving in these accounts.

IRA and 401(k) accounts have attracted a large quantity of savings since they became common in the late 1980s and early 1990s. In fact, the amount of IRAs rose from $\$239$ billion in 1992 to $\$3.7$ trillion in 2005, then to over $\$5$ trillion in 2012, as per the Investment Company Institute, a national association of U.S. investment companies. However, overall U.S. personal savings, as discussed earlier, actually dropped from low to lower in the late 1990s and into the 2000s. Evidently, the larger amounts in these retirement accounts are being offset, in the economy as a whole, either by less savings in other kinds of accounts, or by a larger amount of borrowing (that is, negative savings). The following Clear It Up further explores America's saving rates.

A rise in interest rates makes it easier for people to enjoy higher future consumption. But it also allows them to enjoy higher present consumption, if that is what these individuals desire. Again, a change in prices—in this case, in interest rates—leads to a range of possible outcomes.

How does America's saving rates compare to other countries?

By international standards, Americans do not save a high proportion of their income, as Table 5.5.2 shows. The rate of gross national saving includes saving by individuals, businesses, and government. By this measure, U.S. national savings amount to 17% of the size of the U.S. GDP, which measures the size of the U.S. economy. The comparable world average rate of savings is 22%.

Table 5.5.2: National Savings in Select Countries(Source: http://data.worldbank.org/indicator/NY.GNS.ICTR.ZS)

Country	Gross Domestic Savings as a Percentage of GDP
China	51%
India	30%
Russia	28%
Mexico	22%
Germany	26%
Japan	22%
Canada	21%
France	21%
Brazil	15%
United States	17%
United Kingdom	13%

The Unifying Power of the Utility-Maximizing Budget Set Framework

The choices of households are determined by an interaction between prices, budget constraints, and personal preferences. The flexible and powerful terminology of utility-maximizing gives economists a vocabulary for bringing these elements together.

Not even economists believe that people walk around mumbling about their marginal utilities before they walk into a shopping mall, accept a job, or make a deposit in a savings account. However, economists do believe that individuals seek their own satisfaction or utility and that people often decide to try a little less of one thing and a little more of another. If these assumptions are accepted, then the idea of utility-maximizing households facing budget constraints becomes highly plausible.

Behavioral Economics: An Alternative Viewpoint

As we know, people sometimes make decisions that seem "irrational" and not in their own best interest. People's decisions can seem inconsistent from one day to the next and they even deliberately ignore ways to save money or time. The traditional economic models assume rationality, which means that people take all available information and make consistent and informed decisions that are in their best interest. (In fact, economics professors often delight in pointing out so-called "irrational behavior" each semester to their new students, and present economics as a way to become more rational.)

But a new group of economists, known as behavioral economists, argue that the traditional method leaves out something important: people's state of mind. For example, one can think differently about money if one is feeling revenge, optimism, or loss. These are not necessarily irrational states of mind, but part of a range of emotions that can affect anyone on a given day. And what's more, actions under these conditions are indeed predictable, if the underlying environment is better understood. So, **behavioral economics** seeks to enrich the understanding of decision-making by integrating the insights of psychology into economics. It does this by investigating how given dollar amounts can mean different things to individuals depending on the situation. This can lead to decisions that appear outwardly inconsistent, or irrational, to the outside observer.

The way the mind works, according to this view, may seem inconsistent to traditional economists but is actually far more complex than an unemotional cost-benefit adding machine. For example, a traditional economist would say that if you lost a $10 bill today, and also got an extra $10 in your paycheck, you should feel perfectly neutral. After all, $-\$10+\$10=\$0$. You are the same financially as you were before. However, behavioral economists have done research that shows many people will feel some negative emotion—anger, frustration, and so forth—after those two things happen. We tend to focus more on the loss than the gain.

This is known as **loss aversion**, where a $\$1$ loss pains us 2.25 times more than a $\$1$ gain helps us, according to the economists Daniel Kahneman and Amos Tversky in a famous 1979 article in the journal Econometrica. This insight has implications for investing, as people tend to “overplay” the stock market by reacting more to losses than to gains. Indeed, this behavior looks irrational to traditional economists, but is consistent once we understand better how the mind works, these economists argue.

Traditional economists also assume human beings have complete self-control. But, for instance, people will buy cigarettes by the pack instead of the carton even though the carton saves them money, to keep usage down. They purchase locks for their refrigerators and overpay on taxes to force themselves to save. In other words, we protect ourselves from our worst temptations but pay a price to do so. One way behavioral economists are responding to this is by setting up ways for people to keep themselves free of these temptations. This includes what are called “nudges” toward more rational behavior rather than mandatory regulations from government. For example, up to 20% of new employees do not enroll in retirement savings plans immediately, because of procrastination or feeling overwhelmed by the different choices. Some companies are now moving to a new system, where employees are automatically enrolled unless they “opt out.” Almost no-one opts out in this program and employees begin saving at the early years, which are most critical for retirement.

Another area that seems illogical is the idea of mental accounting, or putting dollars in different mental categories where they take different values. Economists typically consider dollars to be **fungible**, or having equal value to the individual, regardless of the situation.

You might, for instance, think of the $\$25$ you found in the street differently from the $\$25$ you earned from three hours working in a fast food restaurant. The street money might well be treated as “mad money” with little rational regard to getting the best value. This is in one sense strange, since it is still equivalent to three hours of hard work in the restaurant. Yet the “easy come-easy go” mentality replaces the rational economizer because of the situation, or context, in which the money was attained.

In another example of mental accounting that seems inconsistent to a traditional economist, a person could carry a credit card debt of $\$1,000$ that has a 15% yearly interest cost, and simultaneously have a $\$2,000$ savings account that pays only 2% per year. That means she pays $\$150$ a year to the credit card company, while collecting only $\$40$ annually in bank interest, so she loses $\$130$ a year. That doesn’t seem wise.

The “rational” decision would be to pay off the debt, since a $\$1,000$ savings account with $\$0$ in debt is the equivalent net worth, and she would now net $\$20$ per year. But curiously, it is not uncommon for people to ignore this advice, since they will treat a loss to their savings account as higher than the benefit of paying off their credit card. The dollars are not being treated as fungible so it looks irrational to traditional economists.

Which view is right, the behavioral economists’ or the traditional view? Both have their advantages, but behavioral economists have at least shed a light on trying to describe and explain behavior that has historically been dismissed as irrational. If most of us are engaged in some “irrational behavior,” perhaps there are deeper underlying reasons for this behavior in the first place.

"Eeny, Meeny, Miney, Moe"—Making Choices

In what category did consumers worldwide increase their spending during the recession? Higher education. According to the United Nations Educational, Scientific, and Cultural Organization (UNESCO), enrollment in colleges and universities rose one-third in China and almost two-thirds in Saudi Arabia, nearly doubled in Pakistan, tripled in Uganda, and surged by three million—18%—in the United States. Why were consumers willing to spend on education during lean times? Both individuals and countries view higher education as the way to prosperity. Many feel that increased earnings are a significant benefit of attending college.

Bureau of Labor Statistics data from May 2012 supports this view, as shown in Figure $5.5.3$. They show a positive correlation between earnings and education. The data also indicate that unemployment rates fall with higher levels of education and training.

The Impact of Education on Earnings and Unemployment Rates, 2012

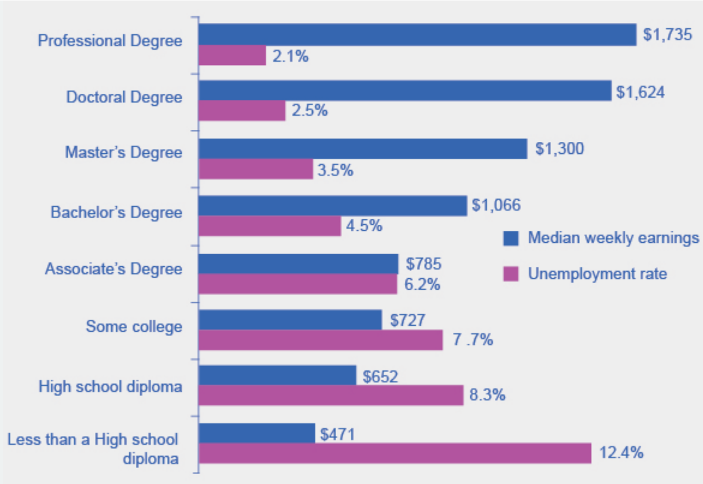


Figure 5.5.3: Those with the highest degrees in 2012 had substantially lower unemployment rates whereas those with the least formal education suffered from the highest unemployment rates. The national median average weekly income was $815, and the nation unemployment average in 2012 was 6.8%. (Source: Bureau of Labor Statistics, May 22, 2013)

Key Concepts and Summary

When making a choice along the intertemporal budget constraint, a household will choose the combination of present consumption, savings, and future consumption that provides the most utility. The result of a higher rate of return (or higher interest rates) can be a higher quantity of saving, the same quantity of saving, or a lower quantity of saving, depending on preferences about present and future consumption. Behavioral economics is a branch of economics that seeks to understand and explain the "human" factors that drive what traditional economists see as people's irrational spending decisions.

References

Holden, Sarah, and Daniel Schrass. 2012. “The rose of IRAs in U.S Households’ Saving for Retirement, 2012.” ICI Research Perspective 18.8 (2012). http://www.ici.org/pdf/per18-08.pdf.

Kahneman, Daniel and Amos Tversky. “Prospect Theory: An Analysis of Decision under Risk.” Econometrica 47.2 (March 1979) 263-291.

Thaler, Richard H. “Shifting Our Retirement Savings into Automatic.” The New York Times, April 6, 2013. www.nytimes.com/2013/04/07/bu...pagewanted=all.

UNESCO Institute for Statistics. “Statistics in Brief / Profiles” Accessed August 2013. http://stats.uis.unesco.org/unesco/T...x?ReportId=121 &IF_Language=en&BR_Country=5580.

Glossary

behavioral economics
: a branch of economics that seeks to enrich the understanding of decision-making by integrating the insights of psychology and by investigating how given dollar amounts can mean different things to individuals depending on the situation.

fungible
: the idea that units of a good, such as dollars, ounces of gold, or barrels of oil are capable of mutual substitution with each other and carry equal value to the individual.

Contributor

- Template:ContribOpenSTAXMicroEcon

- **6.5: Intertemporal Choices in Financial Capital Markets** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/principles-microeconomics.

5.E: Consumer Choices (Exercises)

6.1: Consumption Choices

Self-Check Questions

Q1

Jeremy is deeply in love with Jasmine. Jasmine lives where cell phone coverage is poor, so he can either call her on the land-line phone for five cents per minute or he can drive to see her, at a round-trip cost of $2 in gasoline money. He has a total of $10 per week to spend on staying in touch. To make his preferred choice, Jeremy uses a handy utilimometer that measures his total utility from personal visits and from phone minutes. Using the values given in Table below, figure out the points on Jeremy's consumption choice budget constraint (it may be helpful to do a sketch) and identify his utility-maximizing point.

Round Trips	Total Utility	Phone Minutes	Total Utility
0	0	0	0
1	80	20	200
2	150	40	380
3	210	60	540
4	260	80	680
5	300	100	800
6	330	120	900
7	200	140	980
8	180	160	1040
9	160	180	1080
10	140	200	1100

Q2

Take Jeremy's total utility information in **Q1**, and use the marginal utility approach to confirm the choice of phone minutes and round trips that maximize Jeremy's utility.

Review Questions

Q3

Who determines how much utility an individual will receive from consuming a good?

Q4

Would you expect total utility to rise or fall with additional consumption of a good? Why?

Q5

Would you expect marginal utility to rise or fall with additional consumption of a good? Why?

Q6

Is it possible for total utility to increase while marginal utility diminishes? Explain.

Q7

If people do not have a complete mental picture of total utility for every level of consumption, how can they find their utility-maximizing consumption choice?

Q8

What is the rule relating the ratio of marginal utility to prices of two goods at the optimal choice? Explain why, if this rule does not hold, the choice cannot be utility-maximizing.

Critical Thinking Questions

Q9

Think back to a purchase that you made recently. How would you describe your thinking before you made that purchase?

Q10

The rules of politics are not always the same as the rules of economics. In discussions of setting budgets for government agencies, there is a strategy called “closing the Washington monument.” When an agency faces the unwelcome prospect of a budget cut, it may decide to close a high-visibility attraction enjoyed by many people (like the Washington monument). Explain in terms of diminishing marginal utility why the Washington monument strategy is so misleading. Hint: If you are really trying to make the best of a budget cut, should you cut the items in your budget with the highest marginal utility or the lowest marginal utility? Does the Washington monument strategy cut the items with the highest marginal utility or the lowest marginal utility?

Problems

Q11

Praxilla, who lived in ancient Greece, derives utility from reading poems and from eating cucumbers. Praxilla gets 30 units of marginal utility from her first poem, 27 units of marginal utility from her second poem, 24 units of marginal utility from her third poem, and so on, with marginal utility declining by three units for each additional poem. Praxilla gets six units of marginal utility for each of her first three cucumbers consumed, five units of marginal utility for each of her next three cucumbers consumed, four units of marginal utility for each of the following three cucumbers consumed, and so on, with marginal utility declining by one for every three cucumbers consumed. A poem costs three bronze coins but a cucumber costs only one bronze coin. Praxilla has 18 bronze coins. Sketch Praxilla’s budget set between poems and cucumbers, placing poems on the vertical axis and cucumbers on the horizontal axis. Start off with the choice of zero poems and 18 cucumbers, and calculate the changes in marginal utility of moving along the budget line to the next choice of one poem and 15 cucumbers. Using this step-by-step process based on marginal utility, create a table and identify Praxilla’s utility-maximizing choice. Compare the marginal utility of the two goods and the relative prices at the optimal choice to see if the expected relationship holds.

Hint:

Label the table columns:

1. Choice
2. Marginal Gain from More Poems
3. Marginal Loss from Fewer Cucumbers
4. Overall Gain or Loss
5. Is the previous choice optimal?

Label the table rows:

1. 0 Poems and 18 Cucumbers
2. 1 Poem and 15 Cucumbers
3. 2 Poems and 12 Cucumbers
4. 3 Poems and 9 Cucumbers
5. 4 Poems and 6 Cucumbers
6. 5 Poems and 3 Cucumbers
7. 6 Poems and 0Cucumbers.

Solution

S1

The rows of the table in the problem do not represent the actual choices available on the budget set; that is, the combinations of round trips and phone minutes that Jeremy can afford with his budget. One of the choices listed in the problem, the six round trips, is not even available on the budget set. If Jeremy has only $\$10$ to spend and a round trip costs $\$2$ and phone calls cost $\$0.05$ per minute, he could spend his entire budget on five round trips but no phone calls or 200 minutes of phone calls, but no round trips or any combination of the two in between. It is easy to see all of his budget options with a little algebra. The equation for a budget line is:

$$Budget = P_{RT} \times Q_{RT} + P_{PC} \times Q_{PC} \tag{5.E.1}$$

where P and Q are price and quantity of round trips (RT) and phone calls (PC) (per minute). In Jeremy's case the equation for the budget line is:

$$\begin{gathered} \$10 = \$2 \times Q_{RT} + \$0.05 \times Q_{PC} \\ \frac{\$10}{\$0.05} = \frac{\$2Q_{RT} + \$0.05Q_{PC}}{\$0.05} \\ 200 = 40Q_{RT} + Q_{PC} \\ Q_{PC} = 200 - 40Q_{RT} \end{gathered} \tag{5.E.2}$$

If we choose zero through five round trips (column 1), the table below shows how many phone minutes can be afforded with the budget (column 3). The total utility figures are given in the table below.

Round Trips	Total Utility for Trips	Phone Minutes	Total Utility for Minutes	Total Utility
0	0	200	1100	1100
1	80	160	1040	1120
2	150	120	900	1050
3	210	80	680	890
4	260	40	380	640
5	300	0	0	300

Adding up total utility for round trips and phone minutes at different points on the budget line gives total utility at each point on the budget line. The highest possible utility is at the combination of one trip and 160 minutes of phone time, with a total utility of 1120.

S2

The first step is to use the total utility figures, shown in the table below, to calculate marginal utility, remembering that marginal utility is equal to the change in total utility divided by the change in trips or minutes.

Round Trips	Total Utility	Marginal Utility (per trip)	Phone Minutes	Total Utility	Marginal Utility (per minute)
0	0	-	200	1100	-
1	80	80	160	1040	60/40 = 1.5
2	150	70	120	900	140/40 = 3.5
3	210	60	80	680	220/40 = 5.5
4	260	50	40	380	300/40 = 7.5
5	300	40	0	0	380/40 = 9.5

Note that we cannot directly compare marginal utilities, since the units are trips versus phone minutes. We need a common denominator for comparison, which is price. Dividing MU by the price, yields columns 4 and 8 in the table below.

Round Trips	Total Utility	Marginal Utility (per trip)	MU/P	Phone Minutes	Total Utility	Marginal utility (per minute)	MU/P
0	0	-	-	200	1100	60/40 = 1.5	1.5/$0.05 = 30
1	80	80	80/$2 = 40	160	1040	140/40 = 3.5	3.5/$0.05 = 70
2	150	70	70/$2 = 35	120	900	220/40 = 5.5	5.5/$0.05 = 110
3	210	60	60/$2 = 30	80	680	300/40 =7.5	7.5/$0.05 = 150
4	260	50	50/$2 = 25	40	380	380/40 = 9.5	9.5/$0.05 = 190
5	300	40	40/$2 = 20	0	0	-	-

Start at the bottom of the table where the combination of round trips and phone minutes is $(5, 0)$. This starting point is arbitrary, but the numbers in this example work best starting from the bottom. Suppose we consider moving to the next point up. At $(4, 40)$, the marginal utility per dollar spent on a round trip is 25. The marginal utility per dollar spent on phone minutes is 190.

Since $25 < 190$, we are getting much more utility per dollar spent on phone minutes, so let's choose more of those. At $(3, 80)$, MU/P_{RT} is 30 < 150 (the $MU/_{PM}$), but notice that the difference is narrowing. We keep trading round trips for phone minutes until we get to $(1, 160)$, which is the best we can do. The MU/P comparison is as close as it is going to get (40 $vs.$ 70). Often in the real world, it is not possible to get MU/P exactly equal for both products, so you get as close as you can.

6.2: How Changes in Income and Prices Affect Consumption Choices

Self-Check Questions

Q1

Explain all the reasons why a decrease in the price of a product would lead to an increase in purchases of the product.

Q2

As a college student you work at a part-time job, but your parents also send you a monthly "allowance." Suppose one month your parents forgot to send the check. Show graphically how your budget constraint is affected. Assuming you only buy normal goods, what would happen to your purchases of goods?

Review Questions

Q3

As a general rule, is it safe to assume that a change in the price of a good will always have its most significant impact on the quantity demanded of that good, rather than on the quantity demanded of other goods? Explain.

Q4

Why does a change in income cause a parallel shift in the budget constraint?

Critical Thinking Questions

Q5

Income effects depend on the income elasticity of demand for each good that you buy. If one of the goods you buy has a negative income elasticity, that is, it is an inferior good, what must be true of the income elasticity of the other good you buy?

Problems

Q6

If a 10% decrease in the price of one product that you buy causes an 8% increase in quantity demanded of that product, will another 10% decrease in the price cause another 8% increase (no more and no less) in quantity demanded?

Solution

S1

This is the opposite of the example explained in the text. A decrease in price has a substitution effect and an income effect. The substitution effect says that because the product is cheaper relative to other things the consumer purchases, he or she will tend to buy more of the product (and less of the other things). The income effect says that after the price decline, the consumer could purchase the same goods as before, and still have money left over to purchase more. For both reasons, a decrease in price causes an increase in quantity demanded.

S2

This is a negative income effect. Because your parents' check failed to arrive, your monthly income is less than normal and your budget constraint shifts in toward the origin. If you only buy normal goods, the decrease in your income means you will buy less of every product.

6.3: Labor-Leisure Choices

Self-Check Questions

Q1

Siddhartha has 50 hours per week to devote to work or leisure. He has been working for $8 per hour. Based on the information in Table below, calculate his utility-maximizing choice of labor and leisure time.

Leisure Hours	Total Utility from Leisure	Work Hours	Income	Total Utility from Income
0	0	0	0	0
10	200	10	80	500
20	350	20	160	800
30	450	30	240	1,040
40	500	40	320	1,240
50	530	50	400	1,400

Q2

In Siddhartha's problem, calculate marginal utility for income and for leisure. Now, start off at the choice with 50 hours of leisure and zero income, and a wage of $8 per hour, and explain, in terms of marginal utility how Siddhartha could reason his way to the optimal choice, using marginal thinking only.

Review Questions

Q3

How will a utility-maximizer find the choice of leisure and income that provides the greatest utility?

Q4

As a general rule, is it safe to assume that a higher wage will encourage significantly more hours worked for all individuals? Explain.

Critical Thinking Questions

Q5

In the labor-leisure choice model, what is the price of leisure?

Q6

Think about the backward-bending part of the labor supply curve. Why would someone work less as a result of a higher wage rate?

Q7

What would be the substitution effect and the income effect of a wage increase?

Q8

Visit the BLS website and determine if education level, race/ethnicity, or gender appear to impact labor versus leisure choices.

Solution

S1

This problem is straightforward if you remember leisure hours plus work hours are limited to 50 hours total. If you reverse the order of the last three columns so that more leisure corresponds to less work and income, you can add up columns two and five to find utility is maximized at 10 leisure hours and 40 work hours:

Leisure Hours	Total Utility from Leisure	Work Hours	Income	Total Utility from Income	Total Utility from Both
0	0	50	400	1,400	1,400
10	200	40	320	1,240	1,440
20	350	30	240	1,040	1,390
30	450	20	160	800	1,250
40	500	10	80	500	1,000
50	530	0	0	0	530

S2

Begin from the last table and compute marginal utility from leisure and work:

Leisure Hours	Total Utility from Leisure	MU from Leisure	Work Hours	Income	Total Utility from Income	MU from Income
0	0	-	50	400	1,400	160
10	200	200	40	320	1,240	200
20	350	150	30	240	1,040	240
30	450	100	20	160	800	300
40	500	50	10	80	500	500
50	530	30	0	0	0	-

Suppose Sid starts with 50 hours of leisure and 0 hours of work. As Sid moves up the table, he trades 10 hours of leisure for 10 hours of work at each step. At $(40, 10)$, his $MU_{Leisure} = 50$, which is substantially less than his MU_{Income} of 500. This shortfall signals Sid to keep trading leisure for work/income until at $(10, 40)$ the marginal utility of both is equal at 200. This is the sign that he should stop here, confirming the answer in question 1.

6.4: Intertemporal Choices in Financial Capital Markets

Self-Check Questions

Q1

How would an increase in expected income over one's lifetime affect one's intertemporal budget constraint? How would it affect one's consumption/saving decision?

Q2

How would a decrease in expected interest rates over one's working life affect one's intertemporal budget constraint? How would it affect one's consumption/saving decision?

Review Questions

Q3

According to the model of intertemporal choice, what are the major factors which determine how much saving an individual will do? What factors might a behavioral economist use to explain savings decisions?

Q4

As a general rule, is it safe to assume that a lower interest rate will encourage significantly lower financial savings for all individuals? Explain.

Critical Thinking Questions

Q5

What do you think accounts for the wide range of savings rates in different countries?

Q6

What assumptions does the model of intertemporal choice make that are not likely true in the real world and would make the model harder to use in practice?

Solution

S1

An increase in expected income would cause an outward shift in the intertemporal budget constraint. This would likely increase both current consumption and saving, but the answer would depend on one's time preference, that is, how much one is willing to wait to forgo current consumption. Children are notoriously bad at this, which is to say they might simply consume more, and not save any. Adults, because they think about the future, are generally better at time preference—that is, they are more willing to wait to receive a reward.

S2

Lower interest rates would make lending cheaper and saving less rewarding. This would be reflected in a flatter intertemporal budget line, a rotation around the amount of current income. This would likely cause a decrease in saving and an increase in current consumption, though the results for any individual would depend on time preference.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

6: Production and Cost Structure of the Firm

6.1: Prelude to Production and Cost Structure of the Firm

Learning Objectives

- Explicit and Implicit Costs, and Accounting and Economic Profit
- The Structure of Costs in the Short Run
- The Structure of Costs in the Long Run

Figure 6.1.1: Amazon is an American international electronic commerce company that sells books, among many other things, shipping them directly to the consumer. There is no brick-and-mortar Amazon store. (Credit: modification of work by William Christiansen/Flickr Creative Commons)

Amazon

In less than two decades, Amazon.com has transformed the way books are sold, bought, and even read. Prior to Amazon, books were primarily sold through independent bookstores with limited inventories in small retail locations. There were exceptions, of course; Borders and Barnes & Noble offered larger stores in urban areas. In the last decade, however, independent bookstores have become few and far between, Borders has gone out of business, and Barnes & Noble is struggling. Online delivery and purchase of books has indeed overtaken the more traditional business models. How has Amazon changed the book selling industry? How has it managed to crush its competition?

A major reason for the giant retailer's success is its production model and cost structure, which has enabled Amazon to undercut the prices of its competitors even when factoring in the cost of shipping. Read on to see how firms great (like Amazon) and small (like your corner deli) determine what to sell, at what output and price.

This chapter is the first of four chapters that explore the theory of the firm. This theory explains that firms behave in much the same way as consumers behave. What does that mean? Let's define what is meant by the firm. A firm (or business) combines inputs of labor, capital, land, and raw or finished component materials to produce outputs. If the firm is successful, the outputs are more valuable than the inputs. This activity of production goes beyond manufacturing (i.e., making things). It includes any process or service that creates value, including transportation, distribution, wholesale and retail sales. Production involves a number of important decisions that define the behavior of firms. These decisions include, but are not limited to:

- What product or products should the firm produce?
- How should the products be produced (i.e., what production process should be used)?
- How much output should the firm produce?
- What price should the firm charge for its products?
- How much labor should the firm employ?

The answers to these questions depend on the production and cost conditions facing each firm. The answers also depend on the structure of the market for the product(s) in question. Market structure is a multidimensional concept that involves how competitive the industry is. It is defined by questions such as these:

- How much market power does each firm in the industry possess?

- How similar is each firm's product to the products of other firms in the industry?
- How difficult is it for new firms to enter the industry?
- Do firms compete on the basis of price, advertising, or other product differences?

Figure 6.1.2 illustrates the range of different market structures, which we will explore in Perfect Competition, Monopoly, and Monopolistic Competition and Oligopoly.

The Spectrum of Competition

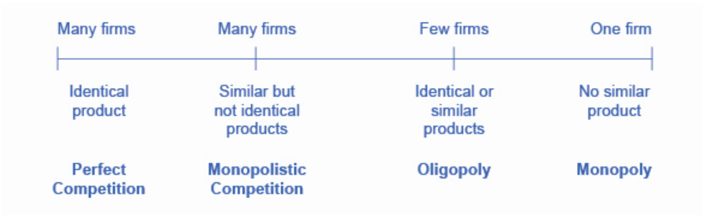


Figure 6.1.2: Firms face different competitive situations. At one extreme—perfect competition—many firms are all trying to sell identical products. At the other extreme—monopoly—only one firm is selling the product, and this firm faces no competition. Monopolistic competition and oligopoly fall between the extremes of perfect competition and monopoly. Monopolistic competition is a situation with many firms selling similar, but not identical, products. Oligopoly is a situation with few firms that sell identical or similar products.

First let's take a look at how firms determine their costs and desired profit levels. Then we will discuss costs in the short run and long run and the factors that can influence each.

Contributor

- Template:ContribOpenSTAXMicroEcon

6.2: Explicit and Implicit Costs, and Accounting and Economic Profit

Learning Objectives

- Explain the difference between explicit costs and implicit costs
- Understand the relationship between cost and revenue

Private enterprise, the ownership of businesses by private individuals, is a hallmark of the U.S. economy. When people think of businesses, often giants like Wal-Mart, Microsoft, or General Motors come to mind. But firms come in all sizes, as shown in Table $6.2.1$. The vast majority of American firms have fewer than 20 employees. As of 2010, the U.S. Census Bureau counted 5.7 million firms with employees in the U.S. economy. Slightly less than half of all the workers in private firms are at the $17,000$ large firms, meaning they employ more than 500 workers. Another 35% of workers in the U.S. economy are at firms with fewer than 100 workers. These small-scale businesses include everything from dentists and lawyers to businesses that mow lawns or clean houses. Indeed, Table $6.2.1$ does not include a separate category for the millions of small "non-employer" businesses where a single owner or a few partners are not officially paid wages or a salary, but simply receive whatever they can earn.

Table $6.2.1$: Range in Size of U.S. Firms(Source: U.S. Census, 2010 www.census.gov)

Number of Employees	Firms (% of total firms)	Number of Paid Employees (% of total employment)
Total	5,734,538	112.0 million
0–9	4,543,315 (79.2%)	12.3 million (11.0%)
10–19	617,089 (10.8%)	8.3 million (7.4%)
20–99	475,125 (8.3%)	18.6 million (16.6%)
100–499	81,773 (1.4%)	15.9 million (14.2%)
500 or more	17,236 (0.30%)	50.9 million (49.8%)

Each of these businesses, regardless of size or complexity, tries to earn a profit:

$$\text{Profit} = \text{Total Revenue} - \text{Total Cost} \tag{6.2.1}$$

Total **revenue** is the income brought into the firm from selling its products. It is calculated by multiplying the price of the product times the quantity of output sold:

$$\text{Total Revenue} = \text{Price} \times \text{Quantity} \tag{6.2.2}$$

We will see in the following chapters that revenue is a function of the demand for the firm's products.

We can distinguish between two types of cost: explicit and implicit. **Explicit costs** are out-of-pocket costs, that is, payments that are actually made. Wages that a firm pays its employees or rent that a firm pays for its office are explicit costs. **Implicit costs** are more subtle, but just as important. They represent the opportunity cost of using resources already owned by the firm. Often for small businesses, they are resources contributed by the owners; for example, working in the business while not getting a formal salary, or using the ground floor of a home as a retail store. Implicit costs also allow for depreciation of goods, materials, and equipment that are necessary for a company to operate. (See the Work it Out feature for an extended example.)

These two definitions of cost are important for distinguishing between two conceptions of profit, accounting profit and economic profit. **Accounting profit** is a cash concept. It means total revenue minus explicit costs—the difference between dollars brought in and dollars paid out. **Economic profit** is total revenue minus total cost, including both explicit and implicit costs. The difference is important because even though a business pays income taxes based on its accounting profit, whether or not it is economically successful depends on its economic profit.

Example 6.2.1: Calculating Implicit Costs

Consider the following example. Fred currently works for a corporate law firm. He is considering opening his own legal practice, where he expects to earn $\$200,000$ per year once he gets established. To run his own firm, he would need an office and a law clerk. He has found the perfect office, which rents for $\$50,000$ per year. A law clerk could be hired for $\$35,000$ per year. If these figures are accurate, would Fred's legal practice be profitable?

Step 1: First you have to calculate the costs. You can take what you know about explicit costs and total them:

$$\begin{array}{cc} \text{Office rental:} & \$50,000 \\ \text{Law clerk's salary:} & +\$35,000 \\ \hline \text{Total explicit costs:} & \$85,000 \end{array} \tag{6.2.3}$$

Step 2: Subtracting the explicit costs from the revenue gives you the accounting profit.

$$\begin{array}{cc} \text{Revenues:} & \$200,000 \\ \text{Explicit costs:} & -\$85,000 \\ \hline \text{Accounting profit:} & \$115,000 \end{array} \tag{6.2.4}$$

But these calculations consider only the explicit costs. To open his own practice, Fred would have to quit his current job, where he is earning an annual salary of $\$125,000$ This would be an implicit cost of opening his own firm.

Step 3: You need to subtract both the explicit and implicit costs to determine the true economic profit:

$$\begin{aligned} \text{Economic profit } &= \text{total revenues} - \text{explicit costs} - \text{implicit costs} \\ &= \$200,000 - \$85,000 - \$125,000 \\ &= -\$10,000 \text{ per year} \end{aligned}$$

Fred would be losing $\$10,000$ per year. That does not mean he would not want to open his own business, but it does mean he would be earning $\$10,000$ less than if he worked for the corporate firm.

Implicit costs can include other things as well. Maybe Fred values his leisure time, and starting his own firm would require him to put in more hours than at the corporate firm. In this case, the lost leisure would also be an implicit cost that would subtract from economic profits.

Now that we have an idea about the different types of costs, let's look at cost structures. A firm's cost structure in the long run may be different from that in the short run. We turn to that distinction in the next section.

Key Concepts and Summary

Privately owned firms are motivated to earn profits. Profit is the difference between revenues and costs. While accounting profit considers only explicit costs, economic profit considers both explicit and implicit costs.

References

2010 U.S. Census. www.census.gov.

Glossary

accounting profit

total revenues minus explicit costs, including depreciation

economic profit

total revenues minus total costs (explicit plus implicit costs)

explicit costs

out-of-pocket costs for a firm, for example, payments for wages and salaries, rent, or materials

firm

an organization that combines inputs of labor, capital, land, and raw or finished component materials to produce outputs.

implicit costs

opportunity cost of resources already owned by the firm and used in business, for example, expanding a factory onto land already owned

private enterprise

the ownership of businesses by private individuals

production

the process of combining inputs to produce outputs, ideally of a value greater than the value of the inputs

revenue

income from selling a firm's product; defined as price times quantity sold

Contributor

- Template:ContribOpenSTAXMicroEcon

6.3: The Structure of Costs in the Short Run

Learning Objectives

- Analyze short-run costs as influenced by total cost, fixed cost, variable cost, marginal cost, and average cost.
- Calculate average profit
- Evaluate patterns of costs to determine potential profit

The cost of producing a firm's output depends on how much **labor** and **physical capital** the firm uses. A list of the costs involved in producing cars will look very different from the costs involved in producing computer software or haircuts or fast-food meals. However, the cost structure of all firms can be broken down into some common underlying patterns. When a firm looks at its total costs of production in the short run, a useful starting point is to divide total costs into two categories: fixed costs that cannot be changed in the short run and variable costs that can be changed.

Fixed and Variable Costs

Fixed costs are expenditures that do not change regardless of the level of production, at least not in the short term. Whether you produce a lot or a little, the fixed costs are the same. One example is the rent on a factory or a retail space. Once you sign the lease, the rent is the same regardless of how much you produce, at least until the lease runs out. Fixed costs can take many other forms: for example, the cost of machinery or equipment to produce the product, research and development costs to develop new products, even an expense like advertising to popularize a brand name. The level of fixed costs varies according to the specific line of business: for instance, manufacturing computer chips requires an expensive factory, but a local moving and hauling business can get by with almost no fixed costs at all if it rents trucks by the day when needed.

Variable costs, on the other hand, are incurred in the act of producing—the more you produce, the greater the variable cost. Labor is treated as a variable cost, since producing a greater quantity of a good or service typically requires more workers or more work hours. Variable costs would also include raw materials.

As a concrete example of fixed and variable costs, consider the barber shop called "The Clip Joint" shown in Figure $6.3.1$. The data for output and costs are shown in Table $6.3.1$. The fixed costs of operating the barber shop, including the space and equipment, are $\$160$ per day. The variable costs are the costs of hiring barbers, which in our example is $\$80$ per barber each day. The first two columns of the table show the quantity of haircuts the barbershop can produce as it hires additional barbers. The third column shows the fixed costs, which do not change regardless of the level of production. The fourth column shows the variable costs at each level of output. These are calculated by taking the amount of labor hired and multiplying by the wage. For example, two barbers cost: $2 \times \$80 = \$160$. Adding together the fixed costs in the third column and the variable costs in the fourth column produces the total costs in the fifth column. So, for example, with two barbers the total cost is: $\$160 + \$160 = \$320$.

Table $6.3.1$: Output and Total Costs

Labor	Quantity	Fixed Cost	Variable Cost	Total Cost
1	16	$160	$80	$240
2	40	$160	$160	$320
3	60	$160	$240	$400
4	72	$160	$320	$480
5	80	$160	$400	$560
6	84	$160	$480	$640
7	82	$160	$560	$720

How Output Affects Total Costs

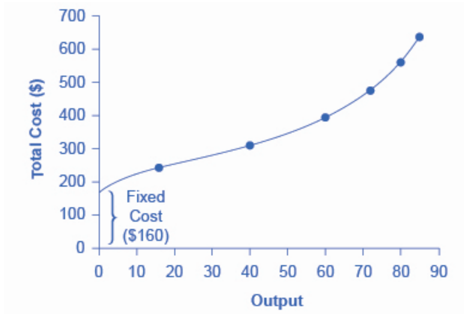


Figure $6.3.1$: At zero production, the fixed costs of $160 are still present. As production increases, variable costs are added to fixed costs, and the total cost is the sum of the two.

The relationship between the quantity of output being produced and the cost of producing that output is shown graphically in the figure. The fixed costs are always shown as the vertical intercept of the total cost curve; that is, they are the costs incurred when output is zero so there are no variable costs.

You can see from the graph that once production starts, total costs and variable costs rise. While variable costs may initially increase at a decreasing rate, at some point they begin increasing at an increasing rate. This is caused by diminishing marginal returns, discussed in the chapter on Choice in a World of Scarcity, which is easiest to see with an example. As the number of barbers increases from zero to one in the table, output increases from 0 to 16 for a marginal gain of 16; as the number rises from one to two barbers, output increases from 16 to 40, a marginal gain of 24. From that point on, though, the marginal gain in output diminishes as each additional barber is added. For example, as the number of barbers rises from two to three, the marginal output gain is only 20; and as the number rises from three to four, the marginal gain is only 12.

To understand the reason behind this pattern, consider that a one-man barber shop is a very busy operation. The single barber needs to do everything: say hello to people entering, answer the phone, cut hair, sweep up, and run the cash register. A second barber reduces the level of disruption from jumping back and forth between these tasks, and allows a greater division of labor and specialization. The result can be greater increasing marginal returns. However, as other barbers are added, the advantage of each additional barber is less, since the specialization of labor can only go so far. The addition of a sixth or seventh or eighth barber just to greet people at the door will have less impact than the second one did. This is the pattern of diminishing marginal returns. As a result, the total costs of production will begin to rise more rapidly as output increases. At some point, you may even see negative returns as the additional barbers begin bumping elbows and getting in each other’s way. In this case, the addition of still more barbers would actually cause output to decrease, as shown in the last row of Table $6.3.1$.

This pattern of **diminishing marginal returns** is common in production. As another example, consider the problem of irrigating a crop on a farmer’s field. The plot of land is the fixed factor of production, while the water that can be added to the land is the key variable cost. As the farmer adds water to the land, output increases. But adding more and more water brings smaller and smaller increases in output, until at some point the water floods the field and actually reduces output. Diminishing marginal returns occur because, at a given level of fixed costs, each additional input contributes less and less to overall production.

Average Total Cost, Average Variable Cost, Marginal Cost

The breakdown of total costs into fixed and variable costs can provide a basis for other insights as well. The first five columns of Table $6.3.2$ duplicate the previous table, but the last three columns show average total costs, average variable costs, and marginal costs. These new measures analyze costs on a per-unit (rather than a total) basis and are reflected in the curves shown in Figure $6.3.1$.

Cost Curves at the Clip Joint

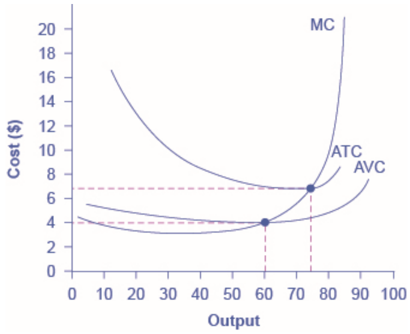


Figure $6.3.2$: The information on total costs, fixed cost, and variable cost can also be presented on a per-unit basis. Average total cost (ATC) is calculated by dividing total cost by the total quantity produced. The average total cost curve is typically U-shaped. Average variable cost (AVC) is calculated by dividing variable cost by the quantity produced. The average variable cost curve lies below the average total cost curve and is typically U-shaped or upward-sloping. Marginal cost (MC) is calculated by taking the change in total cost between two levels of output and dividing by the change in output. The marginal cost curve is upward-sloping.

Table $6.3.2$: Different Types of Costs

Labor	Quantity	Fixed Cost	Variable Cost	Total Cost	Marginal Cost	Average Total Cost	Average Variable Cost
1	16	$160	$80	$240	$5.00	$15.00	$5.00
2	40	$160	$160	$320	$3.30	$8.00	$4.00
3	60	$160	$240	$400	$4.00	$6.60	$4.00
4	72	$160	$320	$480	$6.60	$6.60	$4.40
5	80	$160	$400	$560	$10.00	$7.00	$5.00
6	84	$160	$480	$640	$20.00	$7.60	$5.70

Average total cost (sometimes referred to simply as average cost) is total cost divided by the quantity of output. Since the total cost of producing 40 haircuts is $\$320$, the average total cost for producing each of 40 haircuts is $\$320/40$, or $\$8$ per haircut. Average cost curves are typically U-shaped, as Figure $6.3.2$ shows. Average total cost starts off relatively high, because at low levels of output total costs are dominated by the fixed cost; mathematically, the denominator is so small that average total cost is large. Average total cost then declines, as the fixed costs are spread over an increasing quantity of output. In the average cost calculation, the rise in the numerator of total costs is relatively small compared to the rise in the denominator of quantity produced. But as output expands still further, the average cost begins to rise. At the right side of the average cost curve, total costs begin rising more rapidly as diminishing returns kick in.

Average variable cost obtained when variable cost is divided by quantity of output. For example, the variable cost of producing 80 haircuts is $\$400$, so the average variable cost is $\$400/80$, or $\$5$ per haircut. Note that at any level of output, the average variable cost curve will always lie below the curve for average total cost, as shown in Figure $6.3.2$. The reason is that average total cost includes average variable cost and average fixed cost. Thus, for $Q = 80$ haircuts, the average total cost is $\$8$ per haircut, while the average variable cost is $\$5$ per haircut. However, as output grows, fixed costs become relatively less important (since they do not rise with output), so average variable cost sneaks closer to average cost.

Average total and variable costs measure the average costs of producing some quantity of output. Marginal cost is somewhat different. **Marginal cost** is the additional cost of producing one more unit of output. So it is not the cost per unit of all units being produced, but only the next one (or next few). Marginal cost can be calculated by taking the change in total cost and dividing it by the change in quantity. For example, as quantity produced increases from 40 to 60 haircuts, total costs rise by $400 - 320$, or 80. Thus, the marginal cost for each of those marginal 20 units will be $80/20$, or $\$4$ per haircut. The marginal cost curve is generally upward-sloping, because diminishing marginal returns implies that additional units are more costly to produce. A small range of increasing marginal returns can be seen in the figure as a dip in the marginal cost curve before it starts rising. There is a point at which marginal and average costs meet, as the following Clear it Up feature discusses.

> **Where do marginal and average costs meet?**
>
> The marginal cost line intersects the average cost line exactly at the bottom of the average cost curve—which occurs at a quantity of 72 and cost of $\$6.60$ in Figure $6.3.2$. The reason why the intersection occurs at this point is built into the economic meaning of marginal and average costs. If the marginal cost of production is below the average cost for producing previous units, as it is for the points to the left of where MC crosses ATC, then producing one more additional unit will reduce average costs overall—and the ATC curve will be downward-sloping in this zone. Conversely, if the marginal cost of production for producing an additional unit is above the average cost for producing the earlier units, as it is for points to the right of where MC crosses ATC, then producing a marginal unit will increase average costs overall—and the ATC curve must be upward-sloping in this zone. The point of transition, between where MC is pulling ATC down and where it is pulling it up, must occur at the minimum point of the ATC curve.
>
> This idea of the marginal cost "pulling down" the average cost or "pulling up" the average cost may sound abstract, but think about it in terms of your own grades. If the score on the most recent quiz you take is lower than your average score on previous quizzes, then the marginal quiz pulls down your average. If your score on the most recent quiz is higher than the average on previous quizzes, the marginal quiz pulls up your average. In this same way, low marginal costs of production first pull down average costs and then higher marginal costs pull them up.

The numerical calculations behind average cost, average variable cost, and marginal cost will change from firm to firm. However, the general patterns of these curves, and the relationships and economic intuition behind them, will not change.

Lessons from Alternative Measures of Costs

Breaking down total costs into fixed cost, marginal cost, average total cost, and average variable cost is useful because each statistic offers its own insights for the firm.

Whatever the firm's quantity of production, total revenue must exceed total costs if it is to earn a profit. As explored in the chapter Choice in a World of Scarcity, fixed costs are often sunk costs that cannot be recouped. In thinking about what to do next, sunk costs should typically be ignored, since this spending has already been made and cannot be changed. However, variable costs can be changed, so they convey information about the firm's ability to cut costs in the present and the extent to which costs will increase if production rises.

> **Why are total cost and average cost not on the same graph?**
>
> Total cost, fixed cost, and variable cost each reflect different aspects of the cost of production over the entire quantity of output being produced. These costs are measured in dollars. In contrast, marginal cost, average cost, and average variable cost are costs per unit. In the previous example, they are measured as cost per haircut. Thus, it would not make sense to put all of these numbers on the same graph, since they are measured in different units ($\$$ versus $\$$ per unit of output).
>
> It would be as if the vertical axis measured two different things. In addition, as a practical matter, if they were on the same graph, the lines for marginal cost, average cost, and average variable cost would appear almost flat against the horizontal axis, compared to the values for total cost, fixed cost, and variable cost. Using the figures from the previous example, the total cost of producing 40 haircuts is $\$320$. But the average cost is $\$320/40$, or $\$8$. If you graphed both total and average cost on the same axes, the average cost would hardly show.

Average cost tells a firm whether it can earn profits given the current price in the market. If we divide profit by the quantity of output produced we get **average profit**, also known as the firm's *profit margin*. Expanding the equation for profit gives:

$$\begin{aligned} \text{average profit} &= \frac{\text{profit}}{\text{quantity produced}} \\ &= \frac{\text{total revenue - total cost}}{\text{quantity produced}} \\ &= \frac{\text{total revenue}}{\text{quantity produced}} - \frac{\text{total cost}}{\text{quantity produced}} \\ &= \text{average revenue} - \text{average cost} \end{aligned}$$

But note that:

$$\text{average revenue} = \frac{\text{price} \times \text{quantity produced}}{\text{quantity produced}}$$
$$= \text{price}$$

Thus:

$$\text{average profit} = \text{price} - \text{average cost} \tag{6.3.1}$$

This is the firm's **profit margin**. This definition implies that if the market price is above average cost, average profit, and thus total profit, will be positive; if price is below average cost, then profits will be negative.

The marginal cost of producing an additional unit can be compared with the marginal revenue gained by selling that additional unit to reveal whether the additional unit is adding to total profit—or not. Thus, marginal cost helps producers understand how profits would be affected by increasing or decreasing production.

A Variety of Cost Patterns

The pattern of costs varies among industries and even among firms in the same industry. Some businesses have high fixed costs, but low marginal costs. Consider, for example, an Internet company that provides medical advice to customers. Such a company might be paid by consumers directly, or perhaps hospitals or healthcare practices might subscribe on behalf of their patients. Setting up the website, collecting the information, writing the content, and buying or leasing the computer space to handle the web traffic are all fixed costs that must be undertaken before the site can work. However, when the website is up and running, it can provide a high quantity of service with relatively low variable costs, like the cost of monitoring the system and updating the information. In this case, the total cost curve might start at a high level, because of the high fixed costs, but then might appear close to flat, up to a large quantity of output, reflecting the low variable costs of operation. If the website is popular, however, a large rise in the number of visitors will overwhelm the website, and increasing output further could require a purchase of additional computer space.

For other firms, fixed costs may be relatively low. For example, consider firms that rake leaves in the fall or shovel snow off sidewalks and driveways in the winter. For fixed costs, such firms may need little more than a car to transport workers to homes of customers and some rakes and shovels. Still other firms may find that diminishing marginal returns set in quite sharply. If a manufacturing plant tried to run 24 hours a day, seven days a week, little time remains for routine maintenance of the equipment, and marginal costs can increase dramatically as the firm struggles to repair and replace overworked equipment.

Every firm can gain insight into its task of earning profits by dividing its total costs into fixed and variable costs, and then using these calculations as a basis for average total cost, average variable cost, and marginal cost. However, making a final decision about the profit-maximizing quantity to produce and the price to charge will require combining these perspectives on cost with an analysis of sales and revenue, which in turn requires looking at the market structure in which the firm finds itself. Before we turn to the analysis of market structure in other chapters, we will analyze the firm's cost structure from a long-run perspective.

Key Concepts and Summary

In a short-run perspective, a firm's total costs can be divided into fixed costs, which a firm must incur before producing any output, and variable costs, which the firm incurs in the act of producing. Fixed costs are sunk costs; that is, because they are in the past and cannot be altered, they should play no role in economic decisions about future production or pricing. Variable costs typically show diminishing marginal returns, so that the marginal cost of producing higher levels of output rises.

Marginal cost is calculated by taking the change in total cost (or the change in variable cost, which will be the same thing) and dividing it by the change in output, for each possible change in output. Marginal costs are typically rising. A firm can compare marginal cost to the additional revenue it gains from selling another unit to find out whether its marginal unit is adding to profit.

Average total cost is calculated by taking total cost and dividing by total output at each different level of output. Average costs are typically U-shaped on a graph. If a firm's average cost of production is lower than the market price, a firm will be earning profits.

Average variable cost is calculated by taking variable cost and dividing by the total output at each level of output. Average variable costs are typically U-shaped. If a firm's average variable cost of production is lower than the market price, then the firm would be earning profits if fixed costs are left out of the picture.

Glossary

average profit
profit divided by the quantity of output produced; profit margin

average total cost
total cost divided by the quantity of output

average variable cost
variable cost divided by the quantity of output

fixed cost
expenditure that must be made before production starts and that does not change regardless of the level of production

marginal cost
the additional cost of producing one more unit

total cost
the sum of fixed and variable costs of production

variable cost
cost of production that increases with the quantity produced

Contributor

- Template:ContribOpenSTAXMicroEcon

6.4: The Structure of Costs in the Long Run

Skills to Develop

- Calculate total cost
- Identify economies of scale, diseconomies of scale, and constant returns to scale
- Interpret graphs of long-run average cost curves and short-run average cost curves
- Analyze cost and production in the long run and short run

The long run is the period of time when all costs are variable. The long run depends on the specifics of the firm in question—it is not a precise period of time. If you have a one-year lease on your factory, then the long run is any period longer than a year, since after a year you are no longer bound by the lease. No costs are fixed in the long run. A firm can build new factories and purchase new machinery, or it can close existing facilities. In planning for the long run, the firm will compare alternative **production technologies** (or processes).

In this context, technology refers to all alternative methods of combining inputs to produce outputs. It does not refer to a specific new invention like the tablet computer. The firm will search for the production technology that allows it to produce the desired level of output at the lowest cost. After all, lower costs lead to higher profits—at least if total revenues remain unchanged. Moreover, each firm must fear that if it does not seek out the lowest-cost methods of production, then it may lose sales to competitor firms that find a way to produce and sell for less.

Choice of Production Technology (or Production Function)

Many tasks can be performed with a range of combinations of labor and physical capital. For example, a firm can have human beings answering phones and taking messages, or it can invest in an automated voicemail system. A firm can hire file clerks and secretaries to manage a system of paper folders and file cabinets, or it can invest in a computerized recordkeeping system that will require fewer employees. A firm can hire workers to push supplies around a factory on rolling carts, it can invest in motorized vehicles, or it can invest in robots that carry materials without a driver. Firms often face a choice between buying a many small machines, which need a worker to run each one, or buying one larger and more expensive machine, which requires only one or two workers to operate it. In short, physical capital and labor can often substitute for each other.

Consider the example of a private firm that is hired by local governments to clean up public parks. Three different combinations of labor and physical capital for cleaning up a single average-sized park appear in Table $6.4.1$. The first production technology is heavy on workers and light on machines. This is known as a **labor-intensive production technique**. The next two technologies substitute machines for workers. When there is significantly more capital used the production technique is known as **capital-intensive**. Since all three of these production methods produce the same thing—one cleaned-up park—a profit-seeking firm will choose the production technology that is least expensive, given the prices of labor and machines.

