

Chapter 1: Five Foundations of Economics

August 20, 2023

What is Economics?

Economists study how decisions are made. Examples of economic decision include whether you should buy or lease a car, sublet your apartment, or buy that Gibson guitar you've been eyeing. And just as individuals must choose what to buy within their limits of their income, society as a whole must determine what to produce from its limited set of resources.

Our wants and needs are practically unlimited, but resources available to satisfy these wants and needs are always limited. The term used to describe the limited nature of society's resources is **scarcity**. **Economics** is the study of how individuals and societies allocate their limited resources to satisfy their practically unlimited wants.

Microeconomics and Macroeconomics

Microeconomics (micro) is the study of individual units that make up the economy, such as households and businesses. **Macroeconomics** (macro) is the study of the overall aspects and workings of an economy, such as inflation (an overall increase in prices), growth, employment, interest rates, and the productivity of the economy. Consider a worker who gets laid off and becomes unemployed; is this an issue that would be addressed in microeconomics or macroeconomics? The question seems to fit parts of both definitions. The worker is an individual, which is micro, but employment is one of the broad areas of concern for the economy as a whole, which is macro. However, because one worker is laid off, this is a micro issue. When many workers are laid off, the result is a higher unemployment rate across the entire economy. Then the issue is broad enough to be studied by macroeconomists. However, macroeconomics is more than just an aggregation.

Macroeconomists examine, among other things, government policies regarding the federal budget and money supply, the reasons for inflation and unemployment, economic growth, international trade, and government borrowing – topics that are too complex to be understood using only microeconomic analysis.

What are Five Foundations of Economics?

The study of economics can be complicated, but we can make it very accessible by breaking it down into a set of component parts. The 5 foundations of economics are:

- **Marginal thinking**
- **Opportunity cost**
- **Trade-offs**
- **Incentives**
- **Value created from trade**
- **Economics is everywhere**

Chapter 1: Five Foundations of Economics

August 20, 2023

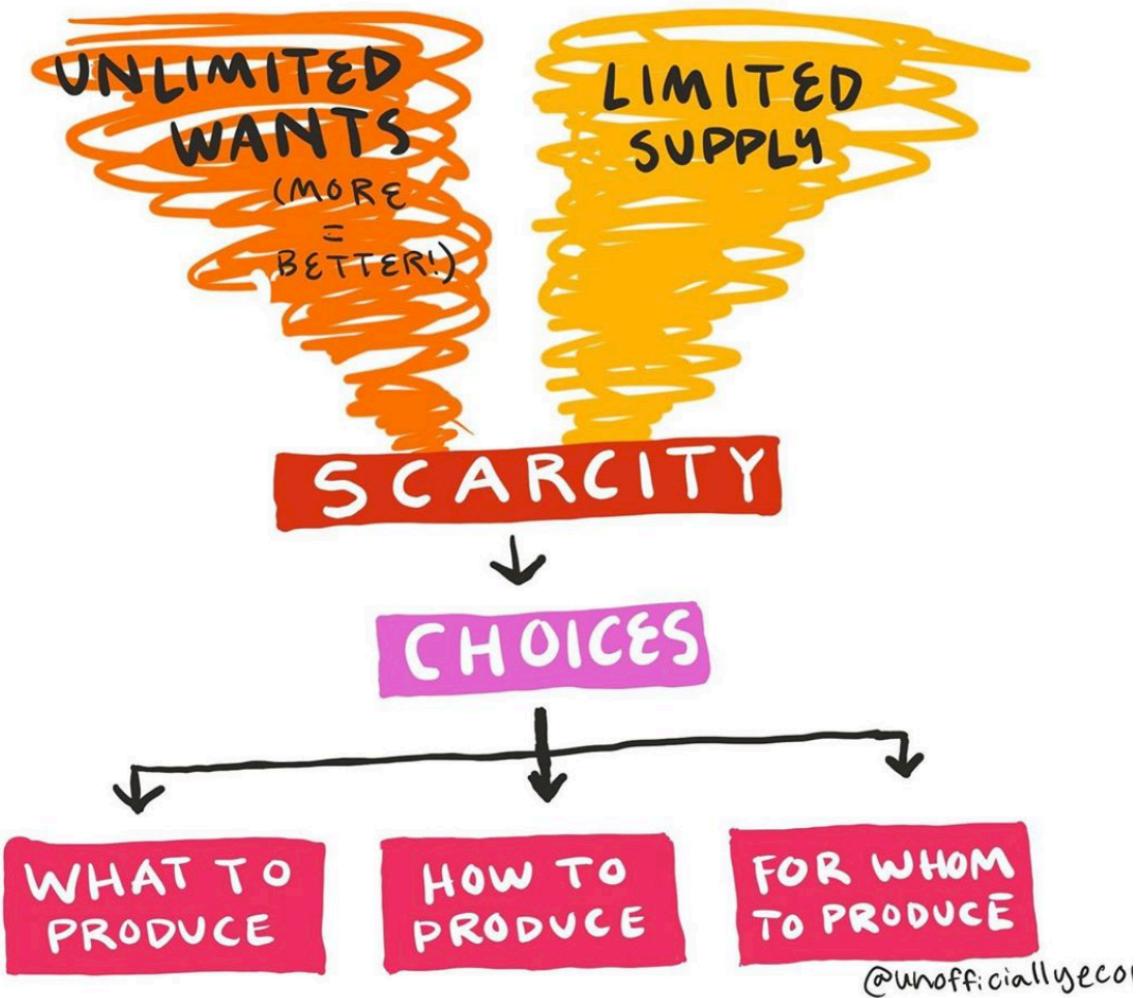
Scarcity

How practically all Americans would like to be everything from “safe” to “romantically fulfilled”, the video segues to interviews with individuals whose “all I want” lists range from endearing (a big, happy dog) to the quirky (a new Wes Anderson movie), the unrealistic (quick and easy weight loss), and the impossible (“I don’t want to die”). We live in a world of scarcity; but that alone doesn’t explain why we’re unable to meet everyone’s wants. Couldn’t we just redistribute goods and services more evenly, to satisfy everyone? No chance, because our wants exceed our needs, and when all our wants are met, we come up with new ones. Many people spend their lives trying to “keep up with the Joneses”. This isn’t all bad, because competitive drive causes people to work longer and harder, which makes the economy more productive. At the same time, when we purchase one good, we have less to spend on other goods we also desire, and therefore we face trade-offs and opportunity costs.

Scarcity exists when the marginal cost of obtaining something is greater than zero. Air and gravity are the only things or items that aren't scarce. How do people allocate their limited resources to satisfy their unlimited wants? Economics can help us overcome scarcity.

Chapter 1: Five Foundations of Economics

August 20, 2023



Incentives

Incentives – factors that motivate you to act or exert effort. For example, your choice to study for an exam you have tomorrow, instead of spending the evening with your friends, is based on your belief that doing well on the exam will provide a greater benefit. You have an incentive to study because you know that an A in the course will raise your grade-point average and make you a more attractive candidate on the job market when you are finished with school. We can further divide incentives into two paired categories: positive and negative, and direct and indirect.

Incentives (both positive and negative) influence our actions in predictable ways. Incentives influence our actions in predictable ways, but also lead to unintended consequences; should you trust your real estate agent to get you the best deal? Let's do the math

Positive and Negative Incentives

Positive Incentives encourage action by offering rewards or payments. For example, end-of-year bonuses motivate employees to work hard throughout the year, higher oil prices cause suppliers to

Chapter 1: Five Foundations of Economics

August 20, 2023

extract more oil, and tax rebates encourage citizens to spend more money. **Negative incentives** discourage action by providing undesirable consequences or punishments. For instance, the fear of receiving a speeding ticket keeps motorists from driving too fast, higher oil prices might spur some consumers to use less oil, and the dread of a trip to the dentist motivates people to brush their teeth regularly. In each case, we see that incentives spur individuals to action.

Direct and Indirect Incentives

Incentives can also be direct or indirect. For instance, if one gas station lowers its prices, it most likely will get business from customers who would not usually stop there. This is a **direct incentive**. Lower gasoline prices also work as an **indirect incentive**, because lower prices might encourage consumers to use more gas.

Direct incentives are easy to recognize. “Cut my grass and I’ll pay you \$30” is an example of a direct incentive. Indirect incentives are more difficult to recognize. But learning to recognize them is one of the keys to mastering economics. For instance, consider the indirect incentives at work in some government assistance programs. In other words, a society that has a direct incentive to alleviate suffering caused by poverty. But how does a society provide this safety net without taking away the incentive to work? If the amount a person receives is higher than the amount that person can hope to make from a job, there is far less incentive to go back to work. In fact, there’s an incentive not to. This situation creates an **unintended consequences**, where assistance originally meant as a safety net could start to be seen as a permanent source of income.

Auto insurance scams

Policymakers have the tough task of deciding how to balance such conflicting incentives. For this reason, many government programs specify limits on the amount of time people can receive benefits. Ideally, this limit allows the welfare programs to continue meeting people’s basic needs while creating incentives that encourage recipients to search for a job and acquire skills that will help them get a job.

Incentives and Innovation

Incentives also play a vital role in innovation, the engine of economic growth. An excellent example is Steve Jobs; he and the company he founded, Apple, held over 300 patents at the time of his death in 2011.

In the United States, the patent system and copyright laws guarantee inventors a specific period of time in which they have the exclusive right to sell their work. This system encourages innovation by creating a powerful financial reward for creativity. Without patents and copyright laws, inventors would bear all the costs, and almost none of the rewards, for their efforts. Why would firms invest in research and development or artists create new music if other could immediately copy and sell their work? To reward the perspiration and inspiration required for innovation, society allows patents and copyrights to create the right incentives for economic growth. In recent years, new forms of technology have made the illegal sharing of copyrighted material quite easy. As a result, illegal downloads of books, music, and movies are widespread. When writers, musicians, actors, and studios cannot effectively protect what they have created, they earn less; so illegal downloads reduce the incentive to produce new content.

Chapter 1: Five Foundations of Economics

August 20, 2023

Maintaining the right rewards, or incentives, for hard work and innovation is essential for making sure that inventors and other creative people are compensated for their creativity and vision. Some see services like Spotify, Apple Music, and SoundCloud as the answer. While streaming services are now very successful, the amount of artists receive is still far lower than it used to be.

Incentives are everywhere

One very powerful incentive is saving time. You can test out your time-savings skills when you walk across campus to a class. An app will give you a detailed route and an estimated time of arrival, but your app won't know the local shortcuts. Sometimes the shortcuts everyone takes are through buildings or along paths...the paths worn into green by students' feet will show you how to get across campus as quickly as possible.

Understanding incentives, from positive to negative and direct to indirect, is the key to understanding economics. If you remember only one concept from this course, it should be that incentives matter.

Trade-offs

In a world of scarcity, each and every decision incurs a cost. Even time is a scarce resource; after all, there are only 24 hours in a day. So deciding to play *Animal Crossing* now means you won't be able to read one of the Harry Potter books until later. More generally, doing one thing often means you will not have the time, resources, or energy to do something else. Similarly, paying for a college education can require spending tens of thousands of dollars that might be used elsewhere instead.

People who don't understand economics sometimes ignore the trade-offs that are natural in a world of scarcity. They unconsciously assume that we can (as individuals or a group) have more of everything we want. But in fact, decision-making generally involves trade-offs. For a recent example, take the Coronavirus Aid, Relief, and Economic Security (CARES) Act, which was passed by Congress with overwhelming bipartisan support and signed on March 27, 2020. This \$2.2 trillion relief package provided financial assistance for American workers, families, small businesses, and preserved jobs in American industries. But what could all that money have bought instead, if we'd spent it differently? Well, we could have created 1.5 million state-of-the-art hospital beds, or we could have built 440 brand-new NFL stadiums, or given every citizen a check for \$6,728 or created 550,000 one-megawatt solar farms, or hired 700,000 K-12 teachers for 50 years each.

Ultimately, thinking about trade-offs means that we will make more informed decisions about how to use our scarce resources.

Trade-offs consist of the entire set of things we might have done; trade-offs are part of life.

Opportunity Cost

The existence of trade-offs require making hard decisions. Trade-offs are about having something up, while opportunity cost quantifies "what" or "how much" is being given up. Choosing one thing means giving up something. No matter what choice you make, there's an opportunity cost, or next-best alternative, that must be sacrificed; ***opportunity cost*** is the highest-valued alternative that must be sacrificed to get something else.

Chapter 1: Five Foundations of Economics

August 20, 2023

Every time we make a choice, we experience an opportunity cost. The key to making the best possible decision is to minimize your opportunity cost by selecting the option that gives you the largest benefit. It takes deliberate effort to see the world through the opportunity cost prism, but it's a worthwhile practice because it will help you make better decisions. For example, imagine you are a small business owner. Your financial officer informs you that you have had a successful year and made a sizeable profit; so everything is good, right? Not so fast, an economist will tell you to ask yourself, "Could you have made **more** profit doing something else?" Good economic thinkers ask this question all the time. "Could I be using my time, talents, or energy on another activity that would be even more profitable for me?"

Profits on an official income statement are only part of the story, because they only measure how well a business does relative to the bottom line. Accountants cannot measure what might have been better, so when economists talk about opportunity cost, they are assessing whether the alternatives are better than what you are currently doing, which considers a larger set of possible outcomes. For example, in the chapter opener, we mused about having a money tree. It turns out that money trees do really exist – in the virtual world. In *Animal Crossing*, Nintendo's breakout title of the 2020 pandemic, the player must bury a bag of Bells (the game's currency) using a golden shovel. After the tree has fully grown, it will bloom with three bags of Bells at amounts dependent on how much was initially buried. In this simple example, there is still an opportunity cost – what might have been done with the bag of Bells instead of growing a Bell tree! Planting a Bell tree today increases your future income, but you forgo the opportunity to spend Bells now.

Opportunity cost is the highest-valued alternative that must given up to engage in an activity.

- ***Value of the next best alternative***

TPS: How many hours would you wait in line on Black Friday to save \$300 on a TV

Marginal Thinking

The process of systematically evaluating a course of action is called economic thinking. **Economic thinking** involves a purposeful evaluation of the available opportunities to make the best decisions possible. In this context, economic thinkers use a process called **marginal analysis** to break down decisions into smaller parts. Often, the choice is not between doing and not doing something, but between doing more or less of something. For instance, if you take on a part-time job while in school, you probably wrestle with the question of how many hours to work. If you work a little more, you can earn additional income. If you work a little less, you have more time to study. Working more has a tangible benefit (more money) and a tangible cost (lower grades). All of this should sound familiar from our earlier discussions about trade-offs; the work-study trade-off affects how much money you have and what kind of grades you earn.

An economist would say that your decision – weighing how much money you want against the grades you want – is a decision at the *margin*. In economics, **marginal thinking** requires decision-makers

Chapter 1: Five Foundations of Economics

August 20, 2023

to evaluate whether the benefit of one more unit of something is greater than its cost. Understanding how to analyze decisions at the margin is essential to thinking like a good economist.

Is the additional benefit greater than the additional cost?

What is the optimal amount of cleanliness?

Trade

Imagine trying to find food in a world without grocery stores. The task of getting what you need to eat each day would require visiting many separate locations. Many centuries ago, this need to bring buyers and sellers together was met by weekly markets, or bazaars, in central locations like town squares. **Markets** bring buyers and sellers together to exchange goods and services. As commerce spread throughout the ancient world, trade routes developed. Markets grew from infrequent gatherings, where exchange involved trading goods and services, into more sophisticated systems that use cash, credit, and other financial instruments. Today, when we think of markets, we often think of eBay or Craigslist. For instance, if you want to find a rare Hot Wheels Black Panther Movie Die-Cast Vehicle, an excellent place to look is eBay, which allows users to search for just about any product, bid on it, and then have it sent directly to their home.

Trade creates value and depends on specialization and comparative advantage.

The Circular Flow

When we consider all the trade that occurs in an economy, it is helpful to use a circular flow diagram. This shows how goods, services, and resources flow through the economy via commerce between households and firms. Households are made up of consumers, as we usually picture them. Firms are businesses. Households desire the goods and services produced by firms, but to produce those goods and services, firms require the resources owned by households. The circular flow diagram illustrates the movement of goods, services, and resources that results when firms and households do business with each other.

Consider a simple example, let's say you spend \$1,000 on a new Dell computer. You trade for your computer in a product market, and Dell gets the \$1,000: this takes place in the top half of the circular flow diagram. Then Dell uses the \$1,000 to pay its workers' wages and other suppliers for the use of resources. This happens in the bottom half of the diagram. In the end, the funds make the complete circuit back to households.

Trade creates value

Trade is the voluntary exchange of goods and services between two or more parties. Voluntary trade among rational individuals creates value for everyone involved.

For example, imagine you are on your way home from class and you want to pick up a gallon of milk. You know that milk will be more expensive at a convenience store than at the grocery store 5 miles away, but you are in a hurry to study for your economics exam and are willing to pay up to \$5 for the convenience of getting the milk quickly. At the store, you find that the price is \$4 and you happily purchase the milk. This ability to buy for less than the price you are willing to pay provides a positive incentive to make the purchase, but what about the seller? If the store owner paid \$3 to buy milk from a

Chapter 1: Five Foundations of Economics

August 20, 2023

supplier, and you are willing to pay the \$4 price she has set in order to make a profit, the store owner has an incentive to sell. The simple voluntary transaction has made both of you better off.

Comparative advantage refers to the situation in which an individual, business, or country can produce at a lower opportunity cost than a competitor can. Comparative advantage harnesses the power of specialization, a topic discussed in further detail in Chapter 2. For instance, Starbucks specializes in making coffee, Honda in making automobiles. You would not want to get your morning cup of joe at Honda anymore than you would want to buy a car from Starbucks! On a broader scale, specialization and trading of services exist at the international level: some countries have highly developed workforces capable of managing and solving complex processes. Other countries have large pools of relatively unskilled labor. As a result, businesses that need skilled labor gravitate to countries where they can easily find the workers they need. Likewise, firms with production processes that rely on unskilled labor look for employees in less developed countries, where workers are paid less. By harnessing the power of increased specialization, global companies and economies create value through increased production and growth. However, globalized trade is not without controversy; when goods and jobs are free to move across borders, not everyone benefits equally, nor should we expect this outcome, but outsourcing is an important component of economic growth in the long run.

Conclusion

Economists ask and answer big questions about life. This is what makes the study of economics so fascinating. Understanding how an entire economy functions may seem like a daunting task, but it is not nearly as difficult as it sounds. Once you have learned the fundamentals of economics, you can use them to analyze almost any problem. In the next chapter, we use the ideas developed in this chapter to explore trade in greater depth.

Value of time

- The value of time determines relative prices of goods and services, investments, productivity, and economic growth, and measures of income inequality.
- Current research suggests society is

CHAPTER 2: MODEL BUILDING AND GAINS FROM TRADE

- Models are simplistic, but they emphasize key concepts
- If we change only one variable at a time and hold other things constant, this is known as *ceteris paribus*
- Judge a model by how well it works in real life
 - Endogenous model – variables inside the model
 - Exogenous model – variables outside the model
- Watch out for faulty assumptions.

When people trade, both sides “normally” win

It's only common sense that trade benefits both parties. After all, if it's voluntary, both sides must be getting something out of it; we can quantify the extent to which the trade makes each side better off. We do this by how much more productive a person can become through trade, given that trade allows people to specialize in what they're good at.

Example #1:

Consider the interaction between a contractor and an architect.

They each have a vital role to play in the building process; the architect designs the plans to the buyer's specifications and the contractor is an expert at bringing the architect's design to fruition by organizing the equipment, supplies, and labor to complete the project on time. The architect is the creative genius and the contractor is a genius at managing the construction workflow; the architect understands how to design plans that pass engineering tests and meet building codes. The contractor understands the supply chain; by specializing, each become more productive and gets their part of the project done faster, trading their expertise and time for monetary payment.

To help understand how trade works, we will develop our first economic model, the production possibilities frontier, so we can explore the more nuanced reasons why trade creates value.

Core Questions:

- *How do economists study the economy?*
 - Economists design hypotheses (proposed explanations) and then test them by collecting real data. The economist's laboratory is the world around us.
 - A good model should be simple, flexible, and useful for making accurate predictions. A model is both more realistic and harder to understand when it involves many variables. To keep models simple, economists often use the concept of *ceteris paribus*, or “all else equal.” Maintaining a positive (as opposed to normative) framework is crucial for economic analysis because it allows decision-makers to observe the facts objectively.
- *What is a production possibilities frontier?*
 - A production possibilities frontier (PPF) is a model that illustrates the combinations of outputs a society can produce if all of its resources are being used efficiently. An outcome is considered efficient when resources are fully utilized, and potential output is maximized. Economists use the PPF to illustrate trade-offs and explain opportunity costs and the role of additional resources and technology in creating economic growth.

- *What are the benefits of specialization and trade?*
 - Society is better off if individuals and firms specialize and trade on the basis of the principle of comparative advantage.
 - Parties that are better at producing goods and services than all their potential trading partner (and thus hold an absolute advantage) still benefit from trade. Trade allows them to specialize and trade what they produce for other goods and services they are relatively less skilled at making.
 - As long as the terms of trade falls between the opportunity costs of both trading partners, the trade benefits both sides.
- *What is the trade-off between more now and having more later?*
 - All societies face a crucial trade-off between the consumption in the short run and economic growth in the long run. Investment in capital goods today help spur economic growth in the future. However, because capital goods are not consumed in the short run, society must be willing to sacrifice how well it lives today in order to have more later.

How do economists study the economy?

Economics is a social science that uses the scientific method to develop **economic models**. To create these models, economists make many assumptions to simplify reality. These models help economists understand the key relationships that drive economic decisions.

The Scientific Method in Economics

Economists use the scientific method to answer questions about observable phenomena and to explain how the world works. The scientific method consist of four steps:

1. Researchers observe a phenomenon that interests them.
2. Based on these observations, researchers develop a *hypothesis*, which is a proposed explanation for the phenomenon.
3. Then they construct a model to test the hypothesis
4. Finally, they look for opportunities to test how well the model (which is based on the hypothesis) works. After collecting data, they use statistical methods to verify, revise, or refute the hypothesis.

The economist's laboratory is the world around us and it ranges from the economy as a whole to the decisions made by firms and individuals. As a result, economists cannot always design experiments to test their hypotheses; often they must gather historical data or wait for real-world events to take place – such as the Great Recession (economic downturn) of 2007 – 2009 or the Coronavirus pandemic of 2020 – to better understand the economy. When real-world events meet the criteria of an experiment designed to test a hypothesis, we have what's called a **natural experiment**.

Positive and Normative Analysis

To be **objective** as possible, economists deploy positive analysis. A **positive statement (analysis)** can bet tested and validated; each positive statement can be though of as a description of “what

is." For instance, the statement "*The unemployment rate is declining*" is a positive statement, because it can be tested by gathering data and shown to be true or false.

In contrast, a statement about "what ought to be" is a **normative statement (analysis)** that cannot be empirically tested or validated. For instance, the statement "*An unemployed worker should receive financial assistance to help make ends meet*" is a matter of opinion. One can reasonably argue that financial assistance to the unemployed is a socially beneficial "anti-poverty" measure. However, some argue that financial unemployment assistance provides the wrong incentives; if the assistance is enough to meet basic needs, workers may end up spending more time unemployed than they otherwise would.

Policy decisions eventually have to be made by somebody, one way or the other. Just remember that when we talk about what is "best" or what "should happen," we are in the realm of differing viewpoints, based on values, beliefs, and opinions.

Economic Models

Thinking like an economist means learning how to analyze complex issues and problems. Many economic topics, such as international trade, Social Security, job loss, and inflation, are complicated; to analyze these phenomena and to determine the effect of various government policy options related to them...economists use models which are simplified versions of reality. Models help us analyze the components of the economy.

Ceteris Paribus

Changing one variable while holding everything else constant involves a concept known as "ceteris paribus", from the Latin meaning "other things being equal" or "all else equal". Economists start with a simplified version of reality; they build models, change one variable at a time, and ask whether the change in the variable has a positive or negative impact on performance.

$$W = f \left(\begin{array}{l} \text{Education (positive)} \\ \text{Age (positive)} \\ \text{Experience (positive)} \\ \text{Skills (positive)} \\ \text{Pleasant conditions (negative)} \\ \text{Female (negative -discrimination)} \end{array} \right)$$

Endogenous vs. Exogenous Factors

Models must account for factors we can control (endogenous) and factors we can't (exogenous). Factors that are accounted for inside the model are **endogenous factors**; factors beyond our control – outside the model – are **exogenous factors**.

We need to be mindful of three factors:

1. What we include in the model
2. The assumptions we make when choosing what to include in the model.
3. The outside conditions that can affect the model's performance.

CHAPTER 2: MODEL BUILDING AND GAINS FROM TRADE

If we add more exogenous variables, or factors we cannot control – for example, wind and rain – to test our model's performance, the test becomes more realistic, but at the same time, the outcome becomes less predictable.

The Importance of Assumptions

When we build a model, we need to make choices about which variables to include; ideally, we would like to include all the important variables inside the model and exclude all the variables that can be safely ignored. Then we have made reasonable simplifying assumptions; excluding the wrong variables, on the other hand, can lead to spectacular failures, so can making false assumptions.

An excellent example is the financial crisis and Great Recession that began in December 2007:

In the years leading up to the crisis, banks sold and repackaged mortgage-backed investments under the faulty assumptions that real estate prices will rise. This assumption seemed perfectly reasonable in a world where real estate prices were rising annually; unfortunately, the assumption turned out to be false. From December 2007 into 2008, real estate prices fell dramatically, because of one faulty assumption, the entire financial market teetered on the edge of collapse. Shocks, be they financial or pandemic-related, often defy the usual logic; contrary to conventional wisdom, the real estate market did not crater during the pandemic, but instead it skyrocketed in 2020 – 2021 as the pandemic caused people to hunker down and invest some more in their homes.

Through this textbook we will assume that firms and households are rational benefit-maximizers who both respond to incentives predictably and thoughtfully consider the costs and benefits of their actions. Rationality is a cornerstone of most economic theory (it's a simplifying assumption).

What is a Production Possibilities Frontier?

The PPF shows the maximum amount of any two products that can be produced from a fixed number of resources. The PPF is non-linear, this is known as the law of increasing opportunity cost.

Graphs are a key tool in economics because they display the relationship between two variables. Your ability to read a graph and understand the model it represents is crucial to learning economics.

A **production possibilities frontier (PPF)** is a model that illustrates the combination of outputs a society can produce if all of its resources are being used efficiently. An outcome is considered **efficient** when resources are fully utilized, and potential output is maximized. To preserve *ceteris paribus*, we assume that the technology available for production and the quantity of resources remain fixed, or constant. These assumptions allow us to model trade-offs more clearly.

Whenever society is producing on the production possibilities frontier, the only way to get more of one good is to accept less of another. Because an economy operating at a point on the frontier will be efficient, every point on the frontier represents full-capacity output; but a society may favor one point over another because it prefers that combination of goods.