Table $6.4.1$***:*** *Three Ways to Clean a Park*

Production technology 1	10 workers	2 machines
Production technology 2	7 workers	4 machines
Production technology 3	3 workers	7 machines

Production technology 1 uses the most labor and least machinery, while production technology 3 uses the least labor and the most machinery. Table $6.4.2$ outlines three examples of how the total cost will change with each production technology as the cost of labor changes. As the cost of labor rises from example A to B to C, the firm will choose to substitute away from labor and use more machinery.

Table $6.4.2$***:*** *Total Cost with Rising Labor Costs*

Example A: Workers cost \$40, machines cost \$80			
	Labor Cost	Machine Cost	Total Cost
Cost of technology 1	10 × \$40 = \$400	2 × \$80 = \$160	\$560
Cost of technology 2	7 × \$40 = \$280	4 × \$80 = \$320	\$600

Cost of technology 3	3 × $40 = $120	7 × $80 = $560	$680
Example B: Workers cost $55, machines cost $80			
	Labor Cost	Machine Cost	Total Cost
Cost of technology 1	10 × $55 = $550	2 × $80 = $160	$710
Cost of technology 2	7 × $55 = $385	4 × $80 = $320	$705
Cost of technology 3	3 × $55 = $165	7 × $80 = $560	$725
Example C: Workers cost $90, machines cost $80			
	Labor Cost	Machine Cost	Total Cost
Cost of technology 1	10 × $90 = $900	2 × $80 = $160	$1,060
Cost of technology 2	7 × $90 = $630	4 × $80 = $320	$950
Cost of technology 3	3 × $90 = $270	7 × $80 = $560	$830

Example A shows the firm's cost calculation when wages are $\$40$ and machines costs are $\$80$. In this case, technology 1 is the low-cost production technology. In example B, wages rise to $\$55$, while the cost of machines does not change, in which case technology 2 is the low-cost production technology. If wages keep rising up to $\$90$, while the cost of machines remains unchanged, then technology 3 clearly becomes the low-cost form of production, as shown in example C.

This example shows that as an input becomes more expensive (in this case, the labor input), firms will attempt to conserve on using that input and will instead shift to other inputs that are relatively less expensive. This pattern helps to explain why the demand curve for labor (or any input) slopes down; that is, as labor becomes relatively more expensive, profit-seeking firms will seek to substitute the use of other inputs. When a multinational employer like Coca-Cola or McDonald's sets up a bottling plant or a restaurant in a high-wage economy like the United States, Canada, Japan, or Western Europe, it is likely to use production technologies that conserve on the number of workers and focuses more on machines. However, that same employer is likely to use production technologies with more workers and less machinery when producing in a lower-wage country like Mexico, China, or South Africa.

Returns to Scale

The concept of returns to scale describes the rate of increase in production relative to the associated increase in the factors of production in the long run. In other words, it describes how effectively and efficiently—or, profitably—a firm is producing its goods or services. In the long run, all factors of production are variable (not fixed) and can scale up. Therefore, the scale of production can be changed by changing the quantity of all factors of production.

Conceptualizing returns to scale is an effort to specifically understand how production increases relative to factors contributing to production. Production functions typically include capital as well as labor. Our production function is noted Q=F(K,L).

The difference between economies of scale and returns to scale is that economies of scale show the effect of an increased output level on unit costs, while the return to scale focus only on the relation between input and output quantities.

Increasing Returns to Scale

Increasing returns to scale take place when all the factors of production increase in a given proportion and the output increases at a greater rate than that of the increase in factors of production.

For example, if the factors of production are doubled, then the output will be **more** than doubled.

Constant Returns to Scale

Constant returns to scale occurs when the output increases in exactly the same proportion as the factors of production. In other words, when inputs (i.e. capital and labor) increase by 50%, outputs likewise increase by 50% as a result.

For example, if the factors of production are doubled, then the output will also be **exactly doubled**.

Diminishing (or Decreasing) Returns to Scale

Diminishing or decreasing returns to scale takes place when all the factors of production increase in a given proportion, but the output increases at a lesser rate than that of the increase in factors of production.

For example, if the factors of production are doubled, then the output will be **less** than doubled.

Economies of Scale

Once a firm has determined the least costly production technology, it can consider the optimal scale of production, or quantity of output to produce. Many industries experience economies of scale. **Economies of scale** refers to the situation where, as the quantity of output goes up, the cost per unit goes down. This is the idea behind "warehouse stores" like Costco or Walmart. In everyday language: a larger factory can produce at a lower average cost than a smaller factory.

Figure $6.4.1$ illustrates the idea of economies of scale, showing the average cost of producing an alarm clock falling as the quantity of output rises. For a small-sized factory like S, with an output level of $1,000$, the average cost of production is $\$12$ per alarm clock. For a medium-sized factory like M, with an output level of $2,000$, the average cost of production falls to $\$8$ per alarm clock. For a large factory like L, with an output of $5,000$, the average cost of production declines still further to $\$4$ per alarm clock.

Economies of Scale

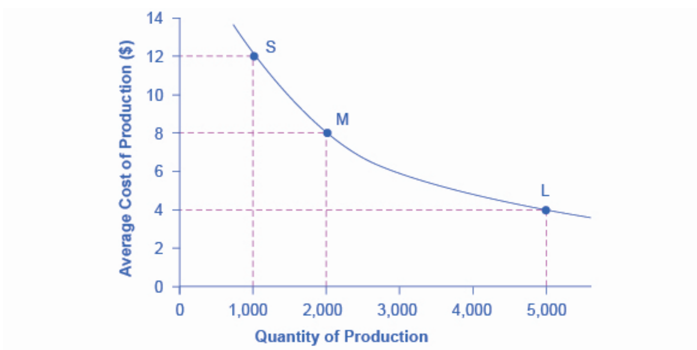


Figure $6.4.1$: *A small factory like S produces 1,000 alarm clocks at an average cost of $12 per clock. A medium factory like M produces 2,000 alarm clocks at a cost of $8 per clock. A large factory like L produces 5,000 alarm clocks at a cost of $4 per clock. Economies of scale exist because the larger scale of production leads to lower average costs.*

The average cost curve in Figure $6.4.1$ may appear similar to the average cost curves presented earlier in this chapter, although it is downward-sloping rather than U-shaped. But there is one major difference. The economies of scale curve is a long-run average cost curve, because it allows all factors of production to change. The short-run average cost curves presented earlier in this chapter assumed the existence of fixed costs, and only variable costs were allowed to change.

One prominent example of economies of scale occurs in the chemical industry. Chemical plants have a lot of pipes. The cost of the materials for producing a pipe is related to the circumference of the pipe and its length. However, the volume of chemicals that can flow through a pipe is determined by the cross-section area of the pipe. The calculations in Table $6.4.3$ show that a pipe which uses twice as much material to make (as shown by the circumference of the pipe doubling) can actually carry four times the volume of chemicals because the cross-section area of the pipe rises by a factor of four (as shown in the Area column).

Table $6.4.3$***:*** *Comparing Pipes: Economies of Scale in the Chemical Industry*

	Circumference ($2\pi r$)	Area (πr^2)
4-inch pipe	12.5 inches	12.5 square inches
8-inch pipe	25.1 inches	50.2 square inches
16-inch pipe	50.2 inches	201.1 square inches

A doubling of the cost of producing the pipe allows the chemical firm to process four times as much material. This pattern is a major reason for economies of scale in chemical production, which uses a large quantity of pipes. Of course, economies of scale in a chemical plant are more complex than this simple calculation suggests. But the chemical engineers who design these plants have long used what they call the "*six-tenths rule*," a rule of thumb which holds that increasing the quantity produced in a chemical plant by a certain percentage will increase total cost by only six-tenths as much.

Shapes of Long-Run Average Cost Curves

While in the short run firms are limited to operating on a single average cost curve (corresponding to the level of fixed costs they have chosen), in the long run when all costs are variable, they can choose to operate on any average cost curve. Thus, the **long-run average cost (LRAC) curve** is actually based on a group of **short-run average cost (SRAC) curves**, each of which represents one specific level of fixed costs. More precisely, the long-run average cost curve will be the least expensive average cost curve for any level of output. Figure $6.4.2$ shows how the long-run average cost curve is built from a group of short-run average cost curves. Five short-run-average cost curves appear on the diagram. Each $SRAC$ curve represents a different level of fixed costs. For example, you can imagine $SRAC_1$ as a small factory, $SRAC_2$ as a medium factory, $SRAC_3$ as a large factory, and $SRAC_4$ and $SRAC_5$ as very large and ultra-large. Although this diagram shows only five $SRAC$ curves, presumably there are an infinite number of other $SRAC$ curves between the ones that are shown. This family of short-run average cost curves can be thought of as representing different choices for a firm that is planning its level of investment in fixed cost physical capital—knowing that different choices about capital investment in the present will cause it to end up with different short-run average cost curves in the future.

From Short-Run Average Cost Curves to Long-Run Average Cost Curves

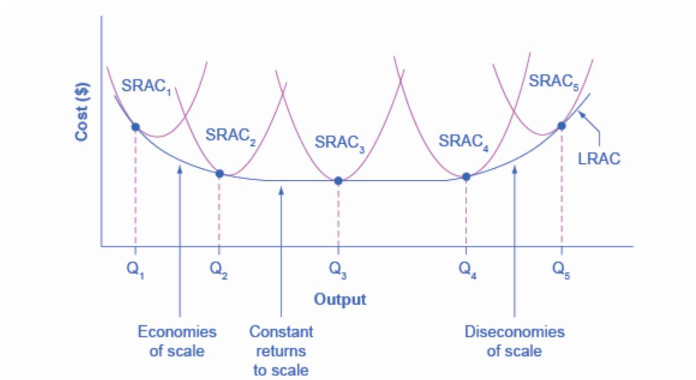


Figure $\mathbf{6.4.2}$***:*** *The five different short-run average cost (SRAC) curves each represents a different level of fixed costs, from the low level of fixed costs at* $SRAC_1$ *to the high level of fixed costs at* $SRAC_5$*. Other* $SRAC$ *curves, not shown in the diagram, lie between the ones that are shown here. The long-run average cost* $(LRAC)$ *curve shows the lowest cost for producing each quantity of output when fixed costs can vary, and so it is formed by the bottom edge of the family of* $SRAC$ *curves. If a firm wished to produce quantity* Q_3*, it would choose the fixed costs associated with* $SRAC_3$*.*

The long-run average cost curve shows the cost of producing each quantity in the long run, when the firm can choose its level of fixed costs and thus choose which short-run average costs it desires. If the firm plans to produce in the long run at an output of Q_3, it should make the set of investments that will lead it to locate on $SRAC_3$, which allows producing Q_3 at the lowest cost. A firm that intends to produce Q_3 would be foolish to choose the level of fixed costs at $SRAC_2$ or $SRAC_4$. At $SRAC_2$ the level of fixed costs is too low for producing Q_3 at lowest possible cost, and producing q3 would require adding a very high level of variable costs and make the average cost very high. At $SRAC_4$, the level of fixed costs is too high for producing Q_3 at lowest possible cost, and again average costs would be very high as a result.

The shape of the long-run cost curve, as drawn in Figure $6.4.2$, is fairly common for many industries. The left-hand portion of the long-run average cost curve, where it is downward- sloping from output levels Q_1 to Q_2 to Q_3, illustrates the case of economies of scale. In this portion of the long-run average cost curve, larger scale leads to lower average costs. This pattern was illustrated earlier in Figure $6.4.1$.

In the middle portion of the long-run average cost curve, the flat portion of the curve around Q_3, economies of scale have been exhausted. In this situation, allowing all inputs to expand does not much change the average cost of production. In this range of the

$LRAC$ curve, the average cost of production does not change much as scale rises or falls. The following Clear it Up feature explains where diminishing marginal returns fit into this analysis.

How do economies of scale compare to diminishing marginal returns?

The concept of economies of scale, where average costs decline as production expands, might seem to conflict with the idea of diminishing marginal returns, where marginal costs rise as production expands. But diminishing marginal returns refers only to the short-run average cost curve, where one variable input (like labor) is increasing, but other inputs (like capital) are fixed. Economies of scale refers to the long-run average cost curve where all inputs are being allowed to increase together. Thus, it is quite possible and common to have an industry that has both diminishing marginal returns when only one input is allowed to change, and at the same time has increasing or constant economies of scale when all inputs change together to produce a larger-scale operation.

Finally, the right-hand portion of the long-run average cost curve, running from output level Q_4 to Q_5, shows a situation where, as the level of output and the scale rises, average costs rise as well. This situation is called **diseconomies** of scale. A firm or a factory can grow so large that it becomes very difficult to manage, resulting in unnecessarily high costs as many layers of management try to communicate with workers and with each other, and as failures to communicate lead to disruptions in the flow of work and materials. It is difficult for overly large factories to exist in the real world, because with their very high production costs, they are unable to compete for long against plants with lower average costs of production.

Diseconomies of scale can also be present across an entire firm, not just a large factory. The **leviathan effect** can hit firms that become too large to run efficiently, across the entirety of the enterprise. Firms that shrink their operations are often responding to finding itself in the diseconomies region, thus moving back to a lower average cost at a lower output level.

The Size and Number of Firms in an Industry

The shape of the long-run average cost curve has implications for how many firms will compete in an industry, and whether the firms in an industry have many different sizes, or tend to be the same size. For example, say that one million dishwashers are sold every year at a price of $\$500$ each and the long-run average cost curve for dishwashers is shown in Figure $6.4.3$ (a). In Figure $6.4.3$ (a), the lowest point of the $LRAC$ curve occurs at a quantity of $10,000$ produced. Thus, the market for dishwashers will consist of 100 different manufacturing plants of this same size. If some firms built a plant that produced $5,000$ dishwashers per year or $25,000$ dishwashers per year, the average costs of production at such plants would be well above $\$500$, and the firms would not be able to compete.

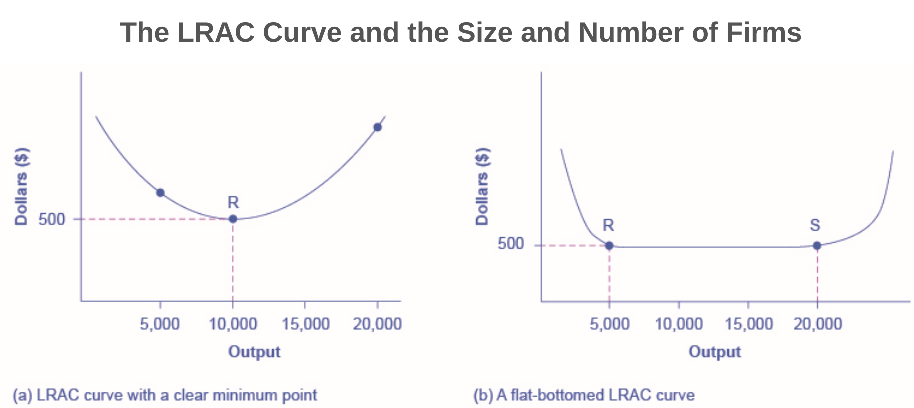


Figure $6.4.3$: *(a) Low-cost firms will produce at output level R. When the $LRAC$ curve has a clear minimum point, then any firm producing a different quantity will have higher costs. In this case, a firm producing at a quantity of 10,000 will produce at a lower average cost than a firm producing, say, 5,000 or 20,000 units. (b) Low-cost firms will produce between output levels R and S. When the $LRAC$ curve has a flat bottom, then firms producing at any quantity along this flat bottom can compete. In this case, any firm producing a quantity between 5,000 and 20,000 can compete effectively, although firms producing less than 5,000 or more than 20,000 would face higher average costs and be unable to compete.*

How can cities be viewed as examples of economies of scale?

Why are people and economic activity concentrated in cities, rather than distributed evenly across a country? The fundamental reason must be related to the idea of economies of scale—that grouping economic activity is more productive in many cases than spreading it out. For example, cities provide a large group of nearby customers, so that businesses can produce at an efficient economy of scale. They also provide a large group of workers and suppliers, so that business can hire easily and purchase whatever specialized inputs they need. Many of the attractions of cities, like sports stadiums and museums, can operate only if they can draw on a large nearby population base. Cities are big enough to offer a wide variety of products, which is what many shoppers are looking for.

These factors are not exactly economies of scale in the narrow sense of the production function of a single firm, but they are related to growth in the overall size of population and market in an area. Cities are sometimes called "*agglomeration economies*."

These agglomeration factors help to explain why every economy, as it develops, has an increasing proportion of its population living in urban areas. In the United States, about 80% of the population now lives in metropolitan areas (which include the suburbs around cities), compared to just 40% in 1900. However, in poorer nations of the world, including much of Africa, the proportion of the population in urban areas is only about 30%. One of the great challenges for these countries as their economies grow will be to manage the growth of the great cities that will arise.

If cities offer economic advantages that are a form of economies of scale, then why don't all or most people live in one giant city? At some point, agglomeration economies must turn into diseconomies. For example, traffic congestion may reach a point where the gains from being geographically nearby are counterbalanced by how long it takes to travel. High densities of people, cars, and factories can mean more garbage and air and water pollution. Facilities like parks or museums may become overcrowded. There may be economies of scale for negative activities like crime, because high densities of people and businesses, combined with the greater impersonality of cities, make it easier for illegal activities as well as legal ones. The future of cities, both in the United States and in other countries around the world, will be determined by their ability to benefit from the economies of agglomeration and to minimize or counterbalance the corresponding diseconomies.

A more common case is illustrated in Figure $6.4.3$ (b), where the $LRAC$ curve has a flat-bottomed area of constant returns to scale. In this situation, any firm with a level of output between $5,000$ and $20,000$ will be able to produce at about the same level of average cost. Given that the market will demand one million dishwashers per year at a price of $\$500$, this market might have as many as 200 producers (that is, one million dishwashers divided by firms making $5,000$ each) or as few as 50 producers (one million dishwashers divided by firms making $20,000$ each). The producers in this market will range in size from firms that make $5,000$ units to firms that make $20,000$ units. But firms that produce below $5,000$ units or more than $20,000$ will be unable to compete, because their average costs will be too high. Thus, if we see an industry where almost all plants are the same size, it is likely that the long-run average cost curve has a unique bottom point as in Figure $6.4.3$ (a). However, if the long-run average cost curve has a wide flat bottom like Figure $6.4.3$ (b), then firms of a variety of different sizes will be able to compete with each other.

The flat section of the long-run average cost curve in Figure $6.4.3$ (b) can be interpreted in two different ways. One interpretation is that a single manufacturing plant producing a quantity of $5,000$ has the same average costs as a single manufacturing plant with four times as much capacity that produces a quantity of $20,000$. The other interpretation is that one firm owns a single manufacturing plant that produces a quantity of $5,000$, while another firm owns four separate manufacturing plants, which each produce a quantity of $5,000$. This second explanation, based on the insight that a single firm may own a number of different manufacturing plants, is especially useful in explaining why the long-run average cost curve often has a large flat segment—and thus why a seemingly smaller firm may be able to compete quite well with a larger firm. At some point, however, the task of coordinating and managing many different plants raises the cost of production sharply, and the long-run average cost curve slopes up as a result.

In the examples to this point, the quantity demanded in the market is quite large (one million) compared with the quantity produced at the bottom of the long-run average cost curve ($5,000$, $10,000$ or $20,000$). In such a situation, the market is set for competition between many firms. But what if the bottom of the long-run average cost curve is at a quantity of 10,000 and the total market demand at that price is only slightly higher than that quantity—or even somewhat lower?

Return to Figure $6.4.3$ (a), where the bottom of the long-run average cost curve is at $10,000$, but now imagine that the total quantity of dishwashers demanded in the market at that price of $\$500$ is only $30,000$. In this situation, the total number of firms in the market would be three. A handful of firms in a market is called an "oligopoly," and the chapter on Monopolistic Competition and Oligopoly will discuss the range of competitive strategies that can occur when oligopolies compete.

Alternatively, consider a situation, again in the setting of Figure $6.4.3$ (a), where the bottom of the long-run average cost curve is $10,000$, but total demand for the product is only $5,000$. (For simplicity, imagine that this demand is highly inelastic, so that it does not vary according to price.) In this situation, the market may well end up with a single firm—a monopoly—producing all $5,000$ units. If any firm tried to challenge this monopoly while producing a quantity lower than $5,000$ units, the prospective competitor firm would have a higher average cost, and so it would not be able to compete in the longer term without losing money. The chapter on Monopoly discusses the situation of a monopoly firm.

Thus, the shape of the long-run average cost curve reveals whether competitors in the market will be different sizes. If the $LRAC$ curve has a single point at the bottom, then the firms in the market will be about the same size, but if the $LRAC$ curve has a flat-bottomed segment of constant returns to scale, then firms in the market may be a variety of different sizes.

The relationship between the quantity at the minimum of the long-run average cost curve and the quantity demanded in the market at that price will predict how much competition is likely to exist in the market. If the quantity demanded in the market far exceeds the quantity at the minimum of the $LRAC$, then many firms will compete. If the quantity demanded in the market is only slightly higher than the quantity at the minimum of the $LRAC$, a few firms will compete. If the quantity demanded in the market is less than the quantity at the minimum of the $LRAC$, a single-producer monopoly is a likely outcome.

Shifting Patterns of Long-Run Average Cost

New developments in production technology can shift the long-run average cost curve in ways that can alter the size distribution of firms in an industry.

For much of the twentieth century, the most common change has been to see alterations in technology, like the assembly line or the large department store, where large-scale producers seemed to gain an advantage over smaller ones. In the long-run average cost curve, the downward-sloping economies of scale portion of the curve stretched over a larger quantity of output.

However, new production technologies do not inevitably lead to a greater average size for firms. For example, in recent years some new technologies for generating electricity on a smaller scale have appeared. The traditional coal-burning electricity plants needed to produce 300 to 600 megawatts of power to exploit economies of scale fully. However, high-efficiency turbines to produce electricity from burning natural gas can produce electricity at a competitive price while producing a smaller quantity of 100 megawatts or less. These new technologies create the possibility for smaller companies or plants to generate electricity as efficiently as large ones. Another example of a technology-driven shift to smaller plants may be taking place in the tire industry. A traditional mid-size tire plant produces about six million tires per year. However, in 2000, the Italian company Pirelli introduced a new tire factory that uses many robots. The Pirelli tire plant produced only about one million tires per year, but did so at a lower average cost than a traditional mid-sized tire plant.

Controversy has simmered in recent years over whether the new information and communications technologies will lead to a larger or smaller size for firms. On one side, the new technology may make it easier for small firms to reach out beyond their local geographic area and find customers across a state, or the nation, or even across international boundaries. This factor might seem to predict a future with a larger number of small competitors. On the other side, perhaps the new information and communications technology will create "winner-take-all" markets where one large company will tend to command a large share of total sales, as Microsoft has done in the production of software for personal computers or Amazon has done in online bookselling. Moreover, improved information and communication technologies might make it easier to manage many different plants and operations across the country or around the world, and thus encourage larger firms. This ongoing battle between the forces of smallness and largeness will be of great interest to economists, businesspeople, and policymakers.

Amazon

Traditionally, bookstores have operated in retail locations with inventories held either on the shelves or in the back of the store. These retail locations were very pricey in terms of rent. Amazon has no retail locations; it sells online and delivers by mail. Amazon offers almost any book in print, convenient purchasing, and prompt delivery by mail. Amazon holds its inventories in huge warehouses in low-rent locations around the world. The warehouses are highly computerized using robots and relatively low-skilled workers, making for low average costs per sale. Amazon demonstrates the significant advantages economies of scale can offer to a firm that exploits those economies.

Key Concepts and Summary

A production technology refers to a specific combination of labor, physical capital, and technology that makes up a particular method of production.

In the long run, firms can choose their production technology, and so all costs become variable costs. In making this choice, firms will try to substitute relatively inexpensive inputs for relatively expensive inputs where possible, so as to produce at the lowest possible long-run average cost.

Economies of scale refers to a situation where as the level of output increases, the average cost decreases. Constant returns to scale refers to a situation where average cost does not change as output increases. Diseconomies of scale refers to a situation where as output increases, average costs increase also.

The long-run average cost curve shows the lowest possible average cost of production, allowing all the inputs to production to vary so that the firm is choosing its production technology. A downward-sloping $LRAC$ shows economies of scale; a flat $LRAC$ shows constant returns to scale; an upward-sloping $LRAC$ shows diseconomies of scale. If the long-run average cost curve has only one quantity produced that results in the lowest possible average cost, then all of the firms competing in an industry should be the same size. However, if the $LRAC$ has a flat segment at the bottom, so that a range of different quantities can be produced at the lowest average cost, the firms competing in the industry will display a range of sizes. The market demand in conjunction with the long-run average cost curve determines how many firms will exist in a given industry.

If the quantity demanded in the market of a certain product is much greater than the quantity found at the bottom of the long-run average cost curve, where the cost of production is lowest, the market will have many firms competing. If the quantity demanded in the market is less than the quantity at the bottom of the $LRAC$, there will likely be only one firm.

Glossary

constant returns to scale

expanding all inputs proportionately does not change the average cost of production

diseconomies of scale

the long-run average cost of producing each individual unit increases as total output increases

long-run average cost (LRAC) curve

shows the lowest possible average cost of production, allowing all the inputs to production to vary so that the firm is choosing its production technology

production technologies

alternative methods of combining inputs to produce output

short-run average cost (SRAC) curve

the average total cost curve in the short term; shows the total of the average fixed costs and the average variable costs

Contributor

- Template:ContribOpenSTAXMicroEcon

6.E: Production and Cost Structure of the Firm (Exercises)

7.1: Explicit and Implicit Costs, and Accounting and Economic Profit

Self-Check Questions

Q1

A firm had sales revenue of $\$1$ million last year. It spent $\$600,000$ on labor, $\$150,000$ on capital and $\$200,000$ on materials. What was the firm's accounting profit?

Q2

Continuing from **Q1**, the firm's factory sits on land owned by the firm that could be rented out for $\$30,000$ per year. What was the firm's economic profit last year?

Review Questions

Q3

What are explicit and implicit costs?

Q4

Would an interest payment on a loan to a firm be considered an explicit or implicit cost?

Q5

What is the difference between accounting and economic profit?

Critical Thinking Questions

Q6

Small "Mom and Pop firms," like inner city grocery stores, sometimes exist even though they do not earn economic profits. How can you explain this?

Problems

Q7

A firm is considering an investment that will earn a 6% rate of return. If it were to borrow the money, it would have to pay 8% interest on the loan, but it currently has the cash, so it will not need to borrow. Should the firm make the investment? Show your work.

Solution

S1

$$\begin{aligned}\text{Accounting profit} &= \text{total revenues minus explicit costs}\\ &= \$1,000,000-(\$600,000+\$150,000+\$200,000)\\ &= \$50,000\end{aligned}$$

S2

$$\begin{aligned}\text{Economic profit} &= \text{accounting profit minus implicit cost}\\ &= \$50,000-\$30,000\\ &= \$20,000\end{aligned}$$

7.2: The Structure of Costs in the Short Run

Self-Check Questions

Q1

The WipeOut Ski Company manufactures skis for beginners. Fixed costs are $30. Fill in Table below for total cost, average variable cost, average total cost, and marginal cost.

Quantity	Variable Cost	Fixed Cost	Total Cost	Average Variable Cost	Average Total Cost	Marginal Cost
0	0	$30				
1	$10	$30				
2	$25	$30				
3	$45	$30				
4	$70	$30				
5	$100	$30				
6	$135	$30				

Q2

Based on your answers to the WipeOut Ski Company in **Q1**, now imagine a situation where the firm produces a quantity of 5 units that it sells for a price of $\$25$ each.

a. What will be the company's profits or losses?
b. How can you tell at a glance whether the company is making or losing money at this price by looking at average cost?
c. At the given quantity and price, is the marginal unit produced adding to profits?

Review Questions

Q3

What is the difference between fixed costs and variable costs?

Q4

Are there fixed costs in the long-run? Explain briefly.

Q5

Are fixed costs also sunk costs? Explain.

Q6

What are diminishing marginal returns as they relate to costs?

Q7

Which costs are measured on per-unit basis: fixed costs, average cost, average variable cost, variable costs, and marginal cost?

Q8

How is each of the following calculated: marginal cost, average total cost, average variable cost?

Critical Thinking Questions

Q9

A common name for fixed cost is “overhead.” If you divide fixed cost by the quantity of output produced, you get average fixed cost. Supposed fixed cost is $\$1,000$. What does the average fixed cost curve look like? Use your response to explain what “spreading the overhead” means.

Q10

How does fixed cost affect marginal cost? Why is this relationship important?

Q11

Average cost curves (except for average fixed cost) tend to be U-shaped, decreasing and then increasing. Marginal cost curves have the same shape, though this may be harder to see since most of the marginal cost curve is increasing. Why do you think that average and marginal cost curves have the same general shape?

Problems

Q12

Return to Figure 7.2.1. What is the marginal gain in output from increasing the number of barbers from 4 to 5 and from 5 to 6? Does it continue the pattern of diminishing marginal returns?

Q13

Compute the average total cost, average variable cost, and marginal cost of producing 60 and 72 haircuts. Draw the graph of the three curves between 60 and 72 haircuts.

Solution

S1

Quantity	Variable Cost	Fixed Cost	Total Cost	Average Variable Cost	Average Total Cost	Marginal Cost
0	0	$30	$30	-	-	

Quantity	Variable Cost	Fixed Cost	Total Cost	Average Variable Cost	Average Total Cost	Marginal Cost
1	\$10	\$30	\$40	\$10.00	\$40.00	\$10
2	\$25	\$30	\$55	\$12.50	\$27.50	\$15
3	\$45	\$30	\$75	\$15.00	\$25.00	\$20
4	\$70	\$30	\$100	\$17.50	\$25.00	\$25
5	\$100	\$30	\$130	\$20.00	\$26.00	\$30
6	\$135	\$30	\$165	\$22.50	\$27.50	\$35

S2

1. Total revenues in this example will be a quantity of five units multiplied by the price of $\$25/unit$, which equals $\$125$. Total costs when producing five units are $\$130$. Thus, at this level of quantity and output the firm experiences losses (or negative profits) of $\$5$.
2. If price is less than average cost, the firm is not making a profit. At an output of five units, the average cost is $\$26/unit$. Thus, at a glance you can see the firm is making losses. At a second glance, you can see that it must be losing $\$1$ for each unit produced (that is, average cost of $\$26/unit$ minus the price of $\$25/unit$). With five units produced, this observation implies total losses of $\$5$.
3. When producing five units, marginal costs are $\$30/unit$. Price is $\$25/unit$. Thus, the marginal unit is not adding to profits, but is actually subtracting from profits, which suggests that the firm should reduce its quantity produced.

7.3: The Structure of Costs in the Long Run

Self-Check Questions

Q1

Return to the problem explained in Table 7.3.1 and Table 7.3.2. If the cost of labor remains at $\$40$, but the cost of a machine decreases to $\$50$, what would be the total cost of each method of production? Which method should the firm use, and why?

Q2

Suppose the cost of machines increases to $\$55$, while the cost of labor stays at $\$40$. How would that affect the total cost of the three methods? Which method should the firm choose now?

Q3

Automobile manufacturing is an industry subject to significant economies of scale. Suppose there are four domestic auto manufacturers, but the demand for domestic autos is no more than 2.5 times the quantity produced at the bottom of the long-run average cost curve. What do you expect will happen to the domestic auto industry in the long run?

Review Questions

Q4

What shapes would you generally expect each of the following cost curves to have: fixed costs, variable costs, marginal costs, average total costs, and average variable costs?

Q5

What is a production technology?

Q6

In choosing a production technology, how will firms react if one input becomes relatively more expensive?

Q7

What is a long-run average cost curve?

Q8

What is the difference between economies of scale, constant returns to scale, and diseconomies of scale?

Q9

What shape of a long-run average cost curve illustrates economies of scale, constant returns to scale, and diseconomies of scale?

Q10

Why will firms in most markets be located at or close to the bottom of the long-run average cost curve?

Critical Thinking Questions

Q11

It is clear that businesses operate in the short run, but do they ever operate in the long run? Discuss.

Q12

How would an improvement in technology, like the high-efficiency gas turbines or Pirelli tire plant, affect the long-run average cost curve of a firm? Can you draw the old curve and the new one on the same axes? How might such an improvement affect other firms in the industry?

Q13

Do you think that the taxicab industry in large cities would be subject to significant economies of scale? Why or why not?

Problems

Q14

A small company that shovels sidewalks and driveways has 100 homes signed up for its services this winter. It can use various combinations of capital and labor: lots of labor with hand shovels, less labor with snow blowers, and still less labor with a pickup truck that has a snowplow on front. To summarize, the method choices are:

Method 1: 50 units of labor, 10 units of capital

Method 2: 20 units of labor, 40 units of capital

Method 3: 10 units of labor, 70 units of capital

If hiring labor for the winter costs $\$100/unit$ and a unit of capital costs $\$400$, what production method should be chosen? What method should be chosen if the cost of labor rises to $\$200/unit$?

Solution

S1

The new table should look like this:

	Labor Cost	Machine Cost	Total Cost
Cost of technology 1	10 × $40 = $400	2 × $50 = $100	$500
Cost of technology 2	7 × $40 = $280	4 × $50 = $200	$480
Cost of technology 3	3 × $40 = $120	7 × $50 = $350	$470

The firm should choose production technology 3 since it has the lowest total cost. This makes sense since, with cheaper machine hours, one would expect a shift in the direction of more machines and less labor.

S2

	Labor Cost	Machine Cost	Total Cost
Cost of technology 1	10 × $40 = $400	2 × $55 = $110	$510
Cost of technology 2	7 × $40 = $280	4 × $55 = $220	$500
Cost of technology 3	3 × $40 = $120	7 × $55 = $385	$505

The firm should choose production technology 2 since it has the lowest total cost. Because the cost of machines increased (relative to the previous question), you would expect a shift toward less capital and more labor.

S3

This is the situation that existed in the United States in the 1970s. Since there is only demand enough for 2.5 firms to reach the bottom of the average cost curve, you would expect one firm will not be around in the long run, and at least one firm will be struggling.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

7: Perfect Competition

7.1: Prelude to Perfect Competition

Learning Objectives

- Perfect Competition and Why It Matters
- How Perfectly Competitive Firms Make Output Decisions
- Entry and Exit Decisions in the Long Run
- Efficiency in Perfectly Competitive Markets

Figure 7.1.1: Depending upon the competition and prices offered, a wheat farmer may choose to grow a different crop. (Credit: modification of work by Daniel X. O'Neil/Flickr Creative Commons)

A Dime a Dozen

When you were younger did you babysit, deliver papers, or mow the lawn for money? If so, you faced stiff competition from a lot of other competitors who offered identical services. There was nothing to stop others from offering their services too.

All of you charged the "going rate." If you tried to charge more, your customers would simply buy from someone else. These conditions are very similar to the conditions agricultural growers face.

Growing a crop may be more difficult to start than a babysitting or lawn mowing service, but growers face the same fierce competition. In the grand scale of world agriculture, farmers face competition from thousands of others because they sell an identical product. After all, winter wheat is winter wheat. But it is relatively easy for farmers to leave the marketplace for another crop. In this case, they do not sell the family farm, they switch crops.

Take the case of the upper Midwest region of the United States—for many generations the area was called "King Wheat." According to the United States Department of Agriculture National Agricultural Statistics Service, statistics by state, in 1997, 11.6 million acres of wheat and 780,000 acres of corn were planted in North Dakota. In the intervening 15 or so years has the mix of crops changed? Since it is relatively easy to switch crops, did farmers change what was planted as the relative crop prices changed? We will find out at chapter's end.

In the meantime, let's consider the topic of this chapter—the perfectly competitive market. This is a market in which entry and exit are relatively easy and competitors are "a dime a dozen."

All businesses face two realities: no one is required to buy their products, and even customers who might want those products may buy from other businesses instead. Firms that operate in perfectly competitive markets face this reality. In this chapter, you will learn how such firms make decisions about how much to produce, how much profit they make, whether to stay in business or not, and many others. Industries differ from one another in terms of how many sellers there are in a specific market, how easy or difficult it is for a new firm to enter, and the type of products that are sold. This is referred to as the **market structure** of the industry. In this chapter, we focus on perfect competition. However, in other chapters we will examine other industry types: Monopoly and Monopolistic Competition and Oligopoly.

Contributor

- Template:ContribOpenSTAXMicroEcon

7.2: Perfect Competition and Why It Matters

Skills to Develop

- Explain the characteristics of a perfectly competitive market
- Discuss how perfectly competitive firms react in the short run and in the long run

Firms are said to be in **perfect competition** when the following conditions occur:

1. many firms produce identical products
2. many buyers are available to buy the product, and many sellers are available to sell the product
3. sellers and buyers have all relevant information to make rational decisions about the product being bought and sold
4. firms can enter and leave the market without any restrictions—in other words, there is free entry and exit into and out of the market.

A **perfectly competitive firm** is known as a **price taker**, because the pressure of competing firms forces them to accept the prevailing equilibrium price in the market. If a firm in a perfectly competitive market raises the price of its product by so much as a penny, it will lose all of its sales to competitors. When a wheat grower, as discussed in the Bring it Home feature, wants to know what the going price of wheat is, he or she has to go to the computer or listen to the radio to check. The market price is determined solely by supply and demand in the entire market and not the individual farmer. Also, a perfectly competitive firm must be a very small player in the overall market, so that it can increase or decrease output without noticeably affecting the overall quantity supplied and price in the market.

A perfectly competitive market is a hypothetical extreme; however, producers in a number of industries do face many competitor firms selling highly similar goods, in which case they must often act as price takers. Agricultural markets are often used as an example. The same crops grown by different farmers are largely interchangeable. According to the United States Department of Agriculture monthly reports, in 2015, U.S. corn farmers received an average price of $\$6.00$ per bushel and wheat farmers received an average price of $\$6.00$ per bushel. A corn farmer who attempted to sell at $\$7.00$ per bushel, or a wheat grower who attempted to sell for $\$8.00$ per bushel, would not have found any buyers. A perfectly competitive firm will not sell below the equilibrium price either. Why should they when they can sell all they want at the higher price? Other examples of agricultural markets that operate in close to perfectly competitive markets are small roadside produce markets and small organic farmers.

This chapter examines how profit-seeking firms decide how much to produce in perfectly competitive markets. Such firms will analyze their costs as discussed in the chapter on Cost and Industry Structure. In the short run, the perfectly competitive firm will seek the quantity of output where profits are highest or, if profits are not possible, where losses are lowest. In this example, the "short run" refers to a situation in which firms are producing with one fixed input and incur fixed costs of production. (In the real world, firms can have many fixed inputs.)

In the long run, perfectly competitive firms will react to profits by increasing production. They will respond to losses by reducing production or exiting the market. Ultimately, a long-run *equilibrium* will be attained when no new firms want to enter the market and existing firms do not want to leave the market, as economic profits have been driven down to zero.

Key Concepts and Summary

A perfectly competitive firm is a price taker, which means that it must accept the equilibrium price at which it sells goods. If a perfectly competitive firm attempts to charge even a tiny amount more than the market price, it will be unable to make any sales. In a perfectly competitive market there are thousands of sellers, easy entry, and identical products. A short-run production period is when firms are producing with some fixed inputs. Long-run equilibrium in a perfectly competitive industry occurs after all firms have entered and exited the industry and seller profits are driven to zero.

Perfect competition means that there are many sellers, there is easy entry and exiting of firms, products are identical from one seller to another, and sellers are price takers.

Glossary

market structure

: the conditions in an industry, such as number of sellers, how easy or difficult it is for a new firm to enter, and the type of products that are sold

perfect competition

each firm faces many competitors that sell identical products

price taker

a firm in a perfectly competitive market that must take the prevailing market price as given

Contributor

- Template:ContribOpenSTAXMicroEcon

7.3: How Perfectly Competitive Firms Make Output Decisions

Learning Objectives

- Calculate profits by comparing total revenue and total cost
- Identify profits and losses with the average cost curve
- Explain the shutdown point
- Determine the price at which a firm should continue producing in the short run

A **perfectly competitive firm** has only one major decision to make—namely, what quantity to produce. To understand why this is so, consider a different way of writing out the basic definition of **profit**:

$$\begin{aligned} \text{Profit } &= \text{Total revenue} - \text{Total cost} \\ &= (\text{Price})(\text{Quantity produced}) - (\text{Average cost})(\text{Quantity produced}) \end{aligned}$$

Since a perfectly competitive firm must accept the price for its output as determined by the product's market demand and supply, it cannot choose the price it charges. This is already determined in the profit equation, and so the perfectly competitive firm can sell any number of units at exactly the same price. It implies that the firm faces a perfectly elastic demand curve for its product: buyers are willing to buy any number of units of output from the firm at the market price. When the perfectly competitive firm chooses what quantity to produce, then this quantity—along with the prices prevailing in the market for output and inputs—will determine the firm's total revenue, total costs, and ultimately, level of profits.

Determining the Highest Profit by Comparing Total Revenue and Total Cost

A perfectly competitive firm can sell as large a quantity as it wishes, as long as it accepts the prevailing market price. Total revenue is going to increase as the firm sells more, depending on the price of the product and the number of units sold. If you increase the number of units sold at a given price, then total revenue will increase. If the price of the product increases for every unit sold, then total revenue also increases. As an example of how a perfectly competitive firm decides what quantity to produce, consider the case of a small farmer who produces raspberries and sells them frozen for $\$4$ per pack. Sales of one pack of raspberries will bring in $\$4$, two packs will be $\$8$, three packs will be $\$12$, and so on. If, for example, the price of frozen raspberries doubles to $\$8$ per pack, then sales of one pack of raspberries will be $\$8$, two packs will be $\$16$, three packs will be $\$24$, and so on.

Total revenue and total costs for the raspberry farm, broken down into fixed and variable costs, are shown in Table $7.3.1$ and also appear in Figure $7.3.1$. The horizontal axis shows the quantity of frozen raspberries produced in packs; the vertical axis shows both total revenue and total costs, measured in dollars. The total cost curve intersects with the vertical axis at a value that shows the level of fixed costs, and then slopes upward. All these cost curves follow the same characteristics as the curves covered in the Cost and Industry Structure chapter.

Total Cost and Total Revenue at the Raspberry Farm

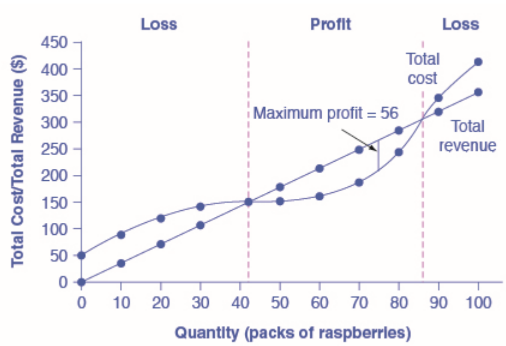


Figure $7.3.1$: Total revenue for a perfectly competitive firm is a straight line sloping up. The slope is equal to the price of the good. Total cost also slopes up, but with some curvature. At higher levels of output, total cost begins to slope upward more steeply because of diminishing marginal returns. The maximum profit will occur at the quantity where the gap of total revenue over total cost is largest.

Table $7.3.1$: Total Cost and Total Revenue at the Raspberry Farm

Quantity (Q)	Total Cost (TC)	Fixed Cost (FC)	Variable Cost (VC)	Total Revenue (TR)	Profit
0	\$62	\$62	-	\$0	−\$62
10	\$90	\$62	\$28	\$40	−\$50
20	\$110	\$62	\$48	\$80	−\$30
30	\$126	\$62	\$64	\$120	−\$6
40	\$144	\$62	\$82	\$160	\$16
50	\$166	\$62	\$104	\$200	\$34
60	\$192	\$62	\$130	\$240	\$48
70	\$224	\$62	\$162	\$280	\$56
80	\$264	\$62	\$202	\$320	\$56
90	\$324	\$62	\$262	\$360	\$36
100	\$404	\$62	\$342	\$400	−\$4

Based on its total revenue and total cost curves, a perfectly competitive firm like the raspberry farm can calculate the quantity of output that will provide the highest level of profit. At any given quantity, total revenue minus total cost will equal profit. One way to determine the most profitable quantity to produce is to see at what quantity total revenue exceeds total cost by the largest amount. On Figure $7.3.1$, the vertical gap between total revenue and total cost represents either profit (if total revenues are greater that total costs at a certain quantity) or losses (if total costs are greater that total revenues at a certain quantity). In this example, total costs will exceed total revenues at output levels from 0 to 40, and so over this range of output, the firm will be making losses. At output levels from 50 to 80, total revenues exceed total costs, so the firm is earning profits. But then at an output of 90 or 100, total costs again exceed total revenues and the firm is making losses. Total profits appear in the final column of Table $7.3.1$. The highest total profits in the table, as in the figure that is based on the table values, occur at an output of 70–80, when profits will be $\$56$.

A higher price would mean that total revenue would be higher for every quantity sold. A lower price would mean that total revenue would be lower for every quantity sold. What happens if the price drops low enough so that the total revenue line is completely below the total cost curve; that is, at every level of output, total costs are higher than total revenues? In this instance, the best the firm can do is to suffer losses. But a profit-maximizing firm will prefer the quantity of output where total revenues come closest to total costs and thus where the losses are smallest.

(Later we will see that sometimes it will make sense for the firm to shutdown, rather than stay in operation producing output.)

Comparing Marginal Revenue and Marginal Costs

Firms often do not have the necessary data they need to draw a complete total cost curve for all levels of production. They cannot be sure of what total costs would look like if they, say, doubled production or cut production in half, because they have not tried it. Instead, firms experiment. They produce a slightly greater or lower quantity and observe how profits are affected. In economic terms, this practical approach to maximizing profits means looking at how changes in production affect marginal revenue and marginal cost.

Figure $7.3.2$ presents the marginal revenue and marginal cost curves based on the total revenue and total cost in Table $7.3.1$. The marginal revenue curve shows the additional revenue gained from selling one more unit. As mentioned before, a firm in perfect competition faces a perfectly elastic demand curve for its product—that is, the firm's demand curve is a horizontal line drawn at the market price level. This also means that the firm's marginal revenue curve is the same as the firm's demand curve: Every time a consumer demands one more unit, the firm sells one more unit and revenue goes up by exactly the same amount equal to the market price. In this example, every time a pack of frozen raspberries is sold, the firm's revenue increases by $\$4$. Table $7.3.2$ shows an example of this. This condition only holds for price taking firms in perfect competition where:

$$\text{marginal revenue} = \text{price} \tag{7.3.1}$$

The formula for marginal revenue is:

$$\text{marginal revenue} = \frac{\text{change in total revenue}}{\text{change in quantity}} \tag{7.3.2}$$

Table 7.3.2: Marginal Revenue at the Raspberry farm

Price	Quantity	Total Revenue	Marginal Revenue
$4	1	$4	-
$4	2	$8	$4
$4	3	$12	$4
$4	4	$16	$4

Notice that marginal revenue does not change as the firm produces more output. That is because the price is determined by supply and demand and does not change as the farmer produces more (keeping in mind that, due to the relative small size of each firm, increasing their supply has no impact on the total market supply where price is determined).

Since a perfectly competitive firm is a price taker, it can sell whatever quantity it wishes at the market-determined price. Marginal cost, the cost per additional unit sold, is calculated by dividing the change in total cost by the change in quantity. The formula for marginal cost is:

$$\text{marginal cost} = \frac{\text{change in total cost}}{\text{change in quantity}} \tag{7.3.3}$$

Ordinarily, marginal cost changes as the firm produces a greater quantity.

In the raspberry farm example, shown in Figure 7.3.2, Figure 7.3.3 and Table 7.3.3, marginal cost at first declines as production increases from 10 to 20 to 30 packs of raspberries—which represents the area of increasing marginal returns that is not uncommon at low levels of production. But then marginal costs start to increase, displaying the typical pattern of diminishing marginal returns. If the firm is producing at a quantity where $MR > MC$, like 40 or 50 packs of raspberries, then it can increase profit by increasing output because the marginal revenue is exceeding the marginal cost. If the firm is producing at a quantity where $MC > MR$, like 90 or 100 packs, then it can increase profit by reducing output because the reductions in marginal cost will exceed the reductions in marginal revenue. The firm's profit-maximizing choice of output will occur where $MR = MC$ (or at a choice close to that point). You will notice that what occurs on the production side is exemplified on the cost side. This is referred to as duality.

Marginal Revenues and Marginal Costs at the Raspberry Farm: Individual Farmer

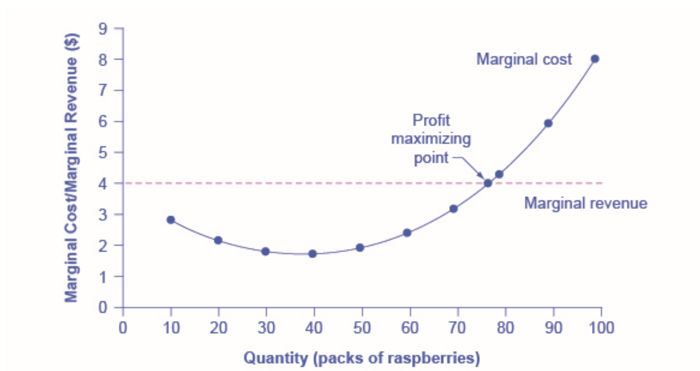


Figure 7.3.3. The marginal cost (MC) curve is sometimes first downward-sloping, if there is a region of increasing marginal returns at low levels of output, but is eventually upward-sloping at higher levels of output as diminishing marginal returns kick in.

Marginal Revenues and Marginal Costs at the Raspberry Farm: Raspberry Market

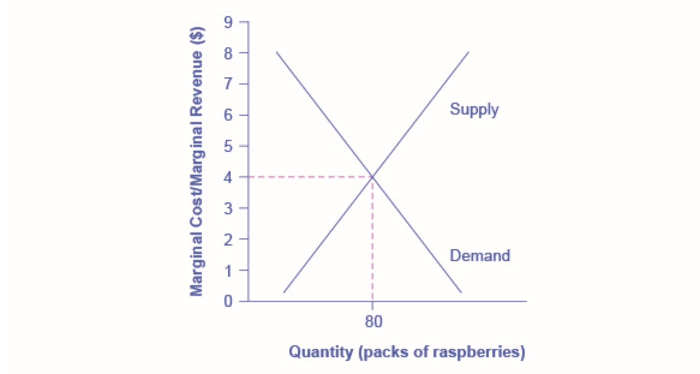


Figure 7.3.3: The equilibrium price of raspberries is determined through the interaction of market supply and market demand at $4.00.

Table 7.3.3: Marginal Revenues and Marginal Costs at the Raspberry Farm

Quantity	Total Cost	Fixed Cost	Variable Cost	Marginal Cost	Total Revenue	Marginal Revenue
0	$62	$62	-	-	-	-
10	$90	$62	$28	$2.80	$40	$4.00
20	$110	$62	$48	$2.00	$80	$4.00
30	$126	$62	$64	$1.60	$120	$4.00
40	$144	$62	$82	$1.80	$160	$4.00
50	$166	$62	$104	$2.20	$200	$4.00
60	$192	$62	$130	$2.60	$240	$4.00
70	$224	$62	$162	$3.20	$280	$4.00
80	$264	$62	$202	$4.00	$320	$4.00
90	$324	$62	$262	$6.00	$360	$4.00
100	$404	$62	$342	$8.00	$400	$4.00

In this example, the marginal revenue and **marginal cost** curves cross at a price of $4 and a quantity of 80 produced. If the farmer started out producing at a level of 60, and then experimented with increasing production to 70, marginal revenues from the increase in production would exceed marginal costs—and so profits would rise. The farmer has an incentive to keep producing. From a level of 70 to 80, marginal cost and marginal revenue are equal so profit doesn't change. If the farmer then experimented further with increasing production from 80 to 90, he would find that marginal costs from the increase in production are greater than marginal revenues, and so profits would decline.

The profit-maximizing choice for a perfectly competitive firm will occur where marginal revenue is equal to marginal cost—that is, where $MR = MC$. A profit-seeking firm should keep expanding production as long as $MR > MC$. But at the level of output where $MR = MC$, the firm should recognize that it has achieved the highest possible level of economic profits. (In the example above, the profit maximizing output level is between 70 and 80 units of output, but the firm will not know they've maximized profit until they reach 80, where $MR = MC$.) Expanding production into the zone where $MR < MC$ will only reduce economic profits. Because the marginal revenue received by a perfectly competitive firm is equal to the price P, so that $P = MR$, the profit-maximizing rule for a perfectly competitive firm can also be written as a recommendation to produce at the quantity where $P = MC$.

Profits and Losses with the Average Cost Curve

Does maximizing profit (producing where $MR = MC$) imply an actual economic profit? The answer depends on the relationship between price and average total cost. If the price that a firm charges is higher than its average cost of production for that quantity produced, then the firm will earn profits. Conversely, if the price that a firm charges is lower than its average cost of production, the firm will suffer losses. You might think that, in this situation, the farmer may want to shut down immediately. Remember, however, that the firm has already paid for fixed costs, such as equipment, so it may continue to produce and incur a loss.

Figure $7.3.4$ illustrates three situations:

a. where price intersects marginal cost at a level above the average cost curve
b. where price intersects marginal cost at a level equal to the average cost curve
c. where price intersects marginal cost at a level below the average cost curve.

Price and Average Cost at the Raspberry Farm

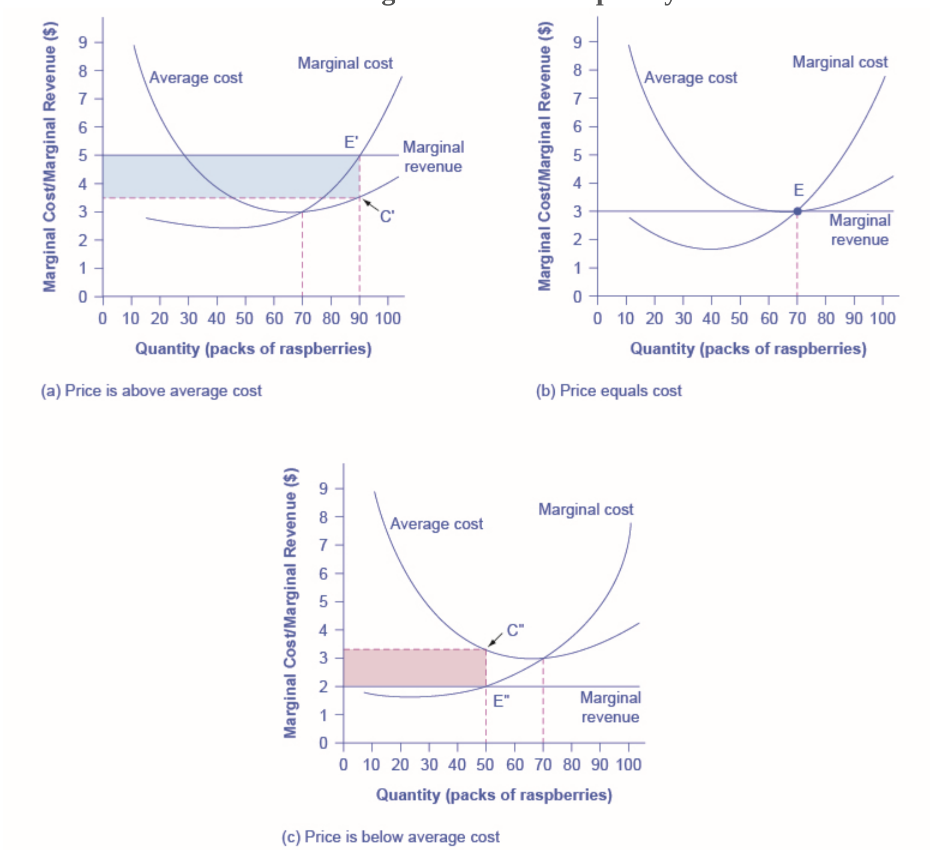


Figure $7.3.4$: In (a), price intersects marginal cost above the average cost curve. Since price is greater than average cost, the firm is making a profit. In (b), price intersects marginal cost at the minimum point of the average cost curve. Since price is equal to average cost, the firm is breaking even. In (c), price intersects marginal cost below the average cost curve. Since price is less than average cost, the firm is making a loss.

First consider a situation where the price is equal to $\$5$ for a pack of frozen raspberries. The rule for a profit-maximizing perfectly competitive firm is to produce the level of output where $\text{Price} = MR = MC$, so the raspberry farmer will produce a quantity of 90, which is labeled as E in Figure $7.3.4$ (a). Remember that the area of a rectangle is equal to its base multiplied by its height. The farm's total revenue at this price will be shown by the large shaded rectangle from the origin over to a quantity of 90 packs (the base) up to point E' (the height), over to the price of $\$5$, and back to the origin. The average cost of producing 80 packs is shown by point C or about $\$3.50$. Total costs will be the quantity of 80 times the average cost of $\$3.50$, which is shown by the area of the rectangle from the origin to a quantity of 90, up to point C, over to the vertical axis and down to the origin. It should be clear from examining the two rectangles that total revenue is greater than total cost. Thus, profits will be the blue shaded rectangle on top.

It can be calculated as:

$$\begin{aligned} \text{profit} &= \text{total revenue} - \text{total cost} \\ &= (90)(\$5.00) - (90)(\$3.50) \\ &= \$135 \end{aligned}$$

Or, it can be calculated as:

$$\begin{aligned} \text{profit} &= (\text{price-average cost}) \times \text{quantity} \\ &= (\$5.00 - \$3.50) \times 90 \\ &= \$135 \end{aligned}$$

Now consider Figure $7.3.4$ (b), where the price has fallen to $\$3.00$ for a pack of frozen raspberries. Again, the perfectly competitive firm will choose the level of output where $\text{Price} = MR = MC$, but in this case, the quantity produced will be 70. At this price and output level, where the marginal cost curve is crossing the average cost curve, the price received by the firm is exactly equal to its average cost of production.

The farm's total revenue at this price will be shown by the large shaded rectangle from the origin over to a quantity of 70 packs (the base) up to point E (the height), over to the price of $\$3$, and back to the origin. The average cost of producing 70 packs is shown by point C'. Total costs will be the quantity of 70 times the average cost of $\$3.00$, which is shown by the area of the rectangle from the origin to a quantity of 70, up to point E, over to the vertical axis and down to the origin. It should be clear from that the rectangles for total revenue and total cost are the same. Thus, the firm is making zero profit. The calculations are as follows:

$$\begin{aligned} \text{profit} &= \text{total revenue} - \text{total cost} \\ &= (70)(\$3.00) - (70)(\$3.00) \\ &= \$0 \end{aligned}$$

Or, it can be calculated as:

$$\begin{aligned} \text{profit} &= (\text{price-average cost}) \times \text{quantity} \\ &= (\$3.00 - \$3.00) \times 70 \\ &= \$0 \end{aligned}$$

In Figure $7.3.4$ (c), the market price has fallen still further to $\$2.00$ for a pack of frozen raspberries. At this price, marginal revenue intersects marginal cost at a quantity of 50. The farm's total revenue at this price will be shown by the large shaded rectangle from the origin over to a quantity of 50 packs (the base) up to point E" (the height), over to the price of $\$2$, and back to the origin. The average cost of producing 50 packs is shown by point C" or about $\$3.30$. Total costs will be the quantity of 50 times the average cost of $\$3.30$, which is shown by the area of the rectangle from the origin to a quantity of 50, up to point C", over to the vertical axis and down to the origin. It should be clear from examining the two rectangles that total revenue is less than total cost. Thus, the firm is losing money and the loss (or negative profit) will be the rose-shaded rectangle.

The calculations are:

$$\begin{aligned} \text{profit} &= \text{total revenue} - \text{total cost} \\ &= (50)(\$2.00) - (50)(\$3.30) \\ &= -\$77.50 \end{aligned}$$

Or:

$$\begin{aligned} \text{profit} &= (\text{price-average cost}) \times \text{quantity} \\ &= (\$1.75 - \$3.30) \times 50 \\ &= -\$77.50 \end{aligned}$$

If the market price received by a perfectly competitive firm leads it to produce at a quantity where the price is greater than average cost, the firm will earn profits. If the price received by the firm causes it to produce at a quantity where price equals average cost, which occurs at the minimum point of the AC curve, then the firm earns zero profits. Finally, if the price received by the firm leads it to produce at a quantity where the price is less than average cost, the firm will earn losses. This is summarized in Table $7.3.4$.

Table 7.3.4: Profit and Loss for Raspberry Farms

If...	Then...
Price > ATC	Firm earns an economic profit
Price = ATC	Firm earns zero economic profit
Price < ATC	Firm earns a loss

The Shutdown Point

The possibility that a firm may earn losses raises a question: Why can the firm not avoid losses by shutting down and not producing at all? The answer is that shutting down can reduce variable costs to zero, but in the short run, the firm has already paid for fixed costs. As a result, if the firm produces a quantity of zero, it would still make losses because it would still need to pay for its fixed costs. So, when a firm is experiencing losses, it must face a question: should it continue producing or should it shut down?

As an example, consider the situation of the Yoga Center, which has signed a contract to rent space that costs $\$10,000$ per month. If the firm decides to operate, its marginal costs for hiring yoga teachers is $\$15,000$ for the month. If the firm shuts down, it must still pay the rent, but it would not need to hire labor. Table $7.3.5$ shows three possible scenarios. In the first scenario, the Yoga Center does not have any clients, and therefore does not make any revenues, in which case it faces losses of $\$10,000$ equal to the fixed costs. In the second scenario, the Yoga Center has clients that earn the center revenues of $\$10,000$ for the month, but ultimately experiences losses of $\$15,000$ due to having to hire yoga instructors to cover the classes. In the third scenario, the Yoga Center earns revenues of $\$20,000$ for the month, but experiences losses of $\$5,000$.

In all three cases, the Yoga Center loses money. In all three cases, when the rental contract expires in the long run, assuming revenues do not improve, the firm should exit this business. In the short run, though, the decision varies depending on the level of losses and whether the firm can cover its variable costs. In scenario 1, the center does not have any revenues, so hiring yoga teachers would increase variable costs and losses, so it should shut down and only incur its fixed costs. In scenario 2, the center's losses are greater because it does not make enough revenue to offset the increased variable costs plus fixed costs, so it should shut down immediately. If price is below the minimum average variable cost, the firm must shut down. In contrast, in scenario 3 the revenue that the center can earn is high enough that the losses diminish when it remains open, so the center should remain open in the short run.

Table $7.3.5$: Should the Yoga Center Shut Down Now or Later?

Scenario 1
If the center shuts down now, revenues are zero but it will not incur any variable costs and would only need to pay fixed costs of \$10,000.
$\begin{aligned} \text{profit} &= (\text{total revenue}) - (\text{fixed costs} + \text{variable cost}) \\ &= 0 - \$10,000 \\ &= -\$10,000 \end{aligned}$
Scenario 2
The center earns revenues of \$10,000, and variable costs are \$15,000. The center should shut down now.
$\begin{aligned} \text{profit} &= (\text{total revenue}) - (\text{fixed costs} + \text{variable cost}) \\ &= \$10,000 - (\$10,000 + \$15,000) \\ &= -\$15,000 \end{aligned}$
Scenario 3
The center earns revenues of \$20,000, and variable costs are \$15,000. The center should continue in business.
$\begin{aligned} \text{profit} &= (\text{total revenue}) - (\text{fixed costs} + \text{variable cost}) \\ &= \$20,000 - (\$10,000 + \$15,000) \\ &= -\$5,000 \end{aligned}$

This example suggests that the key factor is whether a firm can earn enough revenues to cover at least its variable costs by remaining open. Let's return now to our raspberry farm. Figure $7.3.5$ illustrates this lesson by adding the average variable cost curve to the marginal cost and average cost curves. At a price of $\$2.20$ per pack, as shown in Figure $7.3.5$ (a), the farm produces at a level of 50. It is making losses of $\$56$ (as explained earlier), but price is above average variable cost and so the firm continues to operate. However, if the price declined to $\$1.80$ per pack, as shown in Figure $7.3.5$ (b), and if the firm applied its rule of producing where $P = MR = MC$, it would produce a quantity of 40. This price is below average variable cost for this level of output. If the farmer cannot pay workers (the variable costs), then it has to shut down. At this price and output, total revenues would be $\$72$

(quantity of 40 times price of $1.80) and total cost would be $144, for overall losses of $72. If the farm shuts down, it must pay only its fixed costs of $62, so shutting down is preferable to selling at a price of $1.80 per pack.

The Shutdown Point for the Raspberry Farm

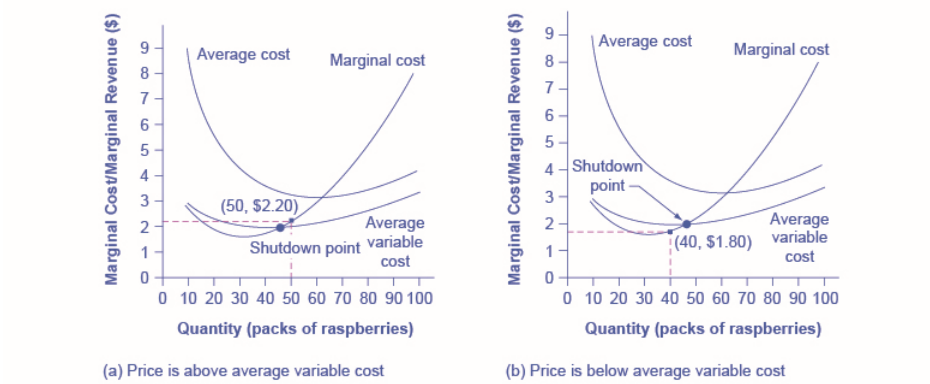


Figure 7.3.5: In (a), the farm produces at a level of 50. It is making losses of $56, but price is above average variable cost, so it continues to operate. In (b), total revenues are $72 and total cost is $144, for overall losses of $72. If the farm shuts down, it must pay only its fixed costs of $62. Shutting down is preferable to selling at a price of $1.80 per pack.

Looking at Table 7.3.6, if the price falls below $2.05, the minimum average variable cost, the firm must shut down.

Table 7.3.6: Cost of Production for the Raspberry Farm

Quantity	Total Cost	Fixed Cost	Variable Cost	Marginal Cost	Average Cost	Average Variable Cost
0	$62	$62	-	-	-	-
10	$90	$62	$28	$2.80	$9.00	$2.80
20	$110	$62	$48	$2.00	$5.50	$2.40
30	$126	$62	$64	$1.60	$4.20	$2.13
40	$144	$62	$82	$1.80	$3.60	$2.05
50	$166	$62	$104	$2.20	$3.32	$2.08
60	$192	$62	$130	$2.60	$3.20	$2.16
70	$224	$62	$162	$3.20	$3.20	$2.31
80	$264	$62	$202	$4.00	$3.30	$2.52
90	$324	$62	$262	$6.00	$3.60	$2.91
100	$404	$62	$342	$8.00	$4.04	$3.42

The intersection of the average variable cost curve and the marginal cost curve, which shows the price where the firm would lack enough revenue to cover its variable costs, is called the **shutdown point**. If the perfectly competitive firm can charge a price above the shutdown point, then the firm is at least covering its average variable costs. It is also making enough revenue to cover at least a portion of fixed costs, so it should limp ahead even if it is making losses in the short run, since at least those losses will be smaller than if the firm shuts down immediately and incurs a loss equal to total fixed costs. However, if the firm is receiving a price below the price at the shutdown point, then the firm is not even covering its variable costs. In this case, staying open is making the firm's losses larger, and it should shut down immediately. To summarize, if:

- price < minimum average variable cost, then firm shuts down
- price = minimum average variable cost, then firm stays in business

Short-Run Outcomes for Perfectly Competitive Firms

The average cost and average variable cost curves divide the marginal cost curve into three segments, as shown in Figure 7.3.6. At the market price, which the perfectly competitive firm accepts as given, the profit-maximizing firm chooses the output level where price or marginal revenue, which are the same thing for a perfectly competitive firm, is equal to marginal cost: $P = MR = MC$.

Profit, Loss, Shutdown

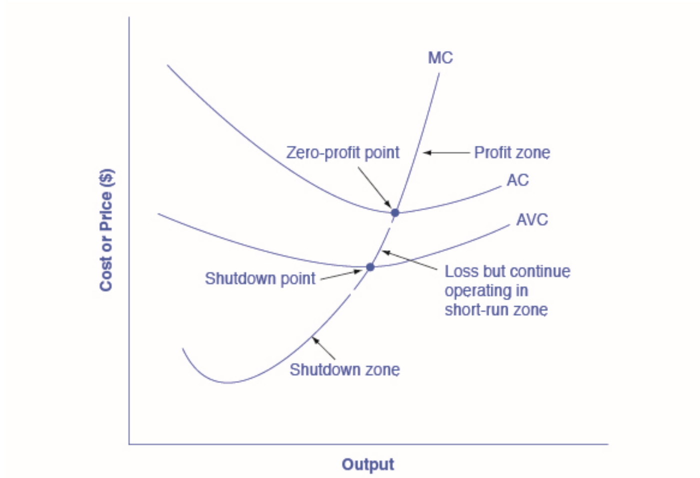


Figure 7.3.6: The marginal cost curve can be divided into three zones, based on where it is crossed by the average cost and average variable cost curves. The point where MC crosses AC is called the zero-profit point. If the firm is operating at a level of output where the market price is at a level higher than the zero-profit point, then price will be greater than average cost and the firm is earning profits. If the price is exactly at the zero-profit point, then the firm is making zero profits. If price falls in the zone between the shutdown point and the zero-profit point, then the firm is making losses but will continue to operate in the short run, since it is covering its variable costs. However, if price falls below the price at the shutdown point, then the firm will shut down immediately, since it is not even covering its variable costs.

First consider the upper zone, where prices are above the level where marginal cost (MC) crosses average cost (AC) at the zero profit point. At any price above that level, the firm will earn profits in the short run. If the price falls exactly on the zero profit point where the MC and AC curves cross, then the firm earns zero profits. If a price falls into the zone between the zero profit point, where MC crosses AC, and the shutdown point, where MC crosses AVC, the firm will be making losses in the short run—but since the firm is more than covering its variable costs, the losses are smaller than if the firm shut down immediately. Finally, consider a price at or below the shutdown point where MC crosses AVC. At any price like this one, the firm will shut down immediately, because it cannot even cover its variable costs.

Marginal Cost and the Firm's Supply Curve

For a perfectly competitive firm, the marginal cost curve is identical to the firm's supply curve starting from the minimum point on the average variable cost curve. To understand why this perhaps surprising insight holds true, first think about what the supply curve means. A firm checks the market price and then looks at its supply curve to decide what quantity to produce. Now, think about what it means to say that a firm will maximize its profits by producing at the quantity where $P = MC$. This rule means that the firm checks the market price, and then looks at its marginal cost to determine the quantity to produce—and makes sure that the price is greater than the minimum average variable cost. In other words, the marginal cost curve above the minimum point on the average variable cost curve becomes the firm's supply curve.

As discussed in the chapter on Demand and Supply, many of the reasons that supply curves shift relate to underlying changes in costs. For example, a lower price of key inputs or new technologies that reduce production costs cause supply to shift to the right; in contrast, bad weather or added government regulations can add to costs of certain goods in a way that causes supply to shift to the left. These shifts in the firm's supply curve can also be interpreted as shifts of the marginal cost curve. A shift in costs of production that increases marginal costs at all levels of output—and shifts MC to the left—will cause a perfectly competitive firm to produce less at any given market price. Conversely, a shift in costs of production that decreases marginal costs at all levels of output will shift MC to the right and as a result, a competitive firm will choose to expand its level of output at any given price. The following Work It Out feature will walk you through an example.

Example 7.3.1: At What Price Should the Firm Continue Producing in the Short Run?

To determine the short-run economic condition of a firm in perfect competition, follow the steps outlined below. Use the data shown in Table 7.3.7.

Table 7.3.7: Economic condition of a firm

Q	P	TFC	TVC	TC	AVC	ATC	MC	TR	Profits
0	$28	$20	$0	-	-	-	-	-	-
1	$28	$20	$20	-	-	-	-	-	-
2	$28	$20	$25	-	-	-	-	-	-
3	$28	$20	$35	-	-	-	-	-	-
4	$28	$20	$52	-	-	-	-	-	-
5	$28	$20	$80	-	-	-	-	-	-

Step 1: Determine the cost structure for the firm. For a given total fixed costs and variable costs, calculate total cost, average variable cost, average total cost, and marginal cost. Follow the formulas given in the Cost and Industry Structure chapter. These calculations are shown in Table 7.3.8.

Table 7.3.8: Calculations

Q	P	TFC	TVC	TC (TFC+TVC)	AVC (TVC/Q)	ATC (TC/Q)	MC (TC2−TC1)/(Q2−Q1)
0	$28	$20	$0	$20+$0=$20	-	-	-
1	$28	$20	$20	$20+$20=$40	$20/1=$20.00	$40/1=$40.00	($40−$20)/(1−0)= $20
2	$28	$20	$25	$20+$25=$45	$25/2=$12.50	$45/2=$22.50	($45−$40)/(2−1)= $5
3	$28	$20	$35	$20+$35=$55	$35/3=$11.67	$55/3=$18.33	($55−$45)/(3−2)= $10
4	$28	$20	$52	$20+$52=$72	$52/4=$13.00	$72/4=$18.00	($72−$55)/(4−3)= $17
5	$28	$20	$80	$20+$80=$100	$80/5=$16.00	$100/5=$20.00	($100−$72)/(5−4)= $28

Step 2: Determine the market price that the firm receives for its product. This should be given information, as the firm in perfect competition is a price taker. With the given price, calculate total revenue as equal to price multiplied by quantity for all output levels produced. In this example, the given price is $28. You can see that in the second column of Table 7.3.9.

Table 7.3.9: Calculations

Quantity	Price	Total Revenue (P × Q)
0	$28	$28×0=$0
1	$28	$28×1=$28
2	$28	$28×2=$56
3	$28	$28×3=$84
4	$28	$28×4=$112

Quantity	Price	Total Revenue (P × Q)
5	$28	$28×5=$140

Step 3: Calculate profits as total cost subtracted from total revenue, as shown in Table $7.3.10$.

Table 7.3.10: Calculations

Quantity	Total Revenue	Total Cost	Profits (TR−TC)
0	$0	$20	$0−$20=−$20
1	$28	$40	$28−$40=−$12
2	$56	$45	$56−$45=$11
3	$84	$55	$84−$55=$29
4	$112	$72	$112−$72=$40
5	$140	$100	$140−$100=$40

Step 4: To find the profit-maximizing output level, look at the Marginal Cost column (at every output level produced), as shown in Table $7.3.11$, and determine where it is equal to the market price. The output level where price equals the marginal cost is the output level that maximizes profits.

Table 7.3.11: Calculations

Q	P	TFC	TVC	TC	AVC	ATC	MC	TR	Profits
0	$28	$20	$0	$20	-	-	-	$0	−$20
1	$28	$20	$20	$40	$20.00	$40.00	$20	$28	−$12
2	$28	$20	$25	$45	$12.50	$22.50	$5	$56	$11
3	$28	$20	$35	$55	$11.67	$18.33	$10	$84	$29
4	$28	$20	$52	$72	$13.00	$18.00	$17	$112	$40
5	$28	$20	$80	$100	$16.40	$20.40	$30	$140	$40

Step 5: Once you have determined the profit-maximizing output level (in this case, output quantity 5), you can look at the amount of profits made (in this case, $\$40$).

Step 6: If the firm is making economic losses, the firm needs to determine whether it produces the output level where price equals marginal revenue and equals marginal cost or it shuts down and only incurs its fixed costs.

Step 7: For the output level where marginal revenue is equal to marginal cost, check if the market price is greater than the average variable cost of producing that output level.

- If $P > AVC$ but $P < ATC$, then the firm continues to produce in the short-run, making economic losses.
- If $P < AVC$, then the firm stops producing and only incurs its fixed costs.

In this example, the price of $\$28$ is greater than the AVC ($\$16.40$) of producing 5 units of output, so the firm continues producing.

Key Concepts and Summary

As a perfectly competitive firm produces a greater quantity of output, its total revenue steadily increases at a constant rate determined by the given market price. Profits will be highest (or losses will be smallest) at the quantity of output where total revenues exceed total costs by the greatest amount (or where total revenues fall short of total costs by the smallest amount). Alternatively, profits will be highest where marginal revenue, which is price for a perfectly competitive firm, is equal to marginal cost. If the market price faced by a perfectly competitive firm is above average cost at the profit-maximizing quantity of output, then the firm is making profits. If the market price is below average cost at the profit-maximizing quantity of output, then the firm is making losses.

If the market price is equal to average cost at the profit-maximizing level of output, then the firm is making zero profits. The point where the marginal cost curve crosses the average cost curve, at the minimum of the average cost curve, is called the “zero profit point.” If the market price faced by a perfectly competitive firm is below average variable cost at the profit-maximizing quantity of output, then the firm should shut down operations immediately. If the market price faced by a perfectly competitive firm is above average variable cost, but below average cost, then the firm should continue producing in the short run, but exit in the long run. The point where the marginal cost curve crosses the average variable cost curve is called the shutdown point.

Glossary

marginal revenue
: the additional revenue gained from selling one more unit

shutdown point
: level of output where the marginal cost curve intersects the average variable cost curve at the minimum point of AVC; if the price is below this point, the firm should shut down immediately

Contributor

- Template:ContribOpenSTAXMicroEcon

7.4: Entry and Exit Decisions in the Long Run

Learning Objectives

- Explain how entry and exit lead to zero profits in the long run
- Discuss the long-run adjustment process

The line between the short run and the long run cannot be defined precisely with a stopwatch, or even with a calendar. It varies according to the specific business. The distinction between the short run and the long run is therefore more technical: in the short run, firms cannot change the usage of fixed inputs, while in the long run, the firm can adjust all factors of production.

In a competitive market, profits are a red cape that incites businesses to charge. If a business is making a profit in the short run, it has an incentive to expand existing factories or to build new ones. New firms may start production, as well. When new firms enter the industry in response to increased industry profits it is called **entry**.

Losses are the black thundercloud that causes businesses to flee. If a business is making losses in the short run, it will either keep limping along or just shut down, depending on whether its revenues are covering its variable costs. But in the long run, firms that are facing losses will shut down at least some of their output, and some firms will cease production altogether. The long-run process of reducing production in response to a sustained pattern of losses is called **exit**. The following Clear It Up feature discusses where some of these losses might come from, and the reasons why some firms go out of business.

Example 7.4.1: Why do firms cease to exist?

Can we say anything about what causes a firm to exit an industry? Profits are the measurement that determines whether a business stays operating or not. Individuals start businesses with the purpose of making profits. They invest their money, time, effort, and many other resources to produce and sell something that they hope will give them something in return. Unfortunately, not all businesses are successful, and many new startups soon realize that their “business adventure” must eventually end.

In the model of perfectly competitive firms, those that consistently cannot make money will “exit,” which is a nice, bloodless word for a more painful process. When a business fails, after all, workers lose their jobs, investors lose their money, and owners and managers can lose their dreams. Many businesses fail. The U.S. Small Business Administration indicates that in 2011, $409,040$ new firms "entered," and $470,376$ firms failed.

Sometimes a business fails because of poor management or workers who are not very productive, or because of tough domestic or foreign competition. Businesses also fail from a variety of causes that might best be summarized as bad luck. For example, conditions of demand and supply in the market shift in an unexpected way, so that the prices that can be charged for outputs fall or the prices that need to be paid for inputs rise. With millions of businesses in the U.S. economy, even a small fraction of them failing will affect many people—and business failures can be very hard on the workers and managers directly involved. But from the standpoint of the overall economic system, business exits are sometimes a necessary evil if a market-oriented system is going to offer a flexible mechanism for satisfying customers, keeping costs low, and inventing new products.

How Entry and Exit Lead to Zero Profits in the Long Run

No **perfectly competitive firm** acting alone can affect the market price. However, the combination of many firms entering or exiting the market will affect overall supply in the market. In turn, a shift in supply for the market as a whole will affect the market price. Entry and exit to and from the market are the driving forces behind a process that, in the long run, pushes the price down to minimum average total costs so that all firms are earning a zero profit.

To understand how short-run profits for a perfectly competitive firm will evaporate in the long run, imagine the following situation. The market is in **long-run equilibrium**, where all firms earn zero economic profits producing the output level where $P = MR = MC$ and $P = AC$. No firm has the incentive to enter or leave the market. Let’s say that the product’s demand increases, and with that, the market price goes up. The existing firms in the industry are now facing a higher price than before, so they will increase production to the new output level where $P = MR = MC$.

This will temporarily make the market price rise above the average cost curve, and therefore, the existing firms in the market will now be earning economic profits. However, these economic profits attract other firms to enter the market. Entry of many new firms causes the market supply curve to shift to the right. As the supply curve shifts to the right, the market price starts decreasing, and with that, economic profits fall for new and existing firms. As long as there are still profits in the market, entry will continue to shift supply to the right. This will stop whenever the market price is driven down to the zero-profit level, where no firm is earning economic profits.

Short-run losses will fade away by reversing this process. Say that the market is in long-run equilibrium. This time, instead, demand decreases, and with that, the market price starts falling. The existing firms in the industry are now facing a lower price than before, and as it will be below the average cost curve, they will now be making economic losses. Some firms will continue producing where the new $P = MR = MC$, as long as they are able to cover their average variable costs. Some firms will have to shut down immediately as they will not be able to cover their average variable costs, and will then only incur their fixed costs, minimizing their losses. Exit of many firms causes the market supply curve to shift to the left. As the supply curve shifts to the left, the market price starts rising, and economic losses start to be lower. This process ends whenever the market price rises to the zero-profit level, where the existing firms are no longer losing money and are at zero profits again. Thus, while a perfectly competitive firm can earn profits in the short run, in the long run the process of entry will push down prices until they reach the zero-profit level. Conversely, while a perfectly competitive firm may earn losses in the short run, firms will not continually lose money. In the long run, firms making losses are able to escape from their fixed costs, and their exit from the market will push the price back up to the zero-profit level. In the long run, this process of entry and exit will drive the price in perfectly competitive markets to the zero-profit point at the bottom of the AC curve, where marginal cost crosses average cost.

The Long-Run Adjustment and Industry Types

Whenever there are expansions in an industry, costs of production for the existing and new firms could either stay the same, increase, or even decrease. Therefore, we can categorize an industry as being (1) a constant cost industry (as demand increases, the cost of production for firms stays the same), (2) an increasing cost industry (as demand increases, the cost of production for firms increases), or (3) a decreasing cost industry (as demand increases the costs of production for the firms decreases).

For a **constant cost industry**, whenever there is an increase in market demand and price, then the supply curve shifts to the right with new firms' entry and stops at the point where the new long-run equilibrium intersects at the same market price as before. But why will costs remain the same? In this type of industry, the supply curve is very elastic. Firms can easily supply any quantity that consumers demand. In addition, there is a perfectly elastic supply of inputs—firms can easily increase their demand for employees, for example, with no increase to wages. Tying in to our Bring it Home discussion, an increased demand for ethanol in recent years has caused the demand for corn to increase. Consequently, many farmers switched from growing wheat to growing corn. Agricultural markets are generally good examples of constant cost industries.

For an **increasing cost industry**, as the market expands, the old and new firms experience increases in their costs of production, which makes the new zero-profit level intersect at a higher price than before. Here companies may have to deal with limited inputs, such as skilled labor. As the demand for these workers rise, wages rise and this increases the cost of production for all firms. The industry supply curve in this type of industry is more inelastic.

For a **decreasing cost industry**, as the market expands, the old and new firms experience lower costs of production, which makes the new zero-profit level intersect at a lower price than before. In this case, the industry and all the firms in it are experiencing falling average total costs. This can be due to an improvement in technology in the entire industry or an increase in the education of employees. High tech industries may be a good example of a decreasing cost market.

Figure $7.4.1$ (a) presents the case of an adjustment process in a constant cost industry. Whenever there are output expansions in this type of industry, the long-run outcome implies more output produced at exactly the same original price. Note that supply was able to increase to meet the increased demand. When we join the before and after long-run equilibriums, the resulting line is the long run supply (LRS) curve in perfectly competitive markets. In this case, it is a flat curve. Figure $7.4.1$ (b) and Figure $7.4.1$ (c) present the cases for an increasing cost and decreasing cost industry, respectively. For an increasing cost industry, the LRS is upward sloping, while for a decreasing cost industry, the LRS is downward sloping.

Adjustment Process in a Constant-Cost Industry

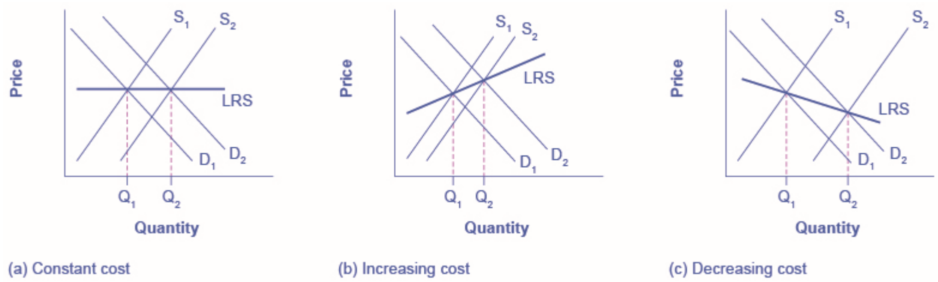


Figure $7.4.1$: In (a), demand increased and supply met it. Notice that the supply increase is equal to the demand increase. The result is that the equilibrium price stays the same as quantity sold increases. In (b), notice that sellers were not able to increase supply as much as demand. Some inputs were scarce, or wages were rising. The equilibrium price rises. In (c), sellers easily increased supply in response to the demand increase. Here, new technology or economies of scale caused the large increase in supply, resulting in declining equilibrium price.

Key Concepts and Summary

In the long run, firms will respond to profits through a process of entry, where existing firms expand output and new firms enter the market. Conversely, firms will react to losses in the long run through a process of exit, in which existing firms reduce output or cease production altogether. Through the process of entry in response to profits and exit in response to losses, the price level in a perfectly competitive market will move toward the zero-profit point, where the marginal cost curve crosses the AC curve, at the minimum of the average cost curve.

The long-run supply curve shows the long-run output supplied by firms in three different types of industries: constant cost, increasing cost, and decreasing cost.

Glossary

entry
: the long-run process of firms entering an industry in response to industry profits

exit
: the long-run process of firms reducing production and shutting down in response to industry losses

long-run equilibrium
: where all firms earn zero economic profits producing the output level where $P = MR = MC$ and $P = AC$

Contributor

- Template:ContribOpenSTAXMicroEcon

7.5: Efficiency in Perfectly Competitive Markets

Learning Objectives

- Apply concepts of productive efficiency and allocative efficiency to perfectly competitive markets
- Compare the model of perfect competition to real-world markets

When profit-maximizing firms in perfectly competitive markets combine with utility-maximizing consumers, something remarkable happens: the resulting quantities of outputs of goods and services demonstrate both productive and allocative efficiency (terms that were first introduced in (Choice in a World of Scarcity) .

Productive efficiency means producing without waste, so that the choice is on the production possibility frontier. In the long run in a perfectly competitive market, because of the process of entry and exit, the price in the market is equal to the minimum of the long-run average cost curve. In other words, goods are being produced and sold at the lowest possible average cost.

Allocative efficiency means that among the points on the production possibility frontier, the point that is chosen is socially preferred—at least in a particular and specific sense. In a perfectly competitive market, price will be equal to the marginal cost of production. Think about the price that is paid for a good as a measure of the social benefit received for that good; after all, willingness to pay conveys what the good is worth to a buyer. Then think about the marginal cost of producing the good as representing not just the cost for the firm, but more broadly as the social cost of producing that good. When perfectly competitive firms follow the rule that profits are maximized by producing at the quantity where price is equal to marginal cost, they are thus ensuring that the social benefits received from producing a good are in line with the social costs of production.

To explore what is meant by **allocative efficiency**, it is useful to walk through an example. Begin by assuming that the market for wholesale flowers is perfectly competitive, and so $P = MC$. Now, consider what it would mean if firms in that market produced a lesser quantity of flowers. At a lesser quantity, marginal costs will not yet have increased as much, so that price will exceed marginal cost; that is, $P > MC$. In that situation, the benefit to society as a whole of producing additional goods, as measured by the willingness of consumers to pay for marginal units of a good, would be higher than the cost of the inputs of labor and physical capital needed to produce the marginal good. In other words, the gains to society as a whole from producing additional marginal units will be greater than the costs.

Conversely, consider what it would mean if, compared to the level of output at the allocatively efficient choice when $P = MC$, firms produced a greater quantity of flowers. At a greater quantity, marginal costs of production will have increased so that $P < MC$. In that case, the marginal costs of producing additional flowers is greater than the benefit to society as measured by what people are willing to pay. For society as a whole, since the costs are outstripping the benefits, it will make sense to produce a lower quantity of such goods.

When perfectly competitive firms maximize their profits by producing the quantity where $P = MC$, they also assure that the benefits to consumers of what they are buying, as measured by the price they are willing to pay, is equal to the costs to society of producing the marginal units, as measured by the marginal costs the firm must pay—and thus that allocative efficiency holds.

The statements that a perfectly competitive market in the long run will feature both productive and allocative efficiency do need to be taken with a few grains of salt. Remember, economists are using the concept of "efficiency" in a particular and specific sense, not as a synonym for "desirable in every way." For one thing, consumers' ability to pay reflects the income distribution in a particular society. Thus, a homeless person may have no ability to pay for housing because they have insufficient income.

Perfect competition, in the long run, is a hypothetical benchmark. For market structures such as monopoly, monopolistic competition, and oligopoly, which are more frequently observed in the real world than perfect competition, firms will not always produce at the minimum of average cost, nor will they always set price equal to marginal cost. Thus, these other competitive situations will not produce productive and allocative efficiency.

Moreover, real-world markets include many issues that are assumed away in the model of perfect competition, including pollution, inventions of new technology, poverty which may make some people unable to pay for basic necessities of life, government programs like national defense or education, discrimination in labor markets, and buyers and sellers who must deal with imperfect and unclear information. These issues are explored in other chapters. However, the theoretical efficiency of perfect competition does provide a useful benchmark for comparing the issues that arise from these real-world problems.

Example 7.5.1: A Dime a Dozen

A quick glance at Table 7.5.1 reveals the dramatic increase in North Dakota corn production—more than double. Taking into consideration that corn typically yields two to three times as many bushels per acre as wheat, it is obvious there has been a significant increase in bushels of corn. Why the increase in corn acreage? Converging prices.

Table 7.5.1: (Source: USDA National Agricultural Statistics Service)

Year	Corn (millions of acres)	Wheat (millions of acres)
2014	91.6	56.82

Historically, wheat prices have been higher than corn prices, offsetting wheat's lower yield per acre. However, in recent years wheat and corn prices have been converging. In April 2013, Agweek reported the gap was just 71 cents per bushel. As the difference in price narrowed, switching to the production of higher yield per acre of corn simply made good business sense. Erik Younggren, president of the National Association of Wheat Growers said in the *Agweek* article, "*I don't think we're going to see mile after mile of waving amber fields [of wheat] anymore.*" (Until wheat prices rise, we will probably be seeing field after field of tasseled corn.)

Key Concepts and Summary

Long-run equilibrium in perfectly competitive markets meets two important conditions: allocative efficiency and productive efficiency. These two conditions have important implications. First, resources are allocated to their best alternative use. Second, they provide the maximum satisfaction attainable by society.

References

Index Mundi. n.d. "Wheat Monthly Price—U.S. Dollars per Metric Ton." Accessed March 11, 2015. http://www.indexmundi.com/commoditie...ommodity=wheat.

Knutson, J. "Wheat on the Defensive in the Northern Plains." Agweek, Associated Press State Wire: North Dakota (ND). April 14, 2013.

SBA Office of Advocacy. 2014. "Frequently Asked Questions: Advocacy: the voice of small business in government." Accessed March 11, 2015. www.sba.gov/sites/default/fi...rch_2014_0.pdf.

Contributor

- Template:ContribOpenSTAXMicroEcon

7.E: Perfect Competition (Exercises)

8.1: Perfect Competition and Why It Matters

Self-Check Questions

Q1

Firms in a perfectly competitive market are said to be “price takers”—that is, once the market determines an equilibrium price for the product, firms must accept this price. If you sell a product in a perfectly competitive market, but you are not happy with its price, would you raise the price, even by a cent?

Q2

Would independent trucking fit the characteristics of a perfectly competitive industry?

Review Questions

Q3

A single firm in a perfectly competitive market is relatively small compared to the rest of the market. What does this mean? How “small” is “small”?

Q4

What are the four basic assumptions of perfect competition? Explain in words what they imply for a perfectly competitive firm.

Q5

What is a “price taker” firm?

Critical Thinking Questions

Q6

Finding a life partner is a complicated process that may take many years. It is hard to think of this process as being part of a very complex market, with a demand and a supply for partners. Think about how this market works and some of its characteristics, such as search costs. Would you consider it a perfectly competitive market?

Q7

Can you name five examples of perfectly competitive markets? Why or why not?

Solution

S1

No, you would not raise the price. Your product is exactly the same as the product of the many other firms in the market. If your price is greater than that of your competitors, then your customers would switch to them and stop buying from you. You would lose all your sales.

S2

Possibly. Independent truckers are by definition small and numerous. All that is required to get into the business is a truck (not an inexpensive asset, though) and a commercial driver's license. To exit, one need only sell the truck. All trucks are essentially the same, providing transportation from point A to point B. (We're assuming we not talking about specialized trucks.) Independent truckers must take the going rate for their service, so independent trucking does seem to have most of the characteristics of perfect competition.

8.2: How Perfectly Competitive Firms Make Output Decisions

Self-Check Questions

Q1

Look at Table below. What would happen to the firm's profits if the market price increases to $6 per pack of raspberries?

Quantity	Total Cost	Fixed Cost	Variable Cost	Total Revenue	Profit
0	$62	$62	-	$0	−$62
10	$90	$62	$28	$60	−$30
20	$110	$62	$48	$120	$10
30	$126	$62	$64	$180	$54
40	$144	$62	$82	$240	$96
50	$166	$62	$104	$300	$134
60	$192	$62	$130	$360	$168
70	$224	$62	$162	$420	$196
80	$264	$62	$202	$480	$216
90	$324	$62	$262	$540	$216
100	$404	$62	$342	$600	$196

Q2

Suppose that the market price increases to $6, as shown in Table below. What would happen to the profit-maximizing output level?

Quantity	Total Cost	Fixed Cost	Variable Cost	Marginal Cost	Total Revenue	Marginal Revenue
0	$62	$62	-	-	$0	-
10	$90	$62	$28	$2.80	$60	$6.00
20	$110	$62	$48	$2.00	$120	$6.00
30	$126	$62	$64	$1.60	$180	$6.00
40	$144	$62	$82	$1.80	$240	$6.00
50	$166	$62	$104	$2.20	$300	$6.00
60	$192	$62	$130	$2.60	$360	$6.00

Quantity	Total Cost	Fixed Cost	Variable Cost	Marginal Cost	Total Revenue	Marginal Revenue
70	$224	$62	$162	$3.20	$420	$6.00
80	$264	$62	$202	$4.00	$480	$6.00
90	$324	$62	$262	$6.00	$540	$6.00
100	$404	$62	$342	$8.00	$600	$6.00

Q3

Explain in words why a profit-maximizing firm will not choose to produce at a quantity where marginal cost exceeds marginal revenue.

Q4

A firm's marginal cost curve above the average variable cost curve is equal to the firm's individual supply curve. This means that every time a firm receives a price from the market it will be willing to supply the amount of output where the price equals marginal cost. What happens to the firm's individual supply curve if marginal costs increase?

Review Questions

Q5

How does a perfectly competitive firm decide what price to charge?

Q6

What prevents a perfectly competitive firm from seeking higher profits by increasing the price that it charges?

Q7

How does a perfectly competitive firm calculate total revenue?

Q8

Briefly explain the reason for the shape of a marginal revenue curve for a perfectly competitive firm.

Q9

What two rules does a perfectly competitive firm apply to determine its profit-maximizing quantity of output?

Q10

How does the average cost curve help to show whether a firm is making profits or losses?

Q11

What two lines on a cost curve diagram intersect at the zero-profit point?

Q12

Should a firm shut down immediately if it is making losses?

Q13

How does the average variable cost curve help a firm know whether it should shut down immediately?

Q14

What two lines on a cost curve diagram intersect at the shutdown point?

Critical Thinking Questions

Q15

Your company operates in a perfectly competitive market. You have been told that advertising can help you increase your sales in the short run. Would you create an aggressive advertising campaign for your product?

Q16

Since a perfectly competitive firm can sell as much as it wishes at the market price, why can the firm not simply increase its profits by selling an extremely high quantity?

Problems

Q17

The AAA Aquarium Co. sells aquariums for $\$20$ each. Fixed costs of production are $\$20$. The total variable costs are $\$20$ for one aquarium, $\$25$ for two units, $\$35$ for the three units, $\$50$ for four units, and $\$80$ for five units. In the form of a table, calculate total revenue, marginal revenue, total cost, and marginal cost for each output level (one to five units). What is the profit-maximizing quantity of output? On one diagram, sketch the total revenue and total cost curves. On another diagram, sketch the marginal revenue and marginal cost curves.

Q18

Perfectly competitive firm Doggies Paradise Inc. sells winter coats for dogs. Dog coats sell for $\$72$ each. The fixed costs of production are $\$100$. The total variable costs are $\$64$ for one unit, $\$84$ for two units, $\$114$ for three units, $\$184$ for four units, and $\$270$ for five units. In the form of a table, calculate total revenue, marginal revenue, total cost and marginal cost for each output level (one to five units). On one diagram, sketch the total revenue and total cost curves. On another diagram, sketch the marginal revenue and marginal cost curves. What is the profit maximizing quantity?

Q19

A computer company produces affordable, easy-to-use home computer systems and has fixed costs of $\$250$. The marginal cost of producing computers is $\$700$ for the first computer, $\$250$ for the second, $\$300$ for the third, $\$350$ for the fourth, $\$400$ for the fifth, $\$450$ for the sixth, and $\$500$ for the seventh.

a. Create a table that shows the company's output, total cost, marginal cost, average cost, variable cost, and average variable cost.
b. At what price is the zero-profit point? At what price is the shutdown point?
c. If the company sells the computers for $\$500$, is it making a profit or a loss? How big is the profit or loss? Sketch a graph with AC, MC, and AVC curves to illustrate your answer and show the profit or loss.
d. If the firm sells the computers for $\$300$, is it making a profit or a loss? How big is the profit or loss? Sketch a graph with AC, MC, and AVC curves to illustrate your answer and show the profit or loss.

Solution

S1

Holding total cost constant, profits at every output level would increase.

S2

When the market price increases, marginal revenue increases. The firm would then increase production up to the point where the new price equals marginal cost, at a quantity of 90.

S3

If marginal costs exceeds marginal revenue, then the firm will reduce its profits for every additional unit of output it produces. Profit would be greatest if it reduces output to where $MR = MC$.

S4

The firm will be willing to supply fewer units at every price level. In other words, the firm's individual supply curve decreases and shifts to the left.

8.3: Entry and Exit Decisions in the Long Run

Self-Check Questions

Q1

If new technology in a perfectly competitive market brings about a substantial reduction in costs of production, how will this affect the market?

Q2

A market in perfect competition is in long-run equilibrium. What happens to the market if labor unions are able to increase wages for workers?

Review Questions

Q3

Why does entry occur?

Q4

Why does exit occur?

Q5

Do entry and exit occur in the short run, the long run, both, or neither?

Q6

What price will a perfectly competitive firm end up charging in the long run? Why?

Critical Thinking Questions

Q7

Many firms in the United States file for bankruptcy every year, yet they still continue operating. Why would they do this instead of completely shutting down?

Q8

Why will profits for firms in a perfectly competitive industry tend to vanish in the long run?

Q9

Why will losses for firms in a perfectly competitive industry tend to vanish in the long run?

Solution

S1

With a technological improvement that brings about a reduction in costs of production, an adjustment process will take place in the market. The technological improvement will result in an increase in supply curves, by individual firms and at the market level. The existing firms will experience higher profits for a while, which will attract other firms into the market. This entry process will stop whenever the market supply increases enough (both by existing and new firms) so profits are driven back to zero.

S2

When wages increase, costs of production increase. Some firms would now be making economic losses and would shut down. The supply curve then starts shifting to the left, pushing the market price up. This process ends when all firms remaining in the market earn zero economic profits. The result is a contraction in the output produced in the market.

8.4: Efficiency in Perfectly Competitive Markets

Self-Check Questions

Q1

Productive efficiency and allocative efficiency are two concepts achieved in the long run in a perfectly competitive market. These are the two reasons why we call them "perfect." How would you use these two concepts to analyze other market structures and label them "imperfect?"

Q2

Explain how the profit-maximizing rule of setting $P = MC$ leads a perfectly competitive market to be allocatively efficient.

Review Questions

Q3

Will a perfectly competitive market display productive efficiency? Why or why not?

Q4

Will a perfectly competitive market display allocative efficiency? Why or why not?

Critical Thinking Questions

Q5

Assuming that the market for cigarettes is in perfect competition, what does allocative and productive efficiency imply in this case? What does it not imply?

Q6

In the argument for why perfect competition is allocatively efficient, the price that people are willing to pay represents the gains to society and the marginal cost to the firm represents the costs to society. Can you think of some social costs or issues that are not included in the marginal cost to the firm? Or some social gains that are not included in what people pay for a good?

Solution

S1

Perfect competition is considered to be “perfect” because both allocative and productive efficiency are met at the same time in a long-run equilibrium. If a market structure results in long-run equilibrium that does not minimize average total costs and/or does not charge a price equal to marginal cost, then either allocative or productive (or both) efficiencies are not met, and therefore the market cannot be labeled “perfect.”

S2

Think of the market price as representing the gain to society from a purchase, since it represents what someone is willing to pay. Think of the marginal cost as representing the cost to society from making the last unit of a good. If $P > MC$, then the benefits from producing more of a good exceed the costs, and society would gain from producing more of the good. If $P < MC$, then the social costs of producing the marginal good exceed the social benefits, and society should produce less of the good. Only if $P = MC$, the rule applied by a profit-maximizing perfectly competitive firm, will society’s costs and benefits be in balance. This choice will be the option that brings the greatest overall benefit to society.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

8: Monopoly

8.1: Introduction

Learning Objectives

- How Monopolies form: Barriers to Entry
- How a Profit-Maximizing Monopoly Chooses Output and Price

Political Power from a Cotton Monopoly

Figure 8.1.1: In the mid-nineteenth century, the United States, specifically the Southern states, had a near monopoly in the cotton supplied to Great Britain. These states attempted to leverage this economic power into political power—trying to sway Great Britain to formally recognize the Confederate States of America. (Credit: modification of work by “ashleylovespizza”/Flickr Creative Commons)

The Rest is History

Many of the opening case studies have focused on current events. This one steps into the past to observe how monopoly, or near monopolies, have helped shape history. In the spring of 1773, the East India Company, a firm that, in its time, was designated ‘too big to fail,’ was continuing to experience financial difficulties. To help shore up the failing firm, the British Parliament authorized the Tea Act. The act continued the tax on teas and made the East India Company the sole legal supplier of tea to the American colonies. By November, the citizens of Boston had had enough. They refused to permit the tea to be unloaded, citing their main complaint: “*No taxation without representation.*” Arriving tea-bearing ships were warned via several newspapers, including The Massachusetts Gazette, “We are prepared, and shall not fail to pay them an unwelcome visit; by The Mohawks.”

Step forward in time to 1860—the eve of the American Civil War—to another near monopoly supplier of historical significance: the U.S. cotton industry. At that time, the Southern states provided the majority of the cotton Britain imported. The South, wanting to secede from the Union, hoped to leverage Britain’s high dependency on its cotton into formal diplomatic recognition of the Confederate States of America.

This leads us to the topic of this chapter: a firm that controls all (or nearly all) of the supply of a good or service—a monopoly. How do monopoly firms behave in the marketplace? Do they have “power?” Does this power potentially have unintended consequences? We’ll return to this case at the end of the chapter to see how the tea and cotton monopolies influenced U.S. history.