The Production Possibilities Frontier and Opportunity Cost

The trade-offs that occur along the production possibilities frontier represent the opportunity cost of producing one good instead of the other; **an opportunity cost is the highest-valued alternative given up to pursue another course of action**. A bowed-out production possibilities frontier reflects the increasing opportunity cost of production. **Figure 2.2** illustrates the **law of increasing opportunity cost**, which states that the opportunity cost of producing a good rises as a society produce more of it. Changes in relative cost mean that a society faces a significant trade-off if it tries to produce an extremely large amount of a single good.

The Production Possibilities Frontier and Economic Growth

So far, we have modeled the production possibilities frontier based on the resources available to society at a particular moment in time. However, most societies hope to create economic growth. **Economic growth is the process that enables a society to produce more output in the future**. We can use the production possibilities frontier to explore economic growth; for example, we can ask what would happen to the PPF if our two-good society develops a new technology that increases productivity.

What are the Benefits of Specialization and Trade

Trade creates value and depends on specialization and comparative advantage

We have seen that improving technology and adding resources make an economy more productive. A third way to create gains for society is through specialization and trade. **Specialization is the limiting of one's work to a particular area (industry)**. Determining what to specialize in is an important part of the process. Every worker, business, or country is relatively good at producing certain products or services.

In the next section, we explore why specializing and exchanging your skilled expertise with others makes gains from trade possible.

Gains from Trade

Absolute Advantage refers to one producer's ability to make more than the another producer with the same quantity of resources.

Comparative Advantage

We have seen that specialization enables workers to enjoy gains from trade. The concept of opportunity cost provides us with a second way of validating the principle that trade creates value. Recall that opportunity cost is the highest-valued alternative that is sacrificed to pursue something else. Applying the concept of opportunity cost helps us see why specialization enables people to produce more.

What is the Trade-Off between having more now and having more later?

So far, we have examined short-run trade-offs; but both individuals and society as a whole must weigh the benefits available today (**the short run**) with those available tomorrow (**the long run**). In the **short run**, we make decisions that reflect our immediate or short-term wants, needs, or limitations; consumers can partially adjust their behavior. In the **long run**, we make decisions that reflect our wants, needs, and

limitations over a much longer time horizon. In the long run, consumers have time to fully adjust to the market conditions.

Many of life's important decisions are about the long run; we must decide where to live, whether and whom to marry, whether and where to go to college, and what type of career to pursue.

Consumer Goods, Capital Goods, and Investment

We have seen that the trade-off between the present and the future is evident in the tension between what we consume now and what we plan to consume later. Any good that is produced for present consumption is a **consumer good** (are produced for personal satisfaction); these goods help to satisfy our needs or wants now. Food, entertainment, and clothing are all examples of consumer goods. **Capital goods** (are used to produce other goods) help in the production of other valuable goods and services in the future. Capital goods are everywhere; roads, factories, trucks, and computers are all capital goods.

Education is a form of capital. The time you spend earning a college degree makes you more attractive to future employers. When you decide to go to college instead of working, you are investing your human capital. **Investment** is the process of using resources to create or buy new capital; because we live in a world with scarce resources, every investment in capital goods has an opportunity cost of forgone consumer goods.

The decision between whether to consume or invest has a significant impact on economic growth in the future, or long run. **What happens when society chooses to produce many consumer goods than capital goods?** When relatively few resources are invested in producing capital goods in the short run, very little capital is created, because new capital is a necessary ingredient for economic growth in the future, the long-run production possibilities curve expand only a small amount. **What happens when society chooses to plan for the future by producing more capital goods than consumer goods in the short run?** With investment in new capital, the long-run production possibilities curve expand outward much more; all societies face the trade-off between spending today and investing tomorrow.

Conclusion

The simple, yet powerful idea that trade creates value has far-reaching consequences for how we should organize our society. Since we all win when voluntary trade takes place, creating opportunities for more trades take place between consumers and producers and across countries enriches all of our lives.

We have developed our first model, the production possibilities frontier; this model illustrates the benefits of trade and also enables us to describe ways to grow the economy. Trade and growth rest on a more fundamental idea – specialization. When producers specialize, they focus their efforts on those goods and services for which they have the lowest opportunity cost, and they trade with others who are good at making something else. To have something valuable to trade, each producer, in effect, must find its comparative advantage. As a result, trade creates value and contributes to an improved standard of living in society.

The Market at Work: Supply and Demand

ECO 304K: Introduction to Microeconomics

Buyers and Sellers together determine the price of the good

Sometimes buyers set the price — through live auctions, on eBay, or at shopgoodwill.com; other times, sellers set the price and how much inventory remains. Buyers and sellers influence both prices and quantities traded, so that these end up being determined by how buyers' and sellers' price-versus-quantity calculations interact.

This chapter describes how markets work and discusses the nature of the competition; to shed light on the process, we introduce the formal model of demand and supply. We begin by looking at demand and supply, then we combine them to see how they interact to establish the market price and determine how much is produced and sold.

Core questions

- **What are the fundamentals of markets?**
 - A market consists of a group of buyers and sellers for a particular product or service
 - A competitive market exists when there are so many buyers and sellers that each has only a small (negligible) impact on the market price and output.
 - Not all market are competitive. When firms have market power, markets are imperfect.
- **What determines demand?**
 - The law of demand states that, all other things being equal, quantity demanded falls when the price rises, and rises when the price falls.
 - The demand curve is downward sloping
 - A price change causes a movement along the demand curve, not a shift of the curve
 - Changes in something other than price (including changes in income, the price of related goods, changes in taste and preferences, price expectations, the number of buyers, and taxes) shift the demand curve.
- **What determines supply?**

- The law of supply states that, all other things being equal, the quantity supplied of good rises, and falls when the price of the good falls.
- The supply curve is upward sloping
- A price change causes movement along the supply curve, not a shift of the curve
- Changes in something other than price (costs of inputs, changes in technology or the production process, taxes and subsidies, the number of firms in the industry, and price expectations) shift the original supply curve.
- **How do supply and demand interact to create equilibrium?**
 - Supply and demand work together in a market-clearing process that leads to equilibrium, the balancing point between two forces. The market-clearing price and output are determined at the equilibrium point.
 - When the price is above the equilibrium point, a surplus exists, and inventories build up. Suppliers lower their price in an effort to sell the unwanted goods. The process continues until the equilibrium price is reached
 - When price is below the equilibrium point, a shortage exists, and inventories are depleted. Suppliers raise the price until the equilibrium point is reached.

What are the Fundamentals of Markets?

Markets bring trading partners together to create order out of chaos. Companies supply goods and services, and customers want to obtain those goods and services that the companies supply. In a **market economy**, resources are allocated among households and firms with little or no government interference.

Adam Smith, the founder of modern economics, described the dynamic best: “*It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest.*” In other words, producers earn a living by selling products consumers want. Consumers are also motivated by self-interest; they must decide how to use their money to select goods they need or want the most. This process, which Adam Smith called the **invisible hand**, guides resources to their highest value use. The exchange of goods and services in a market economy happens through prices that are established in markets.

Why does all of this happen? Supply and demand tell the story; we begin our exploration of supply and demand by looking at where they interact — in markets. A firm’s degree of control over the market price is the distinguishing feature between **competitive markets** and **imperfect markets**.

Competitive Markets

Buyers and sellers of a specific good or service come together to form a market. Formally, a **market** is a collection of buyers and sellers of a particular product or service; the buyers create the demand for the product, while sellers produce the supply. The interaction of the buyers and sellers in a market establishes the price and the quantity produced of a particular good or the amount of a service offered.

Markets exist whenever goods and services are exchanged. Some markets are online and other operate in traditional “brick & mortar” stores. *A competitive market is one in which there are so many buyers and sellers that each has only a small impact on the market price and output.* In fact, the impact is so small that it is negligible; because each buyer and seller is just one small part of the whole market, no single buyer or seller has any influence over the market price. These two characteristics — similar goods and many participants — create a highly competitive market in which the price and quantity sold of a good are determined by the market rather than by one person or business.

Imperfect Markets

Markets are not always fully competitive; British economist Joan Robinson wrote that in imperfect competition, “*a certain difficulty arises [because] the individual demand curve for the product of each of the firms...will depend to some extent upon the price policy of the others.*” Accordingly, we define these **imperfect markets** as markets in which either the buyer or the seller can influence the market price.

When the seller has some control over the price, we say that the market is **imperfect**. Specialized products, such as popular video games, front-row concert tickets, or dinner reservations at a trendy restaurant, give the seller substantial pricing power. **Market power** is a firm’s ability to influence the price of a good or a service by exercising control over its demand, supply or bot.

A monopoly exists when a single company supplies the entire market for a particular good or service; but even in imperfect markets, the forces of supply and demand significantly influence producer and consumer behavior. (For the time being we’ll keep our analysis focused on supply and demand in a competitive market(s)).

What determines demand?

Demand exists when an individual or group wants something badly enough to pay or trade for it. How much an individual or group actually buys depends on the price of the good or service.

In economics, the amount of a good or service that buyers are willing and able to purchase at the current price is known as the **quantity demanded**.

When the price of a good increases, consumers often respond by purchasing less of the good or buying something else. Therefore, as price goes up, quantity demanded goes down; similarly, as price goes down, quantity demanded goes up. This negative (opposite) relationship between price and the quantity demanded is the **law of demand**. **The law of demand** states that, all other things equal, the quantity demanded falls when the price rises when the price falls. **The law of demand** holds true over a wide range of goods and settings.

Market Demand

So far, we have studied individual demand, but a market is composed of many different types of buyers. In this section, we examine the collective demand of all the buyers in a given market. The **market demand** is the sum of all individual quantities demanded by each buyer in a market at each price. Any demand curve shows the law of demand with movements along (up or down) the curve that reflect the effect of a price change on the quantity demanded of the good or service. Only a change in price can cause a movement along a demand curve.

Shifts in the demand curve

We have examined the relationship between price and quantity demanded. This relationship, described by the law of demand, shows us that when price changes, consumers respond by altering the amount they purchase; but in addition to price, many other variables influence how much of a good or a service is purchased. For instance, news about the possible risk(s) or benefit(s) associated with the consumption of a good or service can change overall demand.

Furthermore, many different variables can shift the demand; these include changes in buyers' income, the price of related goods, changes in buyers' taste and preferences, price expectations, the number of buyers, and taxes. The easiest way to keep all of these elements straight is to ask yourself a simple question: **Would this change cause me to buy more or less of the good?** If the change reduces how much you would buy at any given price, you shift the demand curve to the left; if the change increases how much you would buy at any given price, you shift the curve to the right.

Changes in Buyer's income

When your income goes up, you have more to spend. Assuming that prices don't change, individuals with higher incomes are able to buy more of what they want. Similarly, when your income declines, **your purchasing power**, or how much you can afford, falls; in either case, your income affects your overall demand.

When economists look at how consumers spend, they often differentiate between two types of goods: normal and inferior. Consumers will buy more of a **normal good** as their income goes up (assuming all other factors remain constant — *ceteris paribus*). While consumers with an increased income may purchase more of some things, the additional purchasing power will mean they purchase fewer inferior goods. **An inferior good** is one where demand declines as income rises...as income goes up, consumers buy less of an inferior good because they can afford something better. Within a specific product market, you can often find examples of inferior and normal goods in the form of different brands

The Price of Related Goods

Another factor that can shift the demand curve is the price of related goods. Certain goods directly influence the demand for other goods. **Complements** are two goods that are used together; **substitutes** are two goods that are used in place of each other.

Changes in tastes and preferences

While something is popular, demand increases; as soon as it falls out of favor, you can expect demand for it to decrease. Tastes and preferences can change quickly, and this fluctuation alters the demand for a particular good. Though changes in fashion trades are usually purely subjective, other changes in preferences are the result of new information about the goods and services we buy.

Price expectations

Expectations about the future influenced your current demand. If we expect a price to be higher tomorrow, we are likely to buy more today to beat the price increase; the result is an increase in current demand. Likewise, if you expect a price to decline soon, you might delay your purchase to try to get a lower price in the future. An expectation of a lower price in the future will therefore decrease current demand.

The number of buyers

Recall that the market demand curve is the sum of all individual demand curves. Therefore, another way for market demand to increase is for more individual buyers to enter the market. In other words, demographics changes in society are another source of shifts in demand. In many markets, ranging from movie theater attendance to home ownership, population trends play an important role in determining whether the market is expanding or contracting.

Taxes and subsidies

Changes in **excise taxes** (which are taxes on a single product or service) and **sales taxes** (which are general taxes on most goods and services) affect demand as well.

Higher tax in addition to the price they pay for the good. Lower taxes reduce the overall cost to consumers and therefore increase demand. The reverse is true for a **subsidy**, which is a payment made by the government to encourage the consumption or production of a good or service. A lot of times it's a tax break, like mortgage interest tax deduction, or tax credits on eco-friendly cars. In both cases, the tax break encourages consumers to purchase more of the subsidized good.

What determines supply?

Even though we have learned a great deal about demand, our understanding of markets is incomplete without also analyzing supply. We have seen that with demand, price and output are **negatively related**; that is, they move in opposite directions. With supply, however, the price level and quantity supplied is **positively related**; that is, they move in the same direction. **The quantity supplied** is the amount of a good or service that producers are willing and able to sell at the current price. Higher prices cause the quantity supplied to increase; conversely, lower prices cause the quantity supplied to decrease.

When price increases, producers often respond by offering more for sale; as price goes down, quantity supplied also goes down. This direct positive relationship between price and quantity supplied is the **law of supply**; it states that, all other things being equal, the quantity supplied increases when the price rises, and the quantity supplied falls when the price falls. This law holds true over a wide range of goods and settings.

The Supply Curve

A **supply schedule** is a table that shows the relationship between the price of a good and the quantity supplied. A supply curve is a graph of the relationship between the prices in the **supply schedule** and the quantity supplied at those prices.

Market Supply

The **Market supply** is the sum of the quantities supplied by each seller in the market at each price.

Shifts of the supply curve

When a variable other than the price changes, the entire supply curve shifts; an increase in supply shifts the supply curve to the right, but that happens when a variable causes supply to decrease? This decrease in supply shifts the supply curve to the left. Many variables can shift supply, but it also reminds us of what **does not** cause a shift in supply: **the price**. Recall that price is the variable that cause the supply curve to slope upward; **a price change causes a movement along the supply curve, not a shift in the curve.**

Factors that shift the supply curve include the cost of inputs, changes in technology or the production process, taxes and subsidies, the number of firms in the industry, and price expectations. The easiest way to keep them straight is to ask yourself a simple question: Would the change cause a business to produce more of the good or less of the good? If the change would reduce the amount of a good or service a business is willing and able to supply at every given price, the supply curve shifts to the left. If the change would increase the amount of a good or service a business is willing and able to supply at every given price, the supply curve shifts to the right.

The Cost of Inputs

Inputs are resources used in the production process. Inputs may include workers, equipment, raw materials, buildings, and capital goods. Each of these resource are critical to the production process. When the cost of inputs changes, so does the seller's profit; if the cost of inputs declines, the profits improve. Improved profits make the firm more willing to supply the good; conversely, higher input costs reduce profits.

Changes in technology or the production process

Technology encompasses knowledge that produces use to make their products; an improvement in technology enables a producer to increase output with the same resources or to produce a given level of output with fewer resources. If the producers of a good discover a new and improved technology or a better production process, there will be an increase in supply. That is, the supply curve for the good will shift to the right.

Taxes and Subsidies

Taxes placed on suppliers are an added cost of doing business. For example, if property taxes are increased, the cost of doing business goes up. A firm may attempt to pass along the tax to the consumers through higher prices, but higher prices will discourage sales. So, in some cases, the firm will simply have to accept the taxes as an added cost of doing business, either way, a tax makes the firm less profitable. Lower profits make the firm less willing to supply the product; thus, the supply curve shifts to the left and the overall supply declines.

The reverse is true for **a subsidy**, during the COVID-19 pandemic, hospitals received federal subsidies to offset the added costs associated with treating infected patients (more tests, more protective gear, more sterilizing of equipment, more laundry, and so on...) In addition, airlines and small businesses received subsidies to keep workers employed while the lockdown prevented people from traveling and from going about their day-to-day business at work. As a result, more essential workers remained employed, compared to what would have happened without the subsidies.

The number of firms in the industry

We saw that an increase in total buyers (population) shifts the demand curve to the right. A similar dynamic happens with an increase in the number of sellers in an industry. Each additional firm that enters the market increases the available supply of a good. In graphic form, the supply curve shifts to the right to reflect the increased production. By the same reasoning, if the number of firms in the industry decreases, the supply curve shifts to the left. Changes in the number of firms in a market are a regular part of business.

Price expectations

A seller who expects a higher price for a product in the future may wish to delay sales until a time when the product will bring higher price. Likewise, the expectation of lower prices in the future will cause sellers to offer more while prices are still relatively high. This effect is particularly noticeable in the electronics sector, where newer — and much better — products are constantly being developed and released. Sellers know that their current offerings will soon be replaced by something better and that consumer demand for the existing technology will then plummet. This means that prices typically fall when a product has been on the market for a time; because producers know that the price will fall, they supply as many of the current models as possible before the next wave of innovation cuts the price they can charge.

How do Supply and Demand interact to create Equilibrium?

We have examined supply and demand separately. Now it is time to see how the two interact; the real power of supply and demand analysis is in how well it predicts prices and output in the entire market.

Supply, Demand, Equilibrium

The point of equilibrium is where the demand curve and the supply curve intersect. At this point, the two opposing forces of supply and demand are perfectly balanced. The equilibrium price is the quantity supplied that equals to the quantity demanded; it is also called **market-clearing price**, because that is the only price at which no surplus or shortage of the goods exists. Similarly, there's also an equilibrium quantity at which the quantity supplied equals the quantity demanded; when the market is in equilibrium, we sometimes say that **the market clears** or that **the price clears the market**.

The equilibrium point has a special place in economics because movements away from that point throw the market out of balance. The equilibrium process is so powerful that it is often referred to as the law of supply and demand, *the idea that the market prices adjust to bring the quantity supplied and quantity demanded into balance*.

Shortage and Surpluses

When there is more demand for a product than sellers are willing or able to supply, we say there is a shortage. A **shortage**, or **excess demand**, occurs whenever the quantity supplied is less than the quantity demanded. A **surplus**, or **excess supply**, occurs whenever the quantity supplied is greater than the quantity demanded.

Every seller and buyer has a vital role to play in the market, venues like the Pike Place Market bring buyers and sellers together. Amazingly, market equilibrium occurs without the need for government planning to ensure an adequate supply of the goods consumers want or need; you might think that a decentralized system would create chaos, but nothing could be further from the truth. Markets work because buyers and sellers can rapidly adjust to changes in prices, these adjustments bring balance; when markets were suppressed in communist countries during the 20th century, shortages were commonplace, in part because there was no market price system to signal that additional production was needed.

In Appendix 3A, we consider what happens when supply and demand change at the same time. There you will discover the challenges in simultaneously determining price and quantity when more than one variable changes.

Conclusion

Five years from now, if someone asks you what you remember about your first course in economics, you'll probably respond with two words: "supply" and "demand". These two forces allow us to model market behavior through prices. Supply and demand help establish the market equilibrium, or the prices at which quantity demanded are in balance. At the equilibrium point, every good and service produced has a corresponding buyer who wants to purchase it. When the market is out of equilibrium, either a shortage or surplus exists. This condition persists until buyers and sellers have a chance to adjust the quantity they demand and the quantity they supply, respectively.

In the next chapter, we extend our understanding of supply and demand by examining how sensitive, or responsive, consumers, and producers are to price changes. With this knowledge, we can determine whether price changes have a big effect on behavior or not.

Elasticity

[Abdon Morales](#)

Sep 7, 2023

The Rebound Effect

Economists use the idea of elasticity in a similar way, to describe responsiveness to a change in market conditions. In the previous chapter, we learned that demand and supply regulate economic activity by responding to the interests of buyers and sellers through prices. A higher price causes the quantity demanded to fall. A lower price causes the quantity demanded to rise. How strongly the quantity consumed or produced responds to a change in price is its elasticity.

High and Low Elasticity

A high elasticity means a large, very responsive change in quantity, while low elasticity means a small change; so the change in price is like the drop height of a ball, and the change in quantity is like the rebound height.

This is an important concept in economics. Understanding elasticity helps us determine the impact of government policy on the economy, to vote more intelligently, and even to make wiser day-to-day decisions.

BIG QUESTIONS

- What is the price elasticity of demand, and what are its determinants?
 - The price elasticity of demand is a measure of the responsiveness of quantity demanded to a change in price.
 - Demand will generally be more elastic if there are many substitutes available, if the item accounts for a larger share of the consumer's budget, if it is a luxury good, if the market is more narrowly defined, or if the consumer has plenty of time to make a decision.
 - Economists categorize time in three distinct periods: (1) the immediate run, when there is no time for consumers to adjust their behavior; (2) the short run, when consumers can adjust, but only partially; and (3) the long run, when consumers have time to fully adjust to market conditions.
 - The price elasticity of demand is calculated by dividing the percentage change in the quantity demanded by the percentage change in price. A value of zero indicates that the quantity demanded does not respond to a price change; if the price elasticity is zero, demand is said to be perfectly inelastic. When the price elasticity is zero, demand is between 0 and -1, demand is inelastic. If the price elasticity of demand is less than -1, demand is elastic. When price elasticity is exactly -1, the item has unitary elasticity.
- How do changes in income and the prices of other goods affect elasticity?
 - The income elasticity of demand measures how a change in income affects spending. Normal goods have a positive income elasticity; inferior goods have a negative income elasticity.
 - The cross-price elasticity of demand measures the responsiveness of the quantity demanded of one good to a change in the price of a related good.

Elasticity

[Abdon Morales](#)

Sep 7, 2023

Positive values for the cross-price elasticity mean that the goods are substitutes, while negative values indicate that the two goods are complements. If the cross-price elasticity is zero, then the two goods are not related to each other.

- What is the price elasticity of supply?
 - The price elasticity of supply is a measure of the responsiveness of the quantity supplied to a change in price. Supply will generally be more elastic if producers have flexibility in the production process and ample time to adjust production.
 - The price elasticity of supply is calculated by dividing the percentage change in the quantity supplied by the percentage change in price. A value of zero indicates that the quantity supplied does not respond to a price change; if the price elasticity of supply is zero, supply is said to be perfectly inelastic. When the price elasticity of supply is between 0 and 1, demand is relatively inelastic. If the price elasticity of supply is greater than 1, supply is elastic.
- How do the price elasticities of demand and supply relate to each other?
 - The interplay between the price elasticity of demand and the price elasticity of supply determines the magnitude of the resulting price change.

What is the Price Elasticity of Demand, what are its Determinants?

With goods like pasta, where consumers can easily purchase a substitute, we think of demand as being responsive. That is, a small change in price will likely cause many people to switch from one good to another. In contrast, many things in life are irreplaceable or have few good substitutes...consumers are unresponsive, or unwilling to change their behavior, even when the price of the good or service changes.

Elasticity is a measure of the responsiveness of buyers and sellers to changes in price or income. Elasticity is a useful concept because it allows us to measure how much consumers and producers change their behavior when either price or income changes. In the next section, we look at the factors that determine the price elasticity of demand.

Determinants of the Price of Elasticity of Demand

The law of demand tells us that as price goes up, the quantity demanded goes down; and as price goes down, the quantity demanded goes up. In other words, there is a negative relationship between the price of a good and the quantity demanded. Elasticity allows us to measure how much the quantity demanded changes in response to a change in price. If the quantity demanded changes significantly as a result of a price change, then demand is elastic. If the quantity demanded changes a small amount as a result of a price change, then demand is inelastic.

The price elasticity of demand measures the responsiveness of quantity demanded to a change in price. Five determinants play a crucial role in influencing whether demand will be elastic or

Elasticity

[Abdon Morales](#)

Sep 7, 2023

inelastic: the existence of substitutes, the share of the budget spent on a good, whether the good is a necessity or a luxury good, how broadly defined the market is and time.

The existence of substitutes (more substitutes = more elastic (responsive))

The most important determinant of price elasticity is the number of substitutes available. When substitutes are plentiful, market forces tilt in favor of the consumer. When there is no good substitutes? There is no amusement park quite like Disney; nowhere! Because the experience is unique, the number of close substitutes is small, therefore, demand is more inelastic, or less responsive to price changes. To some degree, the price elasticity of demand depends on consumer preferences; ultimately, whether demand is inelastic, or elastic depends on the buyer's preferences and resources.

The share of the budget (bigger share = more elastic (responsive)) spent on the good

Demand is much more inelastic for inexpensive items on sale. For example, if a candy bar is discounted 10%, the price falls by pennies. The savings from switching candy bars is not enough to make a difference in what you can afford elsewhere. Therefore, the incentive to switch is small; most consumers still buy their favorite candy because the savings gained from purchasing a less desirable candy bar are small in comparison to the consumer's budget.

Necessities vs. Luxury Goods (= more elastic (responsive))

When consumers purchase a necessity, they are generally thinking about the need, not the price. When the need trumps the price, we expect demand to be relatively inelastic. Therefore, the demand for things like cars, textbooks, and heating oil all tend to have inelastic demand; but some goods are necessities, such as having to pay your rent and water bill, purchase gasoline (or diesel) for your car, and eat.

Whether the market is broadly or narrowly defined

The more broadly we define a market for a good, the harder it is to live without.

Elasticity

[Abdon Morales](#)

Sep 7, 2023

Time (more time = more elastic (responsive)) and Adjustment Process

When the market price changes, consumers and sellers respond, but that response does not remain the same over time. As time passes, both consumers and producers are able to find substitutes. To understand these different market responses, when considering elasticity economists consider time in three distinct periods: the *immediate run*, the *short run*, and the *long run*.

In the immediate run, there is no time for consumers to adjust behavior; consider the demand for gasoline. When the gas tank is empty, you have to stop at the nearest gas station and pay the posted price. Filling up as soon as possible is more important than driving around and searching for the lowest price. Inelastic demand exists whenever price is secondary to the desire to attain a certain amount of the good; so in the case of an empty tank, the demand for gasoline is inelastic.

There is no time to adjust

But what if your tank is not empty? The ***short run*** is a period of time when consumers can partially adjust their behavior (in this case, we can search for a good deal on gas). In the short run, we can make decisions that reflect our immediate or short-term wants, needs, or limitations. When consumers have some time to make a purchase, they gain flexibility; they can shop for lower prices at the pump, carpool to save gas, or even change how often they drive. In the short run, flexibility reduces the demand for expensive gasoline and makes consumer demand more elastic.

There is time to partially adjust

Finally, if we relax the time constraint completely, it is possible to use even less gasoline. The ***long run*** is a period of time consumers have time to fully adjust to market conditions. **In the long run, we make decisions that reflect our wants, needs, and limitations over a long time horizon.** If gasoline prices are high in the long run, consumers can relocate closer to work and purchase fuel-efficient cars. These changes further reduce the demand for gasoline. As a result of the flexibility that additional time gives the consumer, the demand for gasoline becomes more elastic.

There is time to fully adjust

We have looked at the five determinants of elasticity - substitutes, the share of the budget spent on the good, necessities versus luxury goods, whether the market is broadly or narrowly defined, and time. Each is significant, but the number of substitutes tends to be the most influential factor and dominates the others.