There is a widespread belief that top executives at firms are the strongest supporters of market competition, but this belief is far from the truth. Think about it this way: If you very much wanted to win an Olympic gold medal, would you rather be far better than everyone else, or locked in competition with many athletes just as good as you are? Similarly, if you would like to attain a very high level of profits, would you rather manage a business with little or no competition, or struggle against many tough competitors who are trying to sell to your customers? By now, you might have read the chapter on Perfect Competition. In this chapter, we explore the opposite extreme: monopoly.

If perfect competition is a market where firms have no market power and they simply respond to the market price, monopoly is a market with no competition at all, and firms have complete market power. In the case of monopoly, one firm produces all of the output in a market. Since a monopoly faces no significant competition, it can charge any price it wishes. While a monopoly, by definition, refers to a single firm, in practice the term is often used to describe a market in which one firm merely has a very high market share. This tends to be the definition that the U.S. Department of Justice uses.

Even though there are very few true monopolies in existence, we do deal with some of those few every day, often without realizing it: The U.S. Postal Service, your electric and garbage collection companies are a few examples. Some new drugs are produced by only one pharmaceutical firm—and no close substitutes for that drug may exist.

From the mid-1990s until 2004, the U.S. Department of Justice prosecuted the Microsoft Corporation for including Internet Explorer as the default web browser with its operating system. The Justice Department's argument was that, since Microsoft possessed an extremely high market share in the industry for operating systems, the inclusion of a free web browser constituted unfair competition to other browsers, such as Netscape Navigator. Since nearly everyone was using Windows, including Internet Explorer eliminated the incentive for consumers to explore other browsers and made it impossible for competitors to gain a foothold in the market. In 2013, the Windows system ran on more than 90% of the most commonly sold personal computers. In 2015, a U.S. federal court tossed out antitrust charges that Google had an agreement with mobile device makers to set Google as the default search engine.

This chapter begins by describing how monopolies are protected from competition, including laws that prohibit competition, technological advantages, and certain configurations of demand and supply. It then discusses how a monopoly will choose its profit-maximizing quantity to produce and what price to charge. While a monopoly must be concerned about whether consumers will purchase its products or spend their money on something altogether different, the monopolist need not worry about the actions of other competing firms producing its products. As a result, a monopoly is not a price taker like a perfectly competitive firm, but instead exercises some power to choose its market price.

Contributor

- Template:ContribOpenSTAXMicroEcon

8.2: How Monopolies Form- Barriers to Entry

Skills to Develop

- Distinguish between a natural monopoly and a legal monopoly.
- Explain how economies of scale and the control of natural resources led to the necessary formation of legal monopolies
- Analyze the importance of trademarks and patents in promoting innovation
- Identify examples of predatory pricing

Because of the lack of competition, monopolies tend to earn significant economic profits. These profits should attract vigorous competition as described in Perfect Competition, and yet, because of one particular characteristic of monopoly, they do not. Barriers to entry are the legal, technological, or market forces that discourage or prevent potential competitors from entering a market. **Barriers to entry** can range from the simple and easily surmountable, such as the cost of renting retail space, to the extremely restrictive. For example, there are a finite number of radio frequencies available for broadcasting. Once the rights to all of them have been purchased, no new competitors can enter the market.

In some cases, barriers to entry may lead to monopoly. In other cases, they may limit competition to a few firms. Barriers may block entry even if the firm or firms currently in the market are earning profits. Thus, in markets with significant barriers to entry, it is not true that abnormally high profits will attract new firms, and that this entry of new firms will eventually cause the price to decline so that surviving firms earn only a normal level of profit in the long run.

There are two types of monopoly, based on the types of barriers to entry they exploit. One is **natural monopoly**, where the barriers to entry are something other than legal prohibition. The other is **legal monopoly**, where laws prohibit (or severely limit) competition.

Natural Monopoly

Economies of scale can combine with the size of the market to limit competition. (This theme was introduced in Cost and Industry Structure). Figure $8.2.1$ presents a long-run average cost curve for the airplane manufacturing industry. It shows economies of scale up to an output of $8,000$ planes per year and a price of P_0, then constant returns to scale from $8,000$ to $20,000$ planes per year, and diseconomies of scale at a quantity of production greater than $20,000$ planes per year.

Now consider the market demand curve in the diagram, which intersects the long-run average cost (LRAC) curve at an output level of $6,000$ planes per year and at a price P_1, which is higher than P_0. In this situation, the market has room for only one producer. If a second firm attempts to enter the market at a smaller size, say by producing a quantity of $4,000$ planes, then its average costs will be higher than the existing firm, and it will be unable to compete. If the second firm attempts to enter the market at a larger size, like $8,000$ planes per year, then it could produce at a lower average cost—but it could not sell all $8,000$ planes that it produced because of insufficient demand in the market.

Economies of Scale and Natural Monopoly

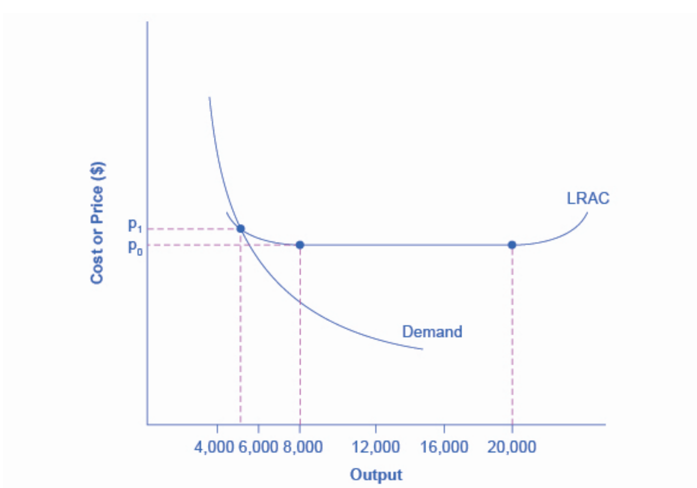


Figure $8.2.1$***:*** *In this market, the demand curve intersects the long-run average cost (LRAC) curve at its downward-sloping part. A natural monopoly occurs when the quantity demanded is less than the minimum quantity it takes to be at the bottom of the long-run average cost curve.*

This situation, when economies of scale are large relative to the quantity demanded in the market, is called a natural monopoly. Natural monopolies often arise in industries where the marginal cost of adding an additional customer is very low, once the fixed costs of the overall system are in place. Once the main water pipes are laid through a neighborhood, the marginal cost of providing water service to another home is fairly low. Once electricity lines are installed through a neighborhood, the marginal cost of providing additional electrical service to one more home is very low. It would be costly and duplicative for a second water company to enter the market and invest in a whole second set of main water pipes, or for a second electricity company to enter the market and invest in a whole new set of electrical wires. These industries offer an example where, because of economies of scale, one producer can serve the entire market more efficiently than a number of smaller producers that would need to make duplicate physical capital investments.

A natural monopoly can also arise in smaller local markets for products that are difficult to transport. For example, cement production exhibits economies of scale, and the quantity of cement demanded in a local area may not be much larger than what a single plant can produce. Moreover, the costs of transporting cement over land are high, and so a cement plant in an area without access to water transportation may be a natural monopoly.

Control of a Scarce Resources

Another type of natural monopoly occurs when a company has control of a scarce physical resource. In the U.S. economy, one historical example of this pattern occurred when ALCOA—the Aluminum Company of America—controlled most of the supply of bauxite, a key mineral used in making aluminum. Back in the 1930s, when ALCOA controlled most of the bauxite, other firms were simply unable to produce enough aluminum to compete.

As another example, the majority of global diamond production is controlled by DeBeers, a multi-national company that has mining and production operations in South Africa, Botswana, Namibia, and Canada. It also has exploration activities on four continents, while directing a worldwide distribution network of rough cut diamonds. Though in recent years they have experienced growing competition, their impact on the rough diamond market is still considerable.

Legal Restrictions

Legal Monopoly

For some products, the government erects barriers to entry by prohibiting or limiting competition. Under U.S. law, no organization but the U.S. Postal Service is legally allowed to deliver first-class mail. Many states or cities have laws or regulations that allow households a choice of only one electric company, one water company, and one company to pick up the garbage. Most legal monopolies are considered utilities—products necessary for everyday life—that are socially beneficial to have. As a consequence, the government allows producers to become regulated monopolies, to insure that an appropriate amount of these products is provided to consumers. Additionally, legal monopolies are often subject to economies of scale, so it makes sense to allow only one provider.

Promoting Innovation

Innovation takes time and resources to achieve. Suppose a company invests in research and development and finds the cure for the common cold. In this world of near ubiquitous information, other companies could take the formula, produce the drug, and because they did not incur the costs of research and development (R&D), undercut the price of the company that discovered the drug. Given this possibility, many firms would choose not to invest in research and development, and as a result, the world would have less innovation. To prevent this from happening, the Constitution of the United States specifies in Article I, Section 8: “The Congress shall have Power . . . To Promote the Progress of Science and Useful Arts, by securing for limited Times to Authors and Inventors the Exclusive Right to their Writings and Discoveries.” Congress used this power to create the U.S. Patent and Trademark Office, as well as the U.S. Copyright Office. A **patent** gives the inventor the exclusive legal right to make, use, or sell the invention for a limited time; in the United States, exclusive patent rights last for 20 years. The idea is to provide limited monopoly power so that innovative firms can recoup their investment in R&D, but then to allow other firms to produce the product more cheaply once the patent expires.

A **trademark** is an identifying symbol or name for a particular good, like Chiquita bananas, Chevrolet cars, or the Nike “swoosh” that appears on shoes and athletic gear. Roughly 1.9 million trademarks are registered with the U.S. government. A firm can renew a trademark over and over again, as long as it remains in active use.

A **copyright**, according to the U.S. Copyright Office, “is a form of protection provided by the laws of the United States for ‘original works of authorship’ including literary, dramatic, musical, architectural, cartographic, choreographic, pantomimic, pictorial, graphic, sculptural, and audiovisual creations.” No one can reproduce, display, or perform a copyrighted work without permission of the author. Copyright protection ordinarily lasts for the life of the author plus 70 years.

Roughly speaking, patent law covers inventions and copyright protects books, songs, and art. But in certain areas, like the invention of new software, it has been unclear whether patent or copyright protection should apply. There is also a body of law known as **trade secrets**. Even if a company does not have a patent on an invention, competing firms are not allowed to steal their secrets. One famous trade secret is the formula for Coca-Cola, which is not protected under copyright or patent law, but is simply kept secret by the company.

Taken together, this combination of patents, trademarks, copyrights, and trade secret law is called **intellectual property**, because it implies ownership over an idea, concept, or image, not a physical piece of property like a house or a car. Countries around the world have enacted laws to protect intellectual property, although the time periods and exact provisions of such laws vary across countries. There are ongoing negotiations, both through the World Intellectual Property Organization (WIPO) and through international treaties, to bring greater harmony to the intellectual property laws of different countries to determine the extent to which patents and copyrights in one country will be respected in other countries.

Government limitations on competition used to be even more common in the United States. For most of the twentieth century, only one phone company—AT&T—was legally allowed to provide local and long distance service. From the 1930s to the 1970s, one set of federal regulations limited which destinations airlines could choose to fly to and what fares they could charge; another set of regulations limited the interest rates that banks could pay to depositors; yet another specified what trucking firms could charge customers.

What products are considered utilities depends, in part, on the available technology. Fifty years ago, local and long distance telephone service was provided over wires. It did not make much sense to have multiple companies building multiple systems of wiring across towns and across the country. AT&T lost its monopoly on long distance service when the technology for providing phone service changed from wires to microwave and satellite transmission, so that multiple firms could use the same transmission mechanism. The same thing happened to local service, especially in recent years, with the growth in cellular phone systems.

The combination of improvements in production technologies and a general sense that the markets could provide services adequately led to a wave of **deregulation**, starting in the late 1970s and continuing into the 1990s. This wave eliminated or reduced government restrictions on the firms that could enter, the prices that could be charged, and the quantities that could be produced in many industries, including telecommunications, airlines, trucking, banking, and electricity.

Around the world, from Europe to Latin America to Africa and Asia, many governments continue to control and limit competition in what those governments perceive to be key industries, including airlines, banks, steel companies, oil companies, and telephone companies.

Intimidating Potential Competition

Businesses have developed a number of schemes for creating barriers to entry by deterring potential competitors from entering the market. One method is known as **predatory pricing**, in which a firm uses the threat of sharp price cuts to discourage competition. Predatory pricing is a violation of U.S. antitrust law, but it is difficult to prove.

Consider a large airline that provides most of the flights between two particular cities. A new, small start-up airline decides to offer service between these two cities. The large airline immediately slashes prices on this route to the bone, so that the new entrant cannot make any money. After the new entrant has gone out of business, the incumbent firm can raise prices again.

After this pattern is repeated once or twice, potential new entrants may decide that it is not wise to try to compete. Small airlines often accuse larger airlines of predatory pricing: in the early 2000s, for example, ValuJet accused Delta of predatory pricing, Frontier accused United, and Reno Air accused Northwest. In 2015, the Justice Department ruled against American Express and Mastercard for imposing restrictions on retailers who encouraged customers to use lower swipe fees on credit transactions.

In some cases, **large advertising budgets** can also act as a way of discouraging the competition. If the only way to launch a successful new national cola drink is to spend more than the promotional budgets of Coca-Cola and Pepsi Cola, not too many companies will try. A firmly established brand name can be difficult to dislodge.

Another way firms can intimidate competition is through the **threat of legal action**. Many large tech companies (e.g. Microsoft and Apple) will sue or threaten a lawsuit against smaller software companies who cannot afford the lawsuit and therefore will leave the market.

Summing Up Barriers to Entry

Table 8.2.1 lists the barriers to entry that have been discussed here. This list is not exhaustive, since firms have proved to be highly creative in inventing business practices that discourage competition. When barriers to entry exist, perfect competition is no longer a reasonable description of how an industry works. When barriers to entry are high enough, monopoly can result.

***Table 8.2.1**: Barriers to Entry*

Barrier to Entry	Government Role?	Example
Natural monopoly	Government often responds with regulation (or ownership)	Water and electric companies
Control of a physical resource	No	DeBeers for diamonds
Legal monopoly	Yes	Post office, past regulation of airlines and trucking
Patent, trademark, and copyright	Yes, through protection of intellectual property	New drugs or software
Intimidating potential competitors	Somewhat	Predatory pricing; well-known brand names

Key Concepts and Summary

Barriers to entry prevent or discourage competitors from entering the market. These barriers include: economies of scale that lead to natural monopoly; control of a physical resource; legal restrictions on competition; patent, trademark and copyright protection; and practices to intimidate the competition like predatory pricing. Intellectual property refers to legally guaranteed ownership of an idea, rather than a physical item. The laws that protect intellectual property include patents, copyrights, trademarks, and trade secrets. A natural monopoly arises when economies of scale persist over a large enough range of output that if one firm supplies the entire market, no other firm can enter without facing a cost disadvantage.

Glossary

barriers to entry

the legal, technological, or market forces that may discourage or prevent potential competitors from entering a market

copyright

a form of legal protection to prevent copying, for commercial purposes, original works of authorship, including books and music

deregulation

removing government controls over setting prices and quantities in certain industries

intellectual property

the body of law including patents, trademarks, copyrights, and trade secret law that protect the right of inventors to produce and sell their inventions

legal monopoly

legal prohibitions against competition, such as regulated monopolies and intellectual property protection

monopoly

a situation in which one firm produces all of the output in a market

natural monopoly

economic conditions in the industry, for example, economies of scale or control of a critical resource, that limit effective competition

patent

a government rule that gives the inventor the exclusive legal right to make, use, or sell the invention for a limited time

predatory pricing

when an existing firm uses sharp but temporary price cuts to discourage new competition

trade secrets

methods of production kept secret by the producing firm

trademark

an identifying symbol or name for a particular good and can only be used by the firm that registered that trademark

Contributor

- Template:ContribOpenSTAXMicroEcon

8.3: How a Profit-Maximizing Monopoly Chooses Output and Price

Skills to Develop

- Explain the perceived demand curve for a perfect competitor and a monopoly
- Analyze a demand curve for a monopoly and determine the output that maximizes profit and revenue
- Calculate marginal revenue and marginal cost
- Explain allocative efficiency as it pertains to the efficiency of a monopoly

Consider a monopoly firm, comfortably surrounded by barriers to entry so that it need not fear competition from other producers. How will this monopoly choose its profit-maximizing quantity of output, and what price will it charge? Profits for the monopolist, like any firm, will be equal to total revenues minus total costs. The pattern of costs for the monopoly can be analyzed within the same framework as the costs of a perfectly competitive firm—that is, by using total cost, fixed cost, variable cost, marginal cost, average cost, and average variable cost. However, because a monopoly faces no competition, its situation and its decision process will differ from that of a **perfectly competitive firm**. (The Clear it Up feature discusses how hard it is sometimes to define "market" in a monopoly situation.)

Demand Curves Perceived by a Perfectly Competitive Firm and by a Monopoly

A perfectly competitive firm acts as a price taker, so its calculation of total revenue is made by taking the given market price and multiplying it by the quantity of output that the firm chooses. The demand curve as *it is perceived by a perfectly competitive* firm appears in Figure $8.3.1$ (a). The flat perceived demand curve means that, from the viewpoint of the perfectly competitive firm, it could sell either a relatively low quantity like Ql or a relatively high quantity like Qh at the market price P.

The Perceived Demand Curve for a Perfect Competitor and a Monopolist

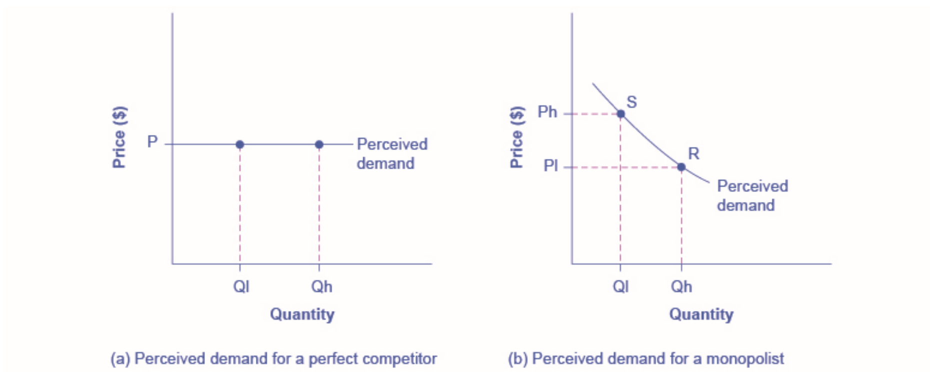


Figure $8.3.1$: *(a) A perfectly competitive firm perceives the demand curve that it faces to be flat. The flat shape means that the firm can sell either a low quantity (Ql) or a high quantity (Qh) at exactly the same price (P). (b) A monopolist perceives the demand curve that it faces to be the same as the market demand curve, which for most goods is downward-sloping. Thus, if the monopolist chooses a high level of output (Qh), it can charge only a relatively low price (Pl); conversely, if the monopolist chooses a low level of output (Ql), it can then charge a higher price (Ph). The challenge for the monopolist is to choose the combination of price and quantity that maximizes profits.*

Example $8.3.1$: What defines the market?

A monopoly is a firm that sells all or nearly all of the goods and services in a given market. But what defines the "market"?

In a famous 1947 case, the federal government accused the DuPont company of having a monopoly in the cellophane market, pointing out that DuPont produced 75% of the cellophane in the United States. DuPont countered that even though it had a 75% market share in cellophane, it had less than a 20% share of the "flexible packaging materials," which includes all other moisture-proof papers, films, and foils. In 1956, after years of legal appeals, the U.S. Supreme Court held that the broader market definition was more appropriate, and the case against DuPont was dismissed.

Questions over how to define the market continue today. True, Microsoft in the 1990s had a dominant share of the software for computer operating systems, but in the total market for all computer software and services, including everything from games to scientific programs, the Microsoft share was only about 14% in 2014. The Greyhound bus company may have a near-monopoly on the market for intercity bus transportation, but it is only a small share of the market for intercity transportation if that market includes private cars, airplanes, and railroad service. DeBeers has a monopoly in diamonds, but it is a much smaller share of the total market for precious gemstones and an even smaller share of the total market for jewelry. A small town in the country may have only one gas station: is this gas station a "monopoly," or does it compete with gas stations that might be 5, 10, or 50 miles away?

In general, if a firm produces a product without close substitutes, then the firm can be considered a monopoly producer in a single market. But if buyers have a range of similar—even if not identical—options available from other firms, then the firm is not a monopoly. Still, arguments over whether substitutes are close or not close can be controversial.

While a monopolist can charge any price for its product, that price is nonetheless constrained by demand for the firm's product. No monopolist, even one that is thoroughly protected by high barriers to entry, can require consumers to purchase its product. Because the monopolist is the only firm in the market, its demand curve is the same as the market demand curve, which is, unlike that for a perfectly competitive firm, downward-sloping.

Figure $8.3.1$ illustrates this situation. The monopolist can either choose a point like R with a low price (Pl) and high quantity (Qh), or a point like S with a high price (Ph) and a low quantity (Ql), or some intermediate point. Setting the price too high will result in a low quantity sold, and will not bring in much revenue. Conversely, setting the price too low may result in a high quantity sold, but because of the low price, it will not bring in much revenue either. The challenge for the monopolist is to strike a profit-maximizing balance between the price it charges and the quantity that it sells. But why isn't the perfectly competitive firm's demand curve also the market demand curve? See the following Clear it Up feature for the answer to this question.

What is the difference between perceived demand and market demand?

The demand curve as perceived by a perfectly competitive firm is not the overall market demand curve for that product. However, the firm's demand curve as perceived by a monopoly is the same as the market demand curve. The reason for the difference is that each perfectly competitive firm perceives the demand for its products in a market that includes many other firms; in effect, the demand curve perceived by a perfectly competitive firm is a tiny slice of the entire market demand curve. In contrast, a monopoly perceives demand for its product in a market where the monopoly is the only producer.

Total Cost and Total Revenue for a Monopolist

Profits for a monopolist can be illustrated with a graph of total revenues and total costs, as shown with the example of the hypothetical HealthPill firm in Figure $8.3.2$. The total cost curve has its typical shape; that is, total costs rise and the curve grows steeper as output increases.

Total Revenue and Total Cost for the HealthPill Monopoly

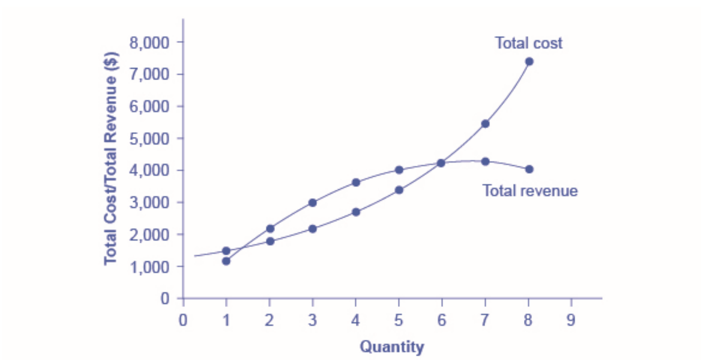


***Figure* $8.3.2$*:* *Total revenue for the monopoly firm called HealthPill first rises, then falls. Low levels of output bring in relatively little total revenue, because the quantity is low. High levels of output bring in relatively less revenue, because the high quantity pushes down the market price. The total cost curve is upward-sloping. Profits will be highest at the quantity of output where total revenue is most above total cost. Of the choices in Table $8.3.1$, the highest profits happen at an output of 4. The profit-maximizing level of output is not the same as the revenue-maximizing level of output, which should make sense, because profits take costs into account and revenues do not.*

***Table* $8.3.1$*:* *Total Costs and Total Revenues of HealthPill*

Quantity	Total Cost	Quantity	Price	Total Revenue	Profit = Total Revenue – Total Cost
1	1,500	1	1,200	1,200	–300
2	1,800	2	1,100	2,200	400
3	2,200	3	1,000	3,000	800
4	2,800	4	900	3,600	800
5	3,500	5	800	4,000	500
6	4,200	6	700	4,200	0
7	5,600	7	600	4,200	–1,400
8	7,400	8	500	4,000	–3,400

To calculate total revenue for a monopolist, start with the demand curve perceived by the monopolist. Table $8.3.1$ shows quantities along the demand curve and the price at each quantity demanded, and then calculates total revenue by multiplying price times quantity at each level of output. (In this example, the output is given as $1, 2, 3, 4$, and so on, for the sake of simplicity. If you prefer a dash of greater realism, you can imagine that these output levels and the corresponding prices are measured per $1,000$ or $10,000$ pills.) As the figure illustrates, total revenue for a monopolist rises, flattens out, and then falls. In this example, total revenue is highest at a quantity of 6 or 7.

Clearly, the total revenue for a monopolist is not a straight upward-sloping line, in the way that total revenue was for a perfectly competitive firm. The different total revenue pattern for a monopolist occurs because the quantity that a monopolist chooses to produce affects the market price, which was not true for a perfectly competitive firm. If the monopolist charges a very high price, then quantity demanded drops, and so total revenue is very low. If the monopolist charges a very low price, then, even if quantity demanded is very high, total revenue will not add up to much. At some intermediate level, total revenue will be highest.

However, the monopolist is not seeking to maximize revenue, but instead to earn the highest possible profit. Profits are calculated in the final row of the table. In the HealthPill example in Figure $8.3.2$, the highest profit will occur at the quantity where total revenue is the farthest above total cost. Of the choices given in the table, the highest profits occur at an output of 4, where profit is 800.

Marginal Revenue and Marginal Cost for a Monopolist

In the real world, a monopolist often does not have enough information to analyze its entire total revenues or total costs curves; after all, the firm does not know exactly what would happen if it were to alter production dramatically. But a monopolist often has fairly reliable information about how changing output by small or moderate amounts will affect its marginal revenues and marginal costs, because it has had experience with such changes over time and because modest changes are easier to extrapolate from current experience. A monopolist can use information on **marginal revenue** and **marginal cost** to seek out the profit-maximizing combination of quantity and price.

The first four columns of Table $8.3.2$ use the numbers on total cost from the HealthPill example in the previous exhibit and calculate marginal cost and average cost. This monopoly faces a typical upward-sloping marginal cost curve, as shown in Figure $8.3.3$. The second four columns of Table $8.3.2$ use the total revenue information from the previous exhibit and calculate marginal revenue.

Notice that marginal revenue is zero at a quantity of 7, and turns negative at quantities higher than 7. It may seem counterintuitive that marginal revenue could ever be zero or negative: after all, does an increase in quantity sold not always mean more revenue? For a perfect competitor, each additional unit sold brought a positive marginal revenue, because marginal revenue was equal to the given market price. But a monopolist can sell a larger quantity and see a decline in **total revenue**. When a monopolist increases sales by one unit, it gains some marginal revenue from selling that extra unit, but also loses some marginal revenue because every other unit must now be sold at a lower price. As the quantity sold becomes higher, the drop in price affects a greater quantity of sales, eventually causing a situation where more sales cause marginal revenue to be negative.

Marginal Revenue and Marginal Cost for the HealthPill Monopoly

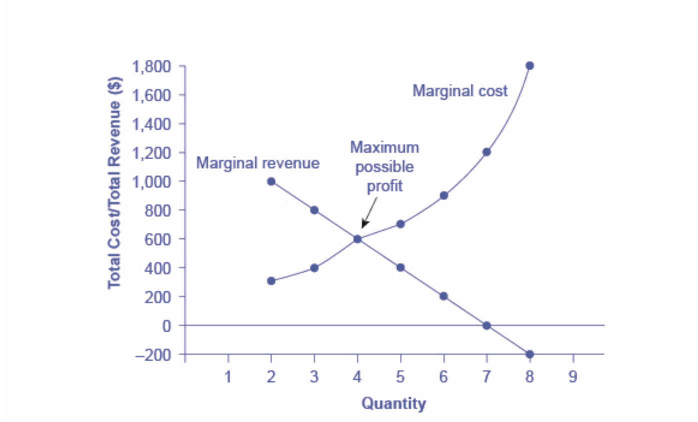


***Figure 8.3.3**: For a monopoly like HealthPill, marginal revenue decreases as additional units are sold. The marginal cost curve is upward-sloping. The profit-maximizing choice for the monopoly will be to produce at the quantity where marginal revenue is equal to marginal cost: that is, $MR = MC$. If the monopoly produces a lower quantity, then $MR > MC$ at those levels of output, and the firm can make higher profits by expanding output. If the firm produces at a greater quantity, then $MC > MR$, and the firm can make higher profits by reducing its quantity of output.*

***Table 8.3.2**: Costs and Revenues of HealthPill*

Cost Information				Revenue Information			
Quantity	Total Cost	Marginal Cost	Average Cost	Quantity	Price	Total Revenue	Marginal Revenue
1	1,500	1,500	1,500	1	1,200	1,200	1,200
2	1,800	300	900	2	1,100	2,200	1,000
3	2,200	400	733	3	1,000	3,000	800
4	2,800	600	700	4	900	3,600	600
5	3,500	700	700	5	800	4,000	400
6	4,200	700	700	6	700	4,200	200
7	5,600	1,400	800	7	600	4,200	0
8	7,400	1,800	925	8	500	4,000	–200

A monopolist can determine its profit-maximizing price and quantity by analyzing the marginal revenue and marginal costs of producing an extra unit. If the marginal revenue exceeds the marginal cost, then the firm should produce the extra unit.

For example, at an output of 3 in Figure 8.3.3, marginal revenue is 800 and marginal cost is 400, so producing this unit will clearly add to overall profits. At an output of 4, marginal revenue is 600 and marginal cost is 600, so producing this unit still means overall profits are unchanged. However, expanding output from 4 to 5 would involve a marginal revenue of 400 and a marginal cost of 700, so that fifth unit would actually reduce profits. Thus, the monopoly can tell from the marginal revenue and marginal cost that of the choices given in the table, the profit-maximizing level of output is 4.

Indeed, the monopoly could seek out the profit-maximizing level of output by increasing quantity by a small amount, calculating marginal revenue and marginal cost, and then either increasing output as long as marginal revenue exceeds marginal cost or reducing output if marginal cost exceeds marginal revenue. This process works without any need to calculate total revenue and total cost. Thus, a profit-maximizing monopoly should follow the rule of producing up to the quantity where marginal revenue is equal to marginal cost—that is, $MR = MC$.

Maximizing Profits

If you find it counterintuitive that producing where marginal revenue equals marginal cost will maximize profits, working through the numbers will help.

Step 1: Remember that marginal cost is defined as the change in total cost from producing a small amount of additional output.

$$MC = \frac{\text{change in total cost}}{\text{change in quantity produced}} \tag{8.3.1}$$

Step 2: Note that in Table $8.3.2$, as output increases from 1 to 2 units, total cost increases from $\$1500$ to $\$1800$. As a result, the marginal cost of the second unit will be:

$$\begin{aligned} MC &= \frac{\$1,800 - \$1,500}{1} \\ &= \$300 \end{aligned}$$

Step 3: Remember that, similarly, marginal revenue is the change in total revenue from selling a small amount of additional output.

$$MR = \frac{\text{change in total revenue}}{\text{change in quantity sold}} \tag{8.3.2}$$

Step 4: Note that in Table $8.3.2$, as output increases from 1 to 2 units, total revenue increases from $\$1200$ to $\$2200$. As a result, the marginal revenue of the second unit will be:

$$\begin{aligned} MC &= \frac{\$2,200 - \$1,200}{1} \\ &= \$1,000 \end{aligned}$$

***Table 8.3.3**: Marginal Revenue, Marginal Cost, Marginal and Total Profit*

Quantity	Marginal Revenue	Marginal Cost	Marginal Profit	Total Profit
1	1,200	1,500	–300	–300
2	1,000	300	700	400
3	800	400	400	800
4	600	600	0	800
5	400	700	–300	500
6	200	700	–500	0
7	0	1,400	–1,400	–1,400

Table $8.3.3$ repeats the marginal cost and marginal revenue data from Table $8.3.2$, and adds two more columns: Marginal profit is the profitability of each additional unit sold. It is defined as marginal revenue minus marginal cost. Finally, total profit is the sum of marginal profits. As long as marginal profit is positive, producing more output will increase total profits. When marginal profit turns negative, producing more output will decrease total profits. Total profit is maximized where marginal revenue equals marginal cost. In this example, maximum profit occurs at 4 units of output.

A perfectly competitive firm will also find its profit-maximizing level of output where $MR = MC$. The key difference with a perfectly competitive firm is that in the case of perfect competition, marginal revenue is equal to price ($MR = P$), while for a monopolist, marginal revenue is not equal to the price, because changes in quantity of output affect the price.

Illustrating Monopoly Profits

It is straightforward to calculate profits of given numbers for total revenue and total cost. However, the size of monopoly profits can also be illustrated graphically with Figure $8.3.4$, which takes the marginal cost and marginal revenue curves from the previous exhibit and adds an average cost curve and the monopolist's perceived demand curve.

Illustrating Profits at the HealthPill Monopoly

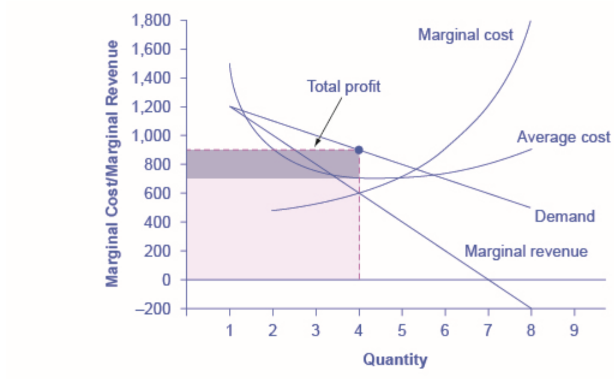


***Figure 8.3.4**: This figure begins with the same marginal revenue and marginal cost curves from the HealthPill monopoly presented in Figure 8.3.3. It then adds an average cost curve and the demand curve faced by the monopolist. The HealthPill firm first chooses the quantity where MR = MC; in this example, the quantity is 4. The monopolist then decides what price to charge by looking at the demand curve it faces. The large box, with quantity on the horizontal axis and marginal revenue on the vertical axis, shows total revenue for the firm. Total costs for the firm are shown by the lighter-shaded box, which is quantity on the horizontal axis and marginal cost of production on the vertical axis. The large total revenue box minus the smaller total cost box leaves the darkly shaded box that shows total profits. Since the price charged is above average cost, the firm is earning positive profits.*

Figure 8.3.5 illustrates the three-step process where a monopolist: selects the profit-maximizing quantity to produce; decides what price to charge; determines total revenue, total cost, and profit.

Step 1: The Monopolist Determines Its Profit-Maximizing Level of Output

The firm can use the points on the demand curve D to calculate total revenue, and then, based on total revenue, calculate its marginal revenue curve. The profit-maximizing quantity will occur where $MR = MC$ —or at the last possible point before marginal costs start exceeding marginal revenue. On Figure 8.3.4, $MR = MC$ occurs at an output of 4.

Step 2: The Monopolist Decides What Price to Charge

The monopolist will charge what the market is willing to pay. A dotted line drawn straight up from the profit-maximizing quantity to the demand curve shows the profit-maximizing price. This price is above the average cost curve, which shows that the firm is earning profits.

Step 3: Calculate Total Revenue, Total Cost, and Profit

Total revenue is the overall shaded box, where the width of the box is the quantity being sold and the height is the price. In Figure 8.3.4, the bottom part of the shaded box, which is shaded more lightly, shows total costs; that is, quantity on the horizontal axis multiplied by average cost on the vertical axis. The larger box of total revenues minus the smaller box of total costs will equal profits, which is shown by the darkly shaded box. In a perfectly competitive market, the forces of entry would erode this profit in the long run. But a monopolist is protected by barriers to entry. In fact, one telltale sign of a possible monopoly is when a firm earns profits year after year, while doing more or less the same thing, without ever seeing those profits eroded by increased competition.

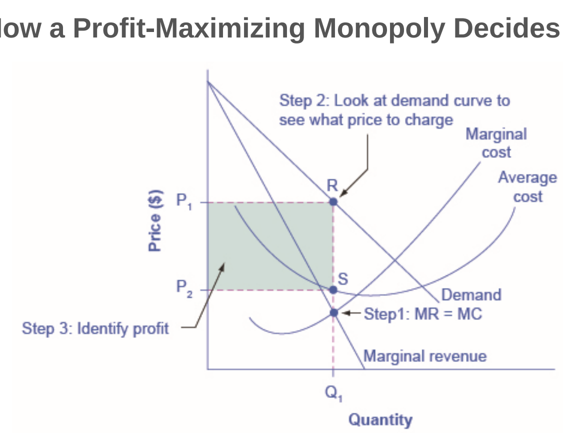


Figure $8.3.5$: In Step 1, the monopoly chooses the profit-maximizing level of output Q_1, by choosing the quantity where $MR = MC$. In Step 2, the monopoly decides how much to charge for output level Q_1 by drawing a line straight up from Q_1 to point R on its perceived demand curve. Thus, the monopoly will charge a price (P_1). In Step 3, the monopoly identifies its profit. Total revenue will be Q_1 multiplied by P_1. Total cost will be Q_1 multiplied by the average cost of producing Q_1, which is shown by point S on the average cost curve to be P_2. Profits will be the total revenue rectangle minus the total cost rectangle, shown by the shaded zone in the figure.

Example $8.3.2$: Why is a monopolist's marginal revenue always less than the price?

The marginal revenue curve for a monopolist always lies beneath the market demand curve. To understand why, think about increasing the quantity along the demand curve by one unit, so that you take one step down the demand curve to a slightly higher quantity but a slightly lower price. A demand curve is not sequential: It is not that first we sell Q_1 at a higher price, and then we sell Q_2 at a lower price. Rather, a demand curve is conditional: If we charge the higher price, we would sell Q_1. If, instead, we charge a lower price (on all the units that we sell), we would sell Q_2.

So when we think about increasing the quantity sold by one unit, marginal revenue is affected in two ways. First, we sell one additional unit at the new market price. Second, all the previous units, which could have been sold at the higher price, now sell for less. Because of the lower price on all units sold, the marginal revenue of selling a unit is less than the price of that unit—and the marginal revenue curve is below the demand curve. *Tip:* For a straight-line demand curve, MR and demand have the same vertical intercept. As output increases, marginal revenue decreases twice as fast as demand, so that the horizontal intercept of MR is halfway to the horizontal intercept of demand. You can see this in the Figure $8.3.6$.

The Monopolist's Marginal Revenue Curve versus Demand Curve

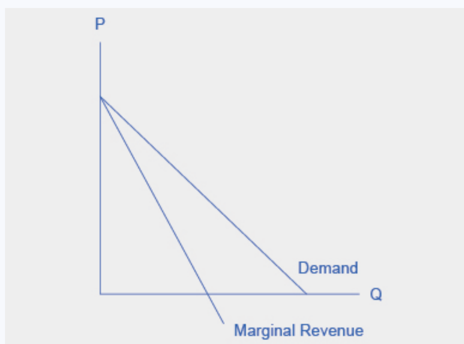


Figure $8.3.6$: Because the market demand curve is conditional, the marginal revenue curve for a monopolist lies beneath the demand curve.

The Inefficiency of Monopoly

Most people criticize monopolies because they charge too high a price, but what economists object to is that monopolies do not supply enough output to be allocatively efficient. To understand why a monopoly is inefficient, it is useful to compare it with the benchmark model of perfect competition.

Allocative efficiency is a social concept. It refers to producing the optimal quantity of some output, the quantity where the marginal benefit to society of one more unit just equals the marginal cost. The rule of profit maximization in a world of perfect competition was for each firm to produce the quantity of output where $P = MC$, where the price (P) is a measure of how much buyers value the good and the marginal cost (MC) is a measure of what marginal units cost society to produce. Following this rule assures allocative efficiency. If $P > MC$, then the marginal benefit to society (as measured by P) is greater than the marginal cost to society of producing additional units, and a greater quantity should be produced. But in the case of monopoly, price is always greater than marginal cost at the profit-maximizing level of output, as can be seen by looking back at Figure $8.3.4$. Thus, consumers will suffer from a monopoly because a lower quantity will be sold in the market, at a higher price, than would have been the case in a perfectly competitive market.

The problem of inefficiency for monopolies often runs even deeper than these issues, and also involves incentives for efficiency over longer periods of time. There are counterbalancing incentives here. On one side, firms may strive for new inventions and new intellectual property because they want to become monopolies and earn high profits—at least for a few years until the competition catches up. In this way, monopolies may come to exist because of competitive pressures on firms. However, once a barrier to entry is in place, a monopoly that does not need to fear competition can just produce the same old products in the same old way—while still ringing up a healthy rate of profit. John Hicks, who won the Nobel Prize for economics in 1972, wrote in 1935: "*The best of all monopoly profits is a quiet life.*" He did not mean the comment in a complimentary way. He meant that monopolies may bank their profits and slack off on trying to please their customers.

When AT&T provided all of the local and long-distance phone service in the United States, along with manufacturing most of the phone equipment, the payment plans and types of phones did not change much. The old joke was that you could have any color phone you wanted, as long as it was black. But in 1982, AT&T was split up by government litigation into a number of local phone companies, a long-distance phone company, and a phone equipment manufacturer. An explosion of innovation followed. Services like call waiting, caller ID, three-way calling, voice mail though the phone company, mobile phones, and wireless connections to the Internet all became available. A wide range of payment plans was offered, as well. It was no longer true that all phones were black; instead, phones came in a wide variety of shapes and colors. The end of the telephone monopoly brought lower prices, a greater quantity of services, and also a wave of innovation aimed at attracting and pleasing customers.

The Rest is History

In the opening case, the East India Company and the Confederate States were presented as a monopoly or near monopoly provider of a good. Nearly every American schoolchild knows the result of the 'unwelcome visit' the 'Mohawks' bestowed upon Boston Harbor's tea-bearing ships—the Boston Tea Party. Regarding the cotton industry, we also know Great Britain remained neutral during the Civil War, taking neither side during the conflict.

Did the monopoly nature of these business have unintended and historical consequences? Might the American Revolution have been deterred, if the East India Company had sailed the tea-bearing ships back to England? Might the southern states have made different decisions had they not been so confident "King Cotton" would force diplomatic recognition of the Confederate States of America? Of course, it is not possible to definitively answer these questions; after all we cannot roll back the clock and try a different scenario. We can, however, consider the monopoly nature of these businesses and the roles they played and hypothesize about what might have occurred under different circumstances.

Perhaps if there had been legal free tea trade, the colonists would have seen things differently; there was smuggled Dutch tea in the colonial market. If the colonists had been able to freely purchase Dutch tea, they would have paid lower prices and avoided the tax.

What about the cotton monopoly? With one in five jobs in Great Britain depending on Southern cotton and the Confederate States nearly the sole provider of that cotton, why did Great Britain remain neutral during the Civil War? At the beginning of the war, Britain simply drew down massive stores of cotton. These stockpiles lasted until near the end of 1862. Why did Britain not recognize the Confederacy at that point? Two reasons: The Emancipation Proclamation and new sources of cotton. Having outlawed slavery throughout the United Kingdom in 1833, it was politically impossible for Great Britain, empty cotton warehouses or not, to recognize, diplomatically, the Confederate States. In addition, during the two years it took to draw down the stockpiles, Britain expanded cotton imports from India, Egypt, and Brazil.

Monopoly sellers often see no threats to their superior marketplace position. In these examples did the power of the monopoly blind the decision makers to other possibilities? Perhaps. But, as they say, the rest is history.

Key Concepts and Summary

A monopolist is not a price taker, because when it decides what quantity to produce, it also determines the market price. For a monopolist, total revenue is relatively low at low quantities of output, because not much is being sold. Total revenue is also relatively low at very high quantities of output, because a very high quantity will sell only at a low price. Thus, total revenue for a monopolist will start low, rise, and then decline. The marginal revenue for a monopolist from selling additional units will decline. Each additional unit sold by a monopolist will push down the overall market price, and as more units are sold, this lower price applies to more and more units.

The monopolist will select the profit-maximizing level of output where $MR = MC$, and then charge the price for that quantity of output as determined by the market demand curve. If that price is above average cost, the monopolist earns positive profits.

Monopolists are not productively efficient, because they do not produce at the minimum of the average cost curve. Monopolists are not allocatively efficient, because they do not produce at the quantity where $P = MC$. As a result, monopolists produce less, at a higher average cost, and charge a higher price than would a combination of firms in a perfectly competitive industry. Monopolists also may lack incentives for innovation, because they need not fear entry.

References

Aboukhadijeh, Feross. "Chapter 20: Girding for War - The North and the South, 1861-1865." StudyNotes, Inc. Accessed July 7, 2013. http://www.apstudynotes.org/us-histo...uth-1861-1865/.

British Parliament. "(28 August 1833). Slavery Abolition Act 1833; Section LXIV." Accessed July 2013. http://www.pdavis.nl/Legis_07.htm.

Dattel, E. (nd). "Cotton and the Civil War." Mississippi Historical Society. Accessed July 2013. http://mshistorynow.mdah.state.ms.us...-the-civil-war.

Gartner. 2015. "Gartner Says Tablet Sales Continue to Be Slow in 2015." Accessed March 12, 2015. http://www.gartner.com/newsroom/id/2954317.

Grogan, David. 2015. "Federal Judge Finds AmEx's Anti-Steering Rule Violates Antitrust Law." American Booksellers Association. Accessed March 12, 2015. http://www.bookweb.org/news/federal-...-antitrust-law.

Massachusetts Historical Society. "The Coming of the American Revolution 1764-1776: The Boston Tea Party." Retrieved from http://www.masshist.org/revolution/teaparty.php.

Massachusetts Historical Society. "Whereas our Nation." The Massachusetts Gazette, p. 2. Accessed July 2013 http://www.masshist.org/revolution/i...k=1&mode=large.

Pelegrin, William. 2015. "Judge Overrules Antitrust Case Against Google , Says Setting Default Search Engines Is Fair." Digital Trends. Accessed March 12, 2015. http://www.digitaltrends.com/mobile/...trust-lawsuit/.

Glossary

allocative efficiency

producing the optimal quantity of some output; the quantity where the marginal benefit to society of one more unit just equals the marginal cost

marginal profit

profit of one more unit of output, computed as marginal revenue minus marginal cost

Contributor

- Template:ContribOpenSTAXMicroEcon

8.E: Monopoly (Exercises)

9.1: How Monopolies Form: Barriers to Entry

Self-Check Questions

Q1

Classify the following as a government-enforced barrier to entry, a barrier to entry that is not government-enforced, or a situation that does not involve a barrier to entry.

a. A patented invention
b. A popular but easily copied restaurant recipe
c. An industry where economies of scale are very small compared to the size of demand in the market
d. A well-established reputation for slashing prices in response to new entry
e. A well-respected brand name that has been carefully built up over many years

Q2

Classify the following as a government-enforced barrier to entry, a barrier to entry that is not government-enforced, or a situation that does not involve a barrier to entry.

a. A city passes a law on how many licenses it will issue for taxicabs
b. A city passes a law that all taxicab drivers must pass a driving safety test and have insurance
c. A well-known trademark
d. Owning a spring that offers very pure water
e. An industry where economies of scale are very large compared to the size of demand in the market

Q3

Suppose the local electrical utility, a legal monopoly based on economies of scale, was split into four firms of equal size, with the idea that eliminating the monopoly would promote competitive pricing of electricity. What do you anticipate would happen to prices?

Q4

If Congress reduced the period of patent protection from 20 years to 10 years, what would likely happen to the amount of private research and development?

Review Questions

Q5

How is monopoly different from perfect competition?

Q6

What is a barrier to entry? Give some examples.

Q7

What is a natural monopoly?

Q8

What is a legal monopoly?

Q9

What are diminishing marginal returns?

Q10

What is predatory pricing?

Q11

How is intellectual property different from other property?

Q12

By what legal mechanisms is intellectual property protected?

Q13

In what sense is a natural monopoly “natural”?

Critical Thinking Questions

Q14

ALCOA does not have the monopoly power it once had. How do you suppose their barriers to entry were weakened?

Q15

Why are generic pharmaceuticals significantly cheaper than name brand ones?

Q16

For many years, the Justice Department has tried to break up large firms like IBM, Microsoft, and most recently Google, on the grounds that their large market share made them essentially monopolies. In a global market, where U.S. firms compete with firms from other countries, would this policy make the same sense as it might in a purely domestic context?

Q17

Intellectual property laws are intended to promote innovation, but some economists, such as Milton Friedman, have argued that such laws are not desirable. In the United States, there is no intellectual property protection for food recipes or for fashion designs. Considering the state of these two industries, and bearing in mind the discussion of the inefficiency of monopolies, can you think of any reasons why intellectual property laws might hinder innovation in some cases?

Problems

Q18

Return to Figure 9.1.1. Suppose P_0 is $\$10$ and P_1 is $\$11$. Suppose a new firm with the same $LRAC$ curve as the incumbent tries to break into the market by selling $4,000$ units of output. Estimate from the graph what the new firm’s average cost of producing output would be. If the incumbent continues to produce $6,000$ units, how much output would be supplied to the market by the two firms? Estimate what would happen to the market price as a result of the supply of both the incumbent firm and the new entrant. Approximately how much profit would each firm earn?

Solution

S1

a. A patent is a government-enforced barrier to entry.
b. This is not a barrier to entry.
c. This is not a barrier to entry.
d. This is a barrier to entry, but it is not government-enforced.
e. This is a barrier to entry, but it is not directly government enforced.

S2

a. This is a government-enforced barrier to entry.
b. This is an example of a government law, but perhaps it is not much of a barrier to entry if most people can pass the safety test and get insurance.
c. Trademarks are enforced by government, and therefore are a barrier to entry.
d. This is probably not a barrier to entry, since there are a number of different ways of getting pure water.
e. This is a barrier to entry, but it is not government-enforced.

S3

Because of economies of scale, each firm would produce at a higher average cost than before. (They would each have to build their own power lines.) As a result, they would each have to raise prices to cover their higher costs. The policy would fail.

S4

Shorter patent protection would make innovation less lucrative, so the amount of research and development would likely decline.

9.2: How a Profit-Maximizing Monopoly Chooses Output and Price

Self-Check Questions

Q1

Suppose demand for a monopoly's product falls so that its profit-maximizing price is below average variable cost. How much output should the firm supply? *Hint:* Draw the graph.

Q2

Imagine a monopolist could charge a different price to every customer based on how much he or she were willing to pay. How would this affect monopoly profits?

Review Questions

Q3

How is the demand curve perceived by a perfectly competitive firm different from the demand curve perceived by a monopolist?

Q4

How does the demand curve perceived by a monopolist compare with the market demand curve?

Q5

Is a monopolist a price taker? Explain briefly.

Q6

What is the usual shape of a total revenue curve for a monopolist? Why?

Q7

What is the usual shape of a marginal revenue curve for a monopolist? Why?

Q8

How can a monopolist identify the profit-maximizing level of output if it knows its total revenue and total cost curves?

Q9

How can a monopolist identify the profit-maximizing level of output if it knows its marginal revenue and marginal costs?

Q10

When a monopolist identifies its profit-maximizing quantity of output, how does it decide what price to charge?

Q11

Is a monopolist allocatively efficient? Why or why not?

Q12

How does the quantity produced and price charged by a monopolist compare to that of a perfectly competitive firm?

Critical Thinking Questions

Q13

Imagine that you are managing a small firm and thinking about entering the market of a monopolist. The monopolist is currently charging a high price, and you have calculated that you can make a nice profit charging 10% less than the monopolist. Before you go ahead and challenge the monopolist, what possibility should you consider for how the monopolist might react?

Q14

If a monopoly firm is earning profits, how much would you expect these profits to be diminished by entry in the long run?

Problems

Q15

Draw the demand curve, marginal revenue, and marginal cost curves from Figure 9.2.4, and identify the quantity of output the monopoly wishes to supply and the price it will charge. Suppose demand for the monopoly's product increases dramatically. Draw the new demand curve. What happens to the marginal revenue as a result of the increase in demand? What happens to the marginal cost curve? Identify the new profit-maximizing quantity and price. Does the answer make sense to you?

Q16

Draw a monopolist's demand curve, marginal revenue, and marginal cost curves. Identify the monopolist's profit-maximizing output level. Now, think about a slightly higher level of output (say $Q_0 + 1$). According to the graph, is there any consumer willing to pay more than the marginal cost of that new level of output? If so, what does this mean?