Computing the Price Elasticity of Demand

Until that point, our discussion of elasticity has been descriptive. However, to apply the concept of elasticity in decision making, we need to view it more quantitatively. If a government is

Elasticity

[Abdon Morales](#)

Sep 7, 2023

considering a new tax, it needs to know how much revenue that tax will generate. These are questions about elasticity that we can evaluate by using a mathematical formula.

The price elasticity of demand formula

$$\text{price elasticity of demand} = E_D = \frac{\text{percentage change in the quantity}}{\text{percentage change in price}}$$

The midpoint method

Economists use the **midpoint method**, which gives the same answer for elasticity no matter what point you begin with. The midpoint method merely specifies how to plugin the initial and ending values for price and quantity to determine the percentage changes. Q_1 and P_1 are the initial values, and Q_2 and P_2 are the ending values.

$$E_D = \frac{\text{change in } Q \div \text{average value of } Q}{\text{change in } P \div \text{average value of } P}$$



Notice the **negative sign**. This is because the law of demand states there is an **inverse relationship** between price and quantity demanded.

$$= \frac{(Q_2 - Q_1) \div [(Q_1 + Q_2) \div 2]}{(P_2 - P_1) \div [(P_1 + P_2) \div 2]}$$

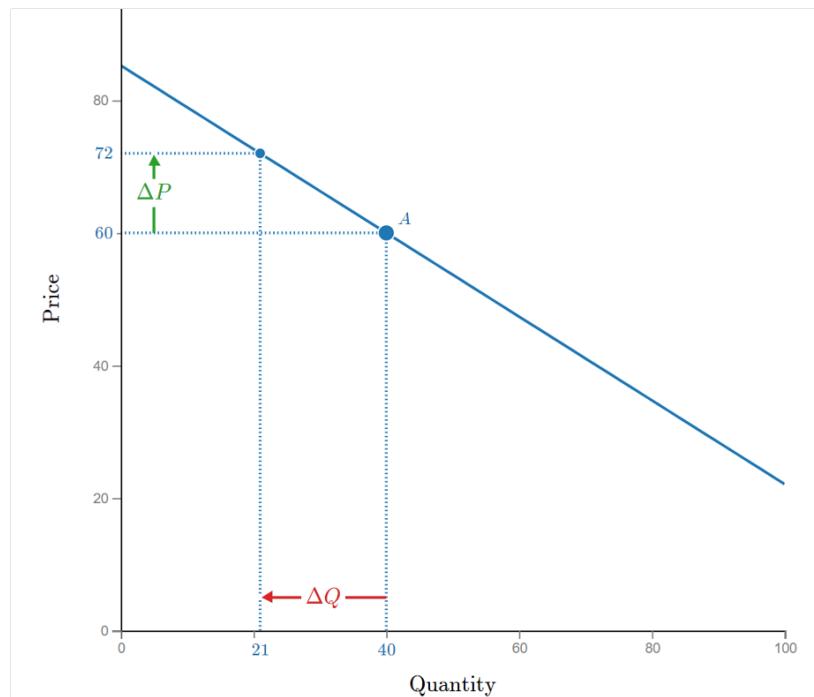
The midpoint method is the preferred method for solving elasticity problems.

When the elasticity coefficient is less than -1 , the opposite is true, and demand is elastic.

Elasticity

[Abdon Morales](#)

Sep 7, 2023



Graphing the Price Elasticity of Demand

Visualizing elasticity graphically helps us understand the relationship between elastic and inelastic demand. As demand becomes increasingly elastic, or responsive to price changes, the demand curve flattens. The range of elasticity runs from perfectly inelastic through perfectly elastic.

Perfectly inelastic demand

Perfectly Inelastic



Elasticity

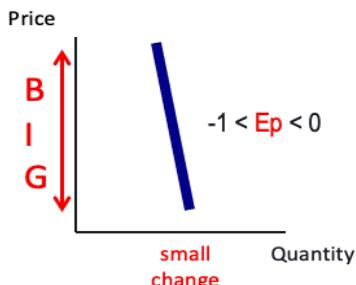
Abdon Morales

Sep 7, 2023

When zero is in the numerator in the formula of price elasticity of demand, we know that the answer will be zero no matter what we find in the denominator. This means that the demand is perfectly inelastic.

Relatively Inelastic demand

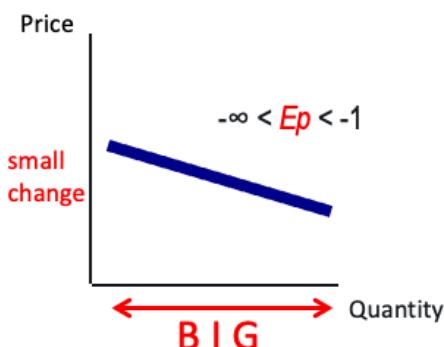
Relatively inelastic



When the change on the quantity axis is small compared with the change on the price axis, the price elasticity is **relatively inelastic**. Plugging in these changes into the elasticity formula, we get $\frac{\text{percentage change in } Q_D}{\text{percentage change in } P} = \frac{\text{small change}}{\text{large change}}$. Recall that the law of demand describes a negative relationship between price and quantity axes always be in opposite directions; a price elasticity of zero tells us that there's no change in the quantity demanded when price changes. So when demand is relatively inelastic, the price elasticity of demand must be closer to zero; therefore, the price elasticity of demand is between 0 and -1 when demand is relatively inelastic.

Relatively elastic demand

Relatively elastic



Because there are many good substitutes for "x" item, the demand for "x" item is relatively elastic. The flexibility of consumer demands for "x" item is illustrated by the degree of responsiveness we see along the quantity axis relative to the change exhibited along the price axis. We can observe this responsiveness by noting that a relatively elastic demand curve is flatter than an inelastic demand curve; so whereas perfectly

Elasticity

Abdon Morales

Sep 7, 2023

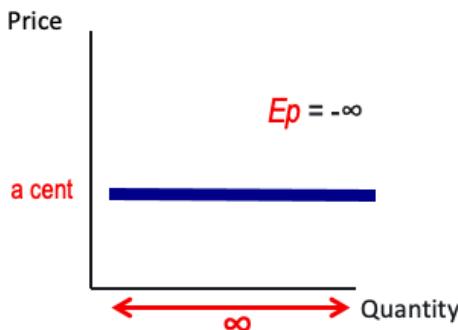
inelastic demand shows no change in demand with an increase in price; and relatively inelastic demand shows a small change in quantity demanded with an increase in price; relatively elastic demand shows a large change in quantity demanded with an increase in price.

Placing this information into the elasticity formula gives us: $E_x = \frac{\text{percentage change in } Q_D}{\text{percentage change in } P} =$

large change over *small change*. Now the numerator – the percentage change in Q_D – is large, and the denominator – the percentage change in P – is small; that means E_D is less than -1. Recall that the sign must be negative, because there is a negative relationship between price and the quantity demanded; as the price elasticity of demand move farther away from zero, the consumer becomes more responsive to a price change, thus a small change in price of an item will have a large effect on the quantity demanded because of many good substitutes for the item.

Perfectly elastic demand

Perfectly Elastic



Interesting example: the demand for a \$10 bill; Would you pay \$11 to get a \$10 bill? No. Would you pay \$10.01 for a \$10? Still no; however, when the price drops to \$10, you will probably become indifferent (that is, you will be equally satisfied with paying \$10.00 for the \$10 bill or not making the trade). The real magic here occurs when the price drops to \$9.99, how many \$10 bills would you buy if you could buy them for \$9.99 or less? The answer: as many as possible!

This is exactly what happens in currency markets, where small differences among currency prices around the global motivate traders to buy and sell large quantities of currency and clear a small profit on the difference in exchange rates. This extreme form of price sensitivity is illustrated by a perfectly horizontal demand curve, which means that demand is **perfectly elastic**.

Solving for the elasticity yields: $E_{bill} = \frac{\text{percentage change in } Q_D}{\text{percentage change in } P} = \frac{\text{nearly infinite change}}{\text{very small } (\$0.01) \text{ change}}$.

We can think of this very small price change, from \$10.00 to \$9.99, as having essentially unlimited effect on the quantity of \$10 bills demanded. Traders go from being

Elasticity

[Abdon Morales](#)

Sep 7, 2023

uninterested in trading at \$10 to seeking to buy as many \$10 bills as possible when the price drops to \$9.99. As a result, the price elasticity of demand approaches negative infinity ($-\infty$).

Unitary elasticity

Unitary elasticity describes the situation in which elasticity is neither elastic nor inelastic. This situation occurs when E_D is exactly -1 , and it happens when the percentage change in price is exactly equal to the percentage change in quantity demanded. This characteristic of unitary elasticity will be important when we discuss the connection between elasticity and total revenue later in this chapter. You're probably wondering what an example of a unitary good would be; it is impossible to find a good that has a price elasticity of exactly -1 at all price points. It is enough to know that unitary demand represents the crossover from elastic to inelastic demand.

Elasticity	E_D coefficient	Interpretation
Perfectly inelastic	$E_D = 0$	Price does not matter
Relatively inelastic	$0 > E_D > -1$	Price is less important than the quantity purchased
Unitary	$E_D = -1$	Price and quantity are equally important
Relatively elastic	$-1 > E_D > -\infty$	Price is more important than the quantity purchased
Perfectly elastic	$E_D = -\infty$	Price is everything

Time, Elasticity, and the demand curve

We have already seen the increased time makes demand more elastic. When price rises from P_1 to P_2 , consumers cannot avoid the price increase in the immediate run, and demand is represented by the perfectly inelastic demand, D_1 .

Elasticity

[Abdon Morales](#)

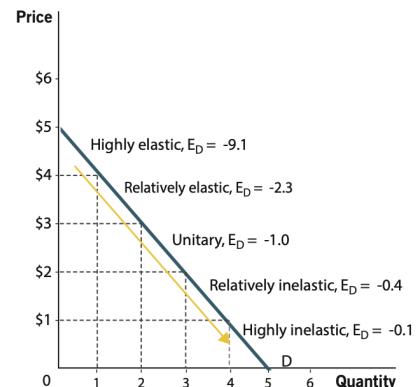
Sep 7, 2023

Slope and Elasticity

Slope and elasticity are NOT the same!

With a linear demand curve: the slope will be the same at all points, elasticity will be different at all points, and elasticity decreases (gets more inelastic) as we move down and right along a linear demand curve.

Along any straight demand curve, the price elasticity of demand (E_D) is not constant, as you can see by noting how the price elasticity changes from highly elastic near the top of the demand curve to highly elastic near the bottom of the curve. In the table, note that all the numbers in the third, fourth, and fifth columns are based on the midpoint formula.



Price (dollars)	Quantity	Percentage change in price	Percentage change in quantity demanded	Elasticity coefficient (midpoint formula)	Interpretation
5	0	-22	200	-9.1	Highly elastic
4	1	-29	67	-2.3	Relatively elastic
3	2	-40	40	-1.0	Unitary
2	3	-67	29	-0.4	Relatively inelastic
1	4	-200	22	-0.1	Highly inelastic
0	5				

Elasticity

[Abdon Morales](#)

Sep 7, 2023

Price elasticity of Supply

Price Elasticity of Supply



Responsiveness of quantity supplied to a change in price

$$E_s = \frac{\% \Delta Q_s}{\% \Delta P}$$

Price Elasticity of Demand and Total Revenue

Understanding the price elasticity of demand for the product you sell is important when running a business. Consumer responsiveness to price changes determines whether a firm would be better off raising or lowering its price for a given product. In this section, we explore the relationship between the price elasticity of demand and a firm's total revenue.

Elasticity

Abdon Morales

Sep 7, 2023

Price (dollars)	Quantity	Total Revenue $P \times Q$	Percentage change in price	Percentage change in quantity demanded	Elasticity coefficient (midpoint formula)	Interpretati on
5	0	0	-22	200	-9.1	Highly elastic
4	1	4	-29	67	-2.3	Relatively elastic
3	2	6	-40	40	-1.0	Unitary
2	3	6	-67	29	-0.4	Relatively inelastic
1	4	4	-200	22	-0.1	Highly inelastic
0	5	0				

But first we need to understand the concept of total revenue: **Total revenue** is the amount that a firm receives from the sale of goods and services. Total revenue for a particular good is calculated by multiplying the price of the good by the quantity of the good that is sold. Table 4.3 reproduces the table from Figure 4.3 (table + graph)(with numbers based on the midpoint formula) and adds a column for the total revenue. We find the total revenue by multiplying the price by the quantity purchased.

After calculating total revenue at each price, we can look at the column of elasticity coefficients to determine the relationship. When we link revenues with the price of elasticity of demand, a trade-off emerges. [this trade-off occurs because total revenue and elasticity relate to the price differently; total revenue involves multiplying the price by the quantity, while elasticity involves dividing the percentage change in price.) Total revenue is zero when the price is too high (\$5 or more) and when the price is too low (\$0). Between these two extremes, prices from \$1 to \$4 generate positive total revenue.

Consider what happens when the price drops from \$5 to \$4...at \$4, the first latte is purchased; the total revenue is $\$4 \times 1 = \4 . This is also the range at which the price elasticity of demand is highly elastic, as result, lowering the price increases revenues; revenue continues to increase when the price drops from \$4 to \$3. Now two lattes are sold, so the total revenue rises to $\$3 \times 2 = \6 , at the same time, demand remains elastic; we thus conclude that when demand is elastic, lowering the price will increase total revenue. This relationship is shown in panel (a) of Figure 4.4; by lowering the price from \$4 to \$3, the business has generated \$2 more in revenue, but to generate extra revenue, the business has lowered the price from \$4 to \$3 and therefore

Elasticity

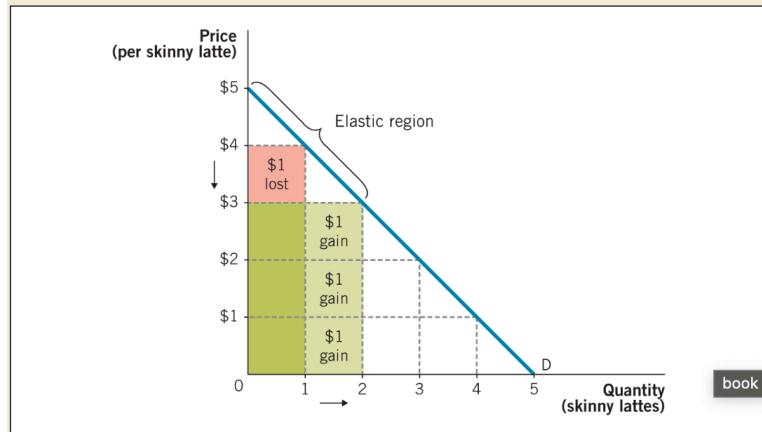
Abdon Morales

Sep 7, 2023

has given up \$1 for each unit it sells. This lost revenue is represented by the red area under the demand curve in panel (a).

(a) The Total Revenue Trade-Off When Demand Is Elastic

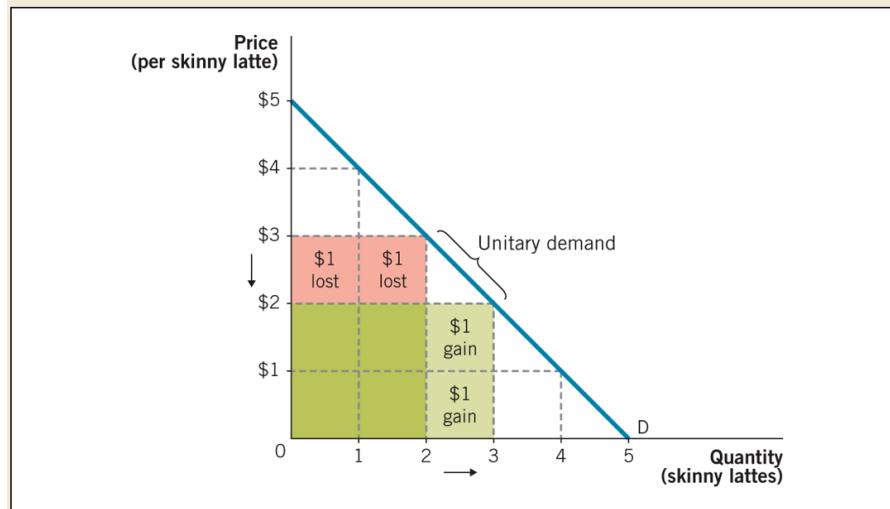
In the elastic region of the demand curve, lowering the price will increase total revenue. The gains from increased purchases, shown in the light green area, are greater than the losses from a lower purchase price, shown in the red area. The green area is part of the total revenue that exists at both prices.



When the price drops from \$3 to \$2, the total revenue stays at \$6; this result occurs because demand is unitary as shown in the panel (b). This special condition exists when the percentage price change is exactly offset by an equal percentage change in the quantity demanded. In this situation, revenue remains constant. At \$2, three lattes are purchased, so the total revenue is $\$2 \times 3$, which is the same as it was when the price was \$3; as a result, we can

(b) . . . When Demand Is Unitary

When demand is unitary, lowering the price will no longer increase total revenue. The gains from increased purchases, shown in the light green area, are equal to the losses from a lower purchase price, shown in the red area.



Elasticity

[Abdon Morales](#)

Sep 7, 2023

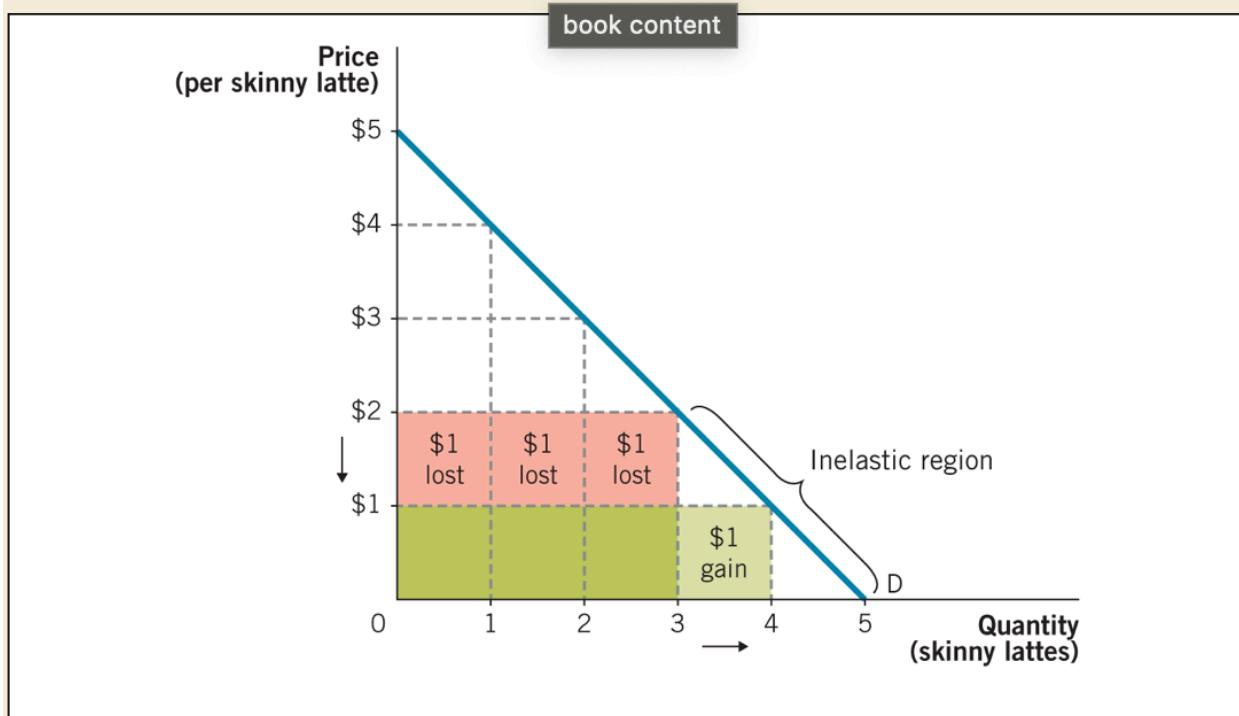
see that total revenue has reached a maximum. Between \$3 and \$2, the price elasticity of demand is unitary, this finding does not necessarily mean that the firm will operate at the unitary point. Maximizing profit, not revenue, is the ultimate goal of a business, and we have not yet accounted for costs in our calculations of profits.

Once we reach a price below unitary demand, we move into the realm of inelastic demand, shown in panel (c). When the price falls to \$1, total revenue declines to \$4; this results occurs because the price elasticity of demand is now relatively inelastic, or price insensitive. Even though the price is declining by \$1, price is increasingly unimportant; as you can see by the light green square, lowering the price to \$1 does not spur a large increase in consumption.

As we see in panel (c), at a price of \$2, three units are sold and total revenue is $\$2 \times 3 = \6 ; when the price falls to \$1, four units are sold, so the total revenue is now $\$4 \times 1 = \4 . By lowering the price from \$2 to \$1, the business has lost \$2 in extra revenue from the lower price; lowering the price from \$2 to \$1 causes a loss of \$3 in existing sales revenue (the red boxes). At the same time, it generates only \$1 in new sales (light green box), so the net change is a loss of \$2.

(c) . . . When Demand Is Inelastic

In the inelastic region of the demand curve, lowering the price will decrease total revenue. The gains from increased purchases, shown in the light green area, are smaller than the losses from a lower purchase price, shown in the red area.



Elasticity

[Abdon Morales](#)

Sep 7, 2023

In this analysis, we see that once the demand curve enters the inelastic area, lowering the price decreases total revenue. This outcome is unambiguously bad for a business; the lower price brings in less revenue and requires business to produce more goods, because making goods is costly, it does not make sense to lower prices into the region where revenues decline. We can make sure that no business will intentionally operate in the inelastic region of the demand curve because it will earn less profit.

How do changes in Income and the Prices of other goods affect elasticity?

We have seen how consumer demand responds to changes in the price of a single good. In this section, we examine how responsive demand is to changes in income and to price changes in other goods.

Income elasticity

Changes in personal income can have a large effect on consumer spending, after all, the money in your pocket influences not only how much you buy, but also the types of purchases you make. A consumer who is low on money may opt to buy a cheap generic product, while someone with a little extra cash can afford to upgrade. For instance, the grocery store aisle reflects different shoppers' budget; store brands and name products compete for shelf space. Lower-income shoppers can choose the store brand to save money, while more affluent shoppers can choose their favorite brand-name product without worrying about the purchase price. **The income elasticity of demand (E)** measures how a change in income affects spending. It is calculated by dividing the percentage change in the quantity demanded by the percentage change in persona income:

$$E_I = \frac{\text{percentage change in the quantity demanded}}{\text{percentage change in income}}$$

Unlike the price elasticity of demand, which is negative, the income elasticity can be positive or negative. When a higher level of income enables the consumer to purchase more, the goods that are purchased are ***normal goods***, because the demand for normal goods goes up with income, they have a positive income elasticity; a rise in income causes a rise in the quantity demanded. Whenever a good is normal, the result is a positive income elasticity of demand, and purchases of the good rise as income expands and purchases of the good fall as income falls.

Normal goods fall into two categories: ***necessities and luxuries***. Goods that people consider necessities generally have income elasticities between 0 and 1; although purchases of necessities will increase as income rises, they do not rise as fast as the increase in income does. Therefore, as income increases, spending on necessities will expand at a slower rate than the increase in income. Rising income

Elasticity

Abdon Morales

Sep 7, 2023

enables consumers to enjoy significantly more luxuries, producing an income elasticity of demand greater than 1 for luxuries.

In chapter 3, we saw that **inferior goods** are those that people will choose not to purchase when their income goes up. Inferior goods have a negative income elasticity, because as income expands, the demand for these goods declines.

Inferior good: $E_I < 0$

Normal good (necessity): $0 < E_I < 1$

Normal good (luxury): $E_I > 1$

Cross-Price Elasticity

Now we will look at how a price change in one good can affect the demand for a related good. **The cross-price elasticity of demand (E_C)** measures the percentage change in the quantity demanded of one good to the percentage change in the price of a related good:

$$E_C = \frac{\text{percentage change in the quantity demanded of one good}}{\text{percentage change in the price of a related good}}$$

Consider how two goods are related; if the goods are substitutes, a price rise in one good will cause the quantity demanded of that good to decline. At the same time, because consumers can purchase the substitute good for the same price as before, the demand for the substitute good will increase.

The opposite is true if the goods are complements; in that case, a price increase in one good makes the joint consumption of both goods more expensive. Therefore, the consumption of both goods will decline.

What if there is no relationship between the two goods? The cross-price elasticity is neither positive nor negative; its zero. **The table below lists cross-price elasticity values according to the type of good.**

Income Elasticity			
Type of good	Subcategory	E_I coefficient	Example
Inferior		$E_I < 0$	Macaroni and cheese
Normal	Necessity	$0 < E_I < 1$	Milk
Normal	Luxury	$E_I > 1$	Diamond ring

Cross-Price Elasticity

Elasticity

Abdon Morales

Sep 7, 2023

Type of good	E_C coefficient	Example
Substitute	$E_C > 0$	Pizza Hut and Dominos
No relationship	$E_C = 0$	A basketball and bedroom slippers
Complements	$E_C < 0$	Turkey and gravy

What is the Price Elasticity of Supply?

Like consumers, sellers are sensitive to price changes. However, the determinants of the price elasticity of supply are substantially different from the determinants of the price elasticity of demand. **The price elasticity of supply** is a measure of the responsiveness of the quantity supplied to a change in price.

In this section, we examine how much sellers respond to price changes. For instance, if the market price of gasoline increases, how will oil companies respond? The answer depends on the elasticity of supply. Oil must be refined into gasoline. If it is difficult for oil companies to increase their output of gasoline supplied will not increase much even if the price increases a lot. In this case, we say that supply is inelastic, or unresponsive. However, if the price increase is small and suppliers respond by offering significantly more gasoline for sale, then supply is elastic. We would expect to observe this outcome if it is easy to refine oil into gasoline.

When supply is not able to respond to a change in price, we say it is inelastic; when the supplier's ability to make quick adjustments is limited, the elasticity of supply is less than 1; Recall the law of supply, which states that there is a direct relationship between price of a good and the quantity that a firm supplies. As a result, the percentage change in the quantity supplied and the percentage change in price move in the same direction. The E_S coefficient reflects this direct relationship with a positive sign.

Determinants of the Price Elasticity of Supply

When we examined the determinants of the price elasticity of demand, we saw that consumers have to consider the number of substitutes, how expensive the item is compared to their overall budget, whether the good is a necessity or a luxury, and the amount of time they have to make a decision. Time and adjustment process are also key elements in determining the price elasticity of supply. However, there is a critical difference: the degree of flexibility that producers have in bringing their product to the market quickly.