Solution

S1

If price falls below AVC, the firm will not be able to earn enough revenues even to cover its variable costs. In such a case, it will suffer a smaller loss if it shuts down and produces no output. By contrast, if it stayed in operation and produced the level of output where $MR = MC$, it would lose all of its fixed costs plus some variable costs. If it shuts down, it only loses its fixed costs.

S2

This scenario is called "perfect price discrimination." The result would be that the monopolist would produce more output, the same amount in fact as would be produced by a perfectly competitive industry. However, there would be no consumer surplus since each buyer is paying exactly what they think the product is worth. Therefore, the monopolist would be earning the maximum possible profits.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

9: Monopolistic Competition and Oligopoly

9.1: Introduction

Learning Objectives

- Monopolistic Competition
- Oligopoly

Competing Brands?

Figure 9.1.1: The laundry detergent market is one that is characterized neither as perfect competition nor monopoly. (Credit: modification of work by Pixel Drip/Flickr Creative Commons)

The Temptation to Defy the Law

Laundry detergent and bags of ice—products of industries that seem pretty mundane, maybe even boring. Hardly! Both have been the center of clandestine meetings and secret deals worthy of a spy novel. In France, between 1997 and 2004, the top four laundry detergent producers (Proctor & Gamble, Henkel, Unilever, and Colgate-Palmolive) controlled about 90 percent of the French soap market. Officials from the soap firms were meeting secretly, in out-of-the-way, small cafés around Paris. Their goals: Stamp out competition and set prices.

Around the same time, the top five Midwest ice makers (Home City Ice, Lang Ice, Tinley Ice, Sisler's Dairy, and Products of Ohio) had similar goals in mind when they secretly agreed to divide up the bagged ice market.

If both groups could meet their goals, it would enable each to act as though they were a single firm—in essence, a monopoly—and enjoy monopoly-size profits. The problem? In many parts of the world, including the European Union and the United States, it is illegal for firms to divide up markets and set prices collaboratively.

These two cases provide examples of markets that are characterized neither as perfect competition nor monopoly. Instead, these firms are competing in market structures that lie between the extremes of monopoly and perfect competition. How do they behave? Why do they exist? We will revisit this case later, to find out what happened.

Perfect competition and monopoly are at opposite ends of the competition spectrum. A perfectly competitive market has many firms selling identical products, who all act as **price takers** in the face of the competition. If you recall, price takers are firms that have no market power. They simply have to take the market price as given.

Monopoly arises when a single firm sells a product for which there are no close substitutes. Microsoft, for instance, has been considered a monopoly because of its domination of the operating systems market.

What about the vast majority of real world firms and organizations that fall between these extremes, firms that could be described as **imperfectly competitive**? What determines their behavior? They have more influence over the price they charge than perfectly competitive firms, but not as much as a monopoly would. What will they do?

One type of imperfectly competitive market is called **monopolistic competition**. Monopolistically competitive markets feature a large number of competing firms, but the products that they sell are not identical. Consider, as an example, the Mall of America in Minnesota, the largest shopping mall in the United States. In 2010, the Mall of America had 24 stores that sold women's "ready-to-wear" clothing (like Ann Taylor and Urban Outfitters), another 50 stores that sold clothing for both men and women (like Banana Republic, J. Crew, and Nordstrom's), plus 14 more stores that sold women's specialty clothing (like Motherhood Maternity and Victoria's Secret). Most of the markets that consumers encounter at the retail level are monopolistically competitive.

The other type of imperfectly competitive market is **oligopoly**. Oligopolistic markets are those dominated by a small number of firms. Commercial aircraft provides a good example: Boeing and Airbus each produce slightly less than 50% of the large commercial aircraft in the world. Another example is the U.S. soft drink industry, which is dominated by Coca-Cola and Pepsi. Oligopolies are characterized by high barriers to entry with firms choosing output, pricing, and other decisions strategically based on the decisions of the other firms in the market. In this chapter, we first explore how monopolistically competitive firms will choose their profit-maximizing level of output. We will then discuss oligopolistic firms, which face two conflicting temptations: to collaborate as if they were a single monopoly, or to individually compete to gain profits by expanding output levels and cutting prices. Oligopolistic markets and firms can also take on elements of monopoly and of perfect competition.

Contributor

- Template:ContribOpenSTAXMicroEcon

9.2: Monopolistic Competition

Skills to Develop

- Explain the significance of differentiated products
- Describe how a monopolistic competitor chooses price and quantity
- Discuss entry, exit, and efficiency as they pertain to monopolistic competition
- Analyze how advertising can impact monopolistic competition

Monopolistic competition involves many **firms** competing against each other, but selling products that are distinctive in some way. Examples include stores that sell different styles of clothing; restaurants or grocery stores that sell different kinds of food; and even products like golf balls or beer that may be at least somewhat similar but differ in public perception because of advertising and brand names. There are over $600,000$ restaurants in the United States. When products are distinctive, each firm has a mini-monopoly on its particular style or flavor or brand name. However, firms producing such products must also compete with other styles and flavors and brand names. The term “monopolistic competition” captures this mixture of mini-monopoly and tough competition, and the following Clear It Up feature introduces its derivation.

Who invented the theory of imperfect competition?

The theory of imperfect competition was developed by two economists independently but simultaneously in 1933. The first was Edward Chamberlin of Harvard University who published *The Economics of Monopolistic Competition*. The second was Joan Robinson of Cambridge University who published The Economics of Imperfect Competition. Robinson subsequently became interested in macroeconomics where she became a prominent Keynesian, and later a post-Keynesian economist. (See the Welcome to Economics! chapters and The Keynesian Perspective for more on Keynes.)

Differentiated Products

A firm can try to make its products different from those of its competitors in several ways: physical aspects of the product, location from which the product is sold, intangible aspects of the product, and perceptions of the product. Products that are distinctive in one of these ways are called **differentiated products**.

Physical aspects of a product include all the phrases you hear in advertisements: unbreakable bottle, nonstick surface, freezer-to-microwave, non-shrink, extra spicy, newly redesigned for your comfort. The location of a firm can also create a difference between producers. For example, a gas station located at a heavily traveled intersection can probably sell more gas, because more cars drive by that corner. A supplier to an automobile manufacturer may find that it is an advantage to locate close to the car factory.

Intangible aspects can differentiate a product, too. Some intangible aspects may be promises like a guarantee of satisfaction or money back, a reputation for high quality, services like free delivery, or offering a loan to purchase the product. Finally, **product differentiation** may occur in the minds of buyers. For example, many people could not tell the difference in taste between common varieties of beer or cigarettes if they were blindfolded but, because of past habits and advertising, they have strong preferences for certain brands. Advertising can play a role in shaping these intangible preferences.

The concept of differentiated products is closely related to the degree of variety that is available. If everyone in the economy wore only blue jeans, ate only white bread, and drank only tap water, then the markets for clothing, food, and drink would be much closer to perfectly competitive. The variety of styles, flavors, locations, and characteristics creates product differentiation and monopolistic competition.

Perceived Demand for a Monopolistic Competitor

A monopolistically competitive firm perceives a demand for its goods that is an intermediate case between monopoly and competition. Figure $9.2.1$ offers a reminder that the **demand curve** as faced by a perfectly competitive firm is **perfectly elastic** or flat, because the perfectly competitive firm can sell any quantity it wishes at the prevailing **market price**. In contrast, the demand curve, as faced by a monopolist, is the market demand curve, since a monopolist is the only firm in the market, and hence is downward sloping.

Perceived Demand for Firms in Different Competitive Settings

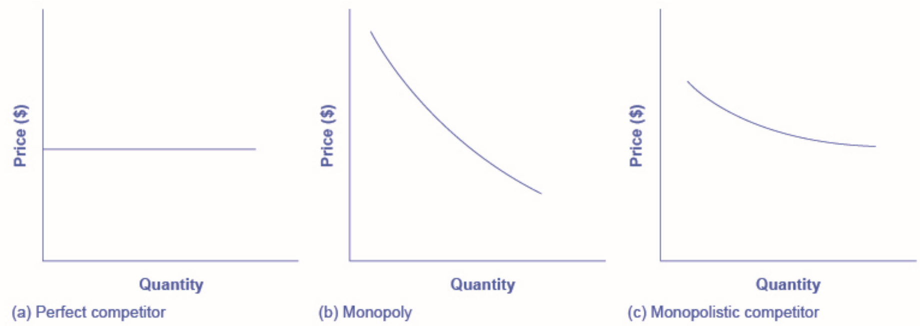


Figure* $9.2.1$*: The demand curve faced by a perfectly competitive firm is perfectly elastic, meaning it can sell all the output it wishes at the prevailing market price. The demand curve faced by a monopoly is the market demand. It can sell more output only by decreasing the price it charges. The demand curve faced by a monopolistically competitive firm falls in between.

The demand curve as faced by a monopolistic competitor is not flat, but rather downward-sloping, which means that the monopolistic competitor can raise its price without losing all of its customers or lower the price and gain more customers. Since there are substitutes, the demand curve facing a monopolistically competitive firm is more elastic than that of a monopoly where there are no close substitutes. If a monopolist raises its price, some consumers will choose not to purchase its product—but they will then need to buy a completely different product. However, when a monopolistic competitor raises its price, some consumers will choose not to purchase the product at all, but others will choose to buy a similar product from another firm. If a monopolistic competitor raises its price, it will not lose as many customers as would a perfectly competitive firm, but it will lose more customers than would a monopoly that raised its prices.

At a glance, the demand curves faced by a monopoly and by a monopolistic competitor look similar—that is, they both slope down. But the underlying economic meaning of these perceived demand curves is different, because a monopolist faces the market demand curve and a monopolistic competitor does not. Rather, a monopolistically competitive firm's demand curve is but one of many firms that make up the "before" market demand curve. Are you following? If so, how would you categorize the market for golf balls? Take a swing, then see the following Clear It Up feature.

Are golf balls really differentiated products?

Monopolistic competition refers to an industry that has more than a few firms, each offering a product which, from the consumer's perspective, is different from its competitors. The U.S. Golf Association runs a laboratory that tests $20,000$ golf balls a year. There are strict rules for what makes a golf ball legal. The weight of a golf ball cannot exceed 1.620 ounces and its diameter cannot be less than 1.680 inches (which is a weight of 45.93 grams and a diameter of 42.67 millimeters, in case you were wondering). The balls are also tested by being hit at different speeds. For example, the distance test involves having a mechanical golfer hit the ball with a titanium driver and a swing speed of 120 miles per hour. As the testing center explains: "*The USGA system then uses an array of sensors that accurately measure the flight of a golf ball during a short, indoor trajectory from a ball launcher. From this flight data, a computer calculates the lift and drag forces that are generated by the speed, spin, and dimple pattern of the ball. ... The distance limit is* 317 *yards.*"

Over 1800 golf balls made by more than 100 companies meet the USGA standards. The balls do differ in various ways, like the pattern of dimples on the ball, the types of plastic used on the cover and in the cores, and so on. Since all balls need to conform to the USGA tests, they are much more alike than different. In other words, golf ball manufacturers are monopolistically competitive.

However, retail sales of golf balls are about $\$500$ million per year, which means that a lot of large companies have a powerful incentive to persuade players that golf balls are highly differentiated and that it makes a huge difference which one you choose. Sure, Tiger Woods can tell the difference. For the average duffer (golf-speak for a "mediocre player") who plays a few times a summer—and who loses a lot of golf balls to the woods and lake and needs to buy new ones—most golf balls are pretty much indistinguishable.

How a Monopolistic Competitor Chooses Price and Quantity

The monopolistically competitive firm decides on its profit-maximizing quantity and price in much the same way as a monopolist. A monopolistic competitor, like a monopolist, faces a downward-sloping demand curve, and so it will choose some combination of price and quantity along its perceived demand curve.

As an example of a **profit-maximizing** monopolistic competitor, consider the Authentic Chinese Pizza store, which serves pizza with cheese, sweet and sour sauce, and your choice of vegetables and meats. Although Authentic Chinese Pizza must compete against other pizza businesses and restaurants, it has a differentiated product. The firm's perceived demand curve is downward sloping, as shown in Figure $9.2.2$ and the first two columns of Table $9.2.1$.

How a Monopolistic Competitor Chooses its Profit Maximizing Output and Price

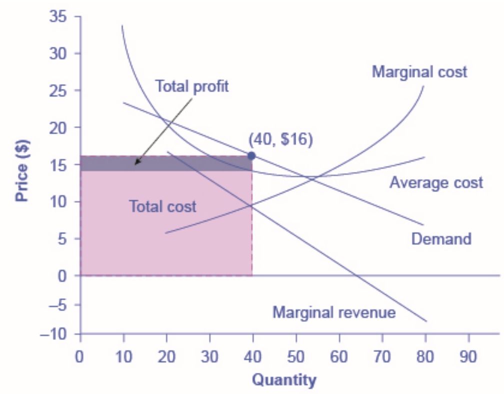


Figure $9.2.2$: *To maximize profits, the Authentic Chinese Pizza shop would choose a quantity where marginal revenue equals marginal cost, or Q where $MR = MC$. Here it would choose a quantity of 40 and a price of $\$16$.*

Table **$9.2.1$*:*** *Revenue and Cost Schedule*

Quantity	Price	Total Revenue	Marginal Revenue	Total Cost	Marginal Cost	Average Cost
10	$23	$230	-	$340	-	$34
20	$20	$400	$17	$400	$6	$20
30	$18	$540	$14	$480	$8	$16
40	$16	$640	$10	$580	$10	$14.50
50	$14	$700	$6	$700	$12	$14
60	$12	$720	$2	$840	$14	$14
70	$10	$700	–$2	$1,020	$18	$14.57
80	$8	$640	–$6	$1,280	$26	$16

The combinations of price and quantity at each point on the demand curve can be multiplied to calculate the total revenue that the firm would receive, which is shown in the third column of Table $9.2.1$. The fourth column, marginal revenue, is calculated as the change in total revenue divided by the change in quantity. The final columns of Table $9.2.1$ show total cost, marginal cost, and average cost. As always, marginal cost is calculated by dividing the change in total cost by the change in quantity, while average cost is calculated by dividing total cost by quantity. The following Work It Out feature shows how these firms calculate how much of its product to supply at what price.

Example $9.2.1$: How a Monopolistic Competitor Determines How Much to Produce and at What Price

The process by which a monopolistic competitor chooses its profit-maximizing quantity and price resembles closely how a monopoly makes these decisions process. First, the firm selects the profit-maximizing quantity to produce. Then the firm decides what price to charge for that quantity.

Step 1: The monopolistic competitor determines its profit-maximizing level of output. In this case, the Authentic Chinese Pizza company will determine the profit-maximizing quantity to produce by considering its marginal revenues and marginal costs.

Two scenarios are possible:

If the firm is producing at a quantity of output where marginal revenue exceeds marginal cost, then the firm should keep expanding production, because each marginal unit is adding to profit by bringing in more revenue than its cost. In this way, the firm will produce up to the quantity where $MR = MC$.

If the firm is producing at a quantity where marginal costs exceed marginal revenue, then each marginal unit is costing more than the revenue it brings in, and the firm will increase its profits by reducing the quantity of output until $MR = MC$.

In this example, MR and MC intersect at a quantity of 40, which is the profit-maximizing level of output for the firm.

Step 2: The monopolistic competitor decides what price to charge. When the firm has determined its profit-maximizing quantity of output, it can then look to its perceived demand curve to find out what it can charge for that quantity of output. On the graph, this process can be shown as a vertical line reaching up through the profit-maximizing quantity until it hits the firm's perceived demand curve. For Authentic Chinese Pizza, it should charge a price of $16 per pizza for a quantity of 40.

Once the firm has chosen price and quantity, it's in a position to calculate total revenue, total cost, and profit. At a quantity of 40, the price of $16 lies above the average cost curve, so the firm is making economic profits. From Table 9.2.1 we can see that, at an output of 40, the firm's total revenue is $640 and its total cost is $580, so profits are $60. In Figure 9.2.2, the firm's total revenues are the rectangle with the quantity of 40 on the horizontal axis and the price of $16 on the vertical axis. The firm's total costs are the light shaded rectangle with the same quantity of 40 on the horizontal axis but the average cost of $14.50 on the vertical axis. Profits are total revenues minus total costs, which is the shaded area above the average cost curve.

Although the process by which a monopolistic competitor makes decisions about quantity and price is similar to the way in which a monopolist makes such decisions, two differences are worth remembering. First, although both a monopolist and a monopolistic competitor face downward-sloping demand curves, the monopolist's perceived demand curve is the market **demand curve**, while the perceived demand curve for a monopolistic competitor is based on the extent of its product differentiation and how many competitors it faces. Second, a monopolist is surrounded by barriers to entry and need not fear entry, but a monopolistic competitor who earns profits must expect the entry of firms with similar, but differentiated, products.

Monopolistic Competitors and Entry

If one monopolistic competitor earns positive economic profits, other firms will be tempted to enter the market. A gas station with a great location must worry that other gas stations might open across the street or down the road—and perhaps the new gas stations will sell coffee or have a carwash or some other attraction to lure customers. A successful restaurant with a unique barbecue sauce must be concerned that other restaurants will try to copy the sauce or offer their own unique recipes. A laundry detergent with a great reputation for quality must be concerned that other competitors may seek to build their own reputations.

The entry of other firms into the same general market (like gas, restaurants, or detergent) shifts the demand curve faced by a monopolistically competitive firm. As more firms enter the market, the quantity demanded at a given price for any particular firm will decline, and the firm's perceived demand curve will shift to the left. As a firm's perceived demand curve shifts to the left, its marginal revenue curve will shift to the left, too. The shift in marginal revenue will change the profit-maximizing quantity that the firm chooses to produce, since marginal revenue will then equal marginal cost at a lower quantity.

Figure 9.2.3 (a) shows a situation in which a monopolistic competitor was earning a profit with its original perceived demand curve (D_0). The intersection of the marginal revenue curve (MR_0) and marginal cost curve (MC) occurs at point S, corresponding to quantity Q_0, which is associated on the demand curve at point T with price P_0. The combination of price P_0 and quantity Q_0 lies above the average cost curve, which shows that the firm is earning positive economic profits.

Monopolistic Competition, Entry, and Exit

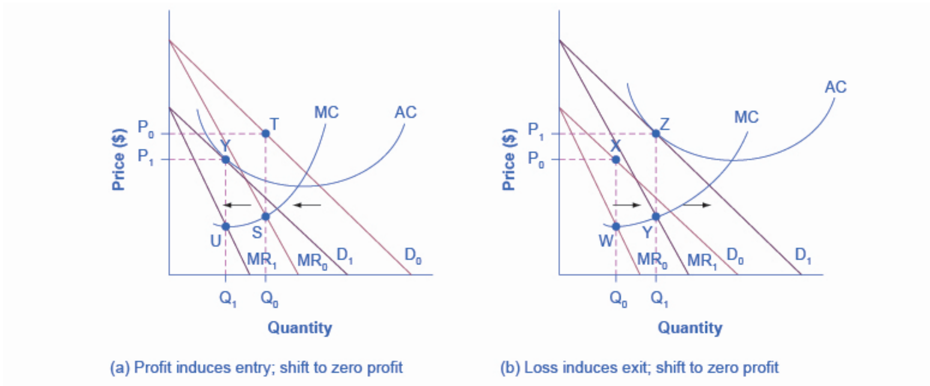


Figure $9.2.3$: *(a) At P_0 and Q_0, the monopolistically competitive firm shown in this figure is making a positive economic profit. This is clear because if you follow the dotted line above Q_0, you can see that price is above average cost. Positive economic profits attract competing firms to the industry, driving the original firm's demand down to D_1. At the new equilibrium quantity (P_1, Q_1), the original firm is earning zero economic profits, and entry into the industry ceases. In (b) the opposite occurs. At P_0 and Q_0, the firm is losing money. If you follow the dotted line above Q_0, you can see that average cost is above price. Losses induce firms to leave the industry. When they do, demand for the original firm rises to D_1, where once again the firm is earning zero economic profit.*

Unlike a monopoly, with its high barriers to entry, a monopolistically competitive firm with positive economic profits will attract competition. When another competitor enters the market, the original firm's perceived demand curve shifts to the left, from D_0 to D_1, and the associated marginal revenue curve shifts from MR_0 to MR_1. The new profit-maximizing output is Q_1, because the intersection of the MR_1 and MC now occurs at point U. Moving vertically up from that quantity on the new demand curve, the optimal price is at P_1.

As long as the **firm** is earning positive economic profits, new competitors will continue to enter the market, reducing the original firm's demand and marginal revenue curves. The long-run **equilibrium** is shown in the figure at point Y, where the firm's perceived demand curve touches the average cost curve. When price is equal to average cost, economic profits are zero. Thus, although a monopolistically competitive firm may earn positive economic profits in the short term, the process of new entry will drive down economic profits to zero in the long run. Remember that zero economic profit is not equivalent to zero accounting profit. A zero economic profit means the firm's **accounting profit** is equal to what its resources could earn in their next best use. Figure $9.2.3$ (b) shows the reverse situation, where a monopolistically competitive firm is originally losing money. The adjustment to long-run equilibrium is analogous to the previous example. The economic losses lead to firms exiting, which will result in increased demand for this particular firm, and consequently lower losses. Firms exit up to the point where there are no more losses in this market, for example when the demand curve touches the average cost curve, as in point Z.

Monopolistic competitors can make an economic profit or loss in the short run, but in the long run, entry and exit will drive these firms toward a zero economic profit outcome. However, the zero economic profit outcome in monopolistic competition looks different from the zero economic profit outcome in perfect competition in several ways relating both to efficiency and to variety in the market.

Monopolistic Competition and Efficiency

The long-term result of entry and exit in a perfectly competitive market is that all firms end up selling at the price level determined by the lowest point on the average cost curve. This outcome is why perfect competition displays **productive efficiency**: goods are being produced at the lowest possible average cost. However, in monopolistic competition, the end result of entry and exit is that firms end up with a price that lies on the downward-sloping portion of the average cost curve, not at the very bottom of the AC curve. Thus, monopolistic competition will not be productively efficient.

In a perfectly competitive market, each firm produces at a quantity where price is set equal to marginal cost, both in the short run and in the long run. This outcome is why perfect competition displays allocative efficiency: the social benefits of additional production, as measured by the marginal benefit, which is the same as the price, equal the marginal costs to society of that production. In a monopolistically competitive market, the rule for maximizing profit is to set $MR = MC$ —and price is higher than marginal revenue, not equal to it because the demand curve is downward sloping. When $P > MC$, which is the outcome in a monopolistically competitive market, the benefits to society of providing additional quantity, as measured by the price that people are willing to pay, exceed the marginal costs to society of producing those units. A monopolistically competitive firm does not produce more, which means that society loses the net benefit of those extra units. This is the same argument we made about monopoly, but in this case to a lesser degree. Thus, a monopolistically competitive industry will produce a lower quantity of a good and charge a higher price for it than would a perfectly competitive industry. See the following Clear It Up feature for more detail on the impact of demand shifts.

Why does a shift in perceived demand cause a shift in marginal revenue?

The combinations of price and quantity at each point on a firm's perceived demand curve are used to calculate total revenue for each combination of price and quantity. This information on total revenue is then used to calculate marginal revenue, which is the change in total revenue divided by the change in quantity. A change in perceived demand will change total revenue at every quantity of output and in turn, the change in total revenue will shift marginal revenue at each quantity of output. Thus, when entry occurs in a monopolistically competitive industry, the perceived demand curve for each firm will shift to the left, because a smaller quantity will be demanded at any given price. Another way of interpreting this shift in demand is to notice that, for each quantity sold, a lower price will be charged. Consequently, the marginal revenue will be lower for each quantity sold—and the marginal revenue curve will shift to the left as well. Conversely, exit causes the perceived demand curve for a monopolistically competitive firm to shift to the right and the corresponding marginal revenue curve to shift right, too.

A monopolistically competitive industry does not display productive and allocative efficiency in either the short run, when firms are making economic profits and losses, nor in the long run, when firms are earning zero profits.

The Benefits of Variety and Product Differentiation

Even though monopolistic competition does not provide productive efficiency or allocative efficiency, it does have benefits of its own. Product differentiation is based on variety and innovation. Many people would prefer to live in an economy with many kinds of clothes, foods, and car styles; not in a world of perfect competition where everyone will always wear blue jeans and white shirts, eat only spaghetti with plain red sauce, and drive an identical model of car. Many people would prefer to live in an economy where firms are struggling to figure out ways of attracting customers by methods like friendlier service, free delivery, guarantees of quality, variations on existing products, and a better shopping experience.

Economists have struggled, with only partial success, to address the question of whether a market-oriented economy produces the optimal amount of variety. Critics of market-oriented economies argue that society does not really need dozens of different athletic shoes or breakfast cereals or automobiles. They argue that much of the cost of creating such a high degree of product differentiation, and then of advertising and marketing this differentiation, is socially wasteful—that is, most people would be just as happy with a smaller range of **differentiated products** produced and sold at a lower price. Defenders of a market-oriented economy respond that if people do not want to buy differentiated products or highly advertised brand names, no one is forcing them to do so. Moreover, they argue that consumers benefit substantially when firms seek short-term profits by providing differentiated products. This controversy may never be fully resolved, in part because deciding on the optimal amount of variety is very difficult, and in part because the two sides often place different values on what variety means for consumers. Read the following Clear It Up feature for a discussion on the role that advertising plays in monopolistic competition.

Example $9.2.2$: How does advertising impact monopolistic competition?

The U.S. economy spent about $\$180.12$ billion on advertising in 2014, according to eMarketer.com. Roughly one third of this was television advertising, and another third was divided roughly equally between Internet, newspapers, and radio. The remaining third was divided up between direct mail, magazines, telephone directory yellow pages, and billboards. Mobile devices are increasing the opportunities for advertisers.

Advertising is all about explaining to people, or making people believe, that the products of one firm are differentiated from the products of another firm. In the framework of monopolistic competition, there are two ways to conceive of how advertising works: either advertising causes a firm's perceived demand curve to become more inelastic (that is, it causes the perceived demand curve to become steeper); or advertising causes demand for the firm's product to increase (that is, it causes the firm's perceived demand curve to shift to the right). In either case, a successful advertising campaign may allow a firm to sell either a greater quantity or to charge a higher price, or both, and thus increase its profits.

However, economists and business owners have also long suspected that much of the advertising may only offset other advertising. Economist A. C. Pigou wrote the following back in 1920 in his book, *The Economics of Welfare*:

"It may happen that expenditures on advertisement made by competing monopolists [that is, what we now call monopolistic competitors] will simply neutralise one another, and leave the industrial position exactly as it would have been if neither had expended anything. For, clearly, if each of two rivals makes equal efforts to attract the favour of the public away from the other, the total result is the same as it would have been if neither had made any effort at all."

Key Concepts and Summary

Monopolistic competition refers to a market where many firms sell differentiated products. Differentiated products can arise from characteristics of the good or service, location from which the product is sold, intangible aspects of the product, and perceptions of the product.

The perceived demand curve for a monopolistically competitive firm is downward-sloping, which shows that it is a price maker and chooses a combination of price and quantity. However, the perceived demand curve for a monopolistic competitor is more elastic than the perceived demand curve for a monopolist, because the monopolistic competitor has direct competition, unlike the pure monopolist. A profit-maximizing monopolistic competitor will seek out the quantity where marginal revenue is equal to marginal cost. The monopolistic competitor will produce that level of output and charge the price that is indicated by the firm's demand curve.

If the firms in a monopolistically competitive industry are earning economic profits, the industry will attract entry until profits are driven down to zero in the long run. If the firms in a monopolistically competitive industry are suffering economic losses, then the industry will experience exit of firms until economic profits are driven up to zero in the long run.

A monopolistically competitive firm is not productively efficient because it does not produce at the minimum of its average cost curve. A monopolistically competitive firm is not allocatively efficient because it does not produce where $P = MC$, but instead produces where $P > MC$. Thus, a monopolistically competitive firm will tend to produce a lower quantity at a higher cost and to charge a higher price than a perfectly competitive firm.

Monopolistically competitive industries do offer benefits to consumers in the form of greater variety and incentives for improved products and services. There is some controversy over whether a market-oriented economy generates too much variety.

References

Kantar Media. "Our Insights: Infographic—U.S. Advertising Year End Trends Report 2012." Accessed October 17, 2013. http://kantarmedia.us/insight-center...ds-report-2012.

Statistica.com. 2015. "Number of Restaurants in the United States from 2011 to 2014." Accessed March 27, 2015. http://www.statista.com/statistics/2...nts-in-the-us/.

Glossary

differentiated product

a product that is perceived by consumers as distinctive in some way

imperfectly competitive

firms and organizations that fall between the extremes of monopoly and perfect competition

monopolistic competition

many firms competing to sell similar but differentiated products

oligopoly

when a few large firms have all or most of the sales in an industry

Contributor

- Template:ContribOpenSTAXMicroEcon

9.3: Oligopoly

Skills to Develop

- Explain why and how oligopolies exist
- Contrast collusion and competition
- Interpret and analyze the prisoner's dilemma diagram
- Evaluate the tradeoffs of imperfect competition

Many purchases that individuals make at the retail level are produced in markets that are neither perfectly competitive, monopolies, nor monopolistically competitive. Rather, they are oligopolies. **Oligopoly** arises when a small number of large firms have all or most of the sales in an industry. Examples of oligopoly abound and include the auto industry, cable television, and commercial air travel. Oligopolistic firms are like cats in a bag. They can either scratch each other to pieces or cuddle up and get comfortable with one another. If oligopolists compete hard, they may end up acting very much like perfect competitors, driving down costs and leading to zero profits for all. If oligopolists collude with each other, they may effectively act like a monopoly and succeed in pushing up prices and earning consistently high levels of profit. Oligopolies are typically characterized by mutual interdependence where various decisions such as output, price, advertising, and so on, depend on the decisions of the other firm(s). Analyzing the choices of oligopolistic firms about pricing and quantity produced involves considering the pros and cons of competition versus collusion at a given point in time.

Why Do Oligopolies Exist?

A combination of the barriers to entry that create monopolies and the product differentiation that characterizes monopolistic competition can create the setting for an oligopoly. For example, when a government grants a patent for an invention to one firm, it may create a monopoly. When the government grants patents to, for example, three different pharmaceutical companies that each has its own drug for reducing high blood pressure, those three firms may become an oligopoly.

Similarly, a **natural monopoly** will arise when the quantity demanded in a market is only large enough for a single firm to operate at the minimum of the long-run average cost curve. In such a setting, the market has room for only one firm, because no smaller firm can operate at a low enough average cost to compete, and no larger firm could sell what it produced given the quantity demanded in the market.

Quantity demanded in the market may also be two or three times the quantity needed to produce at the minimum of the average cost curve—which means that the market would have room for only two or three oligopoly firms (and they need not produce differentiated products). Again, smaller firms would have higher average costs and be unable to compete, while additional large firms would produce such a high quantity that they would not be able to sell it at a profitable price. This combination of economies of scale and market demand creates the barrier to entry, which led to the Boeing-Airbus oligopoly for large passenger aircraft.

The product differentiation at the heart of monopolistic competition can also play a role in creating oligopoly. For example, firms may need to reach a certain minimum size before they are able to spend enough on advertising and marketing to create a recognizable brand name. The problem in competing with, say, Coca-Cola or Pepsi is not that producing fizzy drinks is technologically difficult, but rather that creating a brand name and marketing effort to equal Coke or Pepsi is an enormous task.

Collusion or Competition?

When oligopoly firms in a certain market decide what quantity to produce and what price to charge, they face a temptation to act as if they were a monopoly. By acting together, oligopolistic firms can hold down industry output, charge a higher price, and divide up the profit among themselves. When firms act together in this way to reduce output and keep prices high, it is called **collusion**. A group of firms that have a formal agreement to collude to produce the monopoly output and sell at the monopoly price is called a **cartel**. However, one major issue facing oligopolies is trusting each other, as we will see below.

See the following Clear It Up feature for a more in-depth analysis of the difference between the two.

Collusion versus cartels: How can I tell which is which?

In the United States, as well as many other countries, it is illegal for firms to collude since collusion is anti-competitive behavior, which is a violation of antitrust law. Both the **Antitrust Division of the Justice Department** and the **Federal Trade Commission** have responsibilities for preventing collusion in the United States.

The problem of enforcement is finding hard evidence of collusion. Cartels are formal agreements to collude. Because cartel agreements provide evidence of collusion, they are rare in the United States. Instead, most collusion is tacit, where firms implicitly reach an understanding that competition is bad for profits.

The desire of businesses to avoid competing so that they can instead raise the prices that they charge and earn higher profits has been well understood by economists. **Adam Smith** wrote in *Wealth of Nations* in 1776: "*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.*"

Even when oligopolists recognize that they would benefit as a group by acting like a monopoly, each individual oligopoly faces a private temptation to produce just a slightly higher quantity and earn slightly higher profit—while still counting on the other oligopolists to hold down their production and keep prices high. If at least some oligopolists give in to this temptation and start producing more, then the market price will fall. Indeed, a small handful of oligopoly firms may end up competing so fiercely that they all end up earning zero economic profits—as if they were perfect competitors. Trust is Key!

Game Theory

Game theory is the science of strategic decision-making or **strategic-interaction** in situations that involve more than one actor. These can include actual games, or real-life situations like military battles, business interactions, or managerial decisions. According to game theory, the best strategy for an individual may or may not be the same depending on the stakes of the game and given the likely move of the other player involved.

Sometimes, the best strategy will be the same no matter how other players act. This is known as the dominant strategy. A **dominant strategy** is the best strategy for an individual player regardless of the moves of any other player. An example would be if you are playing Poker. In this case, the only hand that a player could have that would be a dominant strategy would be a Royal Flush. An Ace-high, Royal Flush is the only hand that beats every other hand, no other player can have this hand. Thus, the player will win every time and they should bet!

On the other hand, there exists the so-called **Nash equilibrium**, which does not describe a particular strategy per se, but rather a sort of mutual understanding-- each player understands the other player's optimal strategies and takes those into consideration when optimizing her own strategy.

- According to game theory, the dominant strategy is the optimal move for an individual regardless of how other players act.
- A Nash equilibrium describes the optimal state of the game where both players make optimal moves but now consider the moves of their opponent.
- A well-known example of where the Nash equilibrium plays out in game theory is the prisoner's dilemma.

The Prisoner's Dilemma

Because of the complexity of oligopoly, which is the result of mutual interdependence among firms, there is no single, generally-accepted theory of how oligopolies behave, in the same way that we have theories for all the other market structures. Instead, economists use **game theory**, a branch of mathematics that analyzes situations in which players must make decisions and then receive payoffs based on what other players decide to do. Game theory has found widespread applications in the social sciences, as well as in business, law, and military strategy.

The **prisoner's dilemma** is a scenario in which the gains from cooperation are larger than the rewards from pursuing self-interest. It applies well to oligopoly. The story behind the prisoner's dilemma goes like this:

Two co-conspiratorial criminals are arrested. When they are taken to the police station, they refuse to say anything and are put in separate interrogation rooms. Eventually, a police officer enters the room where Prisoner A is being held and says: "*You know what? Your partner in the other room is confessing. So your partner is going to get a light prison sentence of just one year, and because you're remaining silent, the judge is going to stick you with eight years in prison. Why don't you get smart? If you confess, too, we'll cut your jail time down to five years, and your partner will get five years, also.*" Over in the next room, another police officer is giving exactly the same speech to Prisoner B. What the police officers do not say is that if both prisoners remain silent, the evidence against them is not especially strong, and the prisoners will end up with only two years in jail each.

The game theory situation facing the two prisoners is shown in Table $9.3.1$. To understand the dilemma, first consider the choices from Prisoner A's point of view. If A believes that B will confess, then A ought to confess, too, so as to not get stuck with the eight years in prison. But if A believes that B will not confess, then A will be tempted to act selfishly and confess, so as to serve only one year. The key point is that A has an incentive to confess regardless of what choice B makes! B faces the same set of choices, and thus will have an incentive to confess regardless of what choice A makes. Confess is considered the dominant strategy or the strategy an individual (or firm) will pursue regardless of the other individual's (or firm's) decision. The result is that if prisoners pursue their own self-interest, both are likely to confess, and end up doing a total of 10 years of jail time between them.

***Table 9.3.1**: The Prisoner's Dilemma Problem*

		Prisoner B	
		Remain Silent (cooperate with other prisoner)	Confess (do not cooperate with other prisoner)
Prisoner A	Remain Silent (cooperate with other prisoner)	A gets 2 years, B gets 2 years	A gets 8 years, B gets 1 year
	Confess (do not cooperate with other prisoner)	A gets 1 year, B gets 8 years	A gets 5 years B gets 5 years

The game is called a dilemma because if the two prisoners had cooperated by both remaining silent, they would only have had to serve a total of four years of jail time between them. If the two prisoners can work out some way of cooperating so that neither one will confess, they will both be better off than if they each follow their own individual self-interest, which in this case leads straight into longer jail terms.

The Oligopoly Version of the Prisoner's Dilemma (Duopoly-or Two Oligopolists)

The members of an oligopoly can face a prisoner's dilemma, also. If each of the oligopolists cooperates in holding down output, then high monopoly profits are possible. Each oligopolist, however, must worry that while it is holding down output, other firms are taking advantage of the high price by raising output and earning higher profits. Table 9.3.2 shows the prisoner's dilemma for a two-firm oligopoly—known as a duopoly. If Firms A and B both agree to hold down output, they are acting together as a monopoly and will each earn $\$1,000$ in profits. However, both firms' dominant strategy is to increase output, in which case each will earn $\$400$ in profits.

Table 9.3.2: A Prisoner's Dilemma for Oligopolists

		Firm B	
		Hold Down Output (cooperate with other firm)	Increase Output (do not cooperate with other firm)
Firm A	Hold Down Output (cooperate with other firm)	A gets $1,000, B gets $1,000	A gets $200, B gets $1,500
	Increase Output (do not cooperate with other firm)	A gets $1,500, B gets $200	A gets $400, B gets $400

Can the two firms trust each other? Consider the situation of Firm A:

- If A thinks that B will cheat on their agreement and increase output, then A will increase output, too, because for A the profit of $\$400$ when both firms increase output (the bottom right-hand choice in Table 9.3.2) is better than a profit of only $\$200$ if A keeps output low and B raises output (the upper right-hand choice in the table).
- If A thinks that B will cooperate by holding down output, then A may seize the opportunity to earn higher profits by raising output. After all, if B is going to hold down output, then A can earn $\$1,500$ in profits by expanding output (the bottom left-hand choice in the table) compared with only $\$1,000$ by holding down output as well (the upper left-hand choice in the table).

Thus, firm A will reason that it makes sense to expand output if B holds down output and that it also makes sense to expand output if B raises output. Again, B faces a parallel set of decisions.

The result of this prisoner's dilemma is often that even though A and B could make the highest combined profits by cooperating in producing a lower level of output and acting like a monopolist, the two firms may well end up in a situation where they each increase **output** and earn only $\$400$ each in **profits**. The following Clear It Up feature discusses one cartel scandal in particular.

What is the Lysine cartel?

Lysine, a $600 million-a-year industry, is an amino acid used by farmers as a feed additive to ensure the proper growth of swine and poultry. The primary U.S. producer of lysine is Archer Daniels Midland (ADM), but several other large European and Japanese firms are also in this market. For a time in the first half of the 1990s, the world's major lysine producers met together in hotel conference rooms and decided exactly how much each firm would sell and what it would charge. The U.S. Federal Bureau of Investigation (FBI), however, had learned of the cartel and placed wire taps on a number of their phone calls and meetings.

From FBI surveillance tapes, following is a comment that Terry Wilson, president of the corn processing division at ADM, made to the other lysine producers at a 1994 meeting in Mona, Hawaii:

I wanna go back and I wanna say something very simple. If we're going to trust each other, okay, and if I'm assured that I'm gonna get 67,000 tons by the year's end, we're gonna sell it at the prices we agreed to . . . The only thing we need to talk about there because we are gonna get manipulated by these [expletive] buyers—they can be smarter than us if we let them be smarter. . . . They [the customers] are not your friend. They are not my friend. And we gotta have 'em, but they are not my friends. You are my friend. I wanna be closer to you than I am to any customer. Cause you can make us ... money. ... And all I wanna tell you again is let's—let's put the prices on the board. Let's all agree that's what we're gonna do and then walk out of here and do it.

The price of lysine doubled while the cartel was in effect. Confronted by the FBI tapes, Archer Daniels Midland pled guilty in 1996 and paid a fine of $100 million. A number of top executives, both at ADM and other firms, later paid fines of up to $350,000 and were sentenced to 24–30 months in prison.

In another one of the FBI recordings, the president of Archer Daniels Midland told an executive from another competing firm that ADM had a slogan that, in his words, had "penetrated the whole company." The company president stated the slogan this way: "Our competitors are our friends. Our customers are the enemy." That slogan could stand as the motto of cartels everywhere.

Other Game Theory Concepts

A Repeated Game

These are games that are played over and over again. When this happens, players learn more about the game and the players with each play of the game. Think about when you were young and played tic-tac-toe for the first time. The person who you played with was an experienced player and they beat you. However, you learned quickly about the strategies that you needed to use to keep them from earning three "Xs" or "Os" in a row. This is an example of a repeated game. Now every game ends in "Cats-Eye" or a draw.

Repeated games may induce players to cooperate. If you get to know other players well and know their strategies, this could build trust or at least predictability in their moves.

Maximin Strategy

A strategy that allows players to avoid the largest losses is the **Maximin Strategy** (or the safety first strategy). In other words, this is a strategy that maximizes the minimum payoff for one player. This strategy can be identified by identifying the worst possible outcome for each strategy. Players choose the best "worst case" strategy (where the lowest payoff is the highest).

P1 / P2	Invest	Don't Invest
Invest	\$20, \$20	-\$100, \$0
Don't Invest	\$15, \$15	\$10, \$10

Caption: Firm 1 and Firm 2 can either choose to Not Invest or Invest. If Firm 1 chooses to Invest and Firm 2 chooses to Invest, Firm 1's payoff will be \$20. If Firm 1 chooses to Invest and Firm 2 chooses to Don't Invest, Firm 1's payoff will be -\$100. If Firm

1 chooses to Don't Invest and Firm 2 chooses to Invest, Firm 1's payoff will be $15. If Firm 1 chooses to Don't Invest and Firm 2 chooses to Don't Invest, Firm 1's payoff will be $10. In this case, the best worst case scenario for Firm 1 will be to DON'T INVEST. (Check your Understanding: Firm 2s maximin strategy will be to INVEST.)

How to Enforce Cooperation

How can parties who find themselves in a prisoner's dilemma situation avoid the undesired outcome and cooperate with each other? The way out of a prisoner's dilemma is to find a way to penalize those who do not cooperate.

Perhaps the easiest approach for colluding oligopolists, as you might imagine, would be to sign a contract with each other that they will hold output low and keep prices high. If a group of U.S. companies signed such a contract, however, it would be illegal. Certain international organizations, like the nations that are members of the **Organization of Petroleum Exporting Countries (OPEC)**, have signed international agreements to act like a monopoly, hold down output, and keep prices high so that all of the countries can make high profits from oil exports. Such agreements, however, because they fall in a gray area of international law, are not legally enforceable. If Nigeria, for example, decides to start cutting prices and selling more oil, Saudi Arabia cannot sue Nigeria in court and force it to stop.

Because oligopolists cannot sign a legally enforceable contract to act like a monopoly, the firms may instead keep close tabs on what other firms are producing and charging. Alternatively, oligopolists may choose to act in a way that generates pressure on each firm to stick to its agreed quantity of output.

One example of the pressure these firms can exert on one another is the **kinked demand curve**, in which competing oligopoly firms commit to match price cuts, but not price increases. This situation is shown in Figure $9.3.1$. Say that an oligopoly airline has agreed with the rest of a cartel to provide a quantity of $10,000$ seats on the New York to Los Angeles route, at a price of $\$500$. This choice defines the kink in the firm's perceived demand curve. The reason that the firm faces a kink in its demand curve is because of how the other oligopolists react to changes in the firm's price. If the oligopoly decides to produce more and cut its price, the other members of the cartel will immediately match any price cuts—and therefore, a lower price brings very little increase in quantity sold.

If one **firm** cuts its price to $\$300$, it will be able to sell only $11,000$ seats. However, if the airline seeks to raise prices, the other oligopolists will not raise their prices, and so the firm that raised prices will lose a considerable share of sales. For example, if the firm raises its price to $\$550$, its sales drop to $5,000$ seats sold. Thus, if oligopolists always match price cuts by other firms in the cartel, but do not match price increases, then none of the oligopolists will have a strong incentive to change prices, since the potential gains are minimal. This strategy can work like a silent form of cooperation, in which the cartel successfully manages to hold down output, increase **price**, and share a monopoly level of profits even without any legally enforceable agreement.

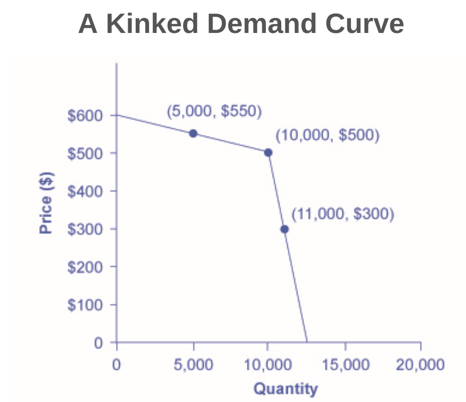


Figure $9.3.1$: *Consider a member firm in an oligopoly cartel that is supposed to produce a quantity of 10,000 and sell at a price of $500. The other members of the cartel can encourage this firm to honor its commitments by acting so that the firm faces a kinked demand curve. If the oligopolist attempts to expand output and reduce price slightly, other firms also cut prices immediately—so if the firm expands output to 11,000, the price per unit falls dramatically, to $300. On the other side, if the oligopoly attempts to raise its price, other firms will not do so, so if the firm raises its price to $550, its sales decline sharply to 5,000. Thus, the members of a* *cartel can discipline each other to stick to the pre-agreed levels of quantity and price through a strategy of matching all price cuts but not matching any price increases.*

Many real-world oligopolies, prodded by economic changes, legal and political pressures, and the egos of their top executives, go through episodes of cooperation and competition. If oligopolies could sustain cooperation with each other on output and pricing, they could earn profits as if they were a single monopoly. However, each firm in an oligopoly has an incentive to produce more and grab a bigger share of the overall market; when firms start behaving in this way, the market outcome in terms of prices and quantity can be similar to that of a highly competitive market.

Tradeoffs of Imperfect Competition

Monopolistic competition is probably the single most common market structure in the U.S. economy. It provides powerful incentives for innovation, as firms seek to earn profits in the short run, while entry assures that firms do not earn economic profits in the long run. However, monopolistically competitive firms do not produce at the lowest point on their average cost curves. In addition, the endless search to impress consumers through product differentiation may lead to excessive social expenses on advertising and marketing.

Oligopoly is probably the second most common market structure. When oligopolies result from patented innovations or from taking advantage of economies of scale to produce at low average cost, they may provide considerable benefit to consumers. Oligopolies are often buffeted by significant barriers to entry, which enable the oligopolists to earn sustained profits over long periods of time. Oligopolists also do not typically produce at the minimum of their average cost curves. When they lack vibrant competition, they may lack incentives to provide innovative products and high-quality service.

The task of public policy with regard to competition is to sort through these multiple realities, attempting to encourage behavior that is beneficial to the broader society and to discourage behavior that only adds to the profits of a few large companies, with no corresponding benefit to consumers. Monopoly and Antitrust Policy discusses the delicate judgments that go into this task.

Example $9.3.1$: The Temptation to Defy the Law

Oligopolistic firms have been called “cats in a bag,” as this chapter mentioned. The French detergent makers chose to “cozy up” with each other. The result? An uneasy and tenuous relationship. When the Wall Street Journal reported on the matter, it wrote: “According to a statement a Henkel manager made to the [French anti-trust] commission, the detergent makers wanted ‘to limit the intensity of the competition between them and clean up the market.’ Nevertheless, by the early 1990s, a price war had broken out among them.” During the soap executives’ meetings, which sometimes lasted more than four hours, complex pricing structures were established. “One [soap] executive recalled ‘chaotic’ meetings as each side tried to work out how the other had bent the rules.” Like many cartels, the soap cartel disintegrated due to the very strong temptation for each member to maximize its own individual profits.

How did this soap opera end? After an investigation, French antitrust authorities fined Colgate-Palmolive, Henkel, and Proctor & Gamble a total of €361 million($\$484$ million). A similar fate befell the icemakers. Bagged ice is a commodity, a perfect substitute, generally sold in 7- or 22-pound bags. No one cares what label is on the bag. By agreeing to carve up the ice market, control broad geographic swaths of territory, and set prices, the icemakers moved from perfect competition to a monopoly model. After the agreements, each firm was the sole supplier of bagged ice to a region; there were profits in both the long run and the short run. According to the courts: “These companies illegally conspired to manipulate the marketplace.” Fines totaled about $\$600,000$—a steep fine considering a bag of ice sells for under $\$3$ in most parts of the United States.

Even though it is illegal in many parts of the world for firms to set prices and carve up a market, the temptation to earn higher profits makes it extremely tempting to defy the law.

Key Concepts and Summary

An oligopoly is a situation where a few firms sell most or all of the goods in a market. Oligopolists earn their highest profits if they can band together as a cartel and act like a monopolist by reducing output and raising price. Since each member of the oligopoly can benefit individually from expanding output, such collusion often breaks down—especially since explicit collusion is illegal.

The prisoner’s dilemma is an example of game theory. It shows how, in certain situations, all sides can benefit from cooperative behavior rather than self-interested behavior. However, the challenge for the parties is to find ways to encourage cooperative behavior.

References

The United States Department of Justice. “Antitrust Division.” Accessed October 17, 2013. http://www.justice.gov/atr/.

eMarketer.com. 2014. “Total US Ad Spending to See Largest Increase Since 2004: Mobile advertising leads growth; will surpass radio, magazines and newspapers this year. Accessed March 12, 2015. http://www.emarketer.com/Article/Tot...e-2004/1010982.

Federal Trade Commission. “About the Federal Trade Commission.” Accessed October 17, 2013. http://www.ftc.gov/ftc/about.shtm.

Glossary

cartel

a group of firms that collude to produce the monopoly output and sell at the monopoly price

collusion

when firms act together to reduce output and keep prices high

duopoly

an oligopoly with only two firms

game theory

a branch of mathematics often used by economists that analyzes situations in which players must make decisions and then receive payoffs based on what decisions the other players make

kinked demand curve

a perceived demand curve that arises when competing oligopoly firms commit to match price cuts, but not price increases

prisoner’s dilemma

a game in which the gains from cooperation are larger than the rewards from pursuing self-interest

Contributor

- Template:ContribOpenSTAXMicroEcon

9.E: 10.E-Monopolistic Competition and Oligopoly (Exercises)

10.1: Monopolistic Competition

Self-Check Questions

Q1

Suppose that, due to a successful advertising campaign, a monopolistic competitor experiences an increase in demand for its product. How will that affect the price it charges and the quantity it supplies?

Q2

Continuing with the scenario outlined in question 1, in the long run, the positive economic profits earned by the monopolistic competitor will attract a response either from existing firms in the industry or firms outside. As those firms capture the original firm's profit, what will happen to the original firm's profit-maximizing price and output levels?

Review Questions

Q3

What is the relationship between product differentiation and monopolistic competition?

Q4

How is the perceived demand curve for a monopolistically competitive firm different from the perceived demand curve for a monopoly or a perfectly competitive firm?

Q5

How does a monopolistic competitor choose its profit-maximizing quantity of output and price?

Q6

How can a monopolistic competitor tell whether the price it is charging will cause the firm to earn profits or experience losses?

Q7

If the firms in a monopolistically competitive market are earning economic profits or losses in the short run, would you expect them to continue doing so in the long run? Why?

Q8

Is a monopolistically competitive firm productively efficient? Is it allocatively efficient? Why or why not?

Critical Thinking Questions

Q9

Aside from advertising, how can monopolistically competitive firms increase demand for their products?

Q10

Make a case for why monopolistically competitive industries never reach long-run equilibrium.

Q11

Would you rather have efficiency or variety? That is, one opportunity cost of the variety of products we have is that each product costs more per unit than if there were only one kind of product of a given type, like shoes. Perhaps a better question is, "*What is the right amount of variety? Can there be too many varieties of shoes, for example?*"

Problems

Q12

Andrea's Day Spa began to offer a relaxing aromatherapy treatment. The firm asks you how much to charge to maximize profits. The demand curve for the treatments is given by the first two columns in Table below; its total costs are given in the third column. For each level of output, calculate total revenue, marginal revenue, average cost, and marginal cost. What is the profit-maximizing level of output for the treatments and how much will the firm earn in profits?

Price	Quantity	TC
$25.00	0	$130
$24.00	10	$275
$23.00	20	$435
$22.50	30	$610
$22.00	40	$800
$21.60	50	$1,005
$21.20	60	$1,225

Solution

S1

An increase in demand will manifest itself as a rightward shift in the demand curve, and a rightward shift in marginal revenue. The shift in marginal revenue will cause a movement up the marginal cost curve to the new intersection between MR and MC at a higher level of output. The new price can be read by drawing a line up from the new output level to the new demand curve, and then over to the vertical axis. The new price should be higher. The increase in quantity will cause a movement along the average cost curve to a possibly higher level of average cost. The price, though, will increase more, causing an increase in total profits.

S2

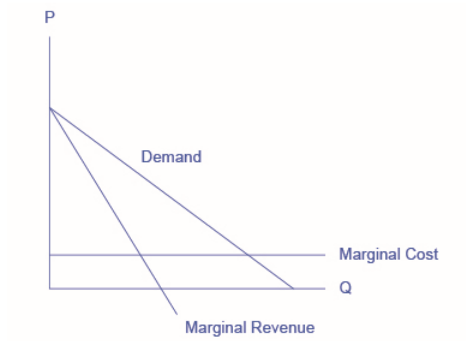
As long as the original firm is earning positive economic profits, other firms will respond in ways that take away the original firm's profits. This will manifest itself as a decrease in demand for the original firm's product, a decrease in the firm's profit-maximizing price and a decrease in the firm's profit-maximizing level of output, essentially unwinding the process described in the answer to question 1. In the long-run equilibrium, all firms in monopolistically competitive markets will earn zero economic profits.

10.2: Oligopoly

Self-Check Questions

Q1

Consider the curve shown in the figure below, which shows the market demand, marginal cost, and marginal revenue curve for firms in an oligopolistic industry. In this example, we assume firms have zero fixed costs.



a. Suppose the firms collude to form a cartel. What price will the cartel charge? What quantity will the cartel supply? How much profit will the cartel earn?
b. Suppose now that the cartel breaks up and the oligopolistic firms compete as vigorously as possible by cutting the price and increasing sales. What will the industry quantity and price be? What will the collective profits be of all firms in the industry?
c. Compare the equilibrium price, quantity, and profit for the cartel and cutthroat competition outcomes.

Q2

Sometimes oligopolies in the same industry are very different in size. Suppose we have a duopoly where one firm (Firm A) is large and the other firm (Firm B) is small, as shown in the prisoner's dilemma box in Table below.

	Firm B colludes with Firm A	Firm B cheats by selling more output
Firm A colludes with Firm B	A gets $1,000, B gets $100	A gets $800, B gets $200
Firm A cheats by selling more output	A gets $1,050, B gets $50	A gets $500, B gets $20

Assuming that the payoffs are known to both firms, what is the likely outcome in this case?

Review Questions

Q3

Will the firms in an oligopoly act more like a monopoly or more like competitors? Briefly explain.

Q4

Does each individual in a prisoner's dilemma benefit more from cooperation or from pursuing self-interest? Explain briefly.

Q5

What stops oligopolists from acting together as a monopolist and earning the highest possible level of profits?

Critical Thinking Questions

Q6

Would you expect the kinked demand curve to be more extreme (like a right angle) or less extreme (like a normal demand curve) if each firm in the cartel produces a near-identical product like OPEC and petroleum? What if each firm produces a somewhat different product? Explain your reasoning.

Q7

When OPEC raised the price of oil dramatically in the mid-1970s, experts said it was unlikely that the cartel could stay together over the long term—that the incentives for individual members to cheat would become too strong. More than forty years later, OPEC still exists. Why do you think OPEC has been able to beat the odds and continue to collude? Hint: You may wish to consider non-economic reasons.

Problems

Q8

Mary and Raj are the only two growers who provide organically grown corn to a local grocery store. They know that if they cooperated and produced less corn, they could raise the price of the corn. If they work independently, they will each earn $\$100$. If they decide to work together and both lower their output, they can each earn $\$150$. If one person lowers output and the other does not, the person who lowers output will earn $\$0$ and the other person will capture the entire market and will earn $\$200$. Table below represents the choices available to Mary and Raj. What is the best choice for Raj if he is sure that Mary will cooperate? If Mary thinks Raj will cheat, what should Mary do and why? What is the prisoner's dilemma result? What is the preferred choice if they could ensure cooperation? A = Work independently; B = Cooperate and Lower Output. (Each results entry lists Raj's earnings first, and Mary's earnings second.)

		Mary	
		A	B
Raj	A	($100, $100)	($200, $0)
	B	($0, $200)	($150, $150)

Q9

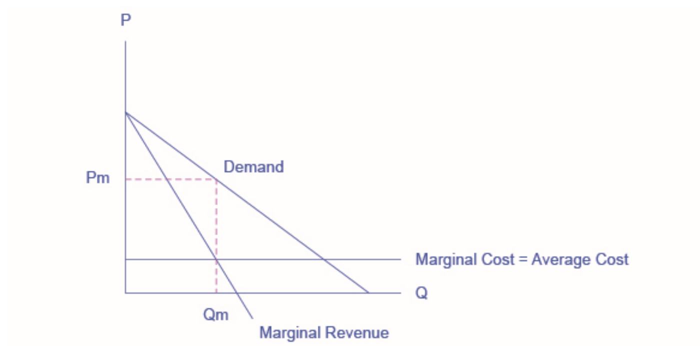
Jane and Bill are apprehended for a bank robbery. They are taken into separate rooms and questioned by the police about their involvement in the crime. The police tell them each that if they confess and turn the other person in, they will receive a lighter sentence. If they both confess, they will be each be sentenced to 30 years. If neither confesses, they will each receive a 20-year sentence. If only one confesses, the confessor will receive 15 years and the one who stayed silent will receive 35 years. Table below represents the choices available to Jane and Bill. If Jane trusts Bill to stay silent, what should she do? If Jane thinks that Bill will confess, what should she do? Does Jane have a dominant strategy? Does Bill have a dominant strategy? A = Confess; B = Stay Silent. (Each results entry lists Jane's sentence first (in years), and Bill's sentence second.)

		Jane	
		A	B
Bill	A	(30, 30)	(15, 35)
	B	(35, 15)	(20, 20)

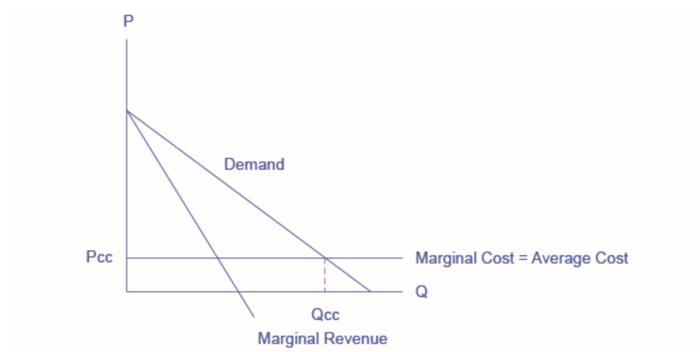
Solution

S1

a. If the firms form a cartel, they will act like a monopoly, choosing the quantity of output where $MR = MC$. Drawing a line from the monopoly quantity up to the demand curve shows the monopoly price. Assuming that fixed costs are zero, and with an understanding of cost and profit, we can infer that when the marginal cost curve is horizontal, average cost is the same as marginal cost. Thus, the cartel will earn positive economic profits equal to the area of the rectangle, with a base equal to the monopoly quantity and a height equal to the difference between price (on the demand above the monopoly quantity) and average cost, as shown in the following figure.



b. The firms will expand output and cut price as long as there are profits remaining. The long-run equilibrium will occur at the point where average cost equals demand. As a result, the oligopoly will earn zero economic profits due to "cutthroat competition," as shown in the next figure.



c. $Pc > Pcc$. $Qc < Qcc$. Profit for the cartel is positive and large. Profit for cutthroat competition is zero.

S2

Firm B reasons that if it cheats and Firm A does not notice, it will double its money. Since Firm A's profits will decline substantially, however, it is likely that Firm A will notice and if so, Firm A will cheat also, with the result that Firm B will lose 90% of what it gained by cheating. Firm A will reason that Firm B is unlikely to risk cheating. If neither firm cheats, Firm A earns $\$1000$. If Firm A cheats, assuming Firm B does not cheat, A can boost its profits only a little, since Firm B is so small. If both firms cheat, then Firm A loses at least 50% of what it could have earned. The possibility of a small gain ($\$50$) is probably not enough to induce Firm A to cheat, so in this case it is likely that both firms will collude.

Contributor

- Template:ContribOpenSTAXMicroEcon

CHAPTER OVERVIEW

10: Market Failure- Externalities

10.1: Introducing Market Failure

Market Failure

learning objectives

- Identify common market failures and governmental responses

Market failure occurs when the price mechanism fails to account for all of the costs and benefits necessary to provide and consume a good. The market will fail by not supplying the socially optimal amount of the good.

Prior to market failure, the supply and demand within the market do not produce quantities of the goods where the price reflects the marginal benefit of consumption. The imbalance causes allocative inefficiency, which is the over- or under-consumption of the good.

The structure of market systems contributes to market failure. In the real world, it is not possible for markets to be perfect due to inefficient producers, externalities, environmental concerns, and lack of public goods. An externality is an effect on a third party which is caused by the production or consumption of a good or service.

***Air pollution**: Air pollution is an example of a negative externality. Governments may enact tradable permits to try and reduce industrial pollution.*

During market failures the government usually responds to varying degrees. Possible government responses include:

- legislation – enacting specific laws. For example, banning smoking in restaurants, or making high school attendance mandatory.
- direct provision of merit and public goods – governments control the supply of goods that have positive externalities. For example, by supplying high amounts of education, parks, or libraries.
- taxation – placing taxes on certain goods to discourage use and internalize external costs. For example, placing a 'sin-tax' on tobacco products, and subsequently increasing the cost of tobacco consumption.
- subsidies – reducing the price of a good based on the public benefit that is gained. For example, lowering college tuition because society benefits from more educated workers. Subsidies are most appropriate to encourage behavior that has positive externalities.
- tradable permits – permits that allow firms to produce a certain amount of something, commonly pollution. Firms can trade permits with other firms to increase or decrease what they can produce. This is the basis behind cap-and-trade, an attempt to reduce of pollution.
- extension of property rights – creates privatization for certain non-private goods like lakes, rivers, and beaches to create a market for pollution. Then, individuals get fined for polluting certain areas.
- advertising – encourages or discourages consumption.
- international cooperation among governments – governments work together on issues that affect the future of the environment.

Causes of Market Failure

Market failure occurs due to inefficiency in the allocation of goods and services.

learning objectives

- Explain some common causes of market failure

Market failure occurs due to inefficiency in the allocation of goods and services. A price mechanism fails to account for all of the costs and benefits involved when providing or consuming a specific good. When this happens, the market will not produce the supply of the good that is socially optimal – it will be over or under produced. To fully understand market failure, it is important to recognize the reasons why a market can fail. Due to the structure of markets, it is impossible for them to be perfect. As a result, most markets are not successful and require forms of intervention.

Reasons for market failure include:

- **Positive and negative externalities**: an externality is an effect on a third party that is caused by the consumption or production of a good or service. A positive externality is a positive spillover that results from the consumption or production of a good or service. For example, although public education may only directly affect students and schools, an educated population may provide positive effects on society as a whole. A negative externality is a negative spillover effect on third parties. For example, secondhand smoke may negatively impact the health of people, even if they do not directly engage in smoking.
- **Environmental concerns:** effects on the environment as important considerations as well as sustainable development.
- **Lack of public goods:** public goods are goods where the total cost of production does not increase with the number of consumers. As an example of a public good, a lighthouse has a fixed cost of production that is the same, whether one ship or one hundred ships use its light. Public goods can be underproduced; there is little incentive, from a private standpoint, to provide a lighthouse because one can wait for someone else to provide it, and then use its light without incurring a cost. This problem – someone benefiting from resources or goods and services without paying for the cost of the benefit – is known as the free rider problem.
- **Underproduction of merit goods:** a merit good is a private good that society believes is under consumed, often with positive externalities. For example, education, healthcare, and sports centers are considered merit goods.
- **Overprovision of demerit goods:** a demerit good is a private good that society believes is over consumed, often with negative externalities. For example, cigarettes, alcohol, and prostitution are considered demerit goods.
- **Abuse of monopoly power:** imperfect markets restrict output in an attempt to maximize profit.

When a market fails, the government usually intervenes depending on the reason for the failure.

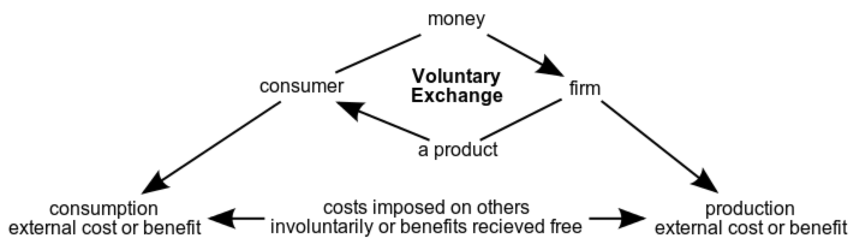
Introducing Externalities

An externality is a cost or benefit that affects an otherwise uninvolved party who did not choose to be subject to the cost or benefit.

learning objectives

- Give examples of externalities that exist in different parts of society

In economics, an externality is a cost or benefit resulting from an activity or transaction, that affects an otherwise uninvolved party who did not choose to be subject to the cost or benefit. An example of an externality is pollution. Health and clean-up costs from pollution impact all of society, not just individuals within the manufacturing industries. In regards to externalities, the cost and benefit to society is the sum of the value of the benefits and costs for all parties involved.



Externality: An externality is a cost or benefit that results from an activity or transaction and that affects an otherwise uninvolved party who did not choose to incur that cost or benefit.

Negative vs. Positive

A negative externality is an result of a product that inflicts a negative effect on a third party. In contrast, positive externality is an action of a product that provides a positive effect on a third party.

Negative Externality: Air pollution caused by motor vehicles is an example of a negative externality.

Externalities originate within voluntary exchanges. Although the parties directly involved benefit from the exchange, third parties can experience additional effects. For those involuntarily impacted, the effects can be negative (pollution from a factory) or positive (domestic bees kept for honey production, pollinate the neighboring crops).

Economic Strain

Neoclassical welfare economics explains that under plausible conditions, externalities cause economic results that are not ideal for society. The third parties who experience external costs from a negative externality do so without consent, while the individuals who receive external benefits do not pay a cost. The existence of externalities can cause ethical and political problems within society.

In regards to externalities, one way to correct the issue is to internalize the third party costs and benefits. However, in many cases, internalizing the costs is not financially possible. Governments may step in to correct such market failures.

Externality Impacts on Efficiency

Economic efficiency is the use resources to maximize the production of goods; externalities are imperfections that limit efficiency.

> **learning objectives**
>
> - Analyze the effects of externalities on efficiency

Economic Efficiency

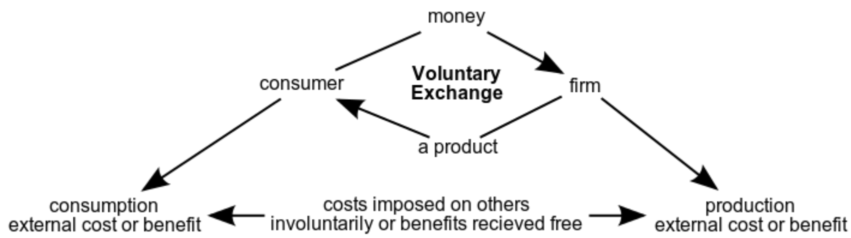
In economics, the term "economic efficiency" is defined as the use of resources in order to maximize the production of goods and services. An economically efficient society can produce more goods or services than another society without using more resources.

A market is said to be economically efficient if:

- No one can be made better off without making someone else worse off.
- No additional output can be obtained without increasing the amounts of inputs.
- Production proceeds at the lowest possible cost per unit.

Externalities

An externality is a cost or benefit that results from an activity or transaction and affects a third party who did not choose to incur the cost or benefit. Externalities are either positive or negative depending on the nature of the impact on the third party. An example of a negative externality is pollution. Manufacturing plants emit pollution which impacts individuals living in the surrounding areas. Third parties who are not involved in any aspect of the manufacturing plant are impacted negatively by the pollution. An example of a positive externality would be an individual who lives by a bee farm. The third parties' flowers are pollinated by the neighbor's bees. They have no cost or investment in the business, but they benefit from the bees.



Externality: This diagram shows the voluntary exchange that takes place within a market system. It also shows the economic costs that are associated with externalities.

Externalities and Efficiency

Positive and negative externalities both impact economic efficiency. Neoclassical welfare economics states that the existence of externalities results in outcomes that are not ideal for society as a whole. In the case of negative externalities, third parties experience negative effects from an activity or transaction in which they did not choose to be involved. In order to compensate for negative externalities, the market as a whole is reducing its profits in order to repair the damage that was caused which decreases efficiency. Positive externalities are beneficial to the third party at no cost to them. The collective social welfare is improved, but the providers of the benefit do not make any money from the shared benefit. As a result, less of the good is produced or profited from which is less optimal society and decreases economic efficiency.

In order to deal with externalities, markets usually internalize the costs or benefits. For costs, the market has to spend additional funds in order to make up for damages incurred. Benefits are also internalized because they are viewed as goods produced and used by third parties with no monetary gain for the market. Internalizing costs and benefits is not always feasible, especially when the monetary value or a good or service cannot be determined.

Externalities directly impact efficiency because the production of goods is not efficient when costs are incurred due to damages. Efficiency also decreases when potential money earned is lost on non-paying third parties.

In order to maximize economic efficiency, regulations are needed to reduce market failures and imperfections, like internalizing externalities. When market imperfections exist, the efficiency of the market declines.

Key Points

- Prior to market failure, the supply and demand within the market do not produce quantities of the goods where the price reflects the marginal benefit of consumption.
- The structure of market systems contributes to market failure. In the real world, it is not possible for markets to be perfect due to inefficient producers, externalities, environmental concerns, and lack of public goods.
- Government responses to market failure include legislation, direct provision of merit goods and public goods, taxation, subsidies, tradable permits, extension of property rights, advertising, and international cooperation among governments.
- A price mechanism fails to account for all of the costs and benefits involved when providing or consuming a specific good. When this happens, the market will not produce the supply of the good that is socially optimal – it will be over or under produced.
- Due to the structure of markets, it may be impossible for them to be perfect.
- Reasons for market failure include: positive and negative externalities, environmental concerns, lack of public goods, underprovision of merit goods, overprovision of demerit goods, and abuse of monopoly power.
- In regards to externalities, the cost and benefit to society is the sum of the benefits and costs for all parties involved.
- Market failure occurs when the price mechanism fails to consider all of the costs and benefits necessary for providing and consuming a good.

- In regards to externalities, one way to correct the issue is to internalize the third party costs and benefits. However, in many cases, internalizing the costs is not feasible. When externalities exist, it is possible that the particular industry will experience market failure.
- In many cases, the government intervenes when there is market failure.
- An economically efficient society can produce more goods or services than another society without using more resources.
- An externality is a cost or benefit that results from an activity or transaction and affects a third party who did not choose to incur the cost or benefit. Externalities are either positive or negative depending on the nature of the impact on the third party.
- Neoclassical welfare economics states that the existence of externalities results in outcomes that are not ideal for society as a whole.
- In order to maximize economic efficiency, regulations are needed to reduce market failures and imperfections, like internalizing externalities. When market imperfections exist, the efficiency of the market declines.
- In order for economic efficiency to be achieved, one defining rule is that no one can be made better off without making someone else worse off. When externalities are present, not everyone benefits from the production of the good or service.

Key Terms

- **public good**: A good that is both non-excludable and non-rivalrous in that individuals cannot be effectively excluded from use and where use by one individual does not reduce availability to others.
- **merit good**: A commodity which is judged that an individual or society should have on the basis of some concept of need, rather than ability and willingness to pay.
- **externality**: An impact, positive or negative, on any party not involved in a given economic transaction or act.
- **public good**: A good that is both non-excludable and non-rivalrous in that individuals cannot be effectively excluded from use and where use by one individual does not reduce availability to others.
- **free rider**: One who obtains benefit from a public good without paying for it directly.
- **monopoly**: A market where one company is the sole supplier.
- **intervene**: To interpose; as, to intervene to settle a quarrel; get involved, so as to alter or hinder an action.
- **externality**: An impact, positive or negative, on any party not involved in a given economic transaction or act.
- **efficient**: Making good, thorough, or careful use of resources; not consuming extra. Especially, making good use of time or energy.

- Free rider problem. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Free_rider_problem**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Boundless. **Provided by**: Boundless Learning. **Located at**: **www.boundless.com//economics/...ition/monopoly**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- public good. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/public%20good**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Pollution de l'air. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/Fi...n_de_l'air.jpg**. **License**: ***Public Domain: No Known Copyright***
- Externalities. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Externalities**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- IB Economics/Microeconomics/Market Failure. **Provided by**: Wikibooks. **Located at**: **en.wikibooks.org/wiki/IB_Econ...Market_Failure**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- intervene. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/intervene**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- externality. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/externality**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Pollution de l'air. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/Fi...n_de_l'air.jpg**. **License**: ***Public Domain: No Known Copyright***
- Externality. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Externality.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Diesel-smoke. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Diesel-smoke.jpg**. **License**: ***Public Domain: No Known Copyright***
- externality. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/externality**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Externalities. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Externalities**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Economic efficiency. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Economic_efficiency**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- efficient. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/efficient**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Pollution de l'air. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/Fi...n_de_l'air.jpg**. **License**: ***Public Domain: No Known Copyright***
- Externality. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Externality.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Diesel-smoke. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Diesel-smoke.jpg**. **License**: ***Public Domain: No Known Copyright***
- Externality. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Externality.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***

10.2: Externalities in Depth

Negative Externalities

Negative externalities are costs caused by an activity that affect an otherwise uninvolved party who did not choose to incur that cost.

Learning Objectives

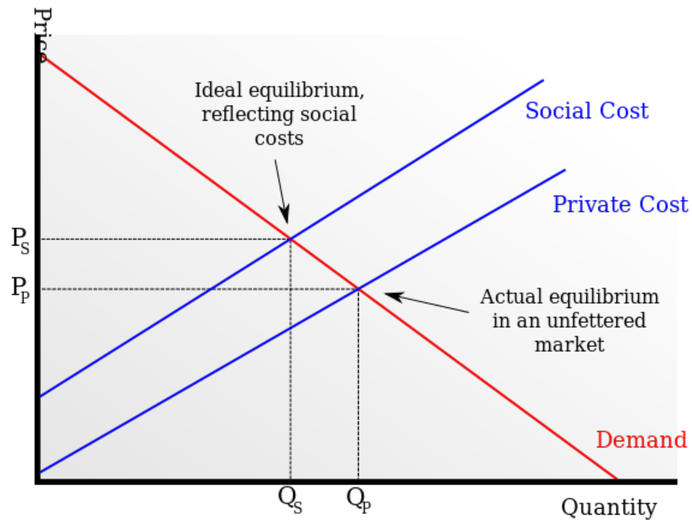
- Describe the impact of a negative externality on society

A negative externality is a cost that results from an activity or transaction and that affects an otherwise uninvolved party who did not choose to incur that cost.

Reasons for Negative Externalities

The reason these negative externalities, otherwise known as social costs, occur is that these expenses are generally not included in calculating the costs of production. Production decisions are generally based on financial data and most social costs are not measured that way. For example, when a firm decides to open up a new factory, it will not account for the cost that residents accrue by drinking water from a river the factory polluted. As a result, a product that shouldn't be produced, because the total expenses exceed the return, are made because social costs were not considered.

In other words, the costs of production represent individual, or private, marginal costs. The private marginal costs are lower than societal marginal costs, which also capture the true costs of the negative externalities. As a result, producers will overestimate the ideal quantity of the good to produce.



Negative Externality: Graphically, negative externalities occur when social costs are lower than private costs, and firms produce more units than is socially optimal. The ideal equilibrium quantity that reflects negative externalities is Qs, but firms may produce at Qp.

Government Solutions for Negative Externalities

In these cases, government intervention is necessary to help "price" negative externalities. Governments can either use regulation (e.g. outlaw an action) or use market solutions. By instituting policies such as pollution penalties, permitting civil lawsuits by private parties to recover damages for negligent actions, and levying environmental taxes, governments can achieve two things. First, these regulations recover funds to help fix the damage caused by negative externalities. Second, these acts help put a financial price on social costs. With that information, businesses can arrive at a more accurate figure for the costs of production. Businesses can then avoid producing products whose financial and social costs exceed the financial return.

Cigarette smoke: Secondhand smoke is an example of a negative externality; a person chooses to smoke, but others who do not choose to smoke are harmed.

Positive Externalities

Positive externalities are benefits caused by activities that affect an otherwise uninvolved party who did not choose to incur that benefit.

Learning Objectives

- Use an example to discuss the concept of a positive externality

Positive externalities are benefits caused by transactions that affect an otherwise uninvolved party who did not choose to incur that benefit. Externalities occur all the time because economic events do not occur within a vacuum. Transactions often require the use of common resources that are shared with parties are not involved with the exchange. The use of these resources, in turn, impacts the uninvolved parties.

In the case of positive externalities, a transaction has *positive* side effects for non-related parties. Let's take a look at some example:

- A homeowner keeps his house maintained, the neighborhood benefits through higher home values. The homeowner's neighbors benefit from a positive externality.
- A person may keep bees for her own enjoyment, but gardeners in the area benefit because their flowers are pollinated. The beekeeper's transaction of purchasing bees ends up positively affecting parties who are not involved in the transaction.
- A person becomes inoculated against a disease, those around him benefit because they cannot catch the disease from him. There was an exchange between the doctor and the patient, but others also benefit.

In each of these cases, the people taking action are presumably not doing it for the sake of the community, but for their own purposes. The people taking the action may also enjoy the additional benefits described above, but initiators of actions are not considered beneficiaries of externalities.

The problem with positive externalities is that the people who create these advantages cannot charge the beneficiaries; the beneficiaries can "free ride," or benefit without paying. For example, assume everyone in a community, except one person, got a flu shot. That one person could choose to abstain from receiving the shot; since everyone else got inoculated, he can't get the disease from the others because they can't catch the flu. That person would be a free rider since he would benefit from inoculations without incurring any cost.

Since parties that create the externality aren't compensated, they do not have any incentive to create more. This results in a suboptimal result, because the producers of the externality will generally create less of the benefit than the larger community needs.

Key Points

- The reason these negative externalities, otherwise known as social costs, occur is that these expenses are generally not included in calculating the costs of production.
- Government intervention is necessary to help " price " negative externalities. They do this through regulations or by instituting market-based policies such as taxes, subsidies, or permit systems.
- Graphically, social costs will be lower than private costs because they do not take into account the additional costs of negative externalities. As a result, firms may produce more units than is optimal from a societal standpoint.
- Graphically, social costs will be lower than private costs because they do not take into account the additional costs of negative externalities. As a result, firms may produce more units than is optimal from a societal standpoint.
- Externalities occur all the time because economic events do not occur within a vacuum. Transactions often require the use of common resources that are shared with parties are not involved with the exchange. The use of these resources in turn impacts the uninvolved parties.
- The problem with positive externalities is that the people who create the externality cannot charge the beneficiaries; the beneficiaries can "free ride," or benefit without paying.
- Free riding results in a suboptimal result, because the producers of the externality will generally create less of the benefit than the larger community needs.

Key Terms

- **externality**: An impact, positive or negative, on any party not involved in a given economic transaction or act.
- **free rider**: One who obtains benefit from a public good without paying for it directly.

10.3: Government Policy Options

Regulation

The government can respond to externalities through command-and-control policies or market-based policies.

> **Learning Objectives**
>
> - Describe the role of government regulation in addressing externalities

The government can respond to externalities in two ways. The government can use command-and-control policies to regulate behavior directly. Alternatively, it can implement market-based policies such as taxes and subsidies to incentivize private decision makers to change their own behavior.

Command-and-control regulation can come in the form of government-imposed standards, targets, process requirements, or outright bans. Such measures make certain behaviors either required or forbidden with the goal of addressing the externality. For example, the government may make it illegal for a company to dump certain chemicals in a river. By doing so, the government hopes to protect the environment or other companies or individuals that use the river that would otherwise suffer a negative impact.

No Smoking: The prohibition of smoking in certain areas is a regulation designed to reduce the negative externalities suffered by non-smokers when they are around smokers.

In practice, implementing regulation effectively is difficult. It requires the regulator to have in-depth knowledge of a certain industry or sphere of economic activity. If done incorrectly, regulation can introduce inefficiency. For example, if the government makes it illegal to dump in the river, the companies and their customers may suffer because the products must be produced using less efficient methods. On the other hand, if the government allows too much to be dumped in the river, they have failed to mitigate the negative externality.

If the government is unsure of how to effectively regulate the market, it should seek other methods of mitigating the externality. Advocates of market-based policies for reducing negative externalities point to the difficulty of creating and enforcing effective regulation for reasons why the government should create systems of incentives and disincentives instead of using the force of regulation.

Tax

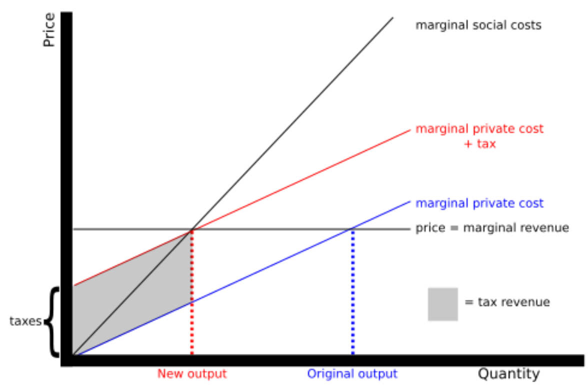
Corrective taxes incentivize economic actors to reduce the production of goods or services generating negative externalities.

Learning Objectives

- Describe the role of taxes in addressing externalities

Taxes are a market-based policy option available to the government to address externalities. A corrective tax (also called a Pigovian tax) is applied to a market activity that is generating negative externalities (costs for a third party). The tax is set equal to the value of the negative externality and provides incentives for allocation of resources closer to the social optimum.

In the case of negative externalities, the social cost of an activity is greater than the private cost of the activity. In such a case, the market outcome is not efficient and may lead to overproduction of the good. Taxes make it more expensive for firms to produce the good or service generating the externality, thus providing an incentive to produce less of it. As the figure demonstrates, a tax shifts the marginal private cost curve up. In response, producers change the output to the socially-optimum level.



Corrective Tax: A tax shifts the marginal private cost curve up by the amount of the tax. This gives producers an incentive to reduce output to the socially optimum level.

Take environmental pollution as an example. The private cost of pollution to a polluter is less than its social cost. If the government levies a tax on pollution, it increases the polluter’s private cost. The polluter now has an incentive to generate less pollution.

The level of the corrective tax is intended to counterbalance the externality. In practice, however, it is extremely difficult for the government to determine the appropriate level for the tax. Moreover, in determining the tax level, the government might come under pressure from various interest groups that would benefit from a higher or lower taxation level. Nevertheless, by introducing corrective taxes in response to negative externalities the government can not only increase efficiency, but raise revenues as well.

Quotas

Tradable permits are a market-based approach allowing the government to limit negative externalities produced by a group of firms.

Learning Objectives

- Evaluate a permit system as a method to address externalities

To address the problem of negative externalities, governments may use a quota system to try and limit them. In a quota system, the negative externality is capped at a certain amount. In the example of pollution, the government may put a quota on the amount of pollution a factory can produce by issuing tradable permits.

Tradable permits are one of the market-based approaches the government can use to address externalities. In the past tradable permits have been primarily used to control pollution.

Emissions Trading: Emissions trading or "cap and trade" is a market-based approach used to control pollution by providing economic incentives for reducing the emissions of pollutants.

When pursuing this approach the government sets a limit or cap on the amount of a pollutant that may be emitted. It then allocates emissions permits up to the specified limit among firms. The permits represent the right to emit or discharge a specific volume of a specified pollutant. Firms are required to hold a number of permits equivalent to their emissions. Firms that need to increase their volume of emissions must buy permits from firms that require fewer of them. This transfer is referred to as a trade. In effect, the buyer is paying a charge for polluting, while the seller is being rewarded for having reduced emissions. The outcome achieved by the market for permits is more efficient, regardless of the initial allocation of permits.

The market for tradable permits creates incentives for firms to produce less pollution. Firms that have a high cost of reducing emissions are willing to pay for the permits, while those that can reduce emissions in the most cost-efficient manner will do so and sell their permits. Tradable permits thus achieve a desired level of the externality by allowing the market to determine which market actors can create the externality.

There are several active trading programs for air pollutants. For greenhouse gases the largest is the European Union Emission Trading Scheme. In the United States there is a national market for sulfur dioxide emissions to reduce acid rain. Markets for other pollutants tend to be smaller and more localized.

Key Points

- Command-and-control regulation requires or forbids certain behaviors with the goal of addressing an externality.
- Regulation is difficult to implement and enforce correctly.

- Command-and-control regulation can come in the form of government-imposed standards, targets, process requirements, or outright bans.
- The allocation of tradable permits is a market-based policy that has been primarily used to combat pollution.
- A corrective tax is a market-based policy option used by the government to address negative externalities.
- Taxes increase the cost of producing goods or services generating the externality, thus encouraging firms to produce less output.
- The tax should be set equal to the value of the negative externality, which is very difficult to do in practice.
- Corrective taxes increase efficiency and provide the government with revenues as well.
- A permit is a right to produce a certain amount of a negative externality, such as pollution.
- Permits are traded among firms. Firms that are able to cheaply reduce production of the externality can sell permits to firms that are unable to make such reductions and are willing to pay for the permits.
- Regardless of the initial allocation of permits, the market for permits achieves an outcome that is more efficient for society.

Key Terms

- **Negative Externality**: A detremental effect suffered by a party due to a transaction it was not a part of.
- **Pigovian tax**: A tax applied to a market activity that is generating negative externalities (costs for somebody else).
- **Permit**: The right to produce a given amount of a negative externality (for example, the right to emit a specific volume of a pollutant).
- **quota**: A restriction on the import of something to a specific quantity.

- Emissions trading. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Emissions_trading**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Market Failure 2. **Provided by**: econ100-powers Wikispace. **Located at**: **econ100-powers.wikispaces.com/Market+Failure+2**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Tradable smoking pollution permits. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Tradabl...lution_permits**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Individual fishing quota. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Individual_fishing_quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Permit. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Permit**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- quota. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- All sizes | No smoking within 5 meters | Flickr - Photo Sharing!. **Provided by**: Flickr. **Located at**: **www.flickr.com/photos/hendry/...n/photostream/**. **License**: ***CC BY: Attribution***
- Social cost with tax. **Provided by**: hhttp//en.Wikipedia.org/wiki/...t_with_tax.svg. **Located at**: **hhttp//en.Wikipedia.org/wiki/...t_with_tax.svg**. **License**: ***Public Domain: No Known Copyright***
- Coal power plant Datteln 2 Crop1. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Co...ln_2_Crop1.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***

10.4: Private Solutions

Types of Private Solutions

Private actors will sometimes effectively address externalities and reach efficient outcomes without government intervention.

Learning Objectives

- Evaluate how effective private solutions may be in solving market failures produced by externalities

Government intervention is not always necessary to address externalities. Private actors will sometimes arrive at their own solutions.

There are several types of private solutions to market failures:

- **Moral codes:** Moral codes guide individuals' behavior. Individuals know that certain actions are simply not "the right thing to do" or would elicit disapproving reactions from others. This is illustrated in the case of littering. The likelihood of being fined may be small, but moral codes provide an incentive to refrain from littering.
- **Charities:** Charities channel donations from private individuals towards fighting to limit behaviors that result in negative externalities or promoting behaviors that generate positive externalities. The former can be seen in the case of organizations that protect the environment, while the latter is exemplified through organizations that raise money for education.
- **Business mergers or contracts in the self interest of relevant parties:** Two businesses that offer positive externalities to each other can merge or enter into a contract that makes both parties better off.

The Coase theorem, which was developed by Ronald Coase, posits that two parties will be able to bargain with each other to reach an agreement that efficiently addresses externalities. However, the theorem notes several conditions in order for such a solution to occur, including low transaction costs (the costs the parties incur by negotiating and coming to agreement) and well-defined property rights. If the conditions are met, the bargaining parties are expected to reach an agreement where everyone is better off. In practice, however, transaction costs do exist, and the bargaining process does not always run smoothly. As a result, private individuals often fail to resolve problems.

The Coase Theorem

The Coase theorem states that private parties can find efficient solutions to externalities without government intervention.

Learning Objectives

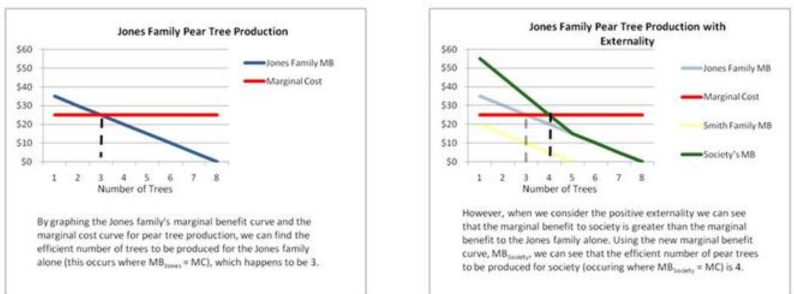
- Explain the usefulness and shortcomings of the Coase Theorem.

The Coase Theorem, named after Nobel laureate Ronald Coase, states that in the presence of an externality, private parties will arrive at an efficient outcome without government intervention. According to the theorem, if trade in an externality is possible and there are no transaction costs, bargaining among private parties will lead to an efficient outcome regardless of the initial allocation of property rights.

Efficient Solution: According to the Coase theorem, two private parties will be able to bargain with each other and find an efficient solution to an externality problem.

Imagine a farm and a ranch next to each other. The rancher's cows occasionally wander over to the farm and damage the farmer's crops. The farmer has an incentive to bargain with the rancher to find a more efficient solution. If it is more efficient to prevent cattle trampling a farmer's field by fencing in the farm, rather than fencing in the cattle, the outcome of the bargaining will be the fence around the farm.

Take another example. The Jones family plants pear trees on their property which is adjacent to the Smith family. The Smith family gets an external benefit from the Jones family's pear trees because they pick up the pears that fall on the ground on their side of the property line (see). This is an externality because the Smith family does not pay the Jones family for the utility received from gathering fallen pears. As a result, the Jones family plants too few pear trees. In response, the Jones family can put up a net that will prevent pears from falling on the Smith's side of the property line, eliminating the externality. Alternatively, the Jones could impose a cost on the Smith family if they want to continue to enjoy the pears from the pear trees. Both parties will be better off if they can agree to the second scenario, as the Smith family will continue to enjoy pears and the Jones family can increase the production of pears.



Effects of Externalities: This graph exemplifies how Coase's Theorem functions in a practical manner, underlining the effects of an externality in an economic model.

In practice, transaction costs are rarely low enough to allow for efficient bargaining and hence the theorem is almost always inapplicable to economic reality.

Key Points

- Private solutions to externalities include moral codes, charities, and business mergers or contracts in the self interest of relevant parties.
- The Coase theorem states that when transaction cost are low, two parties will be able to bargain and reach an efficient outcome in the presence of an externality.
- In practice, private parties often fail to resolve the problem of externalities on their own.
- According to the theorem, the parties affected by an externality will bargain to reach an outcome that will be more efficient.
- Transaction costs must be low in order for parties to arrive at a more efficient outcome.
- In the real world, transaction costs are rarely low, so the Coase theorem is often inapplicable.

Key Terms

- **Transaction cost**: The cost incurred in making an economic exchange, such as the costs required to come to an acceptable agreement with the other party to the transaction, drawing up an appropriate contract and so on.
- **Coase Theorem**: The theorem states that private economic actors can solve the problem of externalities among themselves.

- CHAPTER 10 . EXTERNALITIES. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces....ERNALITIES+%3B)**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Market failure. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Market_failure**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Boundless. **Provided by**: Boundless Learning. **Located at**: **www.boundless.com//economics/.../coase-theorem**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Transaction cost. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Transaction%20cost**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Chapter 10- Externalities. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces....+Externalities**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Externality. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Externality**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Coase Theorem. **Provided by**: uicecon Wikispace. **Located at**: **uicecon.wikispaces.com/Coase+Theorem**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Coase theorem. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Coase_theorem**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Coase theorem. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Coase_theorem**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Coase theorem. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Coase_theorem**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Coase theorem. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Coase_theorem**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Ronald Coase. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Ronald_Coase**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Transaction cost. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Transaction%20cost**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...ree_graphs.JPG**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Handshake. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Handshake.svg**. **License**: ***Public Domain: No Known Copyright***

CHAPTER OVERVIEW

11: Market Failure- Public Goods and Common Resources

11.1: Public Goods

Defining a Good

There are four types of goods in economics, which are defined based on excludability and rivalrousness in consumption.

Learning Objectives

- Define a good

There are four categories of goods in economics, which are defined based on two attributes. The first attribute is excludability, or whether people can be prevented from using the good. The second is whether a good is rival in consumption: whether one person's use of the good reduces another person's ability to use it.

National defense provides an example of a good that is non-excludable. America's national defense establishment offers protection to everyone in the country. Items on sale in a store, on the other hand, are excludable. The store owner can prevent a customer from obtaining a good unless the customer pays for it. National defense also provides an example of a good that is non- rivalrous. One person's protection does not prevent another person from receiving protection. In contrast, shoes are rivalrous. Only one person can wear a pair of shoes at a time.

Combinations of these two attributes create four categories of goods:

	Excludable	Non-Excludable
Rivalrous	**Private Goods** food, clothing, cars, personal electronics	**Common Goods** fish stocks, timber, coal
Non-Rivalrous	**Club Goods** cinemas, private parks, satellite tv	**Public Goods** air, national defense

Four Types of Goods: There are four categories of goods in economics, based on whether the goods are excludable and/or rivalrous in consumption.

- **Private goods:** Private goods are excludable and rival. Examples of private goods include food, clothes, and flowers. There are usually limited quantities of these goods, and owners or sellers can prevent other individuals from enjoying their benefits. Because of their relative scarcity, many private goods are exchanged for payment.
- **Common goods:** Common goods are non-excludable and rival. Because of these traits, common goods are easily over-consumed, leading to a phenomenon called "tragedy of the commons. " In this situation, people withdraw resources to secure short-term gains without regard for the long-term consequences. A classic example of a common good are fish stocks in international waters. No one is excluded from fishing, but as people withdraw fish without limits being imposed, the stocks for later fishermen are depleted.
- **Club goods:** Club goods are excludable but non-rival. This type of good often requires a "membership" payment in order to enjoy the benefits of the goods. Non-payers can be prevented from access to the goods. Cable television is a classic example. It requires a monthly fee, but is non-rival after the payment.
- **Public goods:** Public goods are non-excludable and non-rival. Individuals cannot be effectively excluded from using them, and use by one individual does not reduce the good's availability to others. Examples of public goods include the air we breathe, public parks, and street lights. Public goods may give rise to the "free rider problem. " A free-rider is a person who receives the benefit of a good without paying for it. This may lead to the under-provision of certain goods or services.

Private Goods

A private good is both excludable and rivalrous.

Learning Objectives

- Define a private good

In economics, a private good is defined as an asset that is both excludable and rivalrous. It is excludable in that it is possible to exercise private property rights over it, preventing those who have not paid from using the good or consuming its benefits. For example, person A may have the means and will to pay $20 for a t-shirt. Person B may not wish to pay $20 or may not be able to do so. Person B would not be able to purchase the t-shirt. Additionally, the private good is rivalrous in that its consumption by one person necessarily prevents consumption by another. When person A purchases and drinks a bottle of water, the same bottle of water is not available for person B to purchase and consume.

A private good is a scare economic resource, which causes competition for it. Generally, people have to pay to enjoy the benefits of a private good. Because people have to pay to obtain it, private goods are much less likely to encounter a free-rider problem than public goods. Thus, generally, the market will efficiently allocate resources to produce private goods.

In daily life, examples of private goods abound, including food, clothing, and most other goods that can be purchased in a store. Take an example of an ice cream cone. It is both excludable and rivalrous. It is possible to prevent someone from consuming the ice cream by simply refusing to sell it to them. Additionally, it can be consumed only once, so its consumption by one individual would definitely reduce others’ ability to consume it.

Ice Cream Cone: An ice cream cone is an example of a private good. It is excludable and rival.

Public Goods

Individuals cannot be excluded from using a public good, and one individual's use of it does not limit its availability to others.

Learning Objectives

- Define a public good

A public good is a good that is both non-excludable and non-rivalrous. This means that individuals cannot be effectively excluded from its use, and use by one individual does not reduce its availability to others. Examples of public goods include fresh air, knowledge, lighthouses, national defense, flood control systems, and street lighting.

Streetlight: A streetlight is an example of a public good. It is non-excludable and non-rival in consumption

Public goods can be pure or impure. Pure public goods are those that are perfectly non-rivalrous in consumption and non-excludable. Impure public goods are those that satisfy the two conditions to some extent, but not fully.

The production of public goods results in positive externalities for which producers don't receive full payment. Consumers can take advantage of public goods without paying for them. This is called the "free-rider problem. " If too many consumers decide to "free-ride," private costs to producers will exceed private benefits, and the incentive to provide the good or service through the market will disappear. The market will thus fail to provide enough of the good or service for which there is a need.

For example, a local public radio station relies on support from listeners to operate. The station holds pledge drives several times a year, asking listeners to make contributions or face possible reduction in programming. Yet only a small percentage of the audience makes contributions. Some audience members may even listen to the station for years without ever making a payment. Those listeners who do not make a contribution are "free-riders. " If the station relies solely on funds contributed by listeners, it would under-produce programming. It must obtain additional funding from other sources (such as the government) in order to continue to operate.

Optimal Quantity of a Public Good

The government is providing an efficient quantity of a public good when its marginal benefit equals its marginal cost.

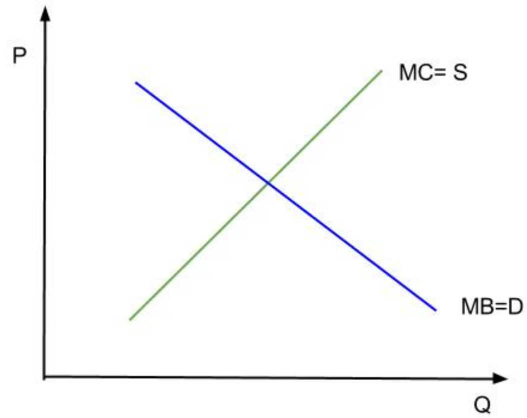
Learning Objectives

- Explain the optimal quantity of a public good

To determine the optimal quantity of a public good, it is necessary to first determine the demand for it. Demand for public goods is represented through price-quantity schedules, which show the price someone is willing to pay for the extra unit of each possible quantity. Unlike the market demand curve for private goods, where individual demand curves are summed horizontally, individual demand curves for public goods are summed vertically to get the market demand curve. As a result, the market demand curve for public goods gives the price *society* is willing to pay for a given quantity. It is equal to the marginal benefit curve. Due to the law of diminishing marginal utility, the demand curve is downward sloping.

Often, the government supplies the public good. The supply curve for a public good is equal to its marginal cost curve. Because of the law of diminishing returns, the marginal cost increases as the quantity of the good produced increases. The supply curve therefore has an upward slope.

As already noted, the demand curve is equal to the marginal benefit curve, while the supply curve is equal to the marginal cost curve. The optimal quantity of the public good occurs where MB (society's marginal benefit) equals MC (provider's marginal cost), or where the two curves intersect. When MB = MC, resources have been allocated efficiently.



Optimal Quantity of a Public Good: The optimal quantity of public good occurs where MB = MC.

The public good provider uses cost-benefit analysis to decide whether to provide a particular good by comparing marginal costs and marginal benefits. Cost-benefit analysis can also help the provider decide the extent to which a project should be pursued. Output activity should be increased as long as the marginal benefit exceeds the marginal cost. An activity should not be pursued when the marginal benefit is less than the marginal cost. An activity should be stopped at the point where MB equals MC. This is the MC=MB rule, by which the provider of the public good can determine which plan, will give society maximum net benefit.

Demand for Public Goods

The aggregate demand curve for a public good is the vertical summation of individual demand curves.

Learning Objectives

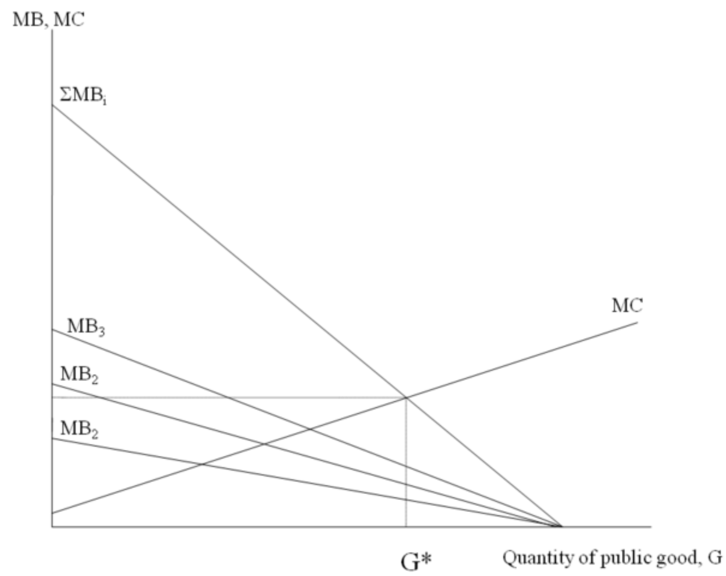
- Analyze the demand for a public good.

The aggregate demand for a public good is derived differently from the aggregate demand for private goods.

To an individual consumer, the total benefit of a public good is the dollar value that he or she places on a given level of provision of the good. The marginal benefit for an individual is the increase in the total benefit that results from a one-unit increase in the quantity provided. The marginal benefit of a public good diminishes as the level of the good provided increases.

Public goods are non-rivalrous, so everyone can consume each unit of a public good. They also have a fixed market quantity: everyone in society must agree on consuming the same amount of the good. However, each individual's willingness to pay for the quantity provided may be different. The individual demand curves show the price someone is willing to pay for an extra unit of each possible quantity of the public good.

The aggregate demand for a public good is the sum of marginal benefits to each person at each quantity of the good provided. The economy's marginal benefit curve (demand curve) for a public good is thus the vertical sum all individual's marginal benefit curves. The vertical summation of individual demand curves for public goods also gives the aggregate willingness to pay for a given quantity of the good.



Demand for a Public Good: The sum of the individual marginal benefit curves (MB) represent the aggregate willingness to pay or aggregate demand (∑MB). The intersection of the aggregate demand and the marginal cost curve (MC) determines the amount of the good provided.

This is in contrast to the aggregate demand curve for a private good, which is the horizontal sum of the individual demand curves at each price. Unlike public goods, society does not have to agree on a given quantity of a private good, and any one person can consume more of the private good than another at a given price.

The efficient quantity of a public good is the quantity that maximizes net benefit (total benefit minus total cost), which is the same as the quantity at which marginal benefit equals marginal cost.

Cost-Benefit Analysis

The government uses cost-benefit analysis to decide whether to provide a public good.