A Closer Look at the Price Elasticity of Supply

Elasticity	E_S coefficient	Example
------------	-------------------	---------

Elasticity

Abdon Morales

Sep 7, 2023

Perfectly inelastic	$E_S = 0$	Oceanfront land
Relatively inelastic	$0 < E_S < 1$	Cell phone tower
Relatively elastic	$E_S > 1$	Hot dog vendor

The Flexibility of Producers

When a producer can quickly ramp up output, supply tends to be elastic. One way to maintain flexibility is to have spare production capacity – extra capacity enables producers to quickly meet changing price conditions, so supply is more responsive, or elastic – the ability to store the good is another way to stay flexible; producers who have stockpiles of their products can respond more quickly to changes in market conditions. However, many businesses cannot adapt to changing market conditions quickly.

Time and the Adjustment Process

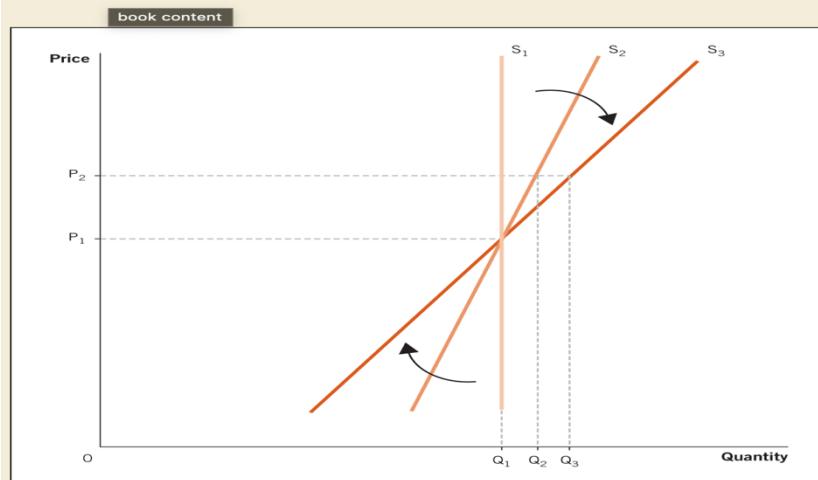
In the immediate run, businesses are stuck with what they have on hand. As we move from immediate run to short run and a price change persists through time, supply – just like demand – becomes more elastic.

Figure 4.5 shows how the two determinants of supply elasticity are mapped onto the supply curve. The supply curve is vertical (S_1). A vertical curve tells us that there is no responsiveness when the price changes. As producers gain additional time to make adjustments, the supply curve rotates from S_1 (the immediate run) to S_2 (the short run) to S_3 (the long run). Like the demand curve, the supply curve rotates clockwise; in contrast, as we saw in

FIGURE 4.5

Elasticity and the Supply Curve

Increased flexibility and more time make supply more elastic. When price rises from P_1 to P_2 , producers are unable to expand output immediately and the supply curve remains at Q_1 in the immediate run. In the short run (S_2), the firm becomes more flexible and output expands to Q_2 . Eventually, in the long run (S_3), the firm is able to produce even more, and it moves to Q_3 in response to higher prices.



Elasticity

[Abdon Morales](#)

Sep 7, 2023

Figure 4.2 (refer to the textbook), the demand curve rotates counterclockwise. With both supply and demand, the most important thing to remember is that more time allows for greater adjustment, so the long run is always more elastic.

Calculating the price elasticity of supply

We can use a simple formula to calculate the price elasticity of supply; doing so is useful when a business owner must decide how much to produce at various prices. The elasticity of supply measures how quickly the producer is able to change production in response to changes in price. When supply is elastic, producers are able to quickly adjust production; if supply is inelastic, production tends to remain roughly constant, despite large swings in price.

Here is the formula for the price elasticity of supply (E_S):

$$E_S = \frac{\text{percentage change in the quantity supplied}}{\text{percentage change in the price}}$$

How do the price elasticities of demand and supply relate to each other?

The interplay between the price elasticity of supply and the price elasticity of demand allows us to explain more fully how the economy operates. With an understanding of elasticity at our disposal, we can conduct a much richer and deeper analysis of the world around us.

The interplay between the price elasticity of demand and the price elasticity of supply determines the magnitude of the resulting price change. We cannot observe demand in isolation without also considering how supply responds. Similarly, we cannot simply think about the short-run consequences of demand and supply shifts; we also must consider how prices and quantity will vary in the long run. Armed with this knowledge, you can begin to see the power of the supply and demand model to explain the world around us.

Conclusion

Do sellers charge the highest price possible? We can now address this misconception definitively: no. Sellers like higher prices in the same way consumers like lower prices, but that does not mean that sellers will charge the highest price possible. At very high prices, consumer demand is quite elastic. Therefore, a seller who charges too high a price will not sell much. As a result, firms learn that they must lower their price to attract more customers and maximize their total revenue.

The ability to determine whether demand and supply are elastic or inelastic enables economists to calculate the effects of personal, business, and policy decisions. When you combine the concept of elasticity with the supply and demand model from Chapter 3 (again refer to the textbook), you get a very powerful tool. In the subsequent chapters, we use elasticity to refine our models of economic behavior and make our results more realistic.

Market Outcomes and Tax Incidence

Econ 304K: Chapter 5

Taxes on Firms Affect Consumers

Many people believe that when the government taxes businesses, consumers catch a break because firms pay the tax; if only life works that way...as this chapter explains, who actually pays the tax often is quite different from the party that is legally responsible for making the tax payment.

Gasoline prices are common and visible sign of the market at work; it is hard not notice when gasoline prices rise or fall, because every gas station posts its prices prominently [but] there are a few things that you might not know. First, in the United States, gasoline taxes are paid by the seller and not displayed at the pump; second, gasoline taxes vary wildly from country to country. For instance, residents of Saudi Arabia pay some of the lowest gasoline prices in the world, while the Netherlands (Dutch) have to put up with the world's third-highest gas prices.

This occurs because the government of certain oil-rich countries, such as Saudi Arabia, subsidize gasoline so that their citizens pay less than the market price; furthermore, in countries where gasoline is subsidized, consumers drive their cars everywhere, mass transportation is largely unavailable, and there is less than the market price. The opposite is true in countries like the Netherlands, where consumers drive less and use public transportation more often, and tend to purchase fuel efficient cars.

What do gasoline taxes and subsidies around the world have in common? They're all folded into price you see at the pump which might lead you to believe that the seller is paying all the tax or receiving the entire subsidy. (Nothing can be further from the truth) The firm will try to pass along the tax to consumers in the form of higher prices, likewise, in countries with subsidies — the firm must pass along lower prices to consumers. After reading this chapter, you will understand how this process works.

We begin this chapter by discussing consumer and producer surplus, two concepts that illustrate gains from trade. These concepts helps us measure the efficiency of markets and the effects of taxation. Then we examine how taxation creates distortions in economic behavior by altering the incentives people and firms face when consuming and producing goods that are taxes

Big Questions

- **What are consumer surplus and producer surplus?**
 - *Consumer surplus is the difference between the willingness to pay for a good or service and the price that is paid to get it. Producer surplus is the difference between the price the seller receives and the price at which the seller is willing to sell the good or service.*
 - *Total surplus (social welfare) is the sum of consumer and producer surplus that exists in a market.*
- **When is a market efficient?**
 - *Markets maximize consumer and producer surplus, provide goods and services to buyers who value them most, and reward sellers who can produce goods and services at the lowest cost. As a result, markets create the largest amount of total surplus possible.*
 - *Whenever an allocation of resources maximize total surplus, the result is said to be efficient. However, economists are also interested in the distribution of the surplus. Equity refers to the fairness of the distribution of the benefits within the society.*
- **Why do taxes create deadweight loss in otherwise efficient markets?**
 - *Deadweight loss occurs because taxes increase the purchase price, which causes consumers to buy less and producers to supply less. Deadweight loss can be lessened by taxing goods or services that have inelastic demand or supply.*
 - *Economists are also concerned about the incidence of taxation: incidence refers to the burden of taxation on the party who pays the tax through higher prices, regardless of whom the tax is actually levied on. The incidence is determined by the balance between the elasticity of supply and elasticity of demand.*

Willing to Pay for a New Economics Textbook	
Buyer	Willingness to pay
Sergio	\$200
Silene	\$150
Raquel	\$100

What are consumer and producer surplus?

Markets create value by bringing together buyers and sellers so that consumers and producers can mutually benefit from trade. **Welfare economics** is the branch of economics that studies how the allocation of resources affects economic well-being. In this section, we develop two concepts that help us measure the value markets create: **consumer surplus** and **producer surplus**.

In competitive markets, the equilibrium price is simultaneously low enough to attract consumers and high enough to encourage producers. This balance between demand and supply enhances the *welfare* (well-being) of society; that is not to say that society's welfare depends solely on markets. People also find satisfaction in many non-market settings; we incorporate aspects of personal satisfaction into our economic models in Chapter 16.

Consumer Surplus

Consumer surplus is the difference between the willingness to pay for a good (or service) and the price paid to get it; whenever the price is greater than the willingness to pay, a rational consumer will not decide to buy.

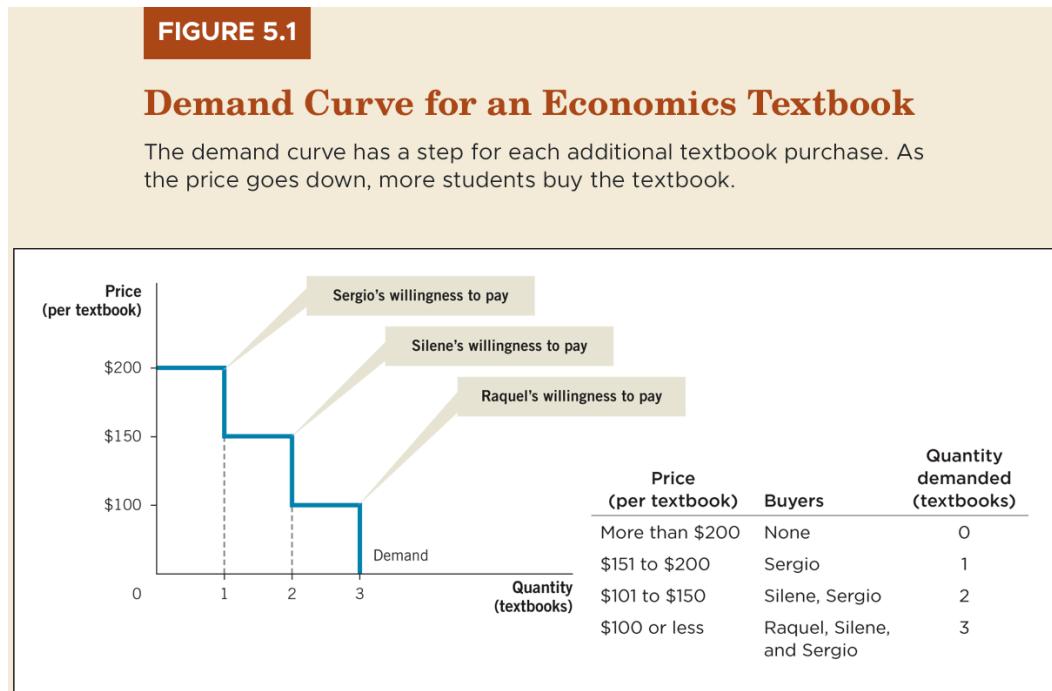
The willingness to pay is the maximum price a consumer will pay for a good or service; the willingness to pay is also known as **reservation price**. In an auction or negotiation, the willingness to pay or reservation price, is beyond which the consumer decides to walk away from the transaction.

Using demand curve to illustrate consumer surplus

In the previous section, we discussed consumer surplus as a dollar amount; we can illustrate it graphically with a demand curve. **Figure 5.1 shows the demand curve drawn from the data in Table 5.1.** Notice that the curve looks like a staircase with three steps — one for each additional market demand curve corresponds to a specific number of units sold.

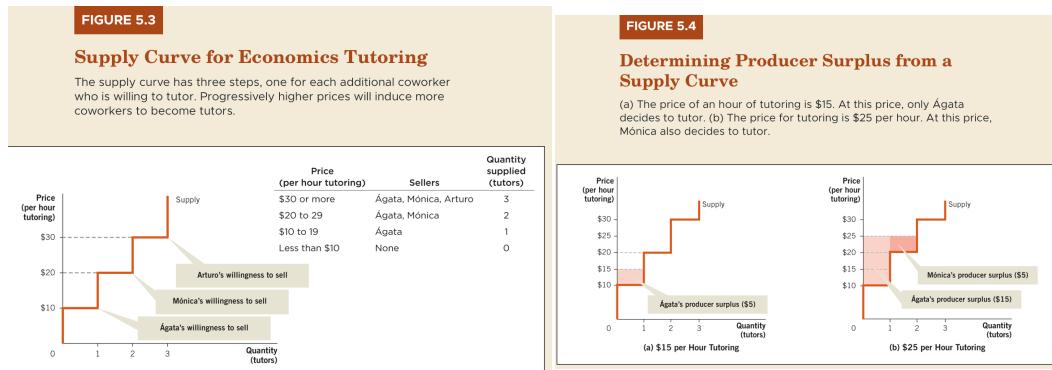
Producer Surplus

Producers surplus is the difference between the willingness to sell a good or service the seller receives. How do producers determine their willingness to sell? They must consider two factors: the direct costs of producing the good and the indirect costs, or opportunity costs.



Sellers also benefit from market transactions; **the willingness to sell** is the minimum price a seller will accept to sell a good or service.

Using supply curves to illustrate producer surplus



When is a market efficient?

We have seen how consumers benefit from lower prices and how producers benefit from higher prices. When we combine the concepts of consumer and producer surplus, we can build a complete picture of the welfare of the buyers and sellers. Adding consumer and producer surplus gives us **total surplus**, also known as **social welfare**, because it measures the well-being of all participants in a market, absent any government intervention. Total surplus is the best way economists have to measure the benefits markets create. The demand and supply curves in this

section are drawn as straight lines (as opposed to the stair steps we have seen so far), since we will now assume there are a large number of buyers and sellers in each market.

When an allocation of resources maximizes total surplus, the result is said to be efficient. To think about why the market creates the largest possible total surplus or social welfare, it is important to recall how the market allocates resources. Consumers who are willing to pay more than the market equilibrium price will buy the good because they will enjoy the consumer surplus. Producers who are willing to sell the good for less than the market equilibrium price will enjoy the producer surplus.

The Efficiency-Equity Debate

When modeling behavior, economists assume that participants in a market are rational decision-makers. That is, we assume that producers will always operate in the region of the triangle representing producer surplus, and that consumers will always operate in the region of the triangle representing consumer surplus. In other words, for the market to work efficiently, voluntary instances of consumer loss must be rare; we assume that self-interest helps to ensure that all participants benefit from the exchange.

However, the fact that both parties benefit from an exchange does not mean that both parties benefit equally. Economists are also interested in the distribution of gains; equity refers to the fairness of the distribution of benefits among the member of a society. In a world where no one cared about equity, only efficiency would matter, and no particular division would be preferred.

Why do taxes create deadweight loss in otherwise efficient markets?

Taxes provide many benefits; taxes help to pay for many of society's needs — public transportation, schools, police, the court system, and the military. Most of us take these services for granted, but without taxes it would be impossible to pay for them; how much do all of these services cost? When you add all the federal, state, and local government budgets in the United States, you get over \$6 trillion a year in taxes!

Spending tax dollars incurs opportunity costs because the money could have been used in other ways. In this section, we use the concepts of consumer and producer surplus to explain the effect of taxation on social welfare and market efficiency. Taxes come in many sizes and shapes: there are taxes on personal income, payroll, property, corporate profits sales and inheritance... fortunately, we do not have to examine the entire tax code all at once. In the pages that follow, we explore the impact of taxes on social welfare by looking at one of the simplest taxes; ***the excise tax***.

Tax incidence

Economists want to know how taxes affect the choices that consumers and producers make; when a tax is imposed on an item, do buyers switch to alternative goods that are not taxed? How do producers respond when the products they sell are taxed? Since taxes cause prices to rise, they can affect how much of a good or service is bought and sold. This outcome is especially evident with **excise taxes**, which are taxes levied on a particular good or service — but because we can isolate changes in consumer behavior that result from taxes on one item, excise taxes help us understand the overall effect of a tax.

In looking at the effect of a tax, economists are also interested in the **incidence of taxation**, which refers to the burden of taxation on the party who pays the tax. To understand this idea, consider a \$1.00 tax on milk purchases; we consider two cases: a tax placed directly on buyers and a tax placed directly on sellers.

Deadweight Loss

Recall that economists measure economic efficiency by looking at total consumer and producer surplus. We have seen that a tax raises the total price consumers pay and lowers the net price producers receive. For this reason, taxes reduce the amount of economic activity. **Deadweight loss** is the decrease in economic activity by market distortions, such as taxes. In the next three sections, we examine how differences in the price elasticity of demand lead to varying amounts of deadweight loss; the tax is placed on the seller, and we evaluate what happens when the demand curve is perfectly inelastic, somewhat elastic, and perfectly elastic.

Tax Revenue and deadweight loss when demand is inelastic

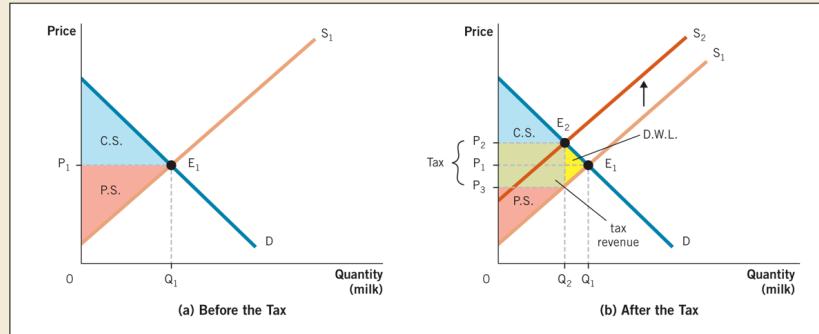
In Chapter 4, we saw that necessary goods and services — for example, water, electricity, and phone service — have highly inelastic demand. These goods and services are often taxed. For example, consider all the taxes associated with your cell phone bill: sales tax, city tax, county tax, federal excise tax, and annual regulatory fees. In addition, many companies add surcharges, including activation fees, local-number portability fees, telephone number pooling charges, emergency 911 service, directory assistance, telecommunications relay service surcharges, and cancellation fees. Cell phone providers and government agencies take advantage of consumers' strongly inelastic demand by tacking on these extra charges.

There are two reasons why the government may favor excise taxes on

FIGURE 5.10

A Tax on Products with More Elastic Demand

(a) Before the tax, the consumer enjoys the consumer surplus (C.S.) shaded in blue, and the producer enjoys the producer surplus (P.S.) shaded in red. (b) A tax on a good for which demand and supply are both somewhat elastic will cause a transfer of welfare from consumers and producers to the government, the revenue shown as the green rectangle. It will also create deadweight loss (D.W.L.), shaded in yellow, because the quantity bought and sold in the market declines (from Q_1 to Q_2).

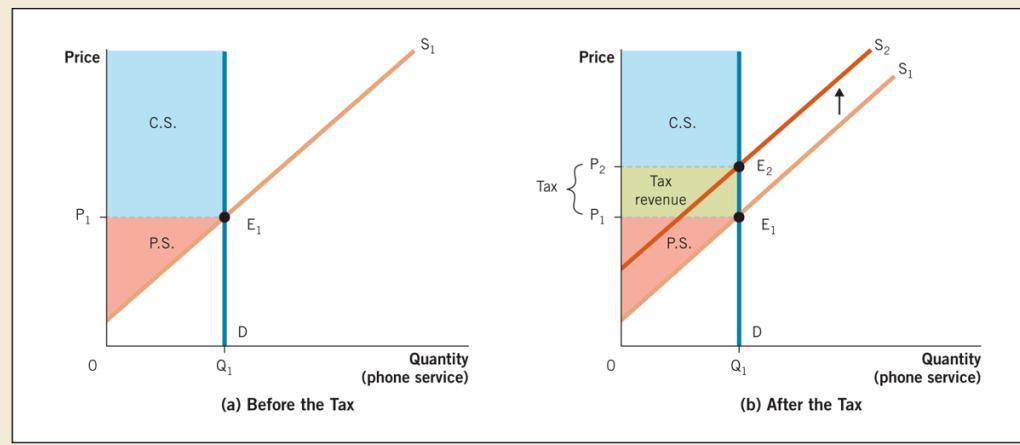


goods with almost perfectly (or highly) inelastic demand. One, because these goods do not have substitutes, the tax will not cause consumers to buy less; thus, the revenue from the tax will remain steady. Second, because the number of transactions or quantity demanded (Q_1) remains constant, there will be no deadweight loss.

FIGURE 5.9

A Tax on Products with Almost Perfectly Inelastic Demand

(a) Before the tax, the consumer enjoys the consumer surplus (C.S.) shaded in blue, and the producer enjoys the producer surplus (P.S.) shaded in red. (b) After the tax, the incidence, or the burden of taxation, is borne entirely by the consumer. A tax on a good with perfectly inelastic demand, such as phone service, represents a transfer of welfare from consumers to the government, as reflected by the reduced size of the blue rectangle in (b) and the creation of the green tax revenue rectangle between P_1 and P_2 .



Tax revenue and deadweight loss when demand is more elastic

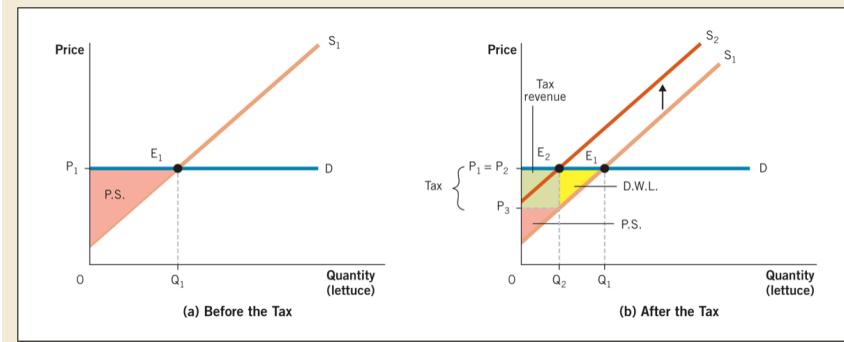
Tax revenue and deadweight loss when demand is more highly elastic

FIGURE 5.11

A Tax on Products with Highly Elastic Demand

$$\text{area of a triangle} = \frac{1}{2} \times \text{base} \times \text{height}$$

(a) Before the tax, the producer enjoys the producer surplus (P.S.) shaded in red. (b) When consumer demand is highly elastic, consumers pay the same price after the tax as before. But they are worse off because less is produced and sold; the quantity produced moves from Q_1 to Q_2 . The tax revenue is the green rectangle between P_1 and P_2 , which is smaller than the yellow triangle. The total surplus, or efficiency of the market, is much smaller than before. The size of the tax revenue (shaded in green) is also noticeably smaller in the market with highly elastic demand.



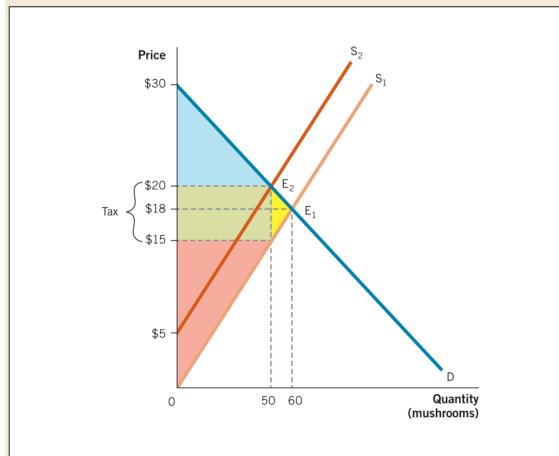
Interaction of demand elasticity and supply elasticity

The incidence of a tax is determined by the relative steepness of the demand curve compared with the supply curve. When the demand curve is steeper (more inelastic) than the supply curve, **consumers bear more of the incidence of the tax**. When the supply curve is steeper (more inelastic) than the demand curve, **suppliers bear more of the incidence of the tax**. Also, **whenever the supply and/or demand curves are relatively steep, deadweight loss is minimized**.

FIGURE 5.12

A Realistic Example

A \$5-per-pound tax is placed on mushroom suppliers, driving the equilibrium price up from E_1 (\$18) to E_2 (\$20). Notice that the price rises by only \$2. Consumers therefore pick up \$2 of the \$5 tax and the seller must pay the remaining \$3. Therefore, most of the incidence is borne by the seller.



TO DETERMINE THE AREA OF THE TRIANGLE FOR DEADWEIGHT LOSS:

TO DETERMINE THE AREA OF THE RECTANGLE FOR TAX REVENUE

Balancing Deadweight loss and Tax revenues

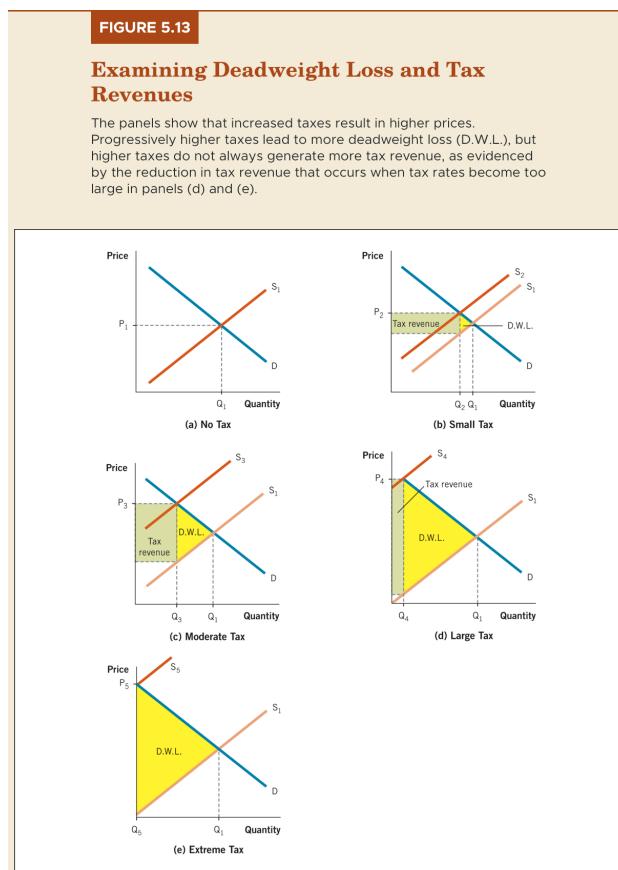
Up to this point, we have kept the size of the tax increase constant, doing so enabled us to examine the impact of the elasticity of demand and supply on deadweight loss and tax revenues — but what happens when a tax is high enough to significantly alter consumer or producer behavior? In this section, we consider how consumers responds to taxes of different sizes, and we determine the relationship between the size of a tax, the deadweight loss, and tax revenues.

Figure 5.13 shows the market response to a variety of tax increases; the five panels in the figure begin with a reference point, **panel (a)**, where no tax is levied, and

progress toward **panel (e)**, where the tax rate becomes so extreme that it curtails all economic activity

As taxes rise, so do prices — you can trace this price rise from **panel (a)**, where there is no tax and the price is P_1 , all the way to **panel (e)**, where the extreme tax causes the price to rise to P_5 . At the same time, deadweight loss (D.W.L.) also rises — you can see this increase by comparing the sizes of the yellow triangles — the trade-off is striking. Without any taxes, deadweight loss does not occur, but soon as taxes are in place; the market equilibrium quantity demanded begins to decline, moving from Q_1 to Q_5 . As the number of transactions (quantity demanded) declines, the area of deadweight loss rapidly expands.

When taxes are smaller, as in **panel (b)**, the tax revenue (green rectangle) is largely relative to the deadweight loss (yellow triangle). However, as we progress through the panels, this relationship slowly reverses. In **panel (c)**, the size of the tax revenue remains larger than the deadweight loss; in **panel (d)**, however, the magnitude of the deadweight loss is far greater than the tax revenue — the size of tax **panel (d)** is creating a



significant cost in terms of economic efficiency. Finally, **panel (e)**, shows an extreme case in which all market activity ceases as a result of the tax; because nothing is produced and sold, there is no tax revenue.

Conclusion

The government largely taxes goods that have inelastic demand, which means that firms are able to transfer most of the tax incidence to consumers through higher prices.

In the first part of this chapter, we learned that society benefits from unregulated markets because they generate the largest possible total surplus. However, society also needs the government to provide an infrastructure for the economy. The taxation of a specific goods and services gives rise to a form of market failure called deadweight loss, which reflects reduced economic activity. Thus, any intervention in the market requires a deep understanding of how society will respond to the incentives created by legislation; in addition, unintended consequences can affect the most well-intentioned tax legislation and, if the process is not well thought through, can cause inefficiencies with far-reaching consequences. None of this means that taxes are undesirable, rather, society must balance (1) the need for tax revenues and the programs those revenues help fund with (2) trade-offs in the market.

Price Controls

Econ 304K: Chapter 6

Price Controls causes Surpluses and Shortages

Price controls are not a new idea — the first recorded attempt to regulate price was 4,000 years ago in ancient Babylon, when King Hammurabi decreed how much corn a farmer could pay for a cow. Similar attempts to control prices occurred in ancient Egypt, Greece, and Rome — each attempt ended badly. History has shown us that price controls generally do not work, why? Because they disrupt the normal functioning of the market; yet, they are tried time and again, as seen in the frequent calls for price controls during the COVID-19 pandemic. By the end of this chapter, you will understand why price controls are rarely the win-win proposition that legislators often claim. To help you understand why price controls lead to disequilibrium in markets, this chapter focuses on the two most common types of price controls: **price ceilings** and **price floors**.

Big Questions

- When do price ceilings matter?
 - A price ceiling is legally imposed maximum price. When the price is set below the equilibrium price, the price quantity demanded will exceed the quantity supplied. The result is a shortage, price ceiling matter when they are binding (below the equilibrium price)
- What effects do price ceilings have on economic activity?
 - Price ceilings create two unintended consequences: a smaller quantity supplied of the good (Q_S) and a higher price for consumers who turn to illegal markets
- When do price floors matter?
 - A price floor is a legal imposed minimum price. The minimum wage is an example of a price floor. If the minimum wage is set above the equilibrium wage, a surplus of labor will develop, however, if the minimum wage is non-binding, it will have no effect on the market wage. Thus, price floors matter when they are set above the equilibrium price
- What effects do price floors have on economic activity?

- Price floors lead to many unintended consequences, including surpluses, the creation of illegal markets, and artificial attempts to bring the market back into balance. For example, proponents of a higher minimum wage are concerned about finding ways to alleviate the resulting surplus of labor or unemployment.

When do price ceilings matter?

Price controls attempt to set prices through government regulations in the market; in most cases, and certainly in the United States, price controls are enacted to ease perceived burdens on society. **A price ceiling** creates a legally established maximum price for a good or service; in the next section, we consider what happens when a price ceiling is in place. *Price ceilings create many unintended effects that policymakers rarely acknowledge.*

Understanding Price Ceilings

To understand how price ceilings work, suppose that most prices are rising as a result of **inflation**, an overall increases in prices. The government is concerned that people with low incomes will not be able to afford to eat; to help the disadvantage, legislators pass a law stating that no one can charge more than \$0.50 for a loaf of bread (**Note that this price ceiling is about one-third the typical price of a loaf of generic white bread**). Does the new law accomplish its goal? What happens?

The law of demand tells us that if the price drops, the quantity that consumers demand will increase. At the same time, the law of supply tells us that the quantity supplied will fall because producers will be receiving lower profits for their efforts. This combination of increased quantity demanded and reduced quantity supplied will cause a shortage of bread.

On the demand side, consumers will want more bread than is available at the legal price; there will be long lines for bread, and many people will not be able to get the bread that they want. On the supply side, producer will look for ways to maintain their profits. They can reduce the size of each loaf they produce; they can also use cheaper ingredients, thereby lowering the quality of their product and they can stop making fancier variants.

In addition, illegal markets will develop; in other words, sellers will go “underground” and charge higher price to customers who want bread.

Table 6.1

Price Ceiling on Bread		
Questions	Answer(s) (Explanations)	Result
Will there be more bread or less bread for sale?	Consumers will want to buy more because the price is lower (the law of demand), but producers will manufacture less (the law of supply). The net result will be a shortage of bread.	Empty Shelves
Will the size of a typical loaf change?	Because the price is capped at \$0.50 per loaf, manufacturers will try to maintain profits by reducing the size of each loaf.	No more giant loaves
Will the quality change?	Because the price is capped, producers will use cheaper ingredients, and many expensive brands and varieties will no longer be profitable to produce. Thus the quality of available bread decline.	Focaccia bread will disappear
Will the opportunity cost of finding bread change?	The opportunity cost of finding bread will rise. Consumers will spend significantly resources going from store to store to see if a bread shipment has arrived and waiting in line for a chance to get some.	Bread lines will become the norm
Will people have to break the law to buy bread?	Because bread will be hard to find and people will still need it, an illegal market will develop. Those selling and buying on these markets will be breaking the law.	Bread dealers in illegal markets will help reduce the shortage

The Effect of Price Ceilings

Now that we have some understanding of how a price ceiling works, we can transfer that knowledge into the supply and demand model for a deeper analysis of how price ceilings affect the market. To explain when price ceilings matter in the short run, we examine two types of price ceilings: **Non-binding and binding**. Both are set by law, but only one actually makes a difference to prices.

Non-biding price ceilings

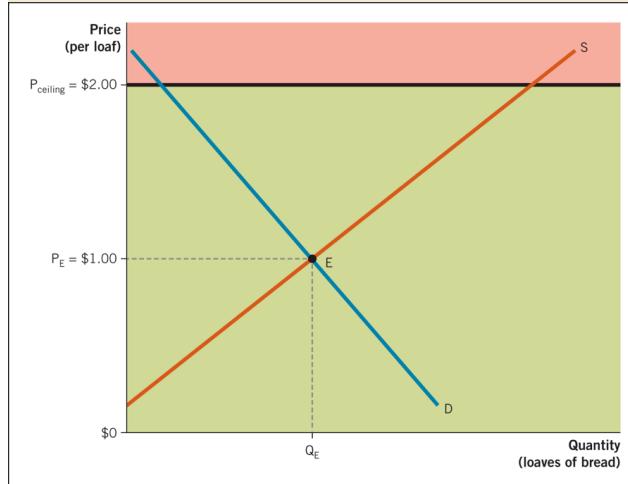
Figure 6.1

The effect of a price ceilings depends on the level at which it is set relative to the equilibrium price. When a price ceiling is above equilibrium, we say it is **non-binding**. Price above the price ceiling are illegal; but because the market equilibrium occurs in the green area, the price ceiling does not influence the market: it is non-binding. As long as the equilibrium price remains below the price ceiling, price will continue to be regulated by supply and demand.

Figure 6.1

A Nonbinding Price Ceiling

The price ceiling (\$2.00) is set above the equilibrium price (\$1.00). Because market prices are set by the intersection of supply (S) and demand (D), as long as the equilibrium price is below the price ceiling, the price ceiling is nonbinding and has no effect.



Binding Price Ceilings

Figure 6.2

When a price ceiling is below the market price, it creates a binding constraint that prevents supply and demand from clearing the market. In Figure 6.2, the price ceiling for bread is set at \$0.50 per loaf because \$0.50 is well below the equilibrium price of \$1.00, the price ceiling is **binding**. Notice that at a price of \$0.50, the quantity demanded (Q_D) is greater than the quantity supplied (Q_S); in other words, a shortage exists. Shortages typically causes prices to rise, but imposed price ceiling prevents that from happening. A price ceiling of \$0.50 allows

only the prices in the green area; the market cannot reach equilibrium point E at \$1.00 per loaf because it is located above the price ceiling, in the red area

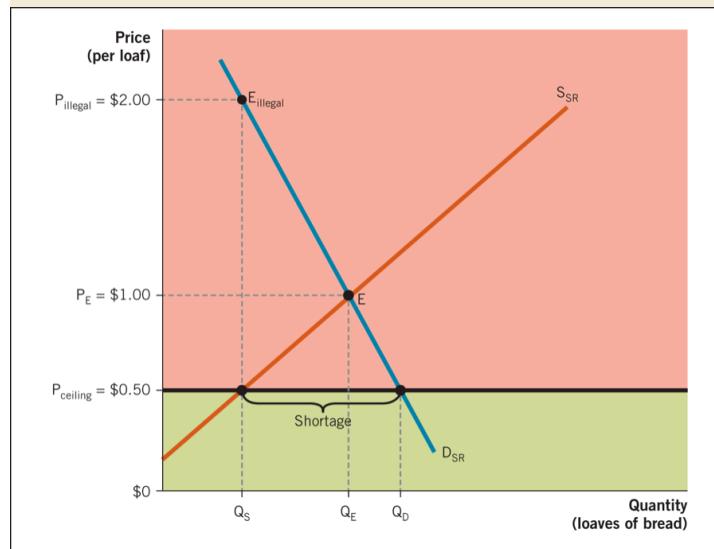
The illegal price is also set by supply and demand, because prices above \$0.50 are illegal, sellers are unwilling to produce more than Q_S . Because a shortage exists, this market will form in response to the shortage; here, purchasers can illegally resell what they have just bought at \$0.50 for far more than what they just paid.

Because the supply of legally produced bread is Q_S , the intersection of the vertical dashed line reflects Q_S with the demand curve D_{SR} at point $E_{illegal}$ establishes a market price ($P_{illegal}$) at \$2.00 per loaf for illegally sold bread. The market price is substantially more than the market price equilibrium (P_E) of \$1.00. As a result, the price in these illegal markets eliminates the shortage caused by the price ceiling, however, the price ceiling has created two unintended consequences: smaller

Figure 6.2

The Effect of a Binding Price Ceiling in the Short Run

A binding price ceiling prevents sellers from increasing the price and causes them to reduce the quantity they offer for sale. As a consequence, prices no longer signal relative scarcity. Consumers desire to purchase the product at the price ceiling level, which creates a shortage in the short run (SR); many will be unable to obtain the good. As a result, those who are shut out of the market will turn to other means to acquire the good, establishing an illegal market for it at a higher, illegal price.



quantity of bread supplied (Q_S is less than Q_E), and higher price for those who purchase it there.

Price Ceilings in the long run

In the long run, supply and demand became more elastic, or flatter; recall that when consumers have additional time to make choices, they find more ways to avoid high-priced goods and more ways to take advantage of low prices. Additional time also gives producers the opportunity to produce more when prices are high and less when prices are low. In this section, we consider what happens if a binding price ceiling on bread remain in effect for a long time; we have already observed that binding price ceilings create shortages and illegal markets in the short run. Are the long-run implications of price ceilings more problematic than the short-run implications?

Figure 6.3 shows the result of a price ceiling that remains in place for a long time; here the supply curve is more elastic than its short-run counterpart in Figure 6.2. The supply curve is flatter because producers respond in the long run by producing less bread and converting their facilities to make similar products that are not subject to price controls and that will bring them a reasonable return on their investments — for example, bagels and rolls. Therefore, in the long run the quantity supplied (Q_S) shrinks even more.

The demand curve is also more elastic (flatter) in the long run; in the long run, more people will attempt to take advantage of the price ceiling by changing their eating habits to consume more bread. Even though consumers will often find empty shelves in the long run, the quantity demanded of cheap bread will increase; the flatter demand curve means that consumers are more flexible. As a result, the quantity demanded (Q_D) expands and bread is harder to find at \$0.50 per loaf. The shortage will become acute (compare Figure 6.3 with Figure 6.2) that consumers will turn to bread substitutes, like bagels and rolls, that are more plentiful because they're not price controlled.

Increased elasticity on the part of producers and consumers magnifies the unintended consequences we observed in the short run. Therefore, products subject to a price ceiling become progressively harder to find in the long run and the illegal market continues to operate. However, in the long run our bread consumers will choose substitutes for this expensive bread, leading to somewhat lower prices.

What effects do price ceilings have on economic activity?

We have seen the logical repercussions of hypothetical price ceiling on bread and the incentives it creates. Now let's use supply and demand analysis to examine two real-world price ceilings: **rent control and price gouging laws.**

Rent control

Under **rent control**, a local government caps the price of apartment rentals to keep housing affordable; while this goal may be laudable, rent control doesn't work. In fact, it doesn't help the low-income residents of the city find affordable housing or gain access to housing at all; in addition, these policies contribute to dangerous living conditions

Figure 6.4

Rent Control in the Short Run and the Long Run

Because rent-controlled apartments are vacated slowly, the quantity supplied contracts in the short run and the supply curve become more elastic in the long run. Demand also becomes more elastic in the long run, causing the quantity demanded to rise. The combination of fewer units available to rent and more consumers looking to find rent-controlled units leads to a larger shortage in the long run.

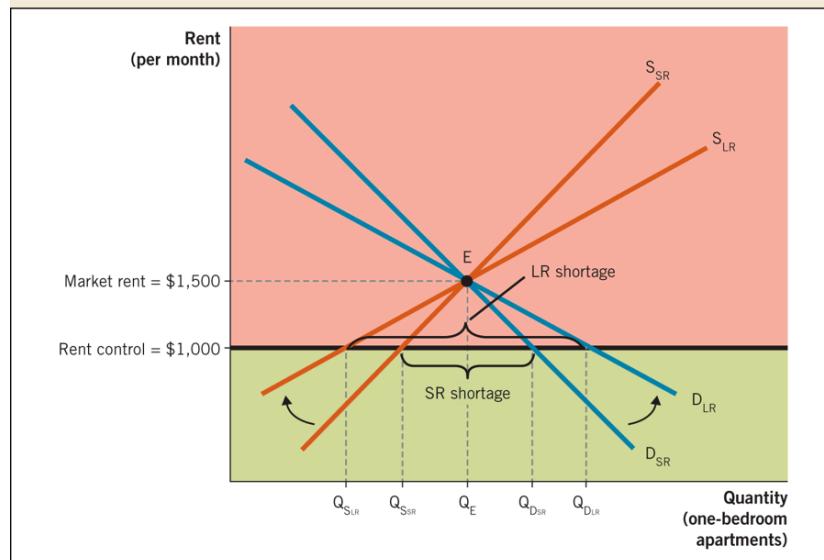


Figure 6.4 shows why rent control fails; as with any binding price ceiling, rent control causes a shortage because the quantity demanded in the short run ($Q_{D_{SR}}$) is greater than

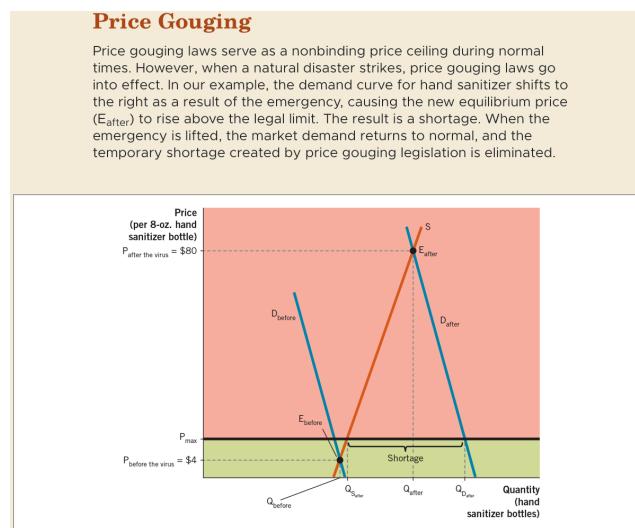
the quantity supplied in the short run (Q_{SSR}). The combination of fewer available units and more consumers looking for rent-controlled units leads to a larger shortage in the long run.

Price Gouging

Another kind of price control, **price gouging laws**, place a temporary ceiling on the prices that sellers can charge during times of emergency. Over 30 U.S states have laws against price gouging; the intent is to keep people in desperate circumstances from being charged an arm and a leg for basic, essential goods. Like all price controls, price gouging laws have unintended consequences.

For a time, then, the normal ability of supply and demand to ration the available hand sanitizer is wiped away because more people demand hand sanitizer after the emergency than before it, those who don't get to the store soon enough are out of luck. When the emergency is lifted and the market returns to normal, the temporary shortage created by price gouging laws will be eliminated. This does not mean that price gouging laws are automatically bad policy; after all, there's a limit to what we are willing to allow in the free market, but from an economic point of view, price going laws are mostly downside.

Figure 6.5



When do Price Floors matter?

A **price floor** creates a legally established minimum price for a good or service; the minimum wage law is an example for a price floor in the market for labor. Like price ceilings, price floors create many unintended effects that policymakers rarely acknowledge, however, unlike price ceilings, price floors result from political pressure of suppliers to keep prices high. Most

consumers prefer lower prices when they shop, so the idea of a law that keeps the prices high may sound like a bad one to you; however, if you are selling a product or service, you might think that legislation to keep prices high is a very good idea.

In this section, we follow the same progression that we did with price ceilings. We begin with a simple thought experiment. Once we understand how price floors work, we use supply and demand analysis to examine the short- and long-run implications for economic activity.

The Effect of Price Floors

We have seen that price floors create unintended consequences. Now we will use the supply and demand model to analyze how price floors affect the market. We look at the short run first.

Non-binding price floors

Like price ceilings, price floors can be binding or non-binding. As you can see in Figure 6.6, at \$2 the price floor is below equilibrium price (P_E), so the price floor is non-binding, because the actual market price is above the legally established minimum price (P_{floor}), the price floor does not prevent the market from reaching equilibrium at Point E. Consequently, the price floor has no impact on the market, as long as the equilibrium price remains above the price floor, price is determined by supply and demand.

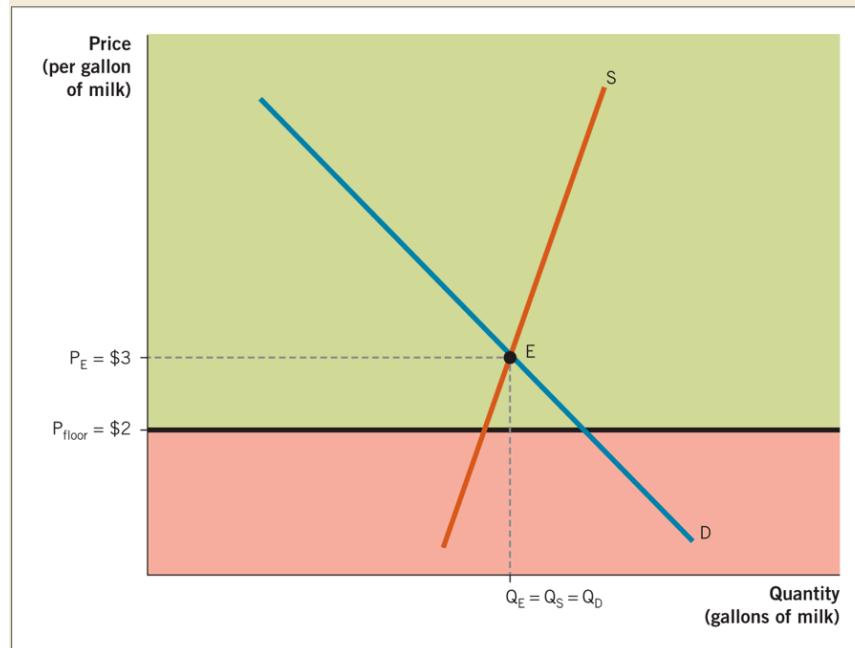
Table 6.2

A Price Floor on Milk		
Questions	Answer(s) (Explanations)	Result
Will the quantity of milk for sale change?	Consumers will purchase less because the price is higher (the law of demand), but producers will manufacture more (the law of supply). The net result will be a surplus of milk	There will be a surplus of milk
Would producers sell below the price floor?	Yes. A surplus of milk would give sellers a strong incentive to undercut the price floor to avoid having to discard leftover milk	Illegal discounts will help reduce the milk surplus.
Will dairy farmer be better off?	Not if they have trouble selling what they produce.	There might be a lot of spoiled milk.

Figure 6.6

A Nonbinding Price Floor

Under a nonbinding price floor, price is regulated by supply and demand. Because the price floor (\$2) is below the equilibrium price (\$3), the market will voluntarily charge more than the legal minimum. Therefore, this nonbinding price floor will have no effect on sales and purchases of milk.

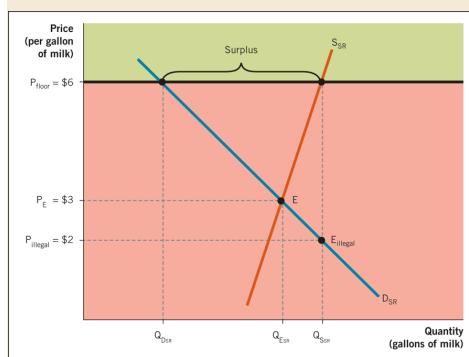


Binding Price Floors

Figure 6.7

A Binding Price Floor in the Short Run

A binding price floor creates a surplus, which has two unintended consequences: (1) a smaller quantity demanded than the equilibrium quantity ($Q_{D,\text{SR}} < Q_{E,\text{SR}}$) and (2) a lower, illegal price to eliminate the glut of the product.



For a price floor to have an impact on the market, it must be set above the market equilibrium price. In this case, it is a binding price floor; with a binding price floor, the quantity supplied will exceed the quantity demanded.

Price Floors in the long run

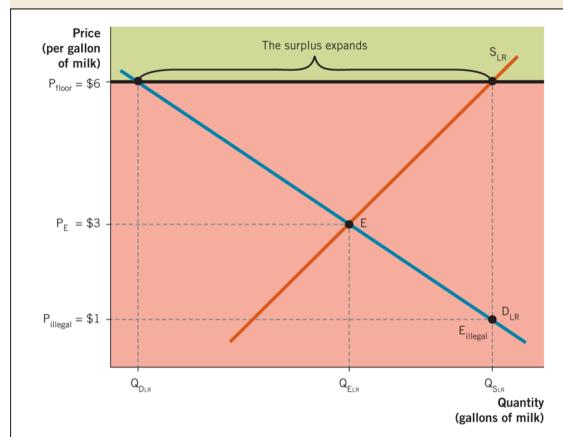
Once price floor legislation is passed, it can be politically difficult to repeal. What happens if a binding price floor on milk stays in effect for a long time? To answer that question, we need to consider elasticity. We have already observed that in the short run, binding price ceilings cause shortages and that illegal markets follow.

What happens to supply? In the long run, producers are more flexible and therefore supply is more elastic. The pool of potential milk producers rises as other closely related businesses retool their operations to supply more milk. The flatter supply curve in Figure 6.8 reflect flexibility; as a result, (Q_{SLR}) expands and becomes much larger than in Figure 6.7. The increased elasticity on the part of both producers and consumers makes the surplus larger in the long run and magnifies the unintended consequences we observed in the short run.

Figure 6.8

The Effect of a Binding Price Floor in the Long Run

When a price floor is left in place over time, supply and demand both become more elastic. The result is a larger surplus ($Q_{SLR} > Q_{DLR}$) in the long run. Because sellers are unable to sell all that they produce at \$6 per gallon, an illegal market develops to eliminate the glut of milk.



What effects do Price floors have on economic activity?

We have seen the logical repercussion of a hypothetical price floor on milk and the incentives it creates. Now let's use supply and demand analysts to examine two real-world price floors: ***minimum wage laws and sugar subsidies***.

The Minimum Wage

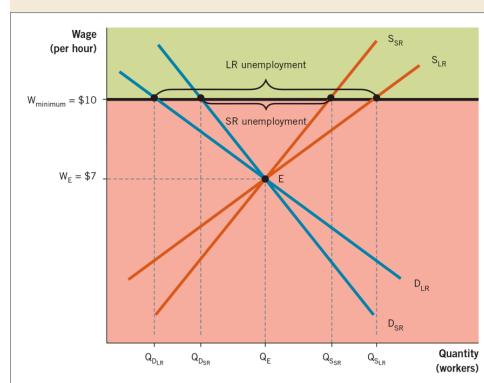
The **minimum wage** is the lowest hourly wage rate that firms may legally pay their workers. Minimum wage workers can be skilled or unskilled and experienced or inexperienced. The common thread is that these workers, for a variety of reasons, lack better prospects.

A minimum wage functions as a floor price; Figure 6.9 shows the effect of a binding minimum wage. Note that the wage, or the cost of labor, on the y-axis (\$10 per hour) is the price that must be paid. However, the market equilibrium wage (\$7) or W_E , is below the minimum wage; the minimum wage prevents the market from reaching W_E at E (the equilibrium point) because only the wages in the green area are legal. The minimum wage raises the cost of hiring workers. Therefore, a higher minimum wage will lower the quantity of labor demanded; at the same time, firms will look for ways to substitute capital for workers. As a result, a binding minimum wage results in unemployment in the short run because ($Q_{SSR} > Q_{DSR}$).

Figure 6.9

Price Floors and a Binding Minimum Wage Market in the Short Run and Long Run

A binding minimum wage is a price floor above the current equilibrium wage, W_E . At \$10 per hour, the number of workers willing to supply their labor (S_{SR}) is greater than the demand for workers (D_{SR}). The result is a surplus of workers (which we recognize as unemployment). Because the supply of workers and demand for workers both become more elastic in the long run, unemployment expands ($Q_{S_{LR}} > Q_{D_{LR}}$).



Businesses generally want to keep costs down, so in the long run they will try to reduce the amount they spend on labor. They often might replace workers with machinery, shorten work hours, offer reduced customer service, or even relocate to countries that do not have minimum wage laws. As we move past the short run, more people will attempt to take advantage of higher minimum wages. Like firms, workers will adjust to the higher minimum wage over time; some workers who might decided to go to school full-time or remain retired or who simply want some extra income will enter the labor market because of minimum wage is now higher. As a result, minimum wage jobs will become progressively harder to find and unemployment will increase, the irony is that in the long run, the minimum wage, just like any other price floor, has created two unintended consequences: a smaller demand for workers by employers (QDLR is significantly less than Q_E) and a larger supply of workers (QSLR) looking for jobs.