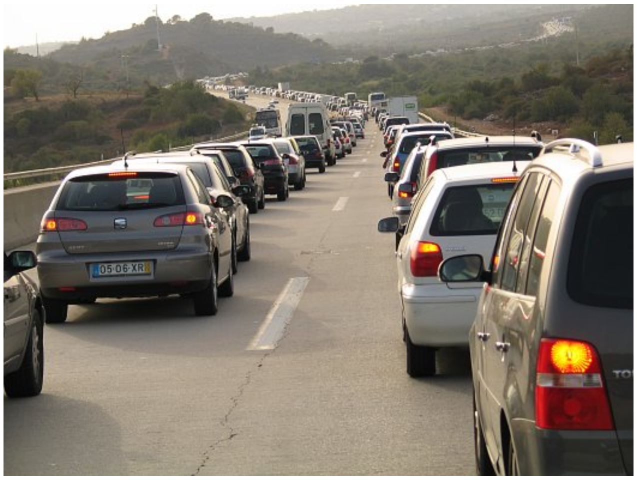
Learning Objectives

- Explain how to determine the net cost/benefit of providing a public good

The government uses cost-benefit analysis to decide whether to provide a particular public good and how much of it to provide. Cost-benefit analysis, which is also sometimes called benefit-cost analysis, is a systematic process for calculating the benefits and costs of a project to society as a whole.

The positive and negative effects captured by cost-benefit analysis may include effects on consumers, effects on non-consumers, externality effects, or other social benefits or costs. The guiding principle is to list all parties affected by a project and add a negative or positive value that they ascribe to the project’s effect on their welfare. Benefits and costs are expressed in monetary terms, and are adjusted for the time value of money, so that all flows of benefits and costs over time are expressed on a common basis in terms of their net present value. Financial costs tend to be most thoroughly represented in cost-benefit analyses due to relatively abundant market data. It is much more difficult to capture non-financial welfare impacts. For example, it is very difficult to place a dollar value on human life, consumers’ time, or environmental impact.

Imagine that the government is considering a project to widen a highway. The benefits side of the analysis might include time savings for passengers who can now avoid traffic, an increase in the number of passenger trips (as more people could now use the road), and lives saved by dint of fewer car accidents. The cost side of the analysis would include the cost of land that must be acquired prior to construction, construction, and maintenance. These costs and benefits will need to be translated into monetary terms for the sake of analysis.



The Highway as a Public Good: The benefits of a highway expansion project might include time savings for passengers, additional passenger trips, and saved lives. Costs might include construction and maintenance.

The procedure for conducting cost-benefit analysis is as follows:

1. Identify project(s) to be analyzed.
2. Estimate all costs and benefits to society associated with the project(s) over a relevant time horizon.
3. Assign a monetary value to all costs and benefits.
4. Calculate the net benefit of the project (total benefit minus total cost).
5. Adjust for inflation and apply the discount rate to calculate present value of the project.
6. Calculate the net present value for the project(s).
7. Make recommendation about project(s). If the benefit outweighs the cost, then the government should proceed with the project.

Key Points

- Private goods are excludable and rival. Examples of private goods include food and clothes.
- Common goods are non-excludable and rival. A classic example is fish stocks in international waters.
- Club goods are excludable but non-rival. Cable television is an example.
- Public goods are non-excludable and non-rival. They include public parks and the air we breathe.
- The owners or sellers of private goods exercise private property rights over them.
- A consumer generally has to pay for a private good.
- Generally, the market will efficiently allocate resources for the production of private goods.
- A public good is both non-excludable and non-rivalrous.
- Pure public goods are perfectly non-rival in consumption and non-excludable. Impure public goods satisfy those conditions to some extent, but not perfectly.
- Public goods provide an example of market failure. Because of the free-rider problem, they may be underpoduced.
- Collective demand for a public good is the vertical summation of individual demand curves. It shows the price society is willing to pay for a given quantity of a public good.
- The demand curve for a public good is downward sloping, due to the law of diminishing marginal utility. The supply curve is upward sloping, due to the law of diminishing returns.
- The optimal quantity of a public good occurs where the demand (marginal benefit) curve intersects the supply (marginal cost) curve.
- The government uses cost-benefit analysis to decide whether to provide a particular good. If MB is greater than MC there is an underallocation of a public good. If MC is greater than MB there is an overallocation of a public good. When MC = MB then there is an optimal allocation of public goods.
- For public goods, aggregate demand is the sum of marginal benefits to each person at each quantity of the good provided.
- As for private goods, the individual demand curves show the price someone is willing to pay for an extra unit of each possible quantity of a good.
- The efficient quantity of a public good is the quantity at which marginal benefit equals marginal cost.
- Cost -benefit analysis is a systematic way of calculating the costs and benefits of a project to society as a whole.

- Benefits and costs are expressed in monetary terms and are adjusted for the time-value of money.
- Financial costs are much easier to capture in the analysis than non-financial welfare impacts, such as impacts on human life or the environment.
- The government should provide a public good if the benefits to society outweigh the costs.

Key Terms

- **Rival**: A good whose consumption by one consumer prevents simultaneous consumption by other consumers
- **Excludable**: A good for which it is possible to prevent consumers who have not paid for it from having access to it.
- **Rivalrous**: A good whose consumption by one consumer prevents simultaneous consumption by other consumers.
- **free rider**: Someone who enjoys the benefits of a good without paying for it
- **Non-excludable**: Non-paying consumers cannot be prevented from accessing a good
- **Non-rivalrous**: A good whose consumption by one consumer does not prevent simultaneous consumption by other consumers
- **Cost-benefit analysis**: A systematic process for calculating and comparing the marginal benefits and marginal costs of a project or activity.
- **public good**: A good that is non-rivalrous and non-excludable.
- **net present value**: The present value of a project determined by summing the discounted incoming and outgoing future cash flows resulting from the decision.

- Excludable. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Excludable**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Goods. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Goods.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Strawberry ice cream cone (5076899310). **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Strawberry_ice_cream_cone_(5076899310).jpg**. **License**: ***CC BY: Attribution***
- Public Goods. **Provided by**: publicecon Wikispace. **Located at**: **publicecon.wikispaces.com/Public+Goods**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Public Goods. **Provided by**: publicecon Wikispace. **Located at**: **publicecon.wikispaces.com/Public+Goods**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Ch.11 Public Goods and Common Resources. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces.com/Ch.11+Public+Goods+and+Common+Resources**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Public good. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Public_good**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Public good. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Public_good**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Non-rivalrous. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Non-rivalrous**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Non-excludable. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Non-excludable**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- free rider. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/free_rider**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Goods. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Goods.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Strawberry ice cream cone (5076899310). **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Strawberry_ice_cream_cone_(5076899310).jpg**. **License**: ***CC BY: Attribution***
- Streetlight pont des Catalans sunset. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Streetlight_pont_des_Catalans_sunset.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Public Goods. **Provided by**: publicecon Wikispace. **Located at**: **publicecon.wikispaces.com/Public+Goods**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- market failure 3.3. **Provided by**: 13ecohghs Wikispace. **Located at**: **13ecohghs.wikispaces.com/market+failure+3.3**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- market failure 3.3. **Provided by**: 13ecohghs Wikispace. **Located at**: **13ecohghs.wikispaces.com/market+failure+3.3**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Market Failure and the Role of Government. **Provided by**: re-econ Wikispace. **Located at**: **re-econ.wikispaces.com/Market...+of+Government**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Cost-benefit analysis. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Cost-benefit%20analysis**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Goods. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Goods.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Strawberry ice cream cone (5076899310). **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Strawberry_ice_cream_cone_(5076899310).jpg**. **License**: ***CC BY: Attribution***
- Streetlight pont des Catalans sunset. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Streetlight_pont_des_Catalans_sunset.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Optimal Quantity of a Public Good. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Optimal_Quantity_of_a_Public_Good.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Public good. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Public_good**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Public Goods. **Provided by**: publicecon Wikispace. **Located at**: **publicecon.wikispaces.com/Public+Goods**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- public goods and common resources.ppt. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces....0resources.ppt**. **License**: ***CC BY-SA: Attribution-ShareAlike***

- public good. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/public%20good**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Goods. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Goods.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Strawberry ice cream cone (5076899310). **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Strawberry_ice_cream_cone_(5076899310).jpg**. **License**: ***CC BY: Attribution***
- Streetlight pont des Catalans sunset. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Streetlight_pont_des_Catalans_sunset.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Optimal Quantity of a Public Good. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Optimal_Quantity_of_a_Public_Good.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Samuelson condition. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Sa..._condition.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Costu2013benefit analysis. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Cost%E2...nefit_analysis**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Cost benefit analysis. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Cost_benefit_analysis**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- market failure 3.3. **Provided by**: 13ecohghs Wikispace. **Located at**: **13ecohghs.wikispaces.com/market+failure+3.3**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- CHAPTER 11- PUBLIC GOODS AND COMMON RESOURCES. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces.com/CHAPTER+11-+PUBLIC+GOODS+AND+COMMON+RESOURCES**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Fundamentals of Transportation/Evaluation. **Provided by**: Wikibooks. **Located at**: **en.wikibooks.org/wiki/Fundame...st_Application**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- net present value. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/net_present_value**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Goods. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Goods.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Strawberry ice cream cone (5076899310). **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Strawberry_ice_cream_cone_(5076899310).jpg**. **License**: ***CC BY: Attribution***
- Streetlight pont des Catalans sunset. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Streetlight_pont_des_Catalans_sunset.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Optimal Quantity of a Public Good. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Optimal_Quantity_of_a_Public_Good.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Samuelson condition. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Sa..._condition.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Auto stoped highway. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Auto_stoped_highway.JPG**. **License**: ***CC BY-SA: Attribution-ShareAlike***

11.2: Common Resources

The Tragedy of the Commons

The tragedy of the commons is the overexploitation of a common good by individual, rational actors.

learning Objectives

- Describe the tragedy of the commons

Common Goods

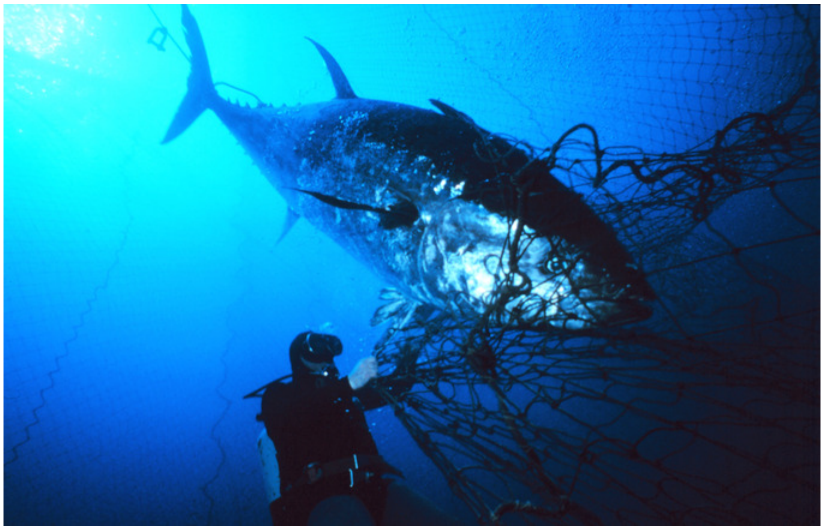
Common goods are goods that are rivalrous and non-excludable. This means that anyone has access to the good, but that the use of the good by one person reduces the ability of someone else to use it. A classic example of a common good are fish stocks in international waters; no one is excluded from fishing, but as people withdraw fish without limits being imposed, the stocks for later fishermen are potentially depleted.

Tragedy of Commons

The tragedy of the commons is the depletion of a common good by individuals who are acting independently and rationally according to each one's self-interest. Consider, the example of fish in international waters. Each individual fisherman, acting independently, will rationally choose to catch some of the fish to sell. This makes sense: there is a resource that the fisherman is able to use to generate a profit. However, when a lot of fishermen, all thinking this way, catch the fish, the total stock of fish may be depleted. When the stock of fish is depleted, none of the fishermen are able to continue fishing, even though, in the long run, each fisherman would have preferred that the fish not be depleted. The tragedy of the commons describes such situations in which people withdraw resources to secure short-term gains without regard for the long-term consequences.

Not all common goods, however, suffer from the tragedy of the commons. If individuals have enlightened self-interest, they will realize the negative long-term effects of their short-term decisions. This would be the same as the fishermen realizing that they should limit their fishing to preserve the stock of fish in the long-term.

In the absence of enlightened self-interest, the government may step in and impose regulations or taxes to discourage the behavior that leads to the tragedy of the commons. This would be like the government imposing limits on the amount of fish that can be caught.



Bluefin Tuna Caught in Net: Fish populations are at risk of becoming fully extinct due to overfishing. The Food and Agriculture Association estimated 70% of the world's fish species are either fully exploited or depleted.

The Free-Rider Problem

The free-rider problem is when individuals benefit from a public good without paying their share of the cost.

learning objectives

- Describe the Free-Rider Problem

It is easy to think about public goods as free. In your everyday life, you benefit from public goods such as roads and bridges even though no transaction occurs when you use them. However, even public goods need to be paid for. In the case of roads and bridges, everyone pays taxes to the government, who then uses the taxes to pay for public goods.

Roads: Free riders are able to use roads without paying their taxes because roads are a non-excludable public good.

Public goods, as you may recall, are both non-rivalrous and non-excludable. It is the second trait- the non-excludability- that leads to what is called the free-rider problem. The free-rider problem is that some people may benefit from a public good without paying their share of the cost.

Since public goods are non-excludable, free-riders not only can't be prevented from using the good, but actually have an incentive to continue to free-ride. If they will be able to use the public good whether they pay their share of the costs, they might as well not pay.

Take the military, for example. National security is a public good: it is both non-rivalrous and non-excludable. In order to have such a public good, everyone pays taxes which are then used by the government to finance the military. However, there are undoubtedly people who have not paid their taxes. These people, without having paid their share of the cost of having a military, still benefit from the protection the military provides. They are free-riders.

Of course, there are commonly regulations that attempt to discourage free-riding. For government-provided public goods, the government makes sure that everyone pays their share of the costs by enforcing tax laws. The threat of fines or jail time are enough of a threat that most people find it more appealing (in the US, at least) to pay their share of public goods via taxes than to free-ride.

Key Points

- Common goods are non-excludable and rivalrous.
- When individuals act independently and rationally, they may collectively trade long-term benefit for short-term gain.
- Enlightened self-interest and government intervention are two ways that the tragedy of the commons may be avoided.
- Public goods are non-excludable, but have a cost, so those who don't pay their share of the cost can still easily benefit from the good.
- Free-riders have an incentive to free ride because they can benefit from a good at a reduced personal cost.
- The providers of public goods often create enforcement mechanisms to mitigate the free-rider problem.

Key Terms

- **Common good**: Goods which are rivalrous and non-excludable.
- **Enlightened Self-Interest**: The ability for individuals to realize when their actions, collectively, will trade long-term benefit for short-term gain.
- **public good**: A good that is non-rivalrous and non-excludable.

CHAPTER OVERVIEW

12: International Trade

12.1: Introduction to International Trade

Reasons for Trade

Countries benefit when they specialize in producing goods for which they have a comparative advantage and engage in trade for other goods.

learning objectives

- Discuss the reasons that international trade may take place

International trade is the exchange of capital, goods, and services across international borders or territories. Trading-partners reap mutual gains when each nation specializes in goods for which it holds a comparative advantage and then engages in trade for other products. In other words, each nation should produce goods for which its domestic opportunity costs are lower than the domestic opportunity costs of other nations and exchange those goods for products that have higher domestic opportunity costs compared to other nations.

International Trade: Countries benefit from producing goods in which they have comparative advantage and trading them for goods in which other countries have the comparative advantage.

In addition to comparative advantage, other reasons for trade include:

- **Differences in factor endowments:** Countries have different amounts of land, labor, and capital. Saudi Arabia may have a lot of oil, but perhaps not enough lumber. It will thus have to trade for lumber. Japan may be able to produce technological goods of superior quality, but it may lack many natural resources. It may trade with Indonesia for inputs.
- **Gains from specialization:** Countries may gain economies of scale from specialization, experiencing long run average cost declines as output increases.
- **Political benefits:** Countries can leverage trade to forge closer cultural and political bonds. International connections also help promote diplomatic (rather than military) solutions to international problems.
- **Efficiency gains:** Domestic firms will be forced to become more efficient in order to be competitive in the global market.
- **Benefits of increased competition**: A greater degree of competition leads to lower prices for consumers, greater responsiveness to consumer wants and needs, and a wider variety of products.

To summarize, international trade benefits mostly all incumbents and generates substantial value for the global economy.

Understanding Production Possibilities

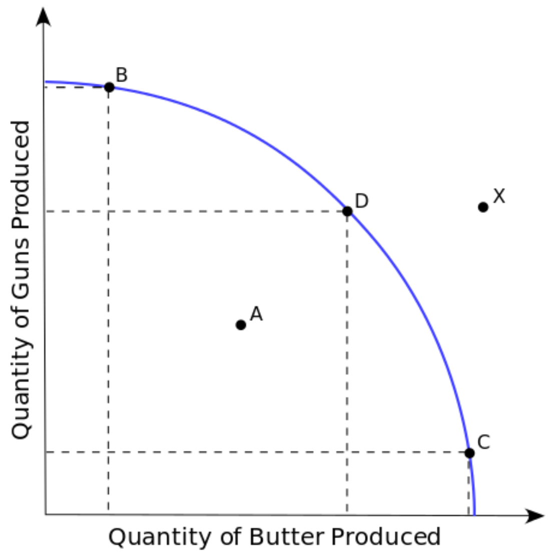
The production possibility frontier shows the combinations of output that could be produced using available inputs.

learning objectives

- Explain the benefits of trade and exchange using the production possibilities frontier (PPF)

In economics, the production possibility frontier (PPF) is a graph that shows the combinations of two commodities that could be produced using the same total amount of the factors of production. It shows the maximum possible production level of one commodity for any production level of another, given the existing levels of the factors of production and the state of technology.

PPFs are normally drawn as extending outward around the origin, but can also be represented as a straight line. An economy that is operating on the PPF is productively efficient, meaning that it would be impossible to produce more of one good without decreasing the production of the other good. For example, if an economy that produces only guns and butter is operating on the PPF, the production of guns would need to be sacrificed in order to produce more butter. If production is efficient, the economy can choose between combinations (i.e., points) on the PPF: B if guns are of interest, C if more butter is needed, or D if an equal mix of butter and guns is required.



Production Possibilities Frontier: If production is efficient, the economy can choose between combinations on the PPF. Point X, however, is unattaible with existing resources and technology if trade does not occur.

If the economy is operating below the curve, it is operating inefficiently, because resources could be reallocated in order to produce more of one or both goods without decreasing the quantity of either. Points outside the curve are unattainable with existing resources and technology if trade does not occur with an outside producer.

The PPF will shift outwards if more inputs (such as capital or labor) become available or if technological progress makes it possible to produce more output with the same level of inputs. An outward shift means that more of one or both outputs can be produced without sacrificing the output of either good. Conversely, the PPF will shift inward if the labor force shrinks, the supply of raw materials is depleted, or a natural disaster decreases the stock of physical capital.

Without trade, each country consumes only what it produces. In this instance, the production possibilities frontier is also the consumption possibilities frontier. Trade enables consumption outside the production possibility frontier. The world PPF is made up by combining countries' PPFs. When countries' autarkic productions are added (when there is no trade), the total quantity of each good produced and consumed is less than the world's PPF under free trade (when nations specialize according to their comparative advantage). This shows that in a free trade system, the absolute quantity of goods available for consumption is higher than the quantity available under autarky.

Defining Absolute Advantage

A country has an absolute advantage in the production of a good when it can produce it more efficiently than other countries.

learning objectives

- Relate absolute advantage, productivity, and marginal cost

Absolute advantage refers to the ability of a country to produce a good more efficiently than other countries. In other words, a country that has an absolute advantage can produce a good with lower marginal cost (fewer materials, cheaper materials, in less time, with fewer workers, with cheaper workers, etc.). Absolute advantage differs from comparative advantage, which refers to the ability of a country to produce specific goods at a lower opportunity cost.

A country with an absolute advantage can sell the good for less than a country that does not have the absolute advantage. For example, the Canadian economy, which is rich in low cost land, has an absolute advantage in agricultural production relative to some other countries. China and other Asian economies export low-cost manufactured goods, which take advantage of their much lower unit labor costs.

China and Consumer Electronics: Many consumer electronics are manufactured in China. China can produce such goods more efficiently, which gives it an absolute advantage relative to many countries.

Imagine that Economy A can produce 5 widgets per hour with 3 workers. Economy B can produce 10 widgets per hour with 3 workers. Assuming that the workers of both economies are paid equally, Economy B has an absolute advantage over Economy A in producing widgets per hour. This is because Economy B can produce twice as many widgets as Economy B with the same number of workers.

Party	Widgets per hour	Number of Employees
A	5	3
B	10	3

Absolute Advantage: Party B has an absolute advantage in producing widgets. It can produce more widgets with the same amount of resources than Party A.

If there is no trade, then each country will consume what it produces. Adam Smith said that countries should specialize in the goods and services in which they have an absolute advantage. When countries specialize and trade, they can move beyond their production possibilities frontiers, and are thus able to consume more goods as a result.

Defining Comparative Advantage

A country has a comparative advantage over another when it can produce a good or service at a lower opportunity cost.

learning objectives

- Analyze the relationship between opportunity cost and comparative advantage

Comparative Advantage

In economics, comparative advantage refers to the ability of a party to produce a particular good or service at a lower marginal and opportunity cost over another. Even if one country is more efficient in the production of all goods (has an absolute advantage in all goods) than another, both countries will still gain by trading with each other. More specifically, countries should import goods if the opportunity cost of importing is lower than the cost of producing them locally.

Specialization according to comparative advantage results in a more efficient allocation of world resources. Larger outputs of both products become available to both nations. The outcome of international specialization and trade is equivalent to a nation having more and/or better resources or discovering improved production techniques.

Determining Comparative Advantage

Imagine that there are two nations, Chiplandia and Entertainia, that currently produce their own computer chips and CD players. Chiplandia uses less time to produce both products, while Entertainia uses more time to produce both products. Chiplandia enjoys and absolute advantage, an ability to produce an item with fewer resources. However, the accompanying table shows that Chiplandia has a comparative advantage in computer chip production, while Entertainia has a comparative advantage in the production of CD players. The nations can benefit from specialization and trade, which would make the allocation of resources more efficient across both countries.

Production without Trade

Product	Chiplandia	Entertainia
1 Computer Chip	5 hours	24 hours
1 CD Player	10 hours	12 hours
Total	15 hours	36 hours

Opportunity Cost of Production

Product	Chiplandia	Entertainia
1 Computer Chip	½ CD Player	2 CD Players
1 CD Player	2 Computer Chips	½ Computer Chip

Production with Trade

Chiplandia		Entertainia	
1 Computer Chip for Chiplandia	5 hours	1 CD Player for Entertainia	12 hours
1 Computer Chip for Entertainia	5 hours	1 CD Player for Chiplandia	12 hours
Total	10 hours		24 hours

Comparative Advantage: Chiplandia has a comparative advantage in producing computer chips, while Entertainia has a comparative advantage in producing CD players. Both nations can benefit from trade.

For another example, if the opportunity cost of producing one more unit of coffee in Brazil is 2/3 units of wheat, while the opportunity cost of producing one more unit of coffee in the United States is 1/3 wheat, then the U.S. should produce coffee, while Brazil should produce wheat (assuming Brazil has the lower opportunity cost of producing wheat).

Comparative vs Competitive Advantage

It is important to distinguish between comparative advantage and competitive advantage. Though they sound similar, they are different concepts. Unlike comparative advantage, competitive advantage refers to a distinguishing attribute of a company or a product. It may or may not have anything to do with opportunity cost or efficiency. For example, having good brand recognition or relationships with suppliers is a competitive advantage, but not a comparative advantage. In the context of international trade, we more often discuss comparative advantage.

Absolute Advantage Versus Comparative Advantage

Absolute advantage refers to differences in productivity of nations, while comparative advantage refers to differences in opportunity costs.

learning objectives

- Differentiate between absolute advantage and comparative advantage

Absolute advantage compares the productivity of different producers or economies. The producer that requires a smaller quantity inputs to produce a good is said to have an absolute advantage in producing that good.

The accompanying figure shows the amount of output Country A and Country B can produce in a given period of time. Country A uses less time than Country B to make either food or clothing. Country A makes 6 units of food while Country B makes one unit, and Country A makes three units of clothing while Country B makes two. In other words, Country A has an absolute advantage in making both food and clothing.

Output per Day of Work

	Food	Clothing
Country A	6	3
Country B	1	2

Absolute Advantage: Country A has an absolute advantage in making both food and clothing, but a comparative advantage only in food.

Comparative advantage refers to the ability of a party to produce a particular good or service at a lower opportunity cost than another. Even if one country has an absolute advantage in producing all goods, different countries could still have different comparative advantages. If one country has a comparative advantage over another, both parties can benefit from trading because each party will receive a good at a price that is lower than its own opportunity cost of producing that good. Comparative advantage drives countries to specialize in the production of the goods for which they have the lowest opportunity cost, which leads to increased productivity.

For example, consider again Country A and Country B in. The opportunity cost of producing 1 unit of clothing is 2 units of food in Country A, but only 0.5 units of food in Country B. Since the opportunity cost of producing clothing is lower in Country B than in Country A, Country B has a comparative advantage in clothing.

Thus, even though Country A has an absolute advantage in both food and clothes, it will specialize in food while Country B specializes clothing. The countries will then trade, and each will gain.

Absolute advantage is important, but comparative advantage is what determines what a country will specialize in.

Benefits of Specialization

Specialization leads to greater economic efficiency and consumer benefits.

learning objectives

- Discuss the effects of specialization on production

Whenever a country has a comparative advantage in production it can benefit from specialization and trade. However, specialization can have both positive and negative effects on a nation's economy. The effects of specialization (and trade) include:

- **Greater efficiency:** Countries specialize in areas that they are naturally good at and also benefit from increasing returns to scale for the production of these goods. They benefit from *economies of scale* , which means that the average cost of producing the good falls (to a certain point) because more goods are being produced. Similarly, countries can benefit from increased learning. They simply are more skilled at making the product because they have specialized in it. These effects both contribute to increased overall efficiency for countries. Countries become better at making the product they specialize in.
- **Consumer benefits:** Specialization means that the opportunity cost of production is lower, which means that globally more goods are produced and prices are lower. Consumers benefit from these lower prices and greater quantity of goods.

- **Opportunities for competitive sectors:** Firms gain access to the whole world market, which allows them to grow bigger and to benefit further from economies of scale.
- **Gains from trade:** Suppose that Britain and Portugal each produce wine and cloth. Britain has a comparative advantage in cloth and Portugal in wine. By specializing and then trading, Britain can get a unit of wine for only 100 units of labor by trading cloth for labor instead of taking 110 units of labor to produce the wine itself (assuming the price of Cloth to Wine is 1). Similarly, Portugal can specialize in wine and get a unit of cloth for only 80 units of labor by trading, instead of the 90 units of labor it would take to produce the cloth domestically. Each country will continue to trade until the price equals the opportunity cost, at which point it will decide to just produce the other good domestically instead of trading. Thus (in this example with no trade costs) both countries benefit from specializing and then trading.

Of course, there are also some potential downsides to specialization:

- **Threats to uncompetitive sectors:** Some parts of the economy may not be able to compete with cheaper or better imports. For example, firms in United States may see demand for their products fall due to cheaper imports from China. This may lead to structural unemployment.
- **Risk of over-specialization:** Global demand may shift, so that there is no longer demand for the good or service produced by a country. For example, the global demand for rubber has fallen due the the availability of synthetic substitutes. Countries may experience high levels of persistent structural unemployment and low GPD because demand for their products has fallen.
- **Strategic vulnerability:** Relying on another country for vital resources makes a country dependent on that country. Political or economic changes in the second country may impact the supply of goods or services available to the first.

As a whole, economists generally support specialization and trade between nations.

Relationship Between Specialization and Trade

Comparative advantage is the driving force of specialization and trade.

learning objectives

- Discuss how countries determine which goods to produce and trade

Specialization refers to the tendency of countries to specialize in certain products which they trade for other goods, rather than producing all consumption goods on their own. Countries produce a surplus of the product in which they specialize and trade it for a different surplus good of another country. The traders decide on whether they should export or import goods depending on comparative advantages.

Imagine that there are two countries and both countries produce only two products. They can both choose to be self-sufficient, because they have the ability to produce both products. However, specializing in the product for which they have a comparative advantage and then trading would allow both countries to consume more than they would on their own.

One might assume that the country that is most efficient at the production of a good would choose to specialize in that good, but this isn’t always the case. Rather than absolute advantage, comparative advantage is the driving force of specialization. When countries decide what products to specialize in, the essential question becomes who could produce the product at a lower opportunity cost. Opportunity cost refers to what must be given up in order to obtain some item. It requires calculating what one could have gotten if one produced another product instead of one unit of the given product.

For example, the opportunity cost to Bob of 1 bottle of ketchup is 1/2 bottle of mustard. This means that in the same amount of time that Bob could produce one bottle of ketchup, he could have produced 1/2 bottle of mustard. Tom could have produced 1/3 bottle of mustard during the time that he was making one bottle of ketchup. Tom will have the comparative advantage in producing ketchup because he has to give up less mustard for the same amount of ketchup. In sum, the producer that has a smaller opportunity cost will have the comparative advantage. It follows that Bob will have a comparative advantage in the production of mustard.

Amount produced in 12 hours

	Ketchup	Mustard
Bob	6 bottles	3 bottles
Tom	12 bottles	4 bottles

Comparative Advantage: Tom has the comparative advantage in producing ketchup, while Bob has the comparative advantage in producing mustard.

There is one case in which countries are not better off trading: when both face the same opportunity costs of production. This doesn't mean that both countries have the same production function – one could still be absolutely more productive than the other – but neither has a comparative advantage over the other. In this case, specialization and trade will result in exactly the same level of consumption as producing all goods domestically.

Key Points

- International trade is the exchange of capital, goods, and services across international borders or territories.
- Each nation should produce goods for which its domestic opportunity costs are lower than the domestic opportunity costs of other nations and exchange those goods for products that have higher domestic opportunity costs compared to other nations.
- Benefits of trade include lower prices and better products for consumers, improved political ties among nations, and efficiency gains for domestic producers.
- The production possibilities curve shows the maximum possible production level of one commodity for any production level of another, given the existing levels of the factors of production and the state of technology.
- Points outside the production possibilities curve are unattainable with existing resources and technology if trade does not occur with an external producer.
- Without trade, each country consumes only what it produces. However, because of specialization and trade, the absolute quantity of goods available for consumption is higher than the quantity that would be available under national economic self-sufficiency.
- A country that has an absolute advantage can produce a good at lower marginal cost.
- A country with an absolute advantage can sell the good for less than the country that does not have the absolute advantage.
- Absolute advantage differs from comparative advantage, which refers to the ability to produce specific goods at a lower opportunity cost.
- Even if one country has an absolute advantage in the production of all goods, it can still benefit from trade.
- Countries should import goods if the opportunity cost of importing is lower than the cost of producing them locally.
- Specialization according to comparative advantage results in a more efficient allocation of world resources. A larger quantity of outputs becomes available to the trading nations.
- Competitive advantage is distinct from comparative advantage because it has to do with distinguishing attributes which are not necessarily related to a lower opportunity cost.
- The producer that requires a smaller quantity inputs to produce a good is said to have an absolute advantage in producing that good.
- Comparative advantage refers to the ability of a party to produce a particular good or service at a lower opportunity cost than another.
- The existence of a comparative advantage allows both parties to benefit from trading, because each party will receive a good at a price that is lower than its opportunity cost of producing that good.
- Whenever countries have different opportunity costs in production they can benefit from specialization and trade.
- Benefits of specialization include greater economic efficiency, consumer benefits, and opportunities for growth for competitive sectors.
- The disadvantages of specialization include threats to uncompetitive sectors, the risk of over-specialization, and strategic vulnerability.

- Nations decide whether they should export or import goods based on comparative advantages.
- Generally, nations can consume more by specializing in a good and trading it for other goods.
- When countries decide which country will specialize in which product, the essential question becomes who could produce the product at a lower opportunity cost.

Key Terms

- **comparative advantage**: The ability of a party to produce a particular good or service at a lower marginal and opportunity cost over another.
- **Production possibilities frontier**: A graph that shows the combinations of two commodities that could be produced using the same total amount of each of the factors of production.
- **Autarky**: National economic self-sufficiency.
- **Absolute advantage**: The capability to produce more of a given product using less of a given resource than a competing entity.
- **Opportunity cost**: The cost of an opportunity forgone (and the loss of the benefits that could be received from that opportunity); the most valuable forgone alternative.
- **competitive advantage**: Something that places a company or a person above the competition

- CMA CGM Balzac. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:CMA_CGM_Balzac.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Production Possibilities Frontier Curve. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Production_Possibilities_Frontier_Curve.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Absolute Advantage and Comparative advantage. **Provided by**: econ100-powers-sectiona Wikispace. **Located at**: **econ100-powers-sectiona.wikispaces.com/Absolute+Advantage+and+Comparative+advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Comparative advantage and trade. **Provided by**: mbaecon Wikispace. **Located at**: **mbaecon.wikispaces.com/Comparative+advantage+and+trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Absolute advantage. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Absolute_advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Absolute advantage. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Absolute_advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Comparative: Absolute advantage notes.doc. **Provided by**: business-tes Wikispace. **Located at**: **business-tes.wikispaces.com/file/detail/Comparative+&+Absolute+advantage+notes.doc**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Absolute advantage. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Absolute+advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- CMA CGM Balzac. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:CMA_CGM_Balzac.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Production Possibilities Frontier Curve. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Production_Possibilities_Frontier_Curve.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Iphone 4G-3 des screen. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Iphone_4G-3_des_screen.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Absolute advantage example 1. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/Fi..._example_1.png**. **License**: ***CC BY: Attribution***
- competitive advantage. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/competitive_advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- UNIT 7 Open EconomY- International Trade and Finance. **Provided by**: mreape Wikispace. **Located at**: **mreape.wikispaces.com/UNIT+7+Open+EconomY-+International+Trade+and+Finance**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Comparative advantage. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Comparative_advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- International Trade. **Provided by**: mchenry Wikispace. **Located at**: **mchenry.wikispaces.com/International+Trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- IB Economics/International Economics/Reasons for trade. **Provided by**: Wikibooks. **Located at**: **en.wikibooks.org/wiki/IB_Economics/International_Economics/Reasons_for_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- International Trade. **Provided by**: mchenry Wikispace. **Located at**: **mchenry.wikispaces.com/International+Trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Competitive advantage. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Competitive_advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Opportunity cost. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Opportunity+cost**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Boundless. **Provided by**: Boundless Learning. **Located at**: **www.boundless.com//business/d...tive-advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- CMA CGM Balzac. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:CMA_CGM_Balzac.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Production Possibilities Frontier Curve. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Production_Possibilities_Frontier_Curve.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Iphone 4G-3 des screen. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Iphone_4G-3_des_screen.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***

- Learning. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Learning**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- CMA CGM Balzac. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:CMA_CGM_Balzac.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Production Possibilities Frontier Curve. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Production_Possibilities_Frontier_Curve.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Iphone 4G-3 des screen. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Iphone_4G-3_des_screen.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Absolute advantage example 1. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Absolute_advantage_example_1.png**. **License**: ***CC BY: Attribution***
- Comparative Advantage. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Comparative_Advantage.jpg**. **License**: ***CC BY: Attribution***
- Absolute advantage. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Absolute_advantage.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- CHAPTER 3 . INTERDEPENDENCE AND THE GAINS FROM TRADE. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces.com/CHAPTER+3+.+INTERDEPENDENCE+AND+THE+GAINS+FROM+TRADE+;)**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Interdependence and the Gains from Trade. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces.com/Chapter+3+-+Interdependence+and+the+Gains+from+Trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Chapter 3 Independence and the Gains from Trade JBS. **Provided by**: mrski-apecon-2008 Wikispace. **Located at**: **mrski-apecon-2008.wikispaces.com/Chapter+3+Independence+and+the+Gains+from+Trade+JBS**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Opportunity cost. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Opportunity+cost**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- CMA CGM Balzac. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:CMA_CGM_Balzac.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Production Possibilities Frontier Curve. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Production_Possibilities_Frontier_Curve.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Iphone 4G-3 des screen. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Iphone_4G-3_des_screen.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Absolute advantage example 1. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Absolute_advantage_example_1.png**. **License**: ***CC BY: Attribution***
- Comparative Advantage. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Comparative_Advantage.jpg**. **License**: ***CC BY: Attribution***
- Absolute advantage. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Absolute_advantage.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Comparative advantage 2. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/F...dvantage_2.jpg**. **License**: ***CC BY-SA: Attribution-ShareAlike***

12.2: Gains from Trade

Exports: The Economic Impacts of Selling Goods to Other Countries

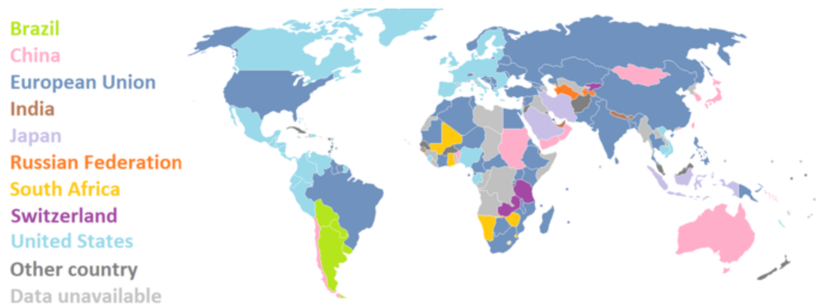
Exporting is a form of international trade which allows for specialization, but can be difficult depending on the transaction.

learning objectives

- Evaluate the effects of international trade on exporting countries

Exports

Export is defined as the act of a country shipping goods and services out of the port of a country. In international trade, an export refers to the selling of goods and services produced in the home country to other markets (other countries). The seller of the goods and services is referred to as the "exporter."



Exports: The map shows the primary exporters for countries around the globe. The colors indicate the leading merchandise export destination for the indicated country (the United States main export destination is the European Union). Exporting is the act of shipping goods and services to other countries.

Protecting Exports

In order to protect exports, commercial goods are subject to customs authorities for both the exporting and importing countries. Legal restrictions and trade barriers are in place internationally to control trade, whether goods are being exported or imported. When legal restrictions and trade barriers are lessened or lifted the producer surplus increases and so does the amount of the goods and services that are exported to other countries.

Impact of Exports

Exporting goods and services has both advantages and disadvantages for countries involved in international trade.

Exporting allows a country's producers to gain ownership advantages and develop low-cost and differentiated products. It is viewed as a low-risk mode of production and trade. Exporters also experience internationalization advantages which are the benefits of retaining a core competence within a company and threading it through the value chain instead of obtaining a license to outsource or sell the goods or services.

Disadvantages of exporting are mainly the result of manufacturers having to sell their goods to importers. In domestic sales, manufacturers sell directly to wholesalers or even directly to the retailer or customer. For exports, manufacturers face and extra layer in the chain of distribution which squeezes the margins. As a result, manufacturers may have to offer lower prices to the importers than to domestic wholesalers in order to move their product and generate business.

Imports: The Economics Impacts of Buying Goods from Other Countries

Imports are critical for many economies; they are the defining financial transactions of international trade and account for a large portion of the GDP.

learning objectives

- Evaluate the effects of international trade on an importing country

Imports

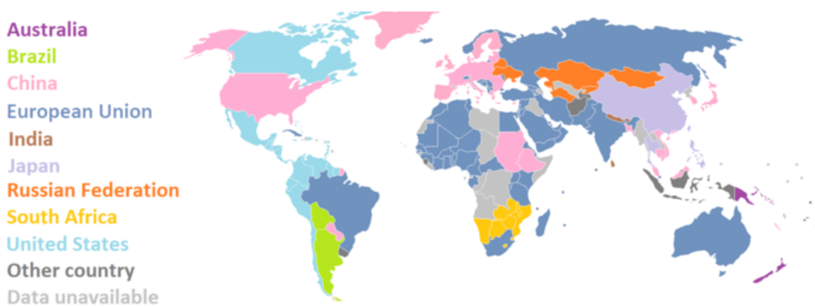
Imports are defined as purchases of good or services by a domestic economy from a foreign economy. The domestic purchaser of the good or service is called an importer. Imports and exports are critical for many economies and they are the defining financial transactions of international trade.

Protecting Imports

Due to the economic importance of imports, countries enact specific laws, barriers, and policies in order to regulate international trade. Protectionism is the economic policy of restraining trade between countries through tariffs on imported goods, restrictive quotas, and government regulations. When trade barriers and policies of protectionism are eliminated, consumer surplus increases. The price of a good or service will decrease while the quantity consumed will increase.

Impacts of Buying Imported Goods

On a national level, in most countries international trade and importing goods represents a significant share of the gross domestic product (GDP). International trade has a significant economic, social, and political importance in many countries. Imports provide countries with access to goods and services from other nations. Without imports, a country would be limited to the goods and services within its own borders.



Imports: The map shows the largest importers on an international scale. The color indicates the leading source of merchandise imports for the indicated country (the United States' imports the largest percentage of its goods from China). Imports account for a significant share in the gross domestic product (GDP) of a country.

International trade is generally less expensive than domestic trade despite additionally imposed costs, taxes, and tariffs. However, the factors of production are usually more mobile domestically than internationally (capital and labor). It is common for countries to import goods rather than a factor of production. For example, the U.S. imports labor-intensive goods from China. Instead of importing Chinese labor, the U.S. imports goods that were produced in China by Chinese labor.

On a business level, companies take part in direct-imports, which occur when a major retailer imports goods that are designed locally from an overseas manufacturer. The direct-import program allows the retailer to bypass the local supplier and purchase the final product directly from the manufacturer. Direct imports save retailers money by eliminating the local supplier.

Costs of Trade

Free trade is a policy where governments do not discriminate against imports and exports; creates a large net gain for society.

learning objectives

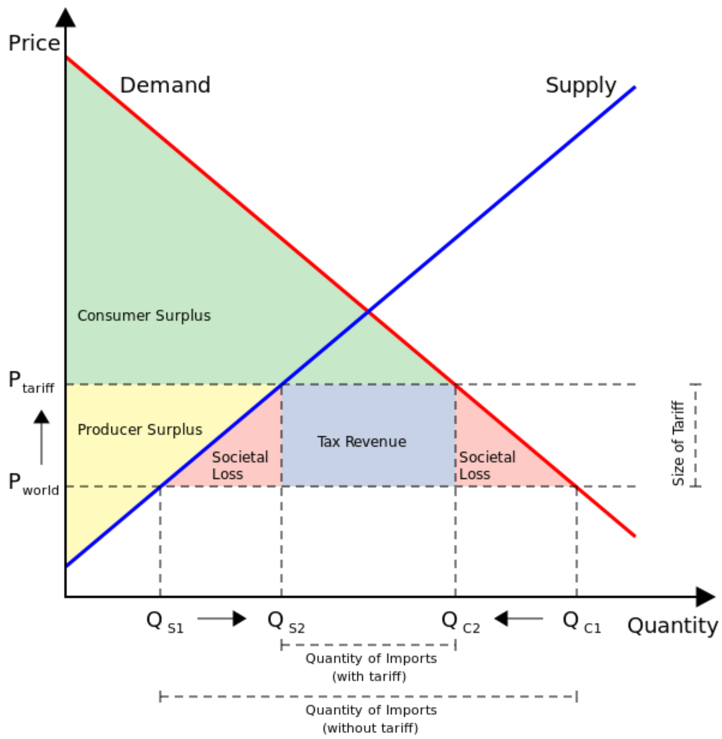
- Identify the groups that benefit and the groups that are harmed by free trade policies

Free Trade

Free trade is a policy where governments do not discriminate against exports and imports. There are few or no restrictions on trade and markets are open to both foreign and domestic supply and demand.

Advantages

Free trade is beneficial to society because it eliminates import and export tariffs. Restricted trade affects the welfare of society because although producers experience increases in surplus and additional revenue, the loss faced by consumers is greater than any benefit obtained. When a country trades freely with the rest of the world, it should theoretically produce a net gain for society and increases social welfare. Free trade policies consist of eliminating export tariffs, import quotas, and export quotas; all of which cause more losses than benefits for a country. With free trade in place, the producers of the exported good in exporting countries and the consumers in importing countries all benefit.



Tariffs: This image shows what happens to societal welfare when free trade is not enacted. Tariffs cause the consumer surplus (green area) to decrease, while the producer surplus (yellow area) and government tax revenue (blue area) increase. The amount of societal loss (pink area) is larger than any benefits experienced by the producers and government. Free trade does not have tariffs and results in net gain for society.

Disadvantages

One of the main disadvantages is the selective application of free trade. Economic inefficiency can be created through trade diversion. It is economically efficient for a good to be produced in the country with the lowest production costs. However, this does not always occur if a high cost producer has a free trade agreement and the low cost producer does not. When free trade is applied to only the high cost producer it can lead to trade diversion to not the most efficient producer, but the one facing the lowest trade barriers, and a net economic loss. Free trade is highly effective and provides society with a net gain, but only if it is applied.

Due to industry specializations, many workers are displaced and do not receive retraining or assistance finding jobs in other sectors. The nature of industries and trade increases economic inequality. As a result of unskilled workers the wages within the various industries may decline.

Another disadvantage is that by increasing returns to scale, can cause certain industries to settle in an geographically area where there is not comparative advantage. Despite this disadvantage, the level of output that is generated by free trade for both the “winner” and the “loser” is increased substantially.

The Results of Free Trade

Economists have studied free trade extensively and although it creates winners and losers, the main consensus is that free trade generates a large net gain for society. In a 2006 survey of American economists, it was found that 85.7% believed that the U.S. should eliminate any remaining tariffs and trade barriers. Economists professor N. Gregory Mankiw explained that, “few propositions command as much consensus among professional economists as that open world trade increases economic growth and raises living standards.”

Key Points

- Export is defined as the act of shipping goods and services out of the port of a country.
- Legal restrictions and trade barriers are in place internationally to control trade, whether goods are being exported or imported.
- When legal restrictions and trade barriers are lessened or lifted the producer surplus increases and so does the amount of the goods and services that are exported from the country.
- Exporting allows a country’s producers to gain ownership advantages and develop low-cost and differentiated products.
- Due to an extra layer in the chain of distribution which squeezes the margins, exporters may have to offer lower prices to the importers than to domestic wholesalers in order to move their product and generate business.
- Imports are defined as purchases of good or services by a domestic economy from a foreign economy.
- Protectionism is the economic policy of restraining trade between countries through tariffs on imported goods, restrictive quotas, and government regulations.
- In most countries, international trade and importing goods represents a significant share of the gross domestic product (GDP).
- International trade is generally more expensive than domestic trade due to additionally imposed costs, taxes, and tariffs.
- On a business level, companies take part in direct-imports; a major retailer imports goods from an overseas manufacturer in order to save money.
- Free trade eliminates export tariffs, import quotas, and export quotas; all of which cause more losses than benefits for a country.
- With free trade in place the producers in exporting countries and the consumers in importing countries all benefit.
- One of the main disadvantages is the selective application of free trade. Economic inefficiency can be created through trade diversion.
- Trade restricts displaces workers, makes overcoming unemployment challenging, increases economic inequality, and can lower wages.
- When free trade is applied to only the high cost producer it can lead to trade diversion and a net economic loss.
- Another disadvantage is that by increasing returns to scale, can cause certain industries to settle in an geographically area where there is not comparative advantage.

Key Terms

- **trade**: Buying and selling of goods and services on a market.
- **export**: Any good or commodity, transported from one country to another country in a legitimate fashion, typically for use in trade.
- **protectionism**: A policy of protecting the domestic producers of a product by imposing tariffs, quotas or other barriers on imports.
- **import**: To bring (something) in from a foreign country, especially for sale or trade.
- **tariff**: A system of government-imposed duties levied on imported or exported goods; a list of such duties, or the duties themselves.
- **welfare**: Health, safety, happiness and prosperity; well-being in any respect.
- **free trade**: International trade free from government interference, especially trade free from tariffs or duties on imports.

- Protectionism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Protectionism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Imports. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Imports**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- International trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/International_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- import. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/import**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- protectionism. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/protectionism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- trade. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Key export markets. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Key_export_markets.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Key import sources. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Ke...rt_sources.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- free trade. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/free_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Free trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Free_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Export. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Export**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Import. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Import**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Free trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Free_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- welfare. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/welfare**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- tariff. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/tariff**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Key export markets. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Key_export_markets.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Key import sources. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Key_import_sources.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- EffectOfTariff. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:EffectOfTariff.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***

12.3: The United States in the Global Economy

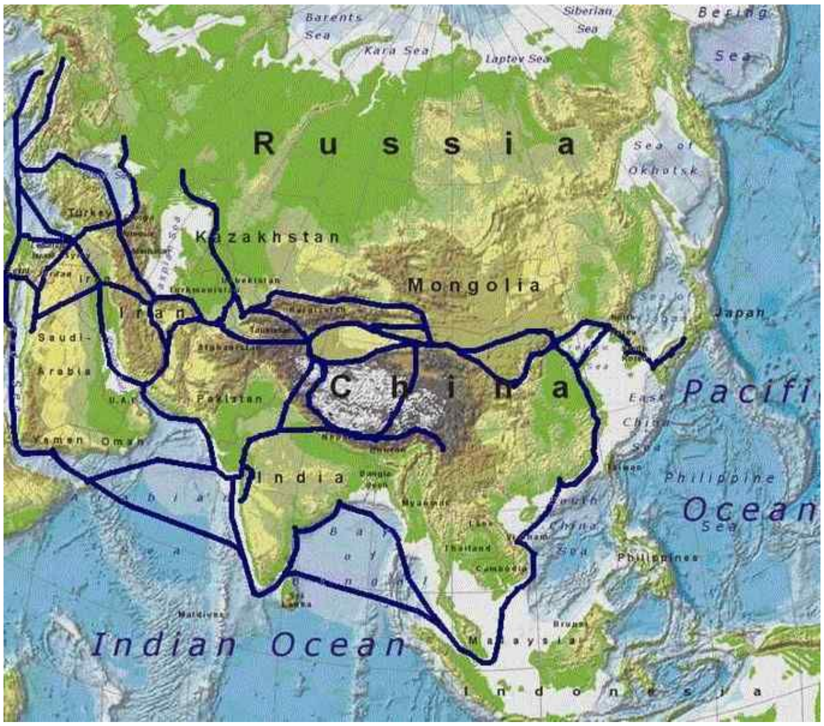
The Importance of Trade

International trade is an integral part of the modern world economy.

learning objectives

- Discuss the reasons of the U.S. increase in international trade participation after World War II

Economists generally support trade because it allows for increased overall utility for both countries. Gains from trade are commonly described as resulting from:



Silk Road Trade: Even in ancient times, people benefited from widespread international trade. The benefits from international trade have increased as costs decline and the international system becomes better integrated.

- specialization in production from division of labor (according to one's comparative advantage), economies of scale, scope, and agglomeration and relative availability of factor resources in types of output by farms, businesses, location and economies
- a resulting increase in total output possibilities
- trade through markets from sale of one type of output for other, more highly valued goods.

The Rise of International Trade

International trade is important, and, over time, has become more important. There have been three primary reasons for this increase in importance.

First, there have been large reductions in the cost of transportation and communication. It is now much cheaper to not only operate internationally and trade with foreign partners, but also to exchange information between potential buys and sellers.

Second, technological advances have made international production and trade easier to coordinate. More efficient telecommunications, from the first transatlantic telephone cable in 1956 to the popularization of the internet in the 1980s and 1990s, have allowed companies to exchange goods more efficiently and lowered the costs of international integration. Technological advances, from the invention of the jet engine to the development of just-in-time manufacturing, have also contributed to the rise in international trade.

Third, trade barriers between countries have fallen and are likely to continue to fall. In particular, the Bretton Woods system of international monetary management has shaped the relationship between the world's major industrial states and has resulted in a much more integrated system of international exchange. Established in 1946 to rebuild the international economic system after World War II, the Bretton Woods Conference set up regulations for production of their individual currencies to maintain fixed exchange rates between countries with the aim of more easily facilitating international trade.This was the foundation of the U.S. vision of postwar world free trade, which also involved lowering tariffs and, among other things, maintaining a balance of trade via fixed exchange rates that would be favorable to the capitalist system. Although the world eventually abolished the system of fixed exchange rates, the goal of more open economies and free international trade remained.