Proponents of minimum wage legislation are aware that it often creates unemployment; to address the problem, they support investment in training, education, and the creation of government jobs program to provide work opportunities. While jobs programs increase the number of minimum wage jobs, training, and additional education enable workers to acquire skills needed for jobs that pay more than the minimum wage. Economists generally believe that education and training programs have longer-lasting benefits to society as a whole because they enable workers to obtain better paying jobs on a permanent basis.

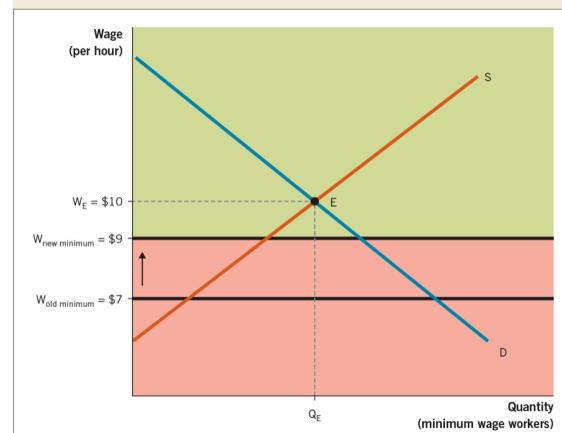
The minimum wage is sometimes non-binding

Most people believe that raising the minimum wage is a simple step that the government can take to improve the standard of living of the working poor. However, in some places the minimum wage is non-binding and therefore has no impact on the market. Why would we have minimum wage if it's largely unbinding?

To help us answer this question, consider the two non-binding minimum wage rates (\$7 and \$9) shown in the Figure 6.10. A minimum wage of \$7 per hour is far below the equilibrium wage of \$10 (W_E), so at that point supply and demand push the

A Nonbinding Minimum Wage

An increase in the minimum wage from \$7 to \$9 remains nonbinding. Therefore, it will not change the demand for labor or the unemployment rate. If the minimum wage rises above the market wage, unemployment will occur.



equilibrium wage up to \$10; suppose that politicians decide to raise the minimum wage to \$9. This new minimum wage of \$9 would remain below the market wage, so there would be no impact on the labor market for workers who are willing to accept the minimum wage. Therefore, an increase in minimum wage from \$7 to \$9 an hour will not create unemployment; unemployment will occur only when the minimum wage rise above \$10

Conclusion

The policies presented in this chapter — rent control, price gouging laws, the minimum wage, and agricultural price controls — create unintended consequences. Attempts to control prices should be viewed cautiously. When the price signal is suppressed through a binding price floor or ceiling, the market's ability to allocate goods and services is diminished, surpluses and shortages develop and expand through time, and obtaining goods and services becomes difficult.

The role of markets in society has many layers, and we've only just begun our analysis. In the next chapter, we consider two cases — externalities and public goods — in which the unregulated market produces an output that is not socially desirable.

Market Inefficiencies: Externalities and Public Goods

Econ 304K: Intro to Microeconomics

Should we eliminate all pollution?

In many preceding chapters, we saw that markets provide many benefits and they typically work because participants pursue their own self-interests, but sometimes markets need a helping hand. To explain why markets do not always operate efficiently, this chapter explores two important concepts: externalities and the difference between private and public goods.

Big Questions

- What are externalities, and how do they affect markets?
- What are private goods and public goods?
- What are the challenges of providing non-excludable goods?

What are externalities and how they Affect Markets?

We have seen that buyers and sellers benefit from trade; but what about the effects trade might have on bystanders? **Externalities**, or the costs and benefits of a market activity that affect a third party, often lead to undesirable consequences. **Market failure** occurs when there is an inefficient allocation of resources in a market; **externalities are type of market failure**.

For a market to work as efficiently as possible, two things must happen; first, each participant must be able to evaluate the **internal costs** of participation — the costs that only the individual participant pays. **External costs** are the costs of a market activity imposed on people who are not participants in that market. In the case of driving, the congestion and pollution our cars create are external costs. Economists define **social costs** as the sum of internal costs and external costs of a market activity.

In this section, we consider some of the mechanisms that encourage consumers and producers to account for the social costs of their actions.

The Third-Party Problem

An externality exists whenever an internal cost (or benefit) diverges from a social cost (or benefit). A **third-party problem** occurs when those not directly involved in a market activity experience negative or positive externalities. If a third party is adversely

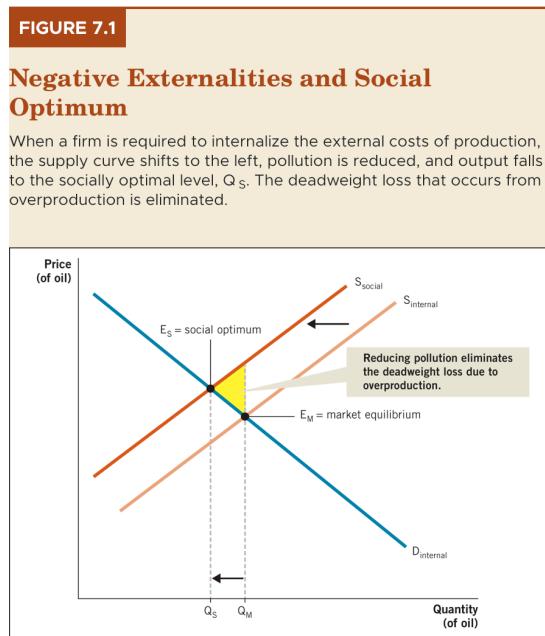
affected, the externality is negative. For examples, a negative externality occurs when the number of vehicles on the roads causes air pollution. Negative externalities present a challenge to society because it is difficult to make consumers and producers take responsibility for the full costs of their actions.

In general, society would benefit if all consumers and producers considered both the internal and external costs of their actions. Given most people feel this expectation is unlikely to happen, government design policies that create incentives for firms and people to limit the amount of pollution they emit.

Correcting for negative externalities

In this section, we explore ways to correct for negative externalities. To do so, we use supply and demand analysis to understand how the externalities affect the market. Let's begin with supply and compare the difference between the market force produce and what is the best for society in the case of an oil refinery. A refinery converts crude oil to gasoline; this complex process generates many negative externalities, including the release of pollutants into the air and the dumping of waste by-products.

The **social optimum** is the price and quantity combination that would exist if there were no externalities. The supply curve $S_{internal}$ represents how much the oil refinery will produce if it does not have to pay for the negative consequences of its activity. In this situation, the market equilibrium, E_M , accounts only for the internal costs of production.



When a negative externality occurs, the government may be able to restore the social optimum by requiring externality-causing market participants to pay for the cost of their actions. In this case, there are three potential solutions. First, the refinery can be required to install pollution abatement equipment or change in production techniques to reduce emissions and waste by-products. Second, the government can levy a tax on the refinery as a disincentive to produce. Finally, the government can require the firm to pay for any environmental damage it causes. Each solution forces the firm to **internalize** the externality, meaning that the firm must take into account the external costs (or benefits) to society that occur as a result of its actions.

Having to pay the costs of imposing pollution on others reduces the amount of the pollution-causing activity. This result is evident in the shift of the supply curve to S_{Social} . The new supply curve reflects a combination of the internal and external costs of producing the good; because each corrective measure requires the refinery to spend money to correct the externality and therefore increases overall costs, the willingness to sell the good declines, or shifts to the left. The result is a social optimum at a lower quantity, Q_S , than at the market equilibrium quantity demanded, Q_M . The trade-off is clear; we can reduce negative externalities by requiring producers to internalize the externality. However, doing so does not come without cost, because the supply curve shifts to the left, the quantity produced is lower and price rises. In the real world, there is always a cost.

Table 7.1

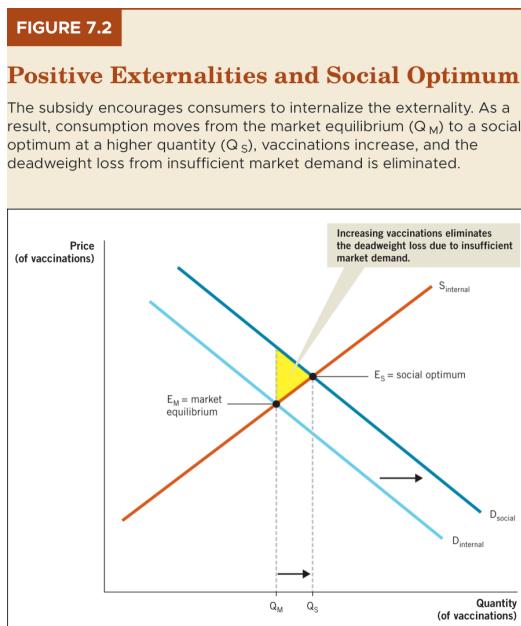
Private and Social Decision-making			
Personal decision	Social optimum	The problem	The Solution
Based on internal costs	Social costs = internal costs + external costs	To get consumers and producers to take responsibility for the externalities they create	Encourage consumers and producers to <i>internalize</i> externalities

Table 7.1 outlines the basic decision-making process that guides private and social decisions. Private decision-makers consider only their internal costs,

but society as a whole experiences both internal and external costs. To align the incentives of private decision-makers with the interests of society, we must find mechanisms that encourage the internalization of externalities.

Correcting for positive externalities

Positive externalities, such as vaccines, are the result of economic activities that have benefits for third parties. As with negative externalities, economists use supply and demand analysis to compare the efficiency of the market with the social optimum. This time we will focus on the demand curve



Markets do not handle externalities well; with a negative externality, the market produces too much of a good; but in the case of a positive externality, the market produces too little. In both cases, the market equilibrium creates deadweight loss. When positive externalities are present, the private market is not efficient because it is not fully capturing the social benefits. In other words, the market equilibrium does not maximize the gains for society as a whole. When positive externalities are internalized, the demand curve shifts outward and output rises to the socially optimal level, Q_S . The deadweight loss that results from insufficient market demand, and therefore underproduction, is eliminated.

Table 7.2 summarizes the key characteristics of positive and negative externalities and presents additional examples of each type.

Before moving on, it is worth noting that not all externalities warrant corrective measures. There are times when the size of the externality is negligible and does not justify the cost of increased regulations, charges, taxes, or subsidies that might achieve the social optimum; because corrective measures have costs, the presence of negligible externalities does not by itself imply that the government should intervene in the market

A summary of externalities		
	Negative externalities	Positive externalities
Definition	Costs borne by third parties	Benefits received by third parties
Examples	Oil refining	Flu shots prevent the spread of disease
	Traffic congestion causes all motorists to spend more time on the road waiting	Education creates a more productive workforce and enables citizens to make more informed decisions for the betterment of society
Corrective measures	Airports create noise pollution	Restored historic buildings enable people to enjoy beautiful architectural details.
	Taxes or charges	Subsidies or government provisions

What are Private Goods and Public Goods?

The presence of externalities reflects a divide between the way markets operate and the social optimum. What creates the divide? The answer is often related to property rights. **Property rights** give the owner the ability to exercise control over a resource; when property rights are not clearly defined, resources can be mistreated.

To understand why firms sometimes overlook their actions' effects on others, we need to examine the role of property rights in market efficiency. When property rights are poorly established or not enforced effectively, the wrong incentives come into play. The difference is apparent when we compare situations in which people do have property rights. Private owners have an incentive to keep their property in good repair because they bear the costs of fixing what they own when it breaks or no longer works properly. The

difference between solving the problem and ignoring it is crucial to understanding why property rights matter.

Private property

One way to minimize externalities is to establish well-defined private property rights. **Private property** provides an exclusive right of ownership that allows for the use, and especially the exchange, of property. This right creates incentives to maintain, protect, and conserve property and to trade with others.

1. The incentive to maintain property
2. The incentive to protect property
3. The incentive to conserve property
4. The incentive to trade with others

The incentives to maintain, protect, and conserve property help to ensure that owners keep their private property in good shape. The fourth incentive, to trade with others, helps to ensure that private property is held by the person with the greatest willingness to pay for it.

The Coase Theorem

In 1960, Nobel Prize-winning economist Ronald Coase argued that establishing private property rights can close the gap between internal costs and social costs. **Coase Theorem** states that if there are no barriers to negotiations and if property rights are fully specified, interested parties will bargain privately to correct externalities. As a result, the assignment of property rights under the law gives each party an incentive to internalize and externalities. If it is difficult to bargain (because the costs of reaching an agreement are too high), private parties will be unable to internalize the externality between themselves, therefore, the Coase Theorem also suggests that private solutions to externality problems are not always possible, implying a prominent role for government in solving complex externality issues.

Private and Public Goods

When we think of public goods, we think of goods provided by the government, like roads, the post office, and the military. The terms “private” and “public” typically imply ownership or production, but that is not the criterion economists use to categorize private and public goods; to understand the

difference between private and public goods, you need to know whether a good is excludable, rival in consumption, or both. An **excludable good** is one for which access can be limited to paying customers. A **rival good** is one that cannot be enjoyed by more than one person at a time.

Private Goods

A **private good** is both excludable and rival in consumption.

Public Goods

A **public good** are consumed by more than one person and it is difficult to exclude non-payers. Public goods are often underproduced because people can get them without paying for them. This causes a phenomenon called **free-rider problem**; occurs whenever people receive a benefit for what they don't pay for.

Most people agree that government should provide certain public goods for society, including national defense, the interstate highway system, and medical and science-related research to fight pandemics. In this case, public-sector provision helps to eliminate the free-rider problem and create the socially optimal level of activity.

Club goods and common-resource goods

There are two additional types of goods we have not yet introduced; because club goods and common-resource goods have characteristics of both private and public goods, the line between private provision and public provision is often blurred.

A **club good** is non-rival in consumption and excludable. A **common-resource good** is rival in consumption but excludable.

TABLE 7.3

The Four Types of Goods

		Consumption	
		Rival	Nonrival
Excludable?	Yes	<i>Private goods</i> are rival and excludable: sandwiches, watches, automobiles. 	<i>Club goods</i> are nonrival and excludable: streaming services, education, country clubs. 
	No	<i>Common-resource goods</i> are rival and nonexcludable: Alaskan king crab, a large shared popcorn at the movies, congested roads. 	<i>Public goods</i> are nonrival and nonex- cludable: street performers, national defense, tsunami warning systems. 

What are the challenges of providing non excludable goods?

Thursday, October 5, 2023

Business Costs and Production

Introduction to Economics (ECON 304K)

Do larger firms always have a cost advantage over their smaller rivals?

Walmart, the nation's largest retailer, leverages its size to get price breaks on bulk purchases from its suppliers. People commonly believe that this kind of leverage enables larger firms to operate at lower costs than smaller firms do. It is true that large firms have broader distribution networks, and they benefit from more specialization and automation compared with their smaller competitors. However, not all industries enjoy lower costs with additional sales the way retailers do. In other words, larger firms do not always have the lowest costs. Some products are especially unsuited to mass production.

We begin the chapter with an examination of costs and how they relate to production. After we understand the basics, we consider how firms can keep their costs low in the long run by choosing a scale of operation that best suits their needs. By the end of this chapter, you will appreciate the importance of cost and understand why smaller and more nimble firms are sometimes able to undercut the prices of larger franchises.

Big Questions

- How are profits and losses calculated?
 - Profits and losses are determined by calculating the difference between expenses (total cost) and total revenue
 - There are two types of profit: economic profit and accounting profit. If a business has an economic profit, its revenue is larger than the combination of its explicit and implicit costs.
 - Economists break costs into two components: explicit costs, which are easy to calculate, and implicit costs, which are hard to calculate. Because economic profit account for implicit costs, the economic profit is always less than the accounting profit
- How much should a firm produce?

- A firm should produce an output that is consistent with the largest possible economic profit.
- To maximize profit, firms must effectively combine land, labor, and capital in the right quantities
- In any short-run production process, a point of diminishing marginal product will occur at which additional units of a variable input no longer generate as much output as before. Diminishing marginal product is a result of fixed inputs (such as capital and land) in the short run.
- Marginal cost (MC) is the key variable in determining a firm's cost structure. The MC curve always leads the average total cost (ATC) and average variable cost (AVC) curves up or down.
- What costs do firms consider in the short-run and the long-run?
 - In the short run, firms consider variable and fixed costs, as well as marginal cost. Firms also consider average variable cost (AVC), average fixed cost (AFC), and average total cost (ATC).
 - With the exception of the average fixed cost (AFC) curve, which always declines, short-run cost curves are U-shaped. All variable costs initially decline due to increased specialization. At a certain point, the advantages of continued specialization give way to diminishing marginal product, and the MC, AVC, and ATC curves begin to rise.
 - Long-run costs are a reflection of scale. Firms can experience diseconomies of scale, economies of scale, or constant returns to scale, depending on the industry.

How are profits and losses calculated?

To determine the potential profits of a business, the first step is to look at how much it will cost to run it. For a company to be profitable, it is not enough to provide products consumers want; it must simultaneously manage its costs. In this section, we discuss how profits and costs are calculated.

Calculating Profit and Loss

The simplest way to determine profit or loss is to calculate the difference between revenue and expenses (costs). The **total revenue** of a business is the amount the firm receives from the sale of goods and services. **Total cost** is the amount a firm spends to produce and/or sell goods and services. To determine total cost, the firm adds the individual costs of the resources used in producing and/or selling the goods. A **profit** occurs whenever total revenue is higher than

total cost. A **loss** occurs whenever total revenue is less than total cost. We can express this relationship as an equation:
 $\text{profit (or loss)} = \text{total revenue} - \text{total cost}$.

To calculate revenue, we look at the dollar amount the business takes in over a specific period.

Calculating costs, however, is a little more complicated than calculating revenue; total cost has two parts — one that is visible and one that is largely invisible. In the next section, we will see that determining total costs is part art and part science.

Explicit and Implicit Costs

Economists break costs into two components: explicit costs and implicit costs. **Explicit costs** are tangible out-of-pocket expenses; to calculate explicit costs, we add every expense incurred to run the business. **Implicit costs** are the costs of resources already owned, for which no out-of-pocket payment is made. Implicit costs are also opportunity costs, because the use of owned resources means that the next-best alternative use is forgone; also, each alternative is an implicit cost (implicit costs are hard to calculate and easy to miss).

To fully account for all the costs of doing business, we must calculate the explicit costs, determine the implicit costs, and add them together:

$$\text{total cost} = \text{explicit costs} + \text{implicit costs}$$

Table 8.1

Examples of a Firm's Explicit and Implicit Costs	
Explicit costs	Implicit costs
The electricity bill	The labor of an owner who works for the company but does not draw a salary
Advertising in the local newspaper	The opportunity costs of the capital in the business
Employee wages	The use of the owner's car, computer, or other personal equipment to conduct company business

Shows examples of a firm's explicit and implicit costs

Accounting Profit vs. Economic Profit

Now that you know about explicit and implicit costs, we can refine our definition of profit. In fact, there are two types of profit — accounting profit and economic profit.

A firm's **accounting profit** is calculated by subtracting only the explicit costs from total revenue. Accounting figures permeate company reports, quarterly and annual statements, and the media.

$$\text{accounting profit} = \text{total revenues} - \text{explicit costs}$$

As you can see, accounting profit does not take into account the implicit costs of doing business. To calculate the full cost of doing business, we need to consider both implicit and explicit costs; doing so yields a firm's economic profit. **Economic profit** is calculated by subtracting both the explicit and the implicit costs from total revenue. Economic profit gives more complete assessment of how a firm is doing.

$$\text{economic profit} = \text{total revenues} - (\text{explicit costs} + \text{implicit costs}) \text{ or } \text{economic profit} = \text{accounting profit} - \text{implicit costs}$$

The difference in accounting profits among various types of firms can be misleading.

Economic profit can be negative, since the negative dollar amount is a loss. If a business has an economic profit, its revenues are larger than the combination of its explicit costs and implicit costs. Likewise, a business has an economic profit, its revenues are larger than the combination of its explicit costs and implicit costs. Likewise, a business has an economic loss when its revenues are smaller than the combination of its explicit and implicit costs. The difficulty in determining economic profit lies in calculating the tangible value of implicit costs

How much should a firm produce?

Every business must decide how much to produce. In this section, we describe the factors that determine output, and we explain how firms use inputs to maximize their production. Because it is possible for a firm to produce too little or too much, we must also consider when a firm should stop production.

The production function

For a firm to earn economic profit, it must produce a product that consumers want. This product is the firm's **output**. A firm should produce an output that is consistent with the largest possible economic profit.

The firm must also control its costs. To do so, the firm must use resources efficiently; there are three primary **factors of production**: labor, land, and capital. Each factor of production is an **input**, or a resource used in the production process to generate the firm's output. **Labor** consists of workers, **land** consists of the geographical location used in production, and **capital** consist of all the resources the workers use to create the final product.

To keep costs down in the production process, a firm needs the right mix of inputs. The **production function** describes the relationship between the inputs a firm uses and the output it creates.

The **marginal product**, which is the change in output associated with one additional unit of an input.

$$\text{marginal product} = (\text{output from } n \text{ inputs}) - (\text{output from } n - 1 \text{ input units})$$

Conversely, for any given number of workers, the total output is the sum of the individual workers' marginal products.

Diminishing Marginal Product

The point at which successive increases in inputs are associated with a slower rise in output is known as the point of **diminishing marginal product**. A common mistake when considering marginal product is to assume that a firm should stop production as soon as marginal product starts to fall. This is not necessarily true; "diminishing" does not mean "negative". There are many times when marginal product is declining but still relatively high.

What costs do firms consider in the short run and the long run?

Production is one part of a firm's decision-making process: the other major component of production is cost.

Every firm, whether just starting out or already well established and profitable, can benefit by assessing how much to produce and how to produce it efficiently. In addition, production and cost considerations are different in the short run and in the long run. We begin with the short run because the majority of firms are most concerned with making the best short-run decisions, and then we extend our analysis to the long run, where planning ahead plays a central role.

Costs in the short run

All firms experience some costs that are unavoidable in the short run. These unavoidable costs are a large part of short-run costs. In the short run, costs can be variable or fixed.

Variable costs change with the rate of output; in **Chapter 9**, we add demand to our analysis to determine how much the firm should produce. For now, we stay focus on the cost side.

Fixed costs are unavoidable; they do not vary with output in the short run. These fixed costs —also know as **overhead**— include rent, insurance, and property taxes.

Interpreting tabular data

Every business must be able to determine how much it costs to provide the products and services it sells. **Table 8.3** lists many different ways to measure the costs associated with business decisions.

When we add fixed cost and variable cost together, we get total cost (TC), $TC = TVC + TFC$. **Average variable cost (AVC)** is the total variable cost divided by the output produced, $AVC = TVC \div Q$.

Average fixed cost (AFC) is calculated by dividing total fixed cost by the output: $AFC = TFC \div Q$. Because total fixed cost is constant, dividing these costs by the output means that as the output rises, the average fixed cost declines. In other words, higher output levels spread out the total fixed cost across more units.

Average total cost (ATC) is calculated by adding the AVC and AFC. It can also be calculated by dividing total cost by quantity

$$ATC = TC \div Q$$

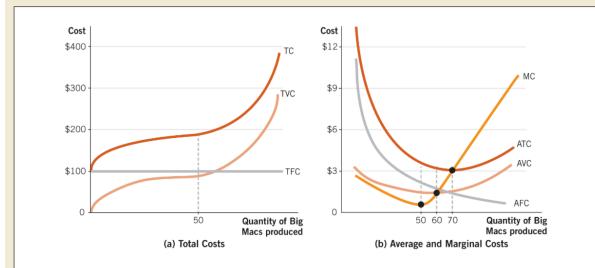
Interpreting data graphically

The **marginal cost (MC)** is the increase in cost that occurs from producing one additional unit of output. (This relationship is shown as the change in TVC divided by the change in quantity produced, where change is indicated by Δ . [Equation is

FIGURE 8.2

The Cost Curves

(a) The total variable cost (TVC) dictates the shape of the total cost (TC) curve. After 50 Big Macs, diminishing marginal product causes the total cost curve to rise at an increasing rate. Notice that the total fixed cost curve (TFC) stays constant, or flat. (b) The marginal cost curve (MC) reaches its minimum before average variable cost (AVC) and average total cost (ATC). Marginals always lead the average variable and average total costs either up or down. Average fixed cost (AFC), which has no variable component, continues to fall with increased quantity, because total fixed costs are spread across more units.



$$\text{marginal cost} = (\text{cost of producing } n \text{ units}) - (\text{cost of producing } n - 1 \text{ units})$$

]

Costs in the Long Run

We have seen that in the short run, businesses have fixed costs and fixed capacities; in the long run, all costs are variable and can be renegotiated. Thus, firms have more control over their costs in the long run, which enables them to reach their desired level of production. One way firms can adjust in the long run is by changing the **scale**, or size, of the production process. If the business is expected to grow, the firm can ramp up production. If the business is faltering, it can scale back its operations. This flexibility enables firms to avoid a situation of negative marginal product. Economists refer to the quantity of output that minimizes the average total cost in the long run as the **efficient scale**.

A long-run time horizon allows a business to choose a scale of operation that best suits its needs. The absence of fixed factors in the long-run production process means that we cannot explain total costs in the long run in the same way we explained short-run costs. Short-run costs are a reflection of diminishing marginal product, whereas long-run costs are a reflection of scale and the cost of providing additional output; because diminishing marginal product is no longer relevant in the long run, one might assume that costs would fall as output expands. However, this is not necessarily the case; depending on the industry and the prevailing economic conditions, long-run costs can rise, fall, or stay approximately the same.

Three types of scale

In this section, we describe three different scenarios for a firm in the long run. A firm may experience **economies of scale**, **diseconomies of scale**, or **constant returns to scale**. Let's consider each of these in turn.

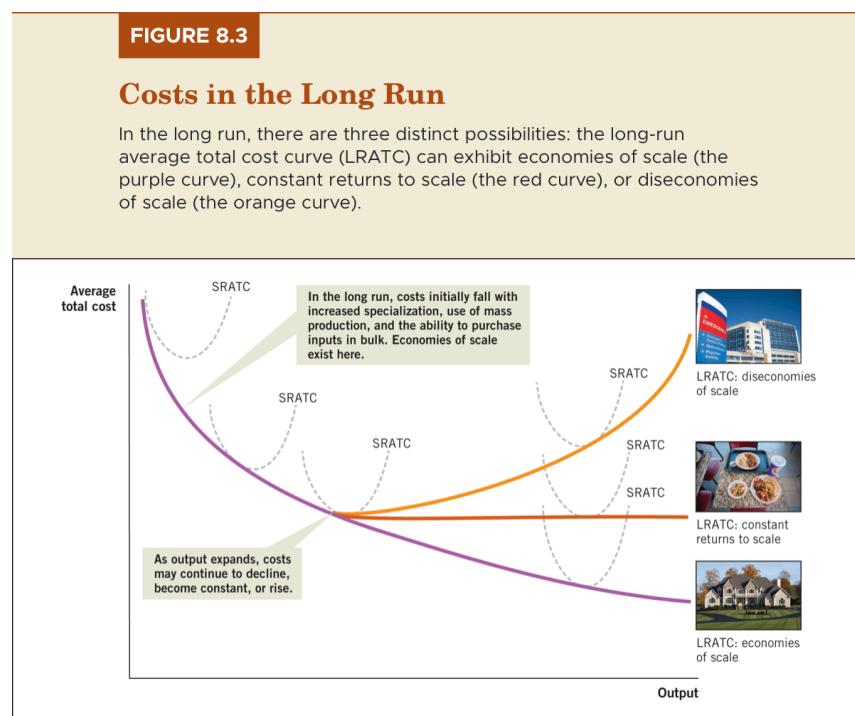
If output expands and long-run average total costs decline in the long run, a business experiences **economies of scale**; but bigger isn't always better! Sometimes a company grows so large that coordination problems make costs rise. A larger firm can become less effective at holding down long-run average total costs and experience **diseconomies of scale**, or higher costs as output expands in the long run.

The problem of diseconomies of scale is especially relevant in the service sector of the economy.

When long-run average total costs remain constant even as output expands in the long run, we say that the firm has **constant returns to scale**.

Long-run cost curves

Now it's time to illustrate the long-run nature of cost curves; we have seen that increased output may not always lead to economies of scale. Average total costs can decline, be constant, or rise with output. The long-run average total cost curve (LRATC) is actually a composite of many short-run average total cost curve (SRATC), which appear as the faint U-shaped dashed curves drawn in gray. By visualizing the short-run cost curves at any given output level, we can develop a composite of them to create the LRATC curve, which comprises all the short-run cost curves the firm may choose to deploy in the long run. In the long run, the firm is free to choose any of its short-run curves, so it always picks the output/cost combination that minimizes costs.



In the long run, there are three distinct possibilities: economies of scale, constant returns to scale, and diseconomies of scale. At first, each LRATC curve exhibits economies of scale as a result of increased specialization, the utilization of mass production, bulk purchasing power, and increased automation. The main question in the long run is whether

the cost curve will continue to decline, level off, or rise. In an industry with economies of scale at high output levels; the cost curve continues to decline, and the most efficient output level is always the largest output: the purple curve in Figure 8.3. In this situation, we would expect only one large firm to dominate the industry because large firms have significant cost advantages. However, in an industry with constant returns to scale; the cost curve flattens out: the red line. Once the curve becomes constant, firms of varying sizes can compete equally with one another because they have the same costs. Finally, in the case of diseconomies of scale — bigger firms have higher costs: the orange curve.

Conclusion

Do larger firms have lower costs? Not always; when diseconomies of scale occur, average total costs will rise with output. This result contradicts the common misconception that bigger firms have lower costs than their smaller competitors. Simply put, sometimes a leaner firm with less overhead can beat its larger rivals on cost.

Friday, October 13, 2023

Firms in a Competitive Market

Econ 304K: Introduction to Economics

Why do firms charge the price they do?

In this chapter and the next four, we look in more detail at how markets work, the profits firms earn, and how market forces determine the price a firm can charge for its product or service. We begin our examination of **market structure**, or how individual firms are interconnected, by looking at the conditions necessary to create a competitive market. Although few real markets achieve this ideal market structure described in this chapter, this model provides a good starting point for understanding other market structures.

Our analysis of competitive markets shows that when competition is widespread, firms have little to no control over the price they can charge, and they make little or no economic profit; thus, in competitive markets, firms are completely at the mercy of market forces that set the price...let's find out why.

Big Questions

- How do competitive markets work?
 - The firms in competitive markets sell similar products. Firms are also free to enter and exit the market whenever they wish.
 - A price taker has no control over the price it receives in the market.
 - In competitive markets, the price and quantity produced are determined by market forces instead of by the firm
- How do firms maximize profits?
 - A firm maximizes profits by expanding output until marginal revenue is equal to marginal cost ($MR = MC$, or the profit-maximizing rule). The profit-maximizing rule is a condition for stopping production at the point where profit opportunities no longer exist.
 - This firm should shut down in the short run if the price it receives does not cover its average variable costs; because variable costs are incurred only when operating, if

a firm can make enough to cover its variable costs in the short run, it will choose to continue to operate,

- In the long run, the firm should go out of business if it cannot cover its average total costs
- What does the supply curve look like in perfectly competitive markets?
 - Profits and losses act as signals for firms to enter or leave a market. As a result, perfectly competitive markets drive economic profit to zero in the long run.
 - The entry and exit of firms ensure that the market supply curve in a competitive market is much more elastic in the long run than in the short run.

How do competitive markets work?

Competitive markets exist when there are so many buyers and sellers that each one has only a small impact on the market price and output. When buyers are willing to purchase a product anywhere, sellers have no control over the price they charge. These two characteristics—similar goods and many participants—create a highly competitive market where the price and quantity sold are determined by the market conditions rather than by any one firm.

Table 9.1

Characteristics of Competitive Markets
Many sellers
Similar products
Free entry and exit
Price taking
Every firm is small

In competitive markets, buyers can expect to find consistently low prices and wide availability of the good they want. Firms that produce goods in competitive markets are known as price takers. A **price taker** has no control over the price set by the market. It “takes”—that is, accept—the price determined by the overall supply and demand conditions that regulate the market. One of the reasons why firms are price takers is that each seller is small compared to the overall market; this means that any individual seller’s decision (to either increase or decrease production) has no impact on the market price.

Competitive markets have another important feature: new competitors can easily enter the market. When barriers to entry into a marketplace are low, new firms are free to compete with existing businesses, which ensures the existence of competitive markets and low prices. **Table 9.1** summarizes the characteristics of competitive markets.

In the next section, we examine the profits competitive firms make. After all, profits motivate firms to produce a product, so knowing how a business can make the profit is central to understanding how competitive markets work.

How do firms maximize profits?

All firms, whether they are active in a competitive market or not, attempt to maximize profits; making a profit requires that a firm have thorough grasp of costs and revenues. In the previous chapter, we learned about the firm's cost structure. In this section, we examine its revenues; combining the firm's revenue with its costs enables us to determine how much profit the firm makes.

Profits are a key goal of almost every firm, but they don't always materialize; sometimes, firms experience losses instead of profits, so we explore whether a firm should shut down or continue to operate in order to minimize its losses. Once we fully understand the firm's decision-making process, we will better comprehend how the entire market functions.

The Profit-Maximizing Rule

The **marginal revenue** is the change (Δ) in total revenue when the firm produces one additional unit of output; in chapter 8, we saw that to understand cost structure, a firm focuses on marginal cost. The same is true on the revenue side; the **profit-maximizing rule** states that profit maximization occurs when a firm expands output as long as marginal revenue is greater than marginal cost, stopping close to $MR = MC$ as practical. According to the $MR = MC$ rule, production should stop at the point at which profit opportunities no longer exist.

Deciding how much to produce in a competitive market

We have observed that a firm in a highly competitive market is a price taker; it has no control over the price set by the market. Because all snow removal companies provide the same service, they must charge the price determined by the overall supply and demand conditions that regulate that particular market.

To better understand these relationships, we look at them visually. In Figure 9.1, we use the MR and MC data from Table 9.3 to illustrate the profit calculation. For reference, we also include the average total cost curve; recall from Chapter 8 that marginal cost curve (MC, shown in orange) always crosses the average total cost curve (ATC) at the lowest point. Figure 9.1 illustrates the relationship between the marginal cost curve (MC) and the marginal revenue curve (MR).

Note that we can use the profit-maximizing rule, $MR = MC$, to identify the most profitable output in a two-step process:

1. Locate the point at which the firm will maximize its profits: $MR = MC$, this is the point labeled A in **Figure 9.1**
2. Look for the profit-maximizing output: move down the vertical dashed line to the x -axis at point Q . Any quantity greater than or less than Q would result in lower profits.

Once we know the profit-maximizing quantity, we can determine the average cost of producing Q units. From Q , we move up along the dashed line until it intersects with the ATC curve.

$$\text{profit} = (\text{price} - \text{ATC} [\text{along the dashed line at quantity } Q]) * Q$$

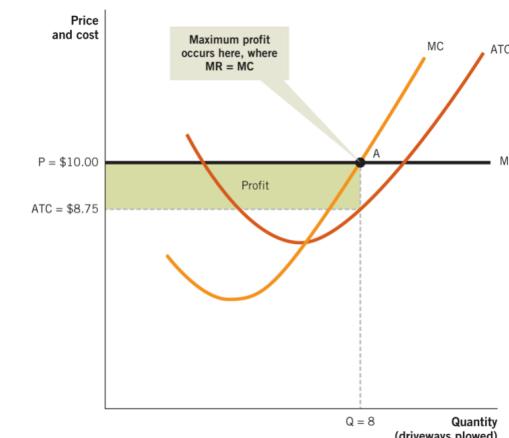
The Firm in the Short Run

Deciding how much to produce in order to maximize profits is the goal of every business in a competitive market. However, there are many times when it is not possible to make a profit; when revenue is sufficient to cover cost, the firm suffers a loss—at which point it must decide whether to operate or temporarily shut down. Successful businesses make this decision all the time.

FIGURE 9.1

Profit Maximization

Mr. Plow uses the profit-maximizing rule to locate the point at which marginal revenue equals marginal cost, or $MR = MC$. This point determines the ideal output level, Q . The firm takes the price from the market; price is shown as the horizontal MR curve at $P = \$10.00$. Because the price charged is higher than the average total cost curve along the dashed line at quantity Q , the firm makes the economic profit shown in the green rectangle.



Fortunately, a firm can use a simple, intuitive rule to decide whether to operate or shutdown in the short run: if the firm would lose less by shutting down than by staying open, it should shut down. Recall that costs are broken into two parts—fixed and variable. Fixed costs must be paid whether the business is open or not because variable costs are incurred only when the business is open, if it can

FIGURE 9.2

When to Operate and When to Shut Down

If the MR (marginal revenue) curve is above the minimum point on the ATC (average total cost) curve, the Ice Cream Float will make a profit (shown in green). If the MR curve is below the minimum point on the ATC curve (\$2.50) but above the minimum point on the AVC (average variable cost) curve (\$2.00), the float will operate at a loss (shown in yellow). If the MR curve is below the minimum point on the AVC curve (\$2.00), the float will temporarily shut down (shown in grey).

Table 9.4

Profit and Loss in the Short Run		
Condition	In words	Outcome
$P > ATC$	The price is greater than the average total cost of production	The firm makes a profit
$ATC > P > AVC$	The average total cost of production is greater than the price the firm charges, but the price is greater than the average variable cost of production.	The firm will operate to minimize loss.
$AVC > P$	The price is less than the average variable cost of production.	The firm will temporarily shut down.

make enough to cover its variable costs—it will choose to remain open. Once the variable costs are covered, any extra money goes toward paying the fixed costs.

A business should operate if it can cover variable costs, and it should shut down if it cannot. **Figure 9.2** illustrates the decision using cost curves, as long as the firm's marginal revenue (MR) is greater than the minimum point on the average variable cost curve (AVC)—the green and yellow areas—the firm will choose to operate. (Note that the MR curve is not shown in **Figure 9.2**; the shaded areas in the figure denote the range of potential MR curves that are profitable and those

that cause a loss.) Finally, if the MR curve falls below the the AVC curve—the red area—the firm should shut down; **Table 9.4** summarizes these decisions.

The Firm's Short-Run Supply Curve

Cost curves provide detailed picture of a firm's willingness to supply a good or service. We have seen that when the MR curve is below the minimum point on the AVC curve, the firm shuts down and production, or output, falls to zero. In other words, when revenues are too low, no supply is produced.

Sunk costs

Unrecoverable costs that have been incurred as a result of past decisions are known as **sunk costs**. They represent money that has already been spent, no matter what comes next, and that therefore should have no bearing on future decisions. This is known as the principle of sunk costs; it's the “What's done is done” rule, applied to economics.

The firm's long-run supply curve

In the long run, a competitive firm's output decision is directly tied to profits; because the firm is flexible in the long run, all costs are variable. As a result, the firm's long-run supply curve exists only when the firm expects to cover its total cost production (because otherwise the firm would go out of business—that is, exit the market).

FIGURE 9.3

The Firm's Short-Run Supply Curve

The short-run supply curve (S_{SR}) and marginal cost curve (MC) are equivalent when the price is above the minimum point on the average variable cost curve (AVC). Below that point, the firm shuts down and no supply exists.

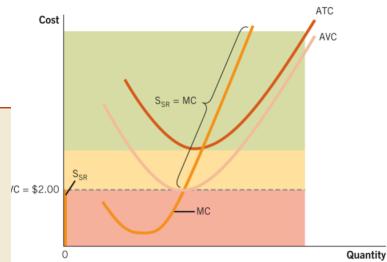


FIGURE 9.4

The Firm's Long-Run Supply Curve

The long-run supply curve (S_{LR}) and marginal cost curve (MC) are equivalent when the price is above the minimum point on the average total cost curve (ATC). Below that point, the firm shuts down and no supply exists.

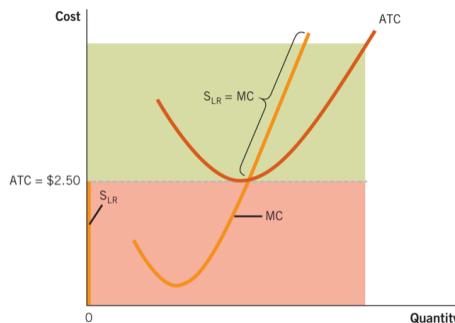


Table 9.5

The Long-Run shutdown criteria		
Condition	In words	Outcome
$P > ATC$	The price is greater than the average total cost of production	The firm makes a profit
$P < ATC$	The price is less than the average total cost of production.	The firm should shut down

So far, we have examined the firms decision-making process in the short run in the context of revenue versus cost, which enabled us to determine the profit each firm makes. However, a single firm represents only a small part of the overall supply in a competitive market. In the next section, we develop the short-run and long-run market supply curves.

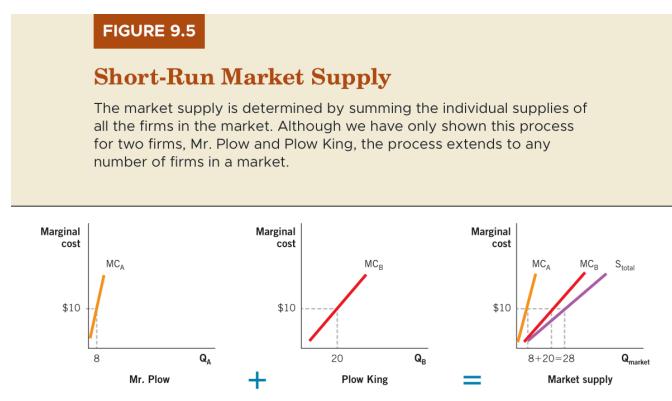
What does the Supply Curve look like in Perfectly Competitive Markets?

We have seen that a firm's willingness to supply a good or service depends on whether the firm is making a short-run or long-run decision. In the short run, a firm may choose to operate at a loss to recover a portion of its fixed costs. In the long run, there are no fixed costs, so a firm is willing to operate only if it expects the price it charges to cover total costs.

However, the supply curve for a single firm represents only a small part of the overall supply in a competitive market. We now turn to market supply and develop the short-run and long-run market supply curves.

The Short-Run Market Supply Curve

Recall that a competitive market is one in which a large number of buyers seek a product many sellers offer. Competitive markets are also characterized by easy entry and exit. Existing firms and entrepreneurs decide whether to enter and exit a market based on incentives. When existing firms are enjoying profits, there is an incentive for them to produce more and also for entrepreneurs to enter the market; the result is an increase in the quantity of the goods supplied. Likewise, when existing firms are



experiencing losses, there is an incentive for them to exit the market; then the quantity supplied decreases.

Entry and exit have the combined effect of regulating the amount of profit a firm can hope to make in the long run. As long as profits exist, the quantity supplied will increase because existing firms expand production or other firms enter the market. When losses exist, the quantity supplied will decrease because existing firms reduce production or other firms exit the market; so both profits and losses signal a need for an adjustment in market supply, in other words, profits and losses act as signals for resources to enter or leave a market. **Signals** convey information about the profitability of various markets.

The only time an adjustment does not take place is when participants in the market make zero economic profit; in that case, the market is in long-run equilibrium. Existing firms and entrepreneurs are not inclined to enter or exit the market; the adjustment process that occurs through price changes ends.

The benefit of a competitive market is that profits guide existing firms and entrepreneurs to produce more goods and services that society values. Losses encourage firms to exit and move somewhere else; without profits and losses acting as signals for firms to enter or exit the market, resources will be misallocated and surpluses and shortages will occur.

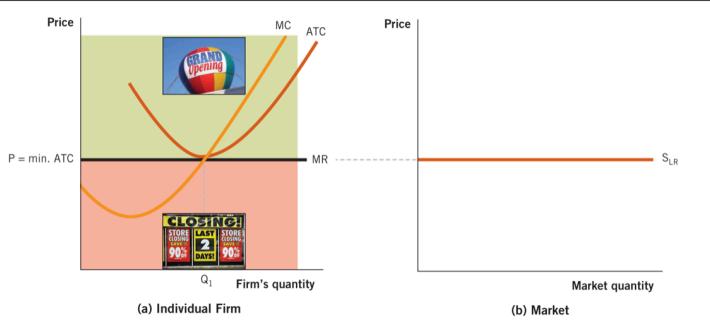
Figure 9.6 captures how entry and exit determine the market supply; the profit-maximizing point of the individual firm in panel

(a), $MR = MC$, is located at the minimum point on the ATC curve. The price ($P = \text{min.ATC}$) that existing firms receive is just enough to cover costs, so profits are zero. As a result, new firms have no incentive to enter the market and existing firms have no reason to leave. At all prices above

FIGURE 9.6

The Long-Run Market Supply Curve and Entry and Exit

Entry into the market and exit from it force the long-run price to be equal to the minimum point on the average total cost curve (ATC). At all prices above $P = \text{min. ATC}$, firms will earn a profit (the green area), and at all prices below $P = \text{min. ATC}$, firms will experience a loss (the red area). For this reason, the long-run supply curve (S_{LR}) must be horizontal at price $P = \text{min. ATC}$. If the price was any higher or lower, firms would enter or exit the market, and the market could not be in a long-run equilibrium.



$P = \min.\text{ATC}$, firms will earn a profit (the green area), and at all prices below $P = \min.\text{ATC}$, firms will experience a loss (the red area). This picture is consistent for all markets with free entry and exit; zero economic profit occurs only one price, and the price is the lowest point of the ATC curve.

At this price, the supply curve in panel (b) must be a higher horizontal line at $P = \min.\text{ATC}$. If the price were any higher, firms would enter, supply would increase, and price would be forced back down to $P = \min.\text{ATC}$; if the price were any lower, firms would exit, supply would decrease, and price would be forced up to $P = \min.\text{ATC}$, because we know that these adjustments will have time to take place in the long run, the long-run supply curve must also be equal to $P = \min.\text{ATC}$ to satisfy demand that exists at this price.

A reminder about economic profit

Now that you have learned how perfect competition affects business profits in the long run, you may not think that a competitive market is a desirable environment for business to seeking to earn profits. After all, if a firm cannot expect to make an economic profit in the long run, why bother? It's easy to forget the distinction between accounting profit and economic profit; firms enter a market when they expect to be reasonably compensated for their investment, and they leave a market when the investment does not yield satisfactory result. Economic profit is determined by deducting the explicit and implicit costs from the total revenue; therefore, firms are willing to stay in perfectly competitive markets in the long run when they're

FIGURE 9.7

A Competitive Market in Long-Run Equilibrium

When a market is in long-run equilibrium, the short-run supply curve (S_{SR}) and short-run demand curve (D_{SR}) intersect along the long-run supply curve (S_{LR}). At this point, the price the firm charges is equal to the minimum point along the average total cost curve (ATC). The existing firms in the market earn zero economic profit, and there is no incentive for firms to enter or exit the market.

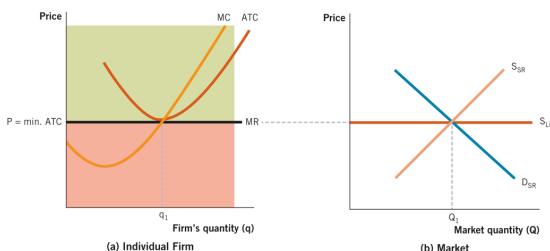
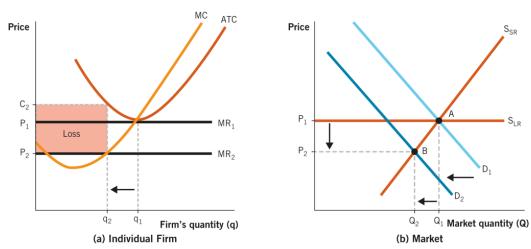


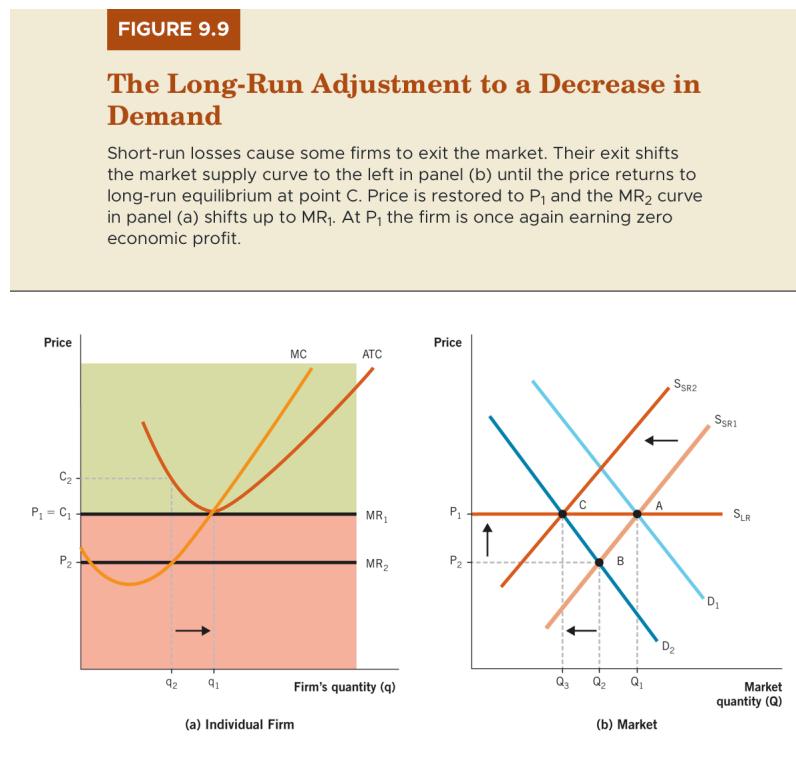
FIGURE 9.8

The Short-Run Adjustment to a Decrease in Demand

A decrease in demand causes the price to fall in the market, as shown by the movement from D_1 to D_2 in panel (b). Because the firm is a price taker, the price it can charge falls to P_2 . As we see in panel (a), the intersection of MR_2 and MC occurs at Q_2 . At this output level, the firm incurs the short-run loss shown by the red area in (a).



breaking even because they are being reasonably compensated for the explicit expenses they have incurred and also for the implicit expenses—like the opportunity costs of other business ventures—they would expect to incur elsewhere.



Conclusion

In competitive markets, where firms are at the mercy of market forces that set the price, individual firms have no control over the price because they sell the same products as their competitors. In addition, profits and losses help regulate economic activity in competitive markets and promote economic efficiency. Profits reward producers for producing a good that is valued more highly than the resources used to produce it. Profits encourage entry into a market. Likewise, losses penalize producers who operate inefficiently or produce goods consumers do not want. Losses encourage exit from the market; the process of entry and exit ensures that resources flow into markets that are undersupplied and away from markets where too many firms exist.

In this chapter, we studied competitive markets to establish a benchmark that will help us understand how other market structure compare with this ideal. In the next few chapters, we explore imperfect markets, which provide a significant contrast with the

results we have just seen. The closer a market is to meeting the criteria of perfect competition, the better the result for consumers and society in general.

Wednesday, October 18, 2023

Understanding Monopoly

Introduction to Microeconomics: Econ 304K

Monopolies have market power

In this chapter, we explore another market structure: monopoly. Monopolists enjoy market power for their specific product, but they cannot force consumers to purchase what they're selling. By the law of demand, when a monopolist charges more, people buy less. If demand is low enough, a monopolist may even experience a loss instead of a profit; but on the other hand, when demand is high, and on top of that is inelastic, a monopolist is sometimes able to earn massive profits.

The typical result of monopoly is higher prices and less output than we find in a competitive market.

In this chapter, we explore the conditions that give rise to monopolies, how they work, and what can be done about the problem they create.

Big Questions

- **What are monopolies created?**
 - Monopoly is a market structure characterized by a single seller that produces a well-defined product with no good substitutes.
 - Monopolies operate in a market with high barriers to entry, the chief source of market power.
 - Monopolies are created when a single seller supplies the entire market for a particular good or service.
- **How much do monopolies charge, and how much do they produce?**
 - Monopolists are price makers who may earn long-run economic profits
 - Like perfectly competitive firms, a monopolist tries to maximize its profits. To do so, it uses the profit-maximizing rule, $MR = MC$, to select the optimal price and quantity combination of a good or service.
- **What are the problems with and solutions for monopoly?**

- From an efficiency standpoint, the monopolist charges too much and produces too little; because the monopolist's output is smaller than the output that would exist in a competitive market, monopolies lead to deadweight loss.
- Government grants of monopoly power encourage rent seeking, or the use of resources to secure monopoly rights through political process.
- There are three potential solutions to the problem of monopoly. First, the government may break up firms that gain too much market power in order to restore a competitive market. Second, the government can promote open markets by reducing trade barriers. Third, the government can regulate a monopolist's ability to charge excessive prices.
- When the costs of government involvement in regulating a monopoly are greater than the efficiency gains that can be realized, it is better to leave the monopolist alone.

How are monopolies created?

As we explained in Chapter 3, a monopoly exists when a single seller supplies the entire market for a particular good or service. Two conditions enable a single seller to become a monopolist. First, the firm must have something unique to sell—that is, something without close substitutes. Second, it must have a way to prevent potential competitors from entering the market.

Monopolies occur in many places and for different reasons. **Monopoly power**, which is a measure of a monopolist's ability to set the price of a good or service.

A monopolist operates in a market with high **barriers to entry**, which are restrictions that make it difficult for new firms to enter a market. As a result, monopolists have no competition nor any immediate threat of competition. High barriers to entry insulate the monopolist from competition, which means that many monopolists enjoy long-run economic profits. There are two types of barriers to entry: natural barriers and government-created barriers.

Natural barriers

Some barriers exist naturally within the market. These include control of resources, problems in raising capital, and economies of scale.

Control of resources

The best way to limit competition is to control a resource that is essential in the production process. This extremely effective barrier to entry is hard to

accomplish; but if you control a scarce resource, other competitors will not be able to find enough of it to compete.

Problems in raising capital

Monopolists are usually very big companies that have grown over an extended period; even if you had a wonderful business plan, it is unlikely that a bank or a venture-capital company would lend you enough money to start a business that could compete effectively with a well-established company. Lenders provide capital for business projects when the chance of success is high, but the chance of a new company successfully competing against an entrenched monopolist is not high. Consequently, raising capital to compete effectively is difficult.