The Balance of Trade

The balance of trade is the difference between the monetary value of exports and imports of output in an economy over a certain period.

learning objectives

- Explain the relationship between the trade balance of a nation and its economic well-being

The balance of trade is the difference between the monetary value of exports and imports of output in an economy over a certain period, measured in the currency of that economy. It is the relationship between a nation's imports and exports. It is measured by finding the country's net exports. A positive balance is known as a trade surplus if it consists of exporting more than is imported; a negative balance is referred to as a trade deficit or, informally, a trade gap.

Factors that can affect the balance of trade include:

- The cost of production (land, labor, capital, taxes, incentives, etc.) in the exporting economy compared to those in the importing economy
- The cost and availability of raw materials, intermediate goods, and other inputs
- Currency exchange rate
- Multilateral, bilateral, and unilateral taxes or restrictions on trade
- Non-tariff barriers such as environmental, health, or safety standards
- The availability of adequate foreign exchange with which to pay for imports
- Prices of goods manufactured at home

In addition, the trade balance is likely to differ across the business cycle. In export-led growth (such as oil and early industrial goods), the balance of trade will improve during an economic expansion. However, with domestic demand led growth (as in the United States and Australia) the trade balance will worsen at the same stage in the business cycle.

Twin Deficits Hypothesis

The twin deficits hypothesis is a concept from macroeconomics that contends that there is a strong link between a national economy's current account balance and its government budget balance. This link can be seen from considering the national accounting model of the economy:

$$Y = C + I + G + (NX) \tag{12.3.1}$$

Y represents national income or GDP, C is consumption, I is investment, G is government spending, and NX stands for net exports (exports minus imports). This represents GDP because all the production in an economy (the left hand side of the equation) is used as consumption (C), investment (I), or government spending (G), and the leftover production is exported (NX).

Another equation defining GDP using alternative terms (which in theory results in the same value) is:

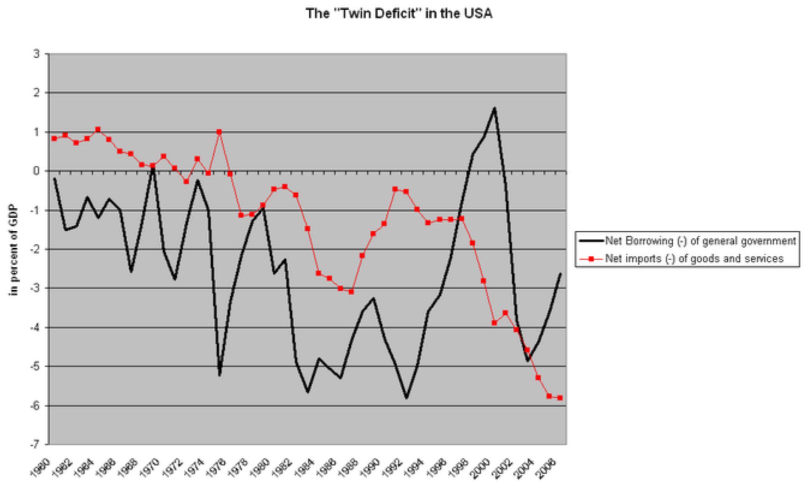
$$Y = C + S + T \tag{12.3.2}$$

Y is again GDP, C is consumption, S is savings, and T is taxes. This is because national income is also equal to output, and all individual income either goes to pay for consumption (C), to pay taxes (T), or becomes savings (S).

Since $Y = C + I + G + NX$, and $Y - C - T = S$, then $S = G - T + NX + I$, which simplifies to:

$$(S - I) + (T - G) = (NX) \tag{12.3.3}$$

If $(T - G)$ is negative, we have a budget deficit. Assuming that the economy is at potential output (meaning Y is fixed), if the budget deficit increases and savings and investment remain the same, then net exports must fall, causing a trade deficit. Thus, budget deficits and trade deficits go hand-in-hand.



Twin Deficits in the US: In the U.S., net borrowing has tended to have a direct relationship with net imports. The red line represents net imports, which is equivalent to the negative balance of trade, and the black line represents net borrowing, which is equivalent to the government budget deficit. Although the two are not identical, a rise in one tends to accompany a rise in the other, and vice versa.

The twin deficits hypothesis implies that as the budget deficit grows, net capital outflow from a country falls. This is because the nation is financing its spending by selling assets to foreigners. The total rate of national savings falls, which may lead to an increase in the interest rate as lending to the country (i.e. buying bonds and other financial assets) becomes more risky.

Key Points

- The international market serves as an important place for the exchange of goods and services.
- Economic theory shows that there are gains from trade for both countries involved.
- Advances in transportation has dramatically reduced the costs of moving goods around the globe.
- Technological advances have made international production and trade easier to coordinate.
- Trade barriers between countries have fallen and are likely to continue to fall.
- A positive balance is known as a trade surplus if it consists of exporting more than is imported; a negative balance is referred to as a trade deficit or, informally, a trade gap.
- Factors that can affect the balance of trade include the currency exchange rate, cost of inputs, barriers to trade such as tariffs and regulations, and the prices of domestic goods.
- The twin deficits hypothesis contends that there is a strong positive relationship between a national economy's current account balance and its government budget balance.

Key Terms

- **comparative advantage**: The ability of a party to produce a particular good or service at a lower margin and opportunity cost over another.
- **production possibilities curve**: The various combinations of amounts of two commodities that could be produced using the same fixed total amount of each of the factors of production
- **net capital outflow**: The net flow of funds being invested abroad by a country during a certain period of time.
- **net exports**: The difference between the monetary value of exports and imports.

CC LICENSED CONTENT, SPECIFIC ATTRIBUTION

- Boundless. **Provided by**: Boundless Learning. **Located at**: **www.boundless.com//business/d...tive-advantage**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Bretton Woods system. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Bretton_Woods_system**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Economic Globalization. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Economic_globalization**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- International Trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/International_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Gains from Trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Gains_from_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Foreign Trade of the United States. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Foreign_trade_of_the_United_States**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- production possibilities curve. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/production%20possibilities%20curve**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Silkroutes. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Silkroutes.jpg**. **License**: ***Public Domain: No Known Copyright***
- Twin deficits hypothesis. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Twin_deficits_hypothesis**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Balance of trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Balance_of_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- net capital outflow. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/net%20capital%20outflow**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- net exports. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/net%20exports**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Silkroutes. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Silkroutes.jpg**. **License**: ***Public Domain: No Known Copyright***
- TwinE. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:TwinE.PNG**. **License**: ***Public Domain: No Known Copyright***

12.4: Barriers to Trade

Tariffs

Tariffs are taxes levied on goods entering or exiting a country, and have consequences for both domestic consumers and producers.

learning objectives

- Discuss the consequences of a tariff for a domestic economy

One barrier to international trade is a tariff. A tariff is a tax that is imposed by a government on imported or exported goods. They are also known as customs duties.

Types of Tariffs

Tariffs can be classified based on what is being taxed:

- Import tariffs: Taxes on goods that are imported into a country. They are more common than export tariffs.
- Export tariffs: Taxes on goods that are leaving a country. This may be done to raise tariff revenue or to restrict world supply of a good.

Tariffs may also be classified by their purpose:

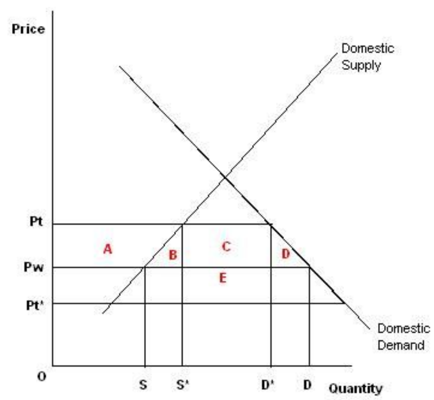
- Protective tariffs: Tariffs levied in order to reduce foreign imports of a product and to protect domestic industries.
- Revenue tariffs: Tariffs levied in order to raise revenue for the government.

Tariffs can also be classified on how the duty amount is valued:

- Specific tariffs: Tariffs that levy a flat rate on each item that is imported. For example, a specific tariff would be a fixed $1,000 duty on every car that is imported into a country, regardless of how much the car costs.
- Ad valorem tariffs: Tariffs based on a percentage of the value of each item. For example, an ad valorem tariff would be a 20% tax on the value of every car imported into a country.
- Compound tariffs: Tariffs that are a combination of specific tariffs and ad valorem tariffs. For example, a compound tariff might consist of a fixed $100 duty plus 10% of the value of every imported car.

Consequences of Levying a Tariff

To see the effects of levying an import tariff, consider the example shown in. Assume that there is an import tax levied on a good in a domestic country, Home. The domestic supply of the good is represented by the diagonal supply curve, and world supply is perfectly elastic and represented by the horizontal line at P_w. Before a tariff is levied, the domestic price is at P_w, and the quantity demanded is at D (with quantity S provided domestically, and quantity $D-S$ imported).



Effects of a Tariff: When a tariff is levied on imported goods, the domestic price of the good rises. This benefits domestic producers by increasing producer surplus, but domestic consumers see a small consumer surplus.

When the tariff is imposed, the domestic price of the good rises to P_t. Now, more of the good is provided domestically; instead of producing S, it now produces S^*. Imports of the good fall, from the quantity $D-S$ to the new quantity $D^* - S^*$. With the higher prices, domestic producers experience a gain in producer surplus (shown as area A). In contrast, because of the higher prices, domestic consumers experience a loss in consumer surplus; consumer surplus shrinks from the area above P_w to the area above P_t (it shrinks by the areas A, B, C, and D).

Because the tariff is a tax, the government gains some revenue. The government charges a tariff amount of $P_t - P_w$ on every imported good. The amount of revenue is equal to the tariff amount times the number of imported goods, or $(P_t - P_w)(D^* - S^*)$. This results in a governmental gain of area C.

In this example, domestic producers and the government both gain from the import tariff, and domestic consumers lose. However, if the world price is higher than the domestic price, a tariff will not change the price or quantity consumed of a good.

Quotas

Quotas are limitations on imported goods, come in an absolute or tariff-rate varieties, and affect supply in the domestic economy.

learning objectives

- Discuss the economic consequences of different kinds of quotas

Barriers to trade exist in many forms. A tariff is a barrier to trade that taxes imports or exports, thus increasing the cost of a good. Another barrier to trade is an import quota, which places a limit on the amount of a good that may enter a country.

Types of Quotas

There are two main types of import quota: the absolute quota and the tariff-rate quota.

An absolute quota is a limit on the quantity of specific goods that may enter a country during a certain time period. Once the quota has been fulfilled, no other goods may be imported into the country. An absolute quota may be set globally, in which case goods may be imported from any country until the goal has been reached. An absolute quota may also be set selectively for certain countries. As an example, suppose an absolute, global quota for pens is set at 50 million. The government is setting a limit that, in total, only 50 million pens can be imported. If there were a selective, absolute quota, only 50 million pens would be able to be imported, but this total would be divided among exporting countries. Country A might only be able to export 10 million pens, Country B might be able to export 25 million pens, and Country C might be able to export 15 million pens. Collectively, the total imports equal 50 million pens, but the proportions of pens from each country are set.

A tariff-rate quota is a two-tier quota system that combines characteristics of tariffs and quotas. Under a tariff-rate quota system, an initial quota of a good is allowed to enter the country at a lower duty rate. Once the initial quota is surpassed, imports are not stopped; instead, more of the good may be imported, but at a higher tariff rate. For example, under a tariff-rate quota system, a country may allow 50 million pens to be imported at the low tariff rate of $1 each. Any pen that is imported after this first-tier quota has been reached would be charged a higher tariff, say $3 each.

Reasons to Implement Quotas

Quotas are often implemented for similar reasons as other trade barriers. Often, quotas are instituted to:

- Protect domestic industries and employment: By reducing the number of foreign imports, domestic suppliers must produce more to meet domestic demand. By producing more, the suppliers must hire more domestic workers, increasing employment. Additionally, setting quotas to reduce foreign competition allows domestic “infant industries,” or young, small industries, to grow and mature to a competitive level.
- Protect against unfair trade practices: Setting a quota helps protect a domestic economy from unfair trade practices such as dumping, the pricing of imports below production cost. By restricting imports, quotas minimize the impact of such activities.
- Protect national security: Import quotas discourage imports and encourage domestic production of goods that may be necessary to the security of the country. By protecting and encouraging the growth of these defense-related industries, a country will not have to be dependent on foreign imports in the event of a war.

Consequences of Quotas

Like other trade barriers, quotas restrict international trade, and thus, have consequences for the domestic market. In particular, quotas restrict competition for domestic commodities, which raises prices and reduces selection. This hurts the domestic consumer, who experiences a loss in consumer surplus. On the other hand, this very action benefits the domestic producer, who sees an increase in producer surplus. Often, the increase in producer surplus is not enough to offset the loss in consumer surplus, so the economy experiences a loss in total surplus.

Quotas may also foster negative economic activities. Import quotas may promote administrative corruption, especially in countries where import quotas are given to selected importers. There are incentives to give the quotas to importers who can provide the most favors or the largest bribes to officials. Quotas may also encourage smuggling. As quotas raise the price of domestic goods, it becomes profitable to try and circumvent the quota by bringing in goods illegally, or in excess of the quota.

Other Barriers

Barriers to trade include specific limitations to trade, customs procedures, governmental participation, and technical barriers to trade.

learning objectives

- Distinguish different barriers to trade

In addition to tariffs and quotas, other barriers to trade exist. They can be divided into four separate categories: specific limitations to trade, customs and administrative procedures, government participation, and technical barriers to trade.

Specific Limitations to Trade

This category of trade barriers stems from regulations on international trade. Some examples include:

- Local content requirements, or domestic content requirements, are rules that mandate how much of a product must be produced domestically in order to qualify for lowered tariffs or other preferential treatment.
- Embargoes are prohibitions on trade ban imports or exports, and may apply to certain categories of products, or strictly to goods supplied by certain countries.

Customs and Administrative Procedures

This category of trade barriers refers to trade impediments that stem from governmental procedures and controls. Some examples include:

- Bureaucratic delays: Delays at ports or other country entrances caused by administrative or bureaucratic red-tape increase uncertainty and the cost of maintaining inventory.
- Anti-dumping duties: In international trade, dumping refers to a form of predatory pricing in which exported products are priced below the cost of production or below the price charged in the home market. Anti-dumping duties are usually extra taxes levied on the product to neutralize the predatory pricing and bring the price closer to the "normal value. "

Government Participation

This category of trade barriers represents direct governmental involvement in international trade. Some examples include:

- Government procurement programs: Public authorities, such as government agencies, are much like private interests in that they must also buy goods and services. Unlike private interests, governments are more likely to buy domestically produced goods and services, rather than the lowest-cost commodities. Because government procurement often represent a significant portion of a country's GDP, foreign suppliers are at a disadvantage to domestic ones when it comes to these programs.
- Export subsidies: Export subsidies are production subsidies granted to exported products, usually by a government. With export subsidies, domestic producers can sell their commodities in foreign markets below cost, which makes them more competitive.
- Countervailing duties: Countervailing duties, or anti-subsidy duties, are extra duties levied on imports in order to neutralize an export subsidy. If a country discovers that a foreign country subsidizes its exports, and domestic producers are injured as a result, a countervailing duty can be imposed in order to reduce the export subsidy advantage. In that respect, countervailing duties are similar to anti-dumping duties in that they both bring a imported product's value closer to the "normal value. "

Technical Barriers to Trade

Technical barriers to trade are non-tariff barriers to trade that refer to standards implemented by countries. Because these standards must be met before goods are allowed to enter or leave a country, they represent international trade barriers. Some examples include:

- Sanitary and phytosanitary measures: These are health standards for plants, animals, and other products, and are designed to protect humans, animals, and plants from pests or diseases.
- Rules for product weights, sizes, or packaging.
- Standards for labeling and testing products.
- Ingredient or identity standards.

Key Points

- Tariffs can be levied on goods being imported in a country (import tariff), or exported from a country (export tariff). They may be levied in order to protect domestic producers (protective tariff), or to raise revenue for the government (revenue tariff).
- Specific tariffs levy a fixed duty on a good. Ad valorem tariffs are based on a percentage of the good's value. Compound tariffs are a combination of specific and ad valorem tariffs.
- Tariffs often increase domestic producer surplus and the quantity of a good supplied domestically, but hurt domestic consumer surplus.
- There are two types of quotas: absolute and tariff -rate. Absolute quotas are quotas that limit the amount of a specific good that may enter a country. Tariff-rate quotas allow a quantity of a good to be imported under a lower duty rate; any amount above this is subject to a higher duty.
- Justifications for the use of quotas include protection for domestic employment and infant industries, protection against unfair foreign trade practices, and protection of national security.
- Quotas often hurt domestic consumers and benefit domestic producers. Quotas may also provide incentives for administrative corruption and smuggling.
- Specific limitations to trade barriers include local content requirements and embargoes. This category of barriers comes from trade regulations.
- Customs and administrative procedure barriers include bureaucratic red tape and anti- dumping practices. This category of barriers comes from government procedures.
- Governmental participation barriers include government procurement programs, export subsidies, and countervailing duties. This category of barriers involves the direct participation of government in trade.
- Technical barriers to trade include sanitary regulations, measurement and labeling standards, and ingredient standards. This category of barriers involves health, safety, and measurement standards.

Key Terms

- **tariff**: A system of government-imposed duties levied on imported or exported goods; a list of such duties, or the duties themselves.
- **absolute quota**: A limitation of the quantity of certain goods that may enter commerce during a specific period.
- **quota**: A restriction on the import of something to a specific quantity.
- **tariff-rate quota**: Allows a specified quantity of imported goods to be entered at a reduced rate of duty during the quota period, with quantities entered in excess of the quota limit subject to a higher duty rate.
- **Dumping**: Selling goods at less than their normal price, especially in the export market.
- **countervailing duty**: A tax levied on an imported article to offset the unfair price advantage it holds due to a subsidy paid to producers or exporters by the government of the exporting country if such imports cause or threaten injury to a domestic industry.
- **embargo**: A ban on trade with another country.

- Tariff. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Tariff**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- tariff. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/tariff**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Tariff. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Tariff.JPG**. **License**: ***Public Domain: No Known Copyright***
- Import quota. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Import_quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Non-tariff barriers to trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Non-tar...trade%23Quotas**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Tariff-rate quota. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Tariff-rate_quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- quota. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- tariff-rate quota. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/tariff-rate%20quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- absolute quota. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/absolute%20quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Tariff. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Tariff.JPG**. **License**: ***Public Domain: No Known Copyright***
- Raw cane sugar light. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/Fi...ugar_light.JPG**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- countervailing duty. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/countervailing_duty**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Technical barriers to trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Technical_barriers_to_trade**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Non-tariff barriers to trade. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Non-tariff_barriers_to_trade%23Quotas**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Dumping (pricing policy). **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Dumping_(pricing_policy)**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Countervailing duty. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Countervailing_duty**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Government procurement. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Government_procurement**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- embargo. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/embargo**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Dumping. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Dumping**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Tariff. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/File:Tariff.JPG**. **License**: ***Public Domain: No Known Copyright***
- Raw cane sugar light. **Provided by**: Wikimedia. **Located at**: **commons.wikimedia.org/wiki/File:Raw_cane_sugar_light.JPG**. **License**: ***CC BY-SA: Attribution-ShareAlike***

12.5: Arguments for and Against Protectionist Policy

National Security Argument

National security protectionist arguments pertain to the risk of dependency upon other nations for economic sustainability.

learning objectives

- Evaluate the arguments in favor of the use of trade protectionism in the security industry

Economic interdependence and globalization has resulted in a system, where each country is largely dependent upon other countries for economic sustainability (though to varying degrees). This results in a substantial national security threat in the form of conflicting or offensive trade strategies between countries. Indeed, economics is often used directly as a weapon of war and conflict via trade sanctions. This highlights a critical protectionist argument pertaining to the very real risk of dependency upon other nations for economic sustainability.

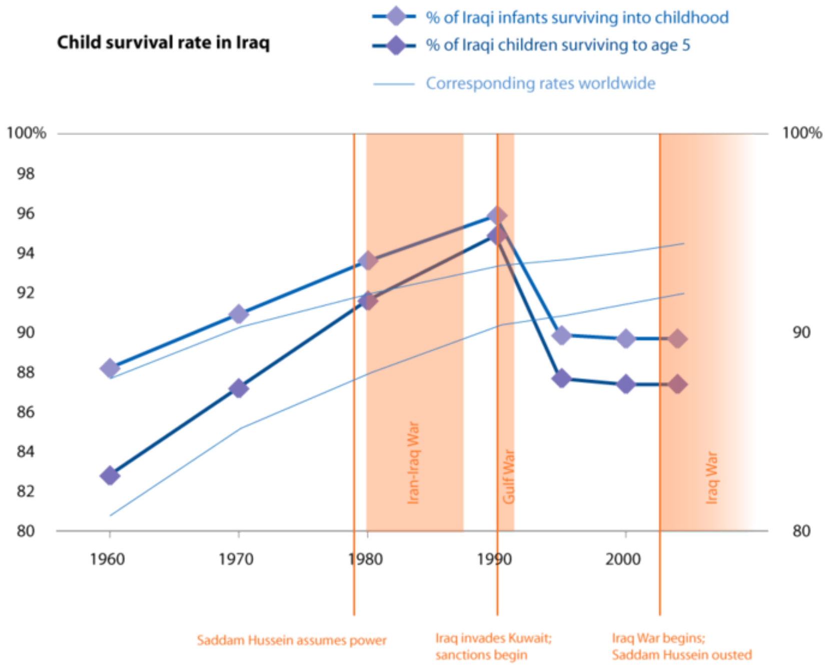
Trade and Conflict

An interesting discussion in economics is the relationship between trade and conflict. It has been noted, somewhat intuitively and empirically, that conflict reduces trade. However, is it also the case that trade reduces conflict? This question is largely unanswered, although the stances are becoming more highly developed. It is hypothesized that trade does not necessarily reduce conflict, but instead changes the nature of the conflict. Economic levers are much more practical than military levers, and are often used for similar reasons. For this reason, it is difficult to separate trade and conflict completely because there is some critical overlap between the two. This is a fundamental foundation for the trade protectionism logic from a national security perspective.

Trade During Conflict

A more specific context for trade and conflict can be the way in which trade is complicated during wartime. Indeed, trade during wartime can be a substantial threat to a nation depending on the scale and scope of the conflict (most notably who is involved). For example, consider World War II. In this scenario Germany was largely isolated in the conflict, and therefore had extremely limited trade partners. Direct conflict will almost always result in a complete cease in trading not only between the country in which the war is occurring, but also any of that country's allies (who may or may not be directly involved). However, some argue self-sufficiency (via protectionism) in war is not necessary, as friendly nations will still provide trade and economic support.

Sanctions also play a dramatic role as an offensive militaristic maneuver. Iran and North Korea are strong modern examples as well as the recent history of the U.S.-Iraq war. In all of these circumstances, either the U.S. alone or along with a number of allies (representing substantial consumption percentages) actively limited the ability for these countries to trade and generate economic value for their nations (and subsequently their people). While this looks purely economic, it has important social and humanitarian implications as well. The chart makes this case quite clearly, pointing out the death toll in wartime if economic levers are utilized.



Infant Mortality in Iraq During Sanctions: This graphic underlines the indirect consequences of employing economic levers (i.e. sanctions) in a militaristic fashion during a conflict. While the justification for these figures is complex, including other war-related factors, the correlation is quite clear. Diminishing a country's economic prospects will in turn result in loss of life, particularly in developing nations.

Protectionism

Combining these ideas, it is clear that there is substantial national security value to trade protectionism. However, the opportunity cost of leveraging the ever-growing global markets make this an unattractive prospect if taken to any extreme, as the benefits of global trade rapidly offset the risk of economic dependency upon hostile nations.

Infant Industry Argument

Economic markets are inherently competitive and newer economies are vulnerable to their more developed counterparts in other countries.

learning objectives

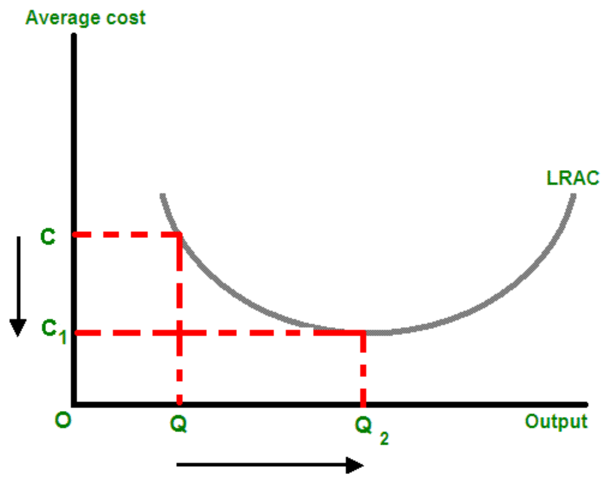
- Discuss the use of trade protectionism to promote new industries

Trade protectionism is defined as national policy restricting international economic trade to alter the balance between imports and goods manufactured domestically, usually executed via policies and governmental regulations such as import quotas, tariffs, taxes, anti-dumping legislation, and other limitations.

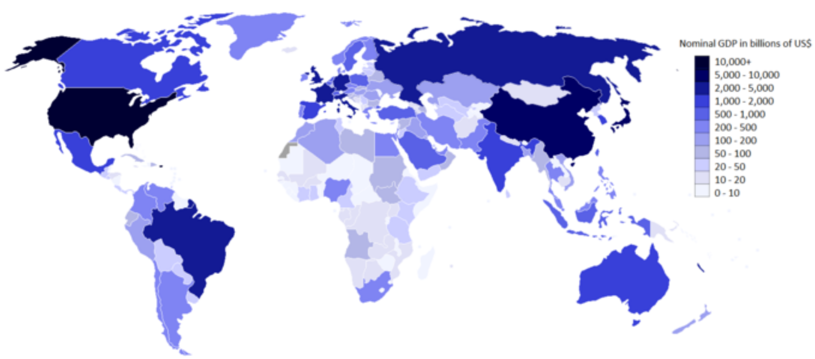
Arguments for Protecting Infant Industry

The primary purpose for this system is as the name implies: protection. Economic markets are inherently competitive, and newer economies are highly vulnerable to their more developed counterparts in other countries for a variety of reasons. The infant industry argument is that new industries need protection until they have become efficient enough to compete in the world market.

Despite the standard argument from mainstream economists postulating that free trade and open markets is the ideal system to allow for capitalistic development, there are many economists who believe that some degree of protectionism is the only way to minimize income gaps and substantial inequity from economy to economy (see). The primary advantage to countries with higher economic power and bigger corporations is simply economies of scale and economies of scope, in addition to being further along the experience curve.



Economies of Scale: The basic premise behind economies of scale is that higher production quantity reduces cost per unit, ultimately allowing for the derivation of economic advantage in the market. Infant industries generally do not have the capacity to do this.



GDP by Country: This map demonstrates the vast difference in overall economic power across the globe, underlining the inequities that need to be addressed in economic policy formulation.

History has proven the value of protection for the countries employing tariff-based international trade policies. Alexander Hamilton first pointed out the inequities of developing economies with young industry in 1790, which was later picked up and developed by Daniel Raymond and Friedrich List in the 19th century. Around this time frame, the United States was employing heavy tariffs to protect their fragile economic system as the economy began to achieve autonomy after British rule. Indeed, Britain employed similarly protectionist policies during this time frame, setting the tone for large economic expansion in the longer term.

Criticism

Of course, protective policy while industry develops domestically is not a cure all. In Brazil in the 1980's there were heavy protective policies in place to defend Brazil's nascent computer industry from highly evolved competitors internationally. While this seemed practical, what ended up happening was quite damaging for Brazil. Technology advanced rapidly, and without strategic alliances on a global scale, Brazil largely missed out on these advances. This protectionism seems to have damaged industry prospects on a global level for Brazil in this scenario.

From a broader and more far-reaching perspective, protectionism as a general principle has been heavily criticized (even in infant industry situations). The reason for this is quite simply the significant jump in prosperity as international trade expanded, and the huge capacity for specialization, economies of scale, technology sharing, and a host of other advantages that have been a direct result of free global markets. The problem still remains, however, that this prosperity is often unregulated and of the greatest benefit to the influential players in established economies, sometimes at the expense of exploitation of developing nations (cheaper labor, reduced governmental oversight, etc.). As a result of this, protecting infant industries can benefit the nation employing them, but generally with the opportunity cost of global value.

Unfair Competition Argument

One of the strongest arguments for trade protectionism is unfair competition emerging due to differences in policy and enforcement ability.

learning objectives

- Examine the use of protectionism as a way of addressing unfair competitive practices

Protectionist policies are a highly charged topic in economic debates, as economies work to attain the optimal balance of free trade and trade protectionism to capture the most value. In many ways, the global markets are torn between pursuing what is best on the global level and what is best at the domestic level, and there is sometimes dissonance between the two. One of the strongest arguments for some degree of trade protectionism is the tendency for unfair competition to emerge, particularly in developing markets without the infrastructure to monitor their businesses and enforce penalties. This is called the unfair competition argument.

Dumping

A popular recent topic is anti-dumping policies directed at international players looking to undercut domestic business through selling at dramatically reduced prices. This can be a substantial threat, particularly from economies where labor laws are lax and workers are exploited to create extremely low cost goods. This is also a risk when governments get too involved in business, a criticism often pointed out in China. Governments can provide subsidies to reduce costs for domestic companies. This can also be a threat in infant industries, where larger and more established players can push out smaller players via undercutting prices, absorbing losses until the competition goes bankrupt.

Offsetting this threat has been an ongoing struggle, with the emergence of international trade agreements and organizations like the World Trade Organization (WTO) playing an increasingly large role. One of the struggles with international trade is the difficulty of enforcement between nations, and the WTO plays a critical role in identifying malpractice and addressing it.

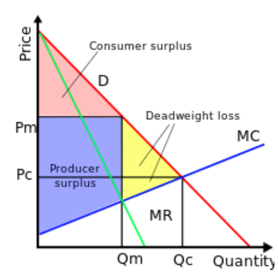
Intellectual Property

Another critical risk in the global market is intellectual property (IP) protection. Patents, in a domestic system, protect the innovator to allow them to generate returns on the substantial time investment required to invent or innovate new products or technologies. On a global scale, however, it is quite common for developing nations to copy new technologies via reverse engineering. This results in copycats violating the patents in an environment where the infrastructure domestically will probably not take legal action. This reduces the desire for innovation and places large economic risks on countries dependent upon this for growth.

This is addressed through international patent laws and trade agreements as well, alongside political pressures such as raising tariffs and placing import quotas on countries suspected to be in violation of patents. The downside to this is that utilizing these measures creates political unrest, global factions, and strained business relationships.

Mergers, Acquisitions, and Market Dominance

Another unfair competition threat is the emergence of global monopolies. Some of the larger ones attain enough global power and geographic diversification to be difficult to break up via domestic antitrust laws. demonstrates the substantial threat of deadweight losses being incurred in economies where consolidation results in a lack of competitive forces to drive down price.



Economic Losses in a Monopoly: This chart highlights the very real risk of lost economic value in a monopolistic situation (deadweight loss in yellow).

On the domestic level monopolies are widely seen as being addressed (though this is hotly debated by many economists in light of the 'too big to fail' and 'too big to jail' banks). On a global scale it is even more difficult to regulate, as the size and scale of these companies often extends beyond the power of the governments where these companies are located. This is addressed through international standards and trade agreements, standardizing governmental policy on a global level to reduce the risk of monopoly and unfair consolidation towards market dominance.

Jobs Argument

Many policy makers who are proponents of trade protectionism argue that limiting imports will create or save more jobs at home.

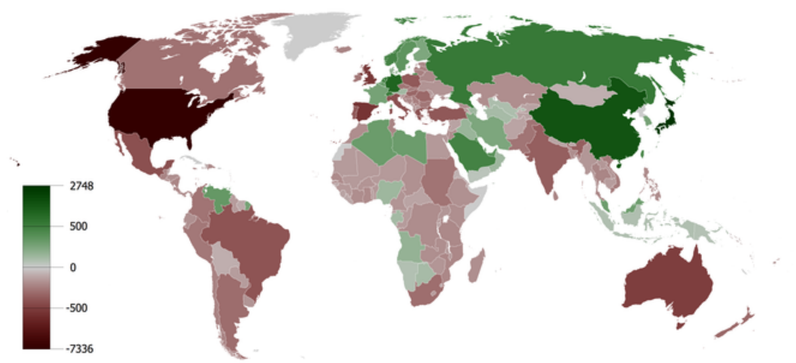
learning objectives

- Analyze the use of trade restrictions for strategic purposes

Many policy makers who are proponents of trade protectionism make the argument that limiting imports will create more jobs at home. This argument is predicated on the idea that buying more domestically will drive up national production, and that this increased production will in turn result in a healthier domestic job market. Domestic industries will not have to compete with foreign producers, and are therefore protected from losing marketshare to cheaper imports.

Trade Balance

It is useful to consider the concept of a trade balance, or net exports, in the context of the jobs argument. It is interesting to look at to assess the extremity to which some nations are 'consumer nations' and others are 'producer nations'. The U.S. and China are a great example of opposite sides of the spectrum, where the trade balance is heavy on one side of the spectrum.



Trade Balances on a Global Scale: It is interesting to look at this graph and assess the extremity to which some nations are 'consumer nations' and others are 'producer nations. ' The U.S. and China are a great example of opposite sides of the spectrum, where the trade balance is heavily on one side of the spectrum.

In the U.S. this has created a dramatic push for trade protectionism policies; something the United States has not actively pursued in quite some time. The disastrous 2008 economic collapse via the clear-cut abuses by the banks, and the resulting drop in employment rates, has created an incredibly tangible social and political agenda to bring production back to domestic jobs from overseas. This sentiment towards protectionism is a direct result of the jobs argument in view of an imbalanced trade ratio, where more exports (production and jobs at home) are required to sustain the ongoing consumption of imports.

Outsourcing

Along similar lines, it is common practice for companies to identify strategic alliances abroad and send much of the production work to these locations. This is often a result of cheaper labor and easier systems of governance in those regions. The obvious perspective, from a policy making context, is that these are jobs lost to overseas competitors. While this perspective is often criticized for being short-sighted and against the modern economic view of free markets, it has resulted in policy makers providing incentives to 'bring jobs back home. '

This idea of limiting outsourcing in light of the protectionist jobs argument has resulted in governmental subsidies that work to offset the costs of manufacturing domestically (in the U.S. particularly). These subsidies are essentially grants or tax breaks for companies operating domestically and creating jobs, driving up employment rates via protectionist strategies.

Trade Restriction Strategies

Offsetting the threats of outsourcing and trade imbalances and driving domestic purchasing, and thus domestic production, is done through a variety of political vehicles. Most notable among them are:

- Import Quotas: This is the act of limiting the number of a certain good that can be purchasing from a given country, ensuring that domestic producers maintain a portion of the market share.
- Tariffs: Tariffs are fairly straight-forward, essentially taxes to bring goods into a given country. High tariffs will raise the cost for foreign producers to sell their goods in a domestic system, providing strategic advantages for local producers. One of the pitfalls of tariffs is the likelihood of retaliation, where the foreign government returns with similar tariffs. This will in turn damage global prospects for domestic suppliers.
- Anti-dumping:Anti-dumping legislation actively offsets the ability of low cost or highly subsidized producers in foreign countries to undercut prices in a domestic system. Dumping is the process of selling goods far below market value to drive out competition, often in pursuit of creating a monopoly.
- Subsidies: On the other end of the spectrum, and as noted above, governments can provide subsidies to domestic producers to lower their costs and drive up competitive ability. This can in turn create jobs.

A Summary of International Trade Agreements

International trade agreements are agreements across national borders that reduce or eliminate trade barriers to promote economic exchange.

learning objectives

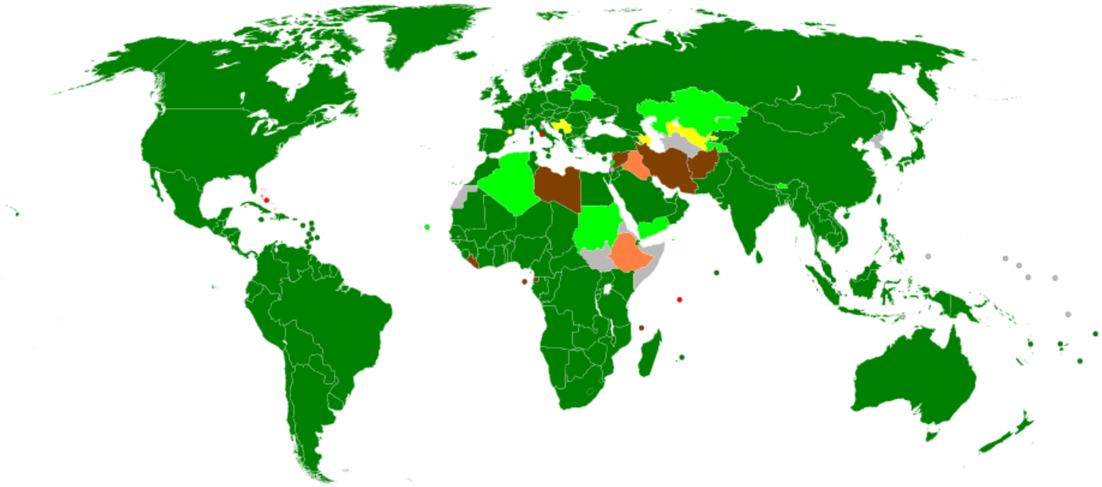
- Identify at least three main international trade agreements

International trade agreements are trade agreements across national borders intended to reduce or eliminate trade barriers to promote economic exchange. International trade encounters a variety of obstacles, some of which pertain to the protectionism identified in other atoms, which reduce trade incentives. This is usually through tariffs, quotas, taxes, and other trade restrictions. It is also useful to create standards and norms across different countries, particularly for things like intellectual property law recognition, which enables businesses to operate across borders.

There are quite a few international trade agreements, some of which are more formal than others. The trade agreements below provide a fairly comprehensive overview of the current international trade environment:

World Trade Organization (WTO)

The WTO is the largest international trade organization, replacing the General Agreement on Tariffs and Trade (GATT) in 1995, designed to enable international trade while reducing unfair practices. In many ways, the WTO is more complex than other international trade agreements because it incorporates a variety of smaller agreements into a larger framework. The WTO includes upwards of 60 different agreements alongside 159 official members and 25 observers. underlines how effective and universal international trade agreements are becoming. The WTO performs several objective functions as well if trade disputes arise, acting as a framework for assessing appropriate international trade practices.



WTO Members: The World Trade Organization (WTO) is an organization designed to oversee and enable international trade. This map shows how successful this has been on a global scale.

The core of the WTO is the most-favored nation (MFN) rule, which states that each WTO member must be charged the lowest tariffs that an importer places on any country. For example, if the US charges Brazil a 5% tariff on imported clothes, and this is the lowest tariff it has placed on any country in the WTO, all other WTO members must also be charged a 5% tariff. Every WTO member gets charged the lowest tariff that an importer charges any other member.

North American Free Trade Agreement (NAFTA)

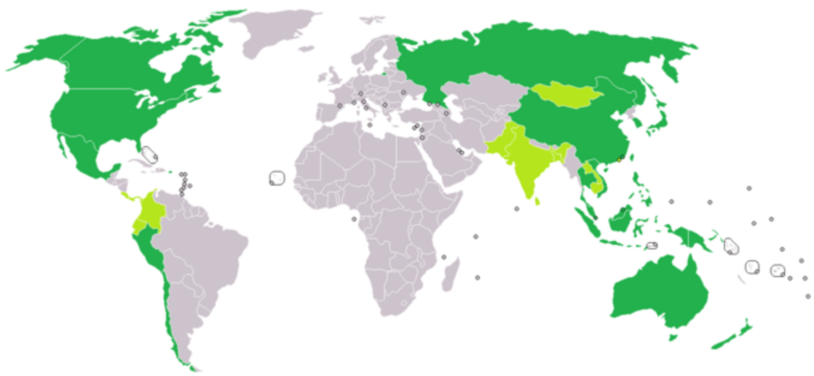
Unlike the WTO, which is an entirely global approach, most international agreements stem from geographic proximity. NAFTA is a trilateral agreement between the United States, Canada and Mexico designed to minimize any trade or investment barriers between any of these countries (primarily in the form of tariffs). Generally speaking, the United States demonstrates a trade deficit with these countries relative to goods and a surplus relative to services. The United States also demonstrates high and fast-growing foreign direct investment (FDI) in both regions.

NAFTA Participants: This map outlines each of the countries involved in the North American Free Trade Agreement, an international trade agreement focused on a geographic proximity.

There has been a great deal of controversy surrounding this trade agreement. Agriculture is not included in this agreement, and is often a tough point of discussion for the WTO as well. Mexico is also a point tension due to the fact that it is developing economically (compared to the U.S. and Canada who are considered already developed). Finally, Canadians have often objected to the NAFTA agreements due to the way in which the United States FDI employs hostile takeovers. These agreements demonstrate some of the validity behind trade protectionism and isolationism (as discussed in other atoms in this chapter).

Asia-Pacific Economic Cooperation (APEC)

The APEC forum is particularly interesting in the context of the above agreements, as it is slightly less formal than the above two (it is referred to as a 'forum'). The APEC forum is a cooperative discussion between 21 countries in the Pacific Rim region promoting free trade, with a focus on newly industrialized economies (NIE). Developing nations gaining access to capital investment and export agreements is the central outcome of APEC, driving economic growth through controlled global expansion. This region represents over half of the world's GDP and 40% of the overall world population, making this a critical region of the world economy.



APEC Participants: The Asia-Pacific Economic Cooperation (APEC) is a forum of 21 countries in the Pacific Rim region, focusing on free trade and economic cooperation.

Key Points

- Economic interdependence and globalization has resulted in a unique capitalistic system, where each country is largely dependent upon other countries for economic sustainability.
- It has been noted, somewhat intuitively and empirically, that conflict reduces trade. This highlights the risk of conflict harming an economy.
- A more specific context for trade and conflict can be the way in which trade is complicated during wartime. Indeed, trade during wartime can be a substantial threat to a nation, as economic levers such as sanctions can be utilized.
- Iran and North Korea are strong modern examples as well as the recent history of the U.S.-Iraq war. All of these economies struggle(d) against harsh economic sanctions.
- Combining these ideas, it is clear that there is substantial national security value to trade protectionism.
- Trade protectionism is national policies restricting international economic trade to alter the balance between imports and goods manufactured domestically through import quotas, tariffs, taxes, anti-dumping legislation, and other limitations.
- The primary advantage to countries with higher economic power and bigger corporations is simply economies of scale, which infant industries in developing countries often protect against.
- The United States was employing heavy tariffs to protect their fragile economic system as the economy began to achieve autonomy after British rule, which proved effective.
- From a broader and more far-reaching perspective, protectionism as a general principle has been heavily criticized (even in infant industry situations). The argument is that free markets add value on a global level, while protectionism confines economic value to the nation employing it.
- Protectionist policies are a highly charged topic in economic debates, as economies work to attain the optimal balance of free trade and trade protectionism to capture the most value.
- A recent topic is anti- dumping policies directed at international players looking to undercut domestic business through selling at dramatically reduced prices.

- Another critical risk in the global market is intellectual property (IP) protection as patents are often ignored globally, particularly by countries which lack the infrastructure to enforce IP laws.
- Another unfair competition threat is the emergence of global monopolies. Some of the larger ones attain enough global power and geographic diversification to be difficult to break up via domestic anti-trust laws.
- This argument is predicated on the simply fact that buying more domestically will drive up national production, and that this increased production will in turn result in a healthier domestic job market.
- Local governments leverage subsidies, tariffs, import quotas, and anti- dumping policies to maximize strategic capacity domestically, thus creating jobs.
- A sentiment towards protectionism has developed in the U.S. due to the jobs argument in view of an imbalanced trade ratio, where more exports (production and jobs at home) is required to sustain the ongoing consumption of imports.
- Along similar lines, it is common practice for companies to identify strategic alliances abroad and send much of the production work to these locations (outsourcing), motivating governments to bring these jobs back home.
- Local governments leverage subsidies, tariffs, import quotas, and anti-dumping policies to maximize strategic capacity domestically, thus creating jobs.
- International trade encounters a variety of obstacles which reduce trade incentives. This is usually through tariffs, quotas, taxes, and other trade restrictions.
- The WTO is the largest international trade organization, replacing the General Agreement on Tariffs and Trade (GATT) in 1995, designed to enable international trade while reducing unfair practices.
- NAFTA is a trilateral agreement between the United States, Canada and Mexico designed to minimize any trade or investment barriers between any of these countries (primarily in the form of tariffs).
- The APEC forum is a cooperative discussion between 21 countries in the Pacific Rim region promoting free trade, with a focus on newly industrialized economies (NIE).

Key Terms

- **sanction**: A penalty, or some coercive measure, intended to ensure compliance; especially one adopted by several nations, or by an international body.
- **Self-sufficiency**: Able to provide for oneself independently of others.
- **Nascent**: Emerging; just coming into existence.
- **Dumping**: Selling goods at less than their normal price, especially in the export market as a means of securing a monopoly.
- **Subsidies**: Financial support or assistance, such as a grant.
- **Reverse engineering**: The process of analyzing the construction and operation of a product in order to manufacture a similar one.
- **Import Quota**: A restriction on the import of something to a specific quantity.
- **Trade Balance**: The difference between the monetary value of exports and imports in an economy over a certain period of time.
- **Foreign direct investment**: Investment into production or business in a country by an individual or company of another country.
- **tariff**: A system of government-imposed duties levied on imported or exported goods; a list of such duties, or the duties themselves.

- International trade and state security. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/International_trade_and_state_security**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Self-sufficiency. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Self-sufficiency**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- sanction. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/sanction**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...-mortality.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Isolationism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Isolationism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Free trade debate. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Free_trade_debate**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Protectionism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Protectionism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Infant industry argument. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Infant_industry_argument**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Nascent. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Nascent**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Dumping. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Dumping**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...-mortality.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...al)_in_US$.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...s_of_scale.PNG**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Isolationism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Isolationism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Protectionism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Protectionism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Free trade debate. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Free_trade_debate**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Commercial law. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Commercial_law**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Competition law. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Competition_law**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Unfair competition. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Unfair_competition**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Subsidies. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Subsidies**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Reverse engineering. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Reverse+engineering**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Dumping. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Dumping**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...-mortality.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/0/08/Map_of_countries_by_GDP_(nominal)_in_US$.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/3/38/Economies_of_scale.PNG**. **License**: ***CC BY-SA: Attribution-ShareAlike***

- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...-surpluses.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Protectionism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Protectionism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Isolationism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Isolationism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Free trade debate. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Free_trade_debate**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Import Quota. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Import+Quota**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Trade Balance. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/Trade+Balance**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...-mortality.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/0/08/Map_of_countries_by_GDP_(nominal)_in_US$.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/3/38/Economies_of_scale.PNG**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/e/ef/Monopoly-surpluses.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/3/32/Cumulative_Current_Account_Balance.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Protectionism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Protectionism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Isolationism. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Isolationism**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- World Trade Organization. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/World_Trade_Organization**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- North American Free Trade Agreement. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/North_American_Free_Trade_Agreement**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- List of free trade agreements. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/List_of_free_trade_agreements**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Asia-Pacific Economic Cooperation. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Asia-Pacific_Economic_Cooperation**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- Foreign direct investment. **Provided by**: Wikipedia. **Located at**: **en.Wikipedia.org/wiki/Foreign%20direct%20investment**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- tariff. **Provided by**: Wiktionary. **Located at**: **en.wiktionary.org/wiki/tariff**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...-mortality.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/0/08/Map_of_countries_by_GDP_(nominal)_in_US$.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/3/38/Economies_of_scale.PNG**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/e/ef/Monopoly-surpluses.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/3/32/Cumulative_Current_Account_Balance.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***

- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/a/a6/North_American_Agreement_(orthographic_projection).svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/wikipedi...nlargement.png**. **License**: ***CC BY-SA: Attribution-ShareAlike***
- **Provided by**: Wikimedia. **Located at**: **upload.wikimedia.org/Wikipedia/commons/9/93/World_Trade_Organization_negotiations.svg**. **License**: ***CC BY-SA: Attribution-ShareAlike***

Index

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Book: Principles of Microeconomics (A. Casolari)

Webpages: 83

All licenses found:

- Undeclared: 83.1% (69 pages)
- CC BY 4.0: 16.9% (14 pages)

By Page

- Book: Principles of Microeconomics (A. Casolari) - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Welcome to Economics! - *Undeclared*
    - 1.1: What Economics Is and Why It's Important - *Undeclared*
    - 1.2: How Economists Use Theories and Models to Understand Economic Issues - *Undeclared*
    - 1.3: How Economies Can Be Organized- An Overview of Economic Systems - *Undeclared*
  - 2: Choice in a World of Scarcity - *CC BY 4.0*
    - 2.1: Welcome to Economics! - *CC BY 4.0*
    - 2.2: Prelude to Choice in a World of Scarcity - *CC BY 4.0*
    - 2.3: How Individuals Make Choices Based on Their Budget Constraint - *CC BY 4.0*
    - 2.4: The Production Possibilities Frontier and Social Choices - *CC BY 4.0*
    - 2.5: Confronting Objections to the Economic Approach - *CC BY 4.0*
    - 2.E: Choice in a World of Scarcity (Exercises) - *CC BY 4.0*
  - 3: Demand and Supply - *Undeclared*
    - 3.1: Prelude to Demand and Supply - *Undeclared*
    - 3.2: Demand, Supply, and Equilibrium in Markets for Goods and Services - *Undeclared*
    - 3.3: Shifts in Demand and Supply for Goods and Services - *Undeclared*
    - 3.4: Changes in Equilibrium Price and Quantity- The Four-Step Process - *Undeclared*
    - 3.5: Price Ceilings and Price Floors - *Undeclared*
    - 3.6: Demand, Supply, and Efficiency - *Undeclared*
    - 3.7: Labor and Financial Markets - *Undeclared*
    - 3.8: Demand and Supply at Work in Labor Markets - *Undeclared*
    - 3.9: Demand and Supply in Financial Markets - *Undeclared*
    - 3.E: Demand and Supply (Exercises) - *Undeclared*
  - 4: Elasticity - *CC BY 4.0*
    - 4.1: Prelude to Elasticity - *CC BY 4.0*
    - 4.2: Price Elasticity of Demand and Price Elasticity of Supply - *CC BY 4.0*
    - 4.3: Polar Cases of Elasticity and Constant Elasticity - *CC BY 4.0*
    - 4.4: Elasticity and Pricing - *CC BY 4.0*
    - 4.5: Elasticity in Areas Other Than Price - *CC BY 4.0*
    - 4.E: Elasticity (Exercises) - *CC BY 4.0*
  - 5: Consumer Choices - *Undeclared*
    - 5.1: Prelude to Consumer Choices - *Undeclared*
    - 5.2: Consumption Choices - *Undeclared*
    - 5.3: How Changes in Income and Prices Affect Consumption Choices - *Undeclared*
    - 5.4: Labor-Leisure Choices - *Undeclared*
    - 5.5: Intertemporal Choices in Financial Capital Markets - *Undeclared*
    - 5.E: Consumer Choices (Exercises) - *Undeclared*
  - 6: Production and Cost Structure of the Firm - *Undeclared*
    - 6.1: Prelude to Production and Cost Structure of the Firm - *Undeclared*
    - 6.2: Explicit and Implicit Costs, and Accounting and Economic Profit - *Undeclared*
    - 6.3: The Structure of Costs in the Short Run - *Undeclared*
    - 6.4: The Structure of Costs in the Long Run - *Undeclared*
    - 6.E: Production and Cost Structure of the Firm (Exercises) - *Undeclared*
  - 7: Perfect Competition - *Undeclared*
    - 7.1: Prelude to Perfect Competition - *Undeclared*