Economies of Scale

In **Chapter 8**, we saw that economies of scale occur when long-run average costs fall as production expands. Low cost units and the low prices that follow give some larger firms the ability to drive rivals out of business. In an industry that enjoys large economies of scale, production costs per unit continue to fall as a firm expands. Smaller rivals then have much higher average costs that

The Characteristics of Monopolies
One seller
A unique product
High barriers to entry
Price making

prevent them from competing with a larger company; as a result, firms in the industry tend to combine over time. These mergers lead to the creation of a **natural monopoly**, which occurs when a single large firm has lower costs than any potential smaller competitor.

Government-created Barriers

The creation of a monopoly can be either intentional or an unintended consequence of a government policy. Government-enforced statutes and regulations, such as laws and regulations covering licenses and patents, limit the scope of competition by creating barriers to entry.

Licensing

In many instances, it makes sense to give single firm the exclusive right to sell a good or service. To minimize negative externalities, governments occasionally establish monopolies, or near monopolies, through licensing requirements. Licensing also creates an opportunity for corruption; in fact, in many parts of the world, bribery is so common that it often determines which companies receive licenses in the first place.

Patents and Copyright Law

Another area in which the government fosters monopoly is that of patents and copyrights. By granting patents and copyrights to developers and inventors, the

government creates monopolies; patents and copyrights create stronger incentives to develop new drugs and produce new music than would exist if market competitors could immediately copy inventions..the social benefits of patent law can be enormous.

After the patent or copyright expires, rivals can mimic the invention; this new competition opens up the market and provides dual benefits: wider access to the innovation and more sellers—both of which are good for consumers in the long run.

Nonetheless, many economists wonder if patents and copyrights are necessary or have unintended consequences. Sometimes copyright holders benefit more from exposure than from exercising their right to charge consumers. The point to remember is that copyright protection gives artists the right to decide how to distribute their work and what price to charge. It also gives them the ability to litigate when their work is stolen, illegally downloaded, or improperly used. Though market-created and government-created barriers occur for different reasons, they have the same effect — they create monopolies. **Table 10.1** summarizes the key characteristics of monopolies. In the next section, we examine how the monopolist determines the price it charges and how much to produce, explaining the term “**price making**” listed in **Table 10.1**.

How much do monopolies charge and how much do they produce?

Both monopolists and firms in a competitive market seek to earn a profit. However, a monopolist is the sole provider of its product and holds market power, thus, monopolists are price makers. A **price maker** has some control over the price it charges. As you learned in **Chapter 9**, a firm in a competitive market is a price taker.

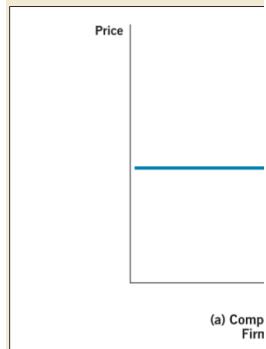
We can see the difference between price takers and price makers graphically in Figure 10.1. The demand curve for the product of a firm in a competitive market, shown in panel (a), is horizontal. When individual firms are price takers, they have no control over what they charge. In other words, demand is perfectly elastic—or horizontal—because every firm sells the same product; demand for an individual firm’s product exists only at the price determined by the market, and each firm is such a small part of the market that it can sell its entire output without lowering the price.

In contrast, because a monopolist is the only firm—the sole provider—in the industry, the demand curve for its product, shown in panel (b), constitutes the market demand curve; but the demand curve is downward sloping, which limits the monopolist’s ability to make a profit. The monopolist would like to exploit its market power by charging a high price to many customers. However, the law of demand, which identifies a negative relationship between price and quantity demanded, dictates otherwise. Unlike the horizontal demand curve of a firm in a competitive market, the downward-sloping demand curve of the monopolist has many price-output combinations. If market

FIGURE 10.1

Comparing the Competitive Firm and the Monopolist

(a) Firms in a competitive market are price takers; the monopolist is the sole seller and its product constitutes the entire demand curve, which is downward sloping. So while the price it charges, the revenue it receives, depends on its price and output.

**TABLE 10.2**

book content

Calculating the Monopolist's Marginal Revenue

	(1) Quantity of customers (Q)	(2) Price of service (P)	(3) Total revenue (TR)	(4) Marginal revenue per 1,000 customers (MR)
	Formula: $Q \times P$			ΔTR
	0	\$100	\$0.00	\$90,000
	1,000	90	90,000	70,000
	2,000	80	160,000	50,000
	3,000	70	210,000	30,000
	4,000	60	240,000	10,000
	5,000	50	250,000	-10,000
	6,000	40	240,000	-30,000
	7,000	30	210,000	-50,000
	8,000	20	160,000	-70,000
	9,000	10	90,000	-90,000
	10,000	0	0.00	

demand is more elastic, a monopolist will choose a lower price; as a result, monopolists must search for the profit-maximizing price and output.

The profit-maximizing rule for the monopolist

A competitive firm can sell all it produces at the existing market price; but a monopolist, because of the downward-sloping demand curve, must search for the most profitable price. To maximize profits, a monopolist can use the profit-maximizing rule we introduced in **Chapter 9**: $MR = MC$. For the price-taking firm, MR is just the market price, full stop; for the monopolist, however, there's a calculation involved.

Table 10.2 shows the marginal revenue for a cable company that serves a small community. Notice the negative relationship between output (quantity of customers) and price in columns 1 and 2: as the price goes, the quantity of consumers goes up. Total revenue is calculated by multiplying output by price ($TR = Q \times P$). At first, total revenue rises as the price falls, once the price becomes too low (\$40), total revenue begins to fall. As a result, the total revenue in column 3 initially rises to \$250,000 before it begins to fall off; the final column, marginal revenue, shows the change (Δ) in total revenue. Here we see positive (thought falling) marginal revenue associated with prices between \$100 and \$50 (see the green dollar amounts in column 4). Below \$50, marginal revenue becomes negative (see the red dollar amounts in column 4).

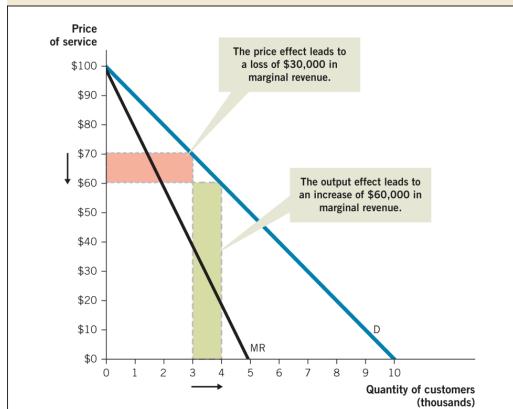
The change in total revenue reflects the trade-off a monopolist encounters in trying to attract additional customers, to gain additional sales, the firm must lower its price, but the lower the price is available to both new and existing customers. The impact on total revenue therefore depends on how many new customers buy the good because of the lower price.

Figure 10.2 uses the linear demand schedule from **Table 10.2** to illustrate the two separate effects that determine marginal revenue. First, there is **price effect**, which reflects how the lower price affects the revenue; but dropping the price also has an **output effect**, which reflects how the lower price affects the number of customers.

FIGURE 10.2

The Marginal Revenue Curve and the Demand Curve

A price drop has two effects. (1) Existing customers now pay less—this is the price effect. (2) New customers decide to purchase the good for the first time—this is the output effect. The relative size of the two effects, as shown by the red and green rectangles, determines whether the firm is able to increase its revenue by lowering its price. In this case, marginal revenue increases by \$30,000.



Lost revenue associated with the price effect are always subtracted from the revenue gains created by the output effect. There is a price effect whenever the price drops, the marginal revenue curve lies below the demand curve, therefore, in **Figure 10.2**, the y intercept is the same for the demand and marginal revenue curves and the x-intercept of the MR curve is half of the demand curve's.

At high price levels — where demand is elastic — the price effect is small relative to the output effect, as the price drops, demand slowly becomes more inelastic. The output effect diminishes and the price effect increases, in other words, as the price falls, it becomes harder for the firm to acquire

enough new customers to make up for the difference in lost revenue. Eventually, the price effect becomes larger than the output effect. This means that the marginal revenue curve will have the same y intercept as the demand curve and be twice as steep; as a result, marginal revenue becomes negative and dips below the x-axis, as shown by the MR curve in Figure 10.2. When the marginal revenue is negative, the firm cannot maximize profit; this outcome puts an upper limit on the amount the firm will produce; this outcome is evident in **Table 10.2**: once the price becomes too low, the firm's marginal revenue is negative.

Deciding how much to produce

In **Chapter 9**, we explored the profit-maximizing rule for a firm in a competitive market. This rule also applies to a monopolist: marginal revenue should equal to marginal cost. However, there is one big difference: a monopolist does not charge a price equal to marginal revenue.

Figure 10.3 illustrates the profit maximizing decision-making process for a monopolist. We use a two-step process to determine the monopolist's profit:

1. Locate the point at which the firm will maximize its profits: $MR = MC$

- Set the price: from the point at which $MR = MC$, determines the profit maximizing output, Q . From Q , move up along the dashed line until it intersects with the demand curve (D). From that point, move horizontally until you come to the y -axis. This point on the y -axis tells us the price (P) the monopolists should charge. Notice that the monopolist's price (P) is greater than MC ($P > MC$); this result differs from the competitive outcome, where $P = MC$.

Using this two-step process, we can determine the monopolist's profits; locate the average total cost, ATC , of making Q units. The difference between the price and the average total cost multiplied by Q tells us the profit (or loss) the firm makes. Any time the U-shaped ATC curve dips below the downward sloping demand curve, D , there is a way to earn positive economic profits.

Table 10.3 summarizes the differences between a competitive market and a monopoly. The competitive firm must take the price established in the market. If it does not operate efficiently, it cannot survive; nor can it make an economic profit in the long run. The monopolist operates very differently; because high barriers to entry limit competition, the monopolist may be able to earn long-run profits by restricting output. It operates inefficiently from society's perspective, and it has significant market power.

FIGURE 10.3

The Monopolist's Profit Maximization

The firm uses the profit-maximizing rule to locate the point at which $MR = MC$. This condition determines the ideal output level, Q . Because the price (which is determined by the demand curve) is higher than the average total cost curve (ATC) along the dashed line at quantity Q , the firm makes the profit shown in the green area.

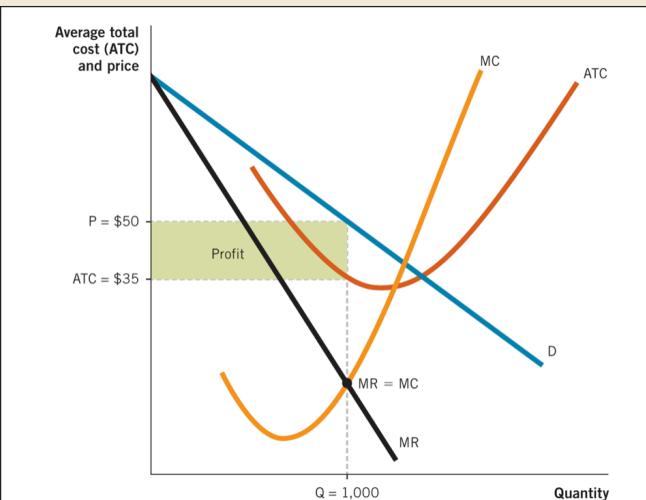


TABLE 10.3

The Major Differences between a Monopoly and a Competitive Market

Competitive market	Monopoly
Many firms	One firm
Cannot earn long-run economic profits	May earn long-run economic profits
Has no market power (is a price taker)	Has significant market power (is a price maker)
Produces an efficient level of output (because $P = MC$)	Produces less than the efficient level of output (because $P > MC$)

What are the problems with and solutions for, monopoly?

Monopolies can adversely affect society by restricting output and charging higher prices than sellers in competitive markets do. This activity causes monopolies to operate inefficiently, provide less choice, promote unhealthy form of competition known as **rent-seeking** (addressed in a later section), and make economic profits that fail to guide resources to their highest-valued use. Recall that market failure occurs when there is an inefficient allocation of resources in a market. Once we have examined the problems with monopoly, we will turn to potential solutions to the problem of monopoly.

The problems with monopoly

Monopolies result in an inefficient level of output, provide fewer choices to consumers, and encourage monopoly firms to lobby for government protection.

Inefficient output and price

From an efficiency standpoint, the monopolist charges too much and produces too little. This result is evident in Figure 10.4, which shows what happens when a competitive market (denoted by subscript C) ends up being controlled by a monopolist (denoted by the subscript M).

FIGURE 10.4

When a Competitive Industry Becomes a Monopoly

(a) In a competitive industry, the intersection of supply and demand determines the price (P_C) and quantity (Q_C). (b) When a monopolist controls an entire industry, the supply curve becomes the monopolist's marginal cost curve. The monopolist uses $MR = MC$ to determine its price (P_M) and quantity (Q_M). As a result, the monopolist charges a higher price and produces a smaller output than when an entire industry is populated with competitive firms.

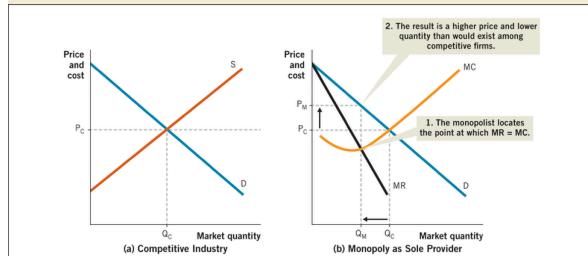
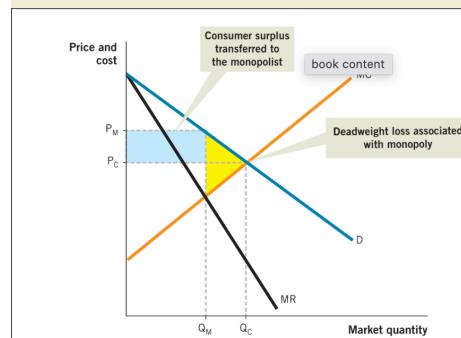


FIGURE 10.5

The Deadweight Loss of Monopoly

Because the profit-maximizing monopolist produces an output of Q_M , an amount that is less than Q_C , the result is the deadweight loss shown in the yellow triangle. The blue rectangle is the consumer surplus transferred to the monopolist.



Few choices for consumers

Another problem associated with monopoly is the lack of choice; have you ever wondered why cable companies offer their services in bundles? You can buy basic, digital, and premium packages, but one thing you cannot do is buy just the cable channels you want. This situation prevails because cable companies function like monopolies, and monopolies limit consumer choice; because the monopolist sells a good with a few close substitutes, it can leverage its market power to offer product features that benefit the monopolist at the expense of consumer choice. With a monopolist, there's only one outlet: if you do not like the design, features, price, or any other aspect of the good provided, you have few [or no] other option.

Rent seeking

The term **rent seeking** was coined by Anne Krueger in 1974; it refers to the attempt to gain monopoly power through the political process, by using lobbying and other means to secure legal monopoly rights.

Throughout this text, we have seen the desirable effects of competition: lower prices, increased efficiency, and enhanced service and quality. However, rent seeking in a form of competition that produces an undesirable result. When firms compete to become monopolists, there is one winner, without any of these benefits usually associated with competition.

This is a general rule: when lobbying is more profitable than creativity and productivity, typically companies that can afford to will prefer to lobby. If the lobbying succeeds, society can be adversely affected.

Solutions to the problem of monopoly

We have learned that monopolies do not produce as much social welfare as competitive markets do. As a result, public policy approaches attempt to address this problem. The policy solutions include breaking up the monopoly, reducing trade barriers, and regulating markets.

Breaking up the monopoly

Eliminating deadweight loss and restoring efficiency can be as simple as promoting competition. The government can help limit monopoly outcomes and restore a competitive balance; the government can accomplish this goal through anti-trust legislation. Anti-trust laws are designed to prevent monopoly practices and promote competition; the government has exercised control over monopoly practices since the passage of the Sherman Act in 1890, and the task currently falls to the Department of Justice.

Reducing trade barriers

Countries use **tariffs**, which are taxes on imported goods, as a trade barrier to prevent competition, and protect domestic business. However, any barrier—whether a tariff, a quota, or a prohibition—limits the possible gains from trade. For monopolists, trade barriers prevent rivals from entering territory; in the other hand, reducing trade barriers creates more competition, lessens the influence of monopoly and promotes efficient use of resources.

Regulating markets

In the case of a natural monopoly, it is not practical to harness the benefits of competition. When a natural monopoly exists, the government may choose to use the marginal cost pricing rule, $P = MC$, to generate the greatest welfare for society; because the price is determined along the demand curve, setting $P = MC$ guarantees that the good or service will be produced as long as the willingness to pay exceeds the additional cost of production. **Figure 10.6** shows the difference in pricing and profits for a regulated monopoly and an unregulated natural monopoly; recall that a natural monopoly is characterized by economies of scale, which we

can idealize as a constant marginal cost that leads to a steadily dropping ATC curve, as in **Figure 10.6**

To maximize profits, an unregulated monopolist sets $MR = MC$ and produces quantity QM at a price of PM ; because PM is greater than the average total cost of producing QM units, or CM , the monopolist earns the profit shown in the green rectangle. If the firm is regulated and the price is set at marginal cost, regulators can set $P = MC$, and the output expands to QR (the subscript R denotes the regulated monopolist.)

A caveat about government oversight

Firms with a profit motive have an incentive to minimize the costs of production, because lower costs translate directly into higher profits. The government oversight and management of monopolies is problematic because there are fewer incentives to keep costs in check.

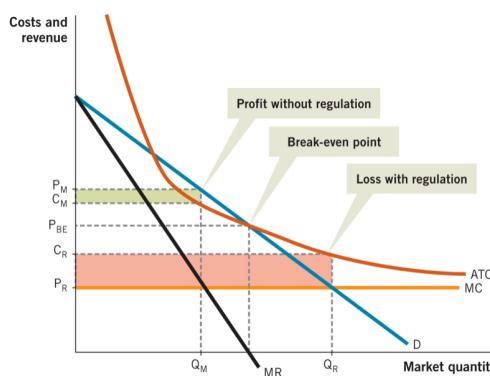
Consequently, the marginal cost pricing rule is not as effective as it first seems. Regulated firms and government-owned businesses do not have the same incentives to keep costs down; without the correct incentives in place, we would expect cost inefficiencies to develop.

Public policy can mitigate the power of monopolies, but, this outcome is not guaranteed. While monopolies are not as efficient as firms in competitive markets, this comparison is not always relevant; we need to ask how the inefficiency of monopoly compares with the inefficiencies associated with government involvement in the market. Good economists assess the benefits as well as the costs, so when the costs of the government involvement are greater than the efficiency gains that can be realized, the best solution to the problem of monopoly might be to do nothing.

FIGURE 10.6

The Regulatory Solution for Natural Monopoly

An unregulated monopolist uses the profit-maximizing rule $MR = MC$ and earns a small profit, shown in the green rectangle. If the monopolist is regulated using the marginal cost pricing rule, $P = MC$, it will experience the loss shown in the red rectangle.



Conclusion

It's tempting to believe that monopolies always earn a profit, but profit is not a guaranteed outcome. This monopolist controls the supply, not the demand, so monopolies occasionally suffer losses despite the advantages they enjoy. Still, many monopolies do make economic profit.

In this chapter, we examined the monopoly model and, along the way, compared the results under monopoly with results of the competitive model that we developed in the previous chapter. While competitive markets generally yield welfare-enhancing outcomes for society, monopolies often do the opposite; because monopolists do not produce an efficient outcome, government often seeks to limit monopoly outcomes and promote competitive markets.

Competitive markets and monopoly are market structures at opposite extremes. Indeed, we rarely encounter the conditions necessary for either a pure monopoly or a perfectly competitive market. Most economic activity takes place between these two alternatives. In the upcoming chapters, we examine monopolistic competition and oligopoly—two markets that constitute the bulk of the economy. Fortunately, if you understand the market structure at the extremes, understanding the middle ground is straightforward. As we move forward, we will deploy the same tools we used to examine monopoly in order to understand monopolistic competition and oligopoly.

Price Discrimination

ECO 304K

Charging different prices to different people isn't automatically a bad thing.

Two-tiered pricing usually works to the advantage of people who need a financial break, and in the big picture there are net social benefits, in the form of greater market efficiency. Here's a rundown of companies that will give you a price break if you can prove you're a student:

- Clothing retailers like J. Crew, Eastern Mountain Sports, and TopShop
- Tech companies such as Apple, Microsoft, Adobe, and MathWorks (makers of MATLAB)
- Entertainment companies such as Cinemark, Major League Baseball, museums, and most ski resorts and amusement parks
- Travel and transportation companies such as Greyhound and Amtrak, many hotel brand and some air carriers
- Your favorite news publications like the New York Times or the Economist (!)
- Auto insurance (if you have any good grades)
- All the major cell phone carriers
- Many national restaurant chains and almost all local restaurants

In this chapter, we examine many real-life pricing situations and how businesses can make additional profits if they charge different price to different groups of customers. The study of **price discrimination** adds a layer of complexity to the simple models of perfect competition and monopoly. A thorough understanding of how price discrimination works is especially useful as we complete our study of market structure with monopolistic competition and oligopoly in the next two chapters.

Big questions

- What is price discrimination?

- How is price discrimination practiced?

What is price discrimination?

Price discrimination occurs when a firm sells the same good or service at different prices to different groups of customers. The difference in price is not related to differences in cost. Although “price discrimination” sound like something illegal, in fact is it beneficial to both sellers and buyers. When a firm can charge more than one price, markets work more efficiently; because price-discriminating firms typically charge a “high” and a “low” price, she consumers are able to buy the product at a low price. Of course, firms are not in business to provide goods at low prices; they want to make a profit. Price discrimination enables them to make more money by dividing their customers into at least two groups: those who get a discount and others who pay more.

We have seen that in competitive markets, firms are **price takers**. If a competitive firm attempts to charge a higher price, its customers will likely buy elsewhere. To practice price discrimination, a firm must be a **price maker**; it must have some market power before it can charge more than one price. Both monopolies and non-monopolistic companies use price discrimination to earn higher profits.

Conditions for Price Discrimination

For price discrimination to take place, two conditions must be met. First, there must be at least two different types of buyers; second, the firm must be able to prevent resale of the product or service.

Distinguishing groups of buyers

To price-discriminate, a firm must be able to distinguish groups of buyers with different price elasticities of demand. Under those conditions, the firm faces a downward-sloping demand curve for its product instead of a horizontal one. This gives the firm market power; firms can generate additional revenue by charging more to customers with inelastic demand and less to customers with elastic demand.

Preventing resale

For price discrimination to be a viable strategy, a firm must also be able to prevent resale of the product or service; in some cases, preventing resale is easy.

One Price vs. Price Discrimination

A business that practices price discrimination would prefer to differentiate every customer by selling the same good or service at a price unique to that customer — a situation known as **perfect price discrimination**. To achieve this result, a business would have to know exactly what any particular customer would be willing to pay and charge him or her exactly that price. In practice, perfect price discrimination is hard to implement, so many firms instead settle for charging two or three prices based on sorting customers into a few easily identified groups.

Price discrimination boosts profits under normal circumstances, but it also gives a firm more flexibility when things get crazy.

The Welfare Effects of Price Discrimination

Price discrimination is profitable for the companies that practice it; but it also increases the welfare of society. How, you might ask, can companies make more profit and also benefit consumers? The answer: because a price discriminator charges a high price to some and a low price to others, more consumers are able to buy the good.

First, in the long run, a perfectly competitive firm would charge a price just equal to the marginal cost. Second, this outcome mirrors the result of a government-regulated monopolist that uses the marginal cost pricing rule, $P = MC$, to enhance social welfare. This strategy provides the firm with the opportunity to convert the area consisting of the two green triangles into more profit. In other words, the process maximizes the quantity sold; the efficiency of the market improves, and the firm generates more profit.

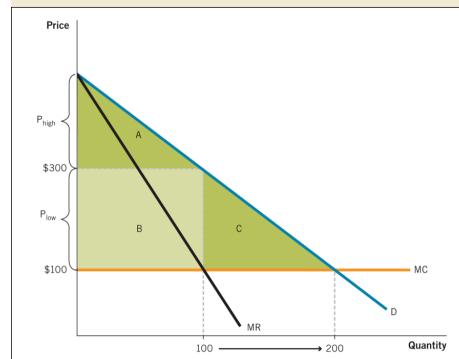
Comparing perfect price discrimination with perfect competition and monopoly

To understand the welfare effects of perfect price discrimination, we can compare the consumer and producer surplus in three scenarios: a

FIGURE 11.2

Perfect Price Discrimination

If the firm charges one price, the most it can earn is the profit in the light green rectangle. However, if a firm is able to perfectly price-discriminate, it can pick up the additional profit represented by the green triangles.



competitive market, a market in which a monopolist charges a single price, and a market characterized by perfect price discrimination.

In a perfectly competitive market, there are no barriers to entry and no firm has market power. In the long run, the price will be equal to the marginal cost. Under perfect competition, the market structure clearly favors consumers.

A monopoly holds substantial market power, so the firm in this scenario sets a price using the profit-maximizing rule, $MR = MC$, without having to worry about competition driving the price down to marginal cost.

Conclusion

The word “discrimination” has negative connotations, but not when combined with the word “price”. Charging different prices to different groups of customers results in more economic activity and is more efficient than charging a single price across the board. Under price discrimination, many customers pay less than they would if a firm had charged a single price, while other consumers will pay more because their demand is more inelastic; but overall, total social welfare increases, and the amount of deadweight loss in society is reduced.

Price discrimination also helps us understand how many markets actually function, because instances of perfectly competitive markets and monopoly are rare.

Monopolistic Competition and Advertising

ECO 304K: Introduction to Microeconomics

Advertising and Product Differentiation are notable features of Monopolistic Competition

If you drive down a busy street, you will find many competing businesses, often right next to one another. These competing firms advertise heavily; the temptation is to see advertising as driving up the price of a product without any benefit to the consumer. However, in markets where competitors sell slightly differentiated products, advertising enables firms to inform their customers about new products and services. Yes, costs rise, but consumers also gain information to help make purchasing decisions.

Consumers also benefit from added variety, and we all get a product that's pretty close to our vision of a perfect good—no other market structure delivers that outcome.

In this chapter, we will look at ***monopolistic competition***, a widespread market structure that has features of both competitive markets and monopoly. We also explore the benefits and disadvantages of advertising, which is prevalent with monopolistic competition.

Big Questions

- What is monopolistic competition?
 - **Monopolistic competition is a market structure characterized by low barriers to entry and many firms selling differentiated products.**
 - **Differentiation of products takes three forms**
 - ***Differentiation by style or type***
 - ***Location***
 - ***Quality***
 - What are the differences among monopolistic competition, competitive markets, and monopoly?

- Monopolistic competitors, like monopolists, are price makers with downward-sloping demand curves. Whenever the demand curve is downward sloping, the firm is able to mark up the price above marginal cost. The results are excess capacity and an inefficient level of output.
- In the long run, barriers to entry enable a monopoly to earn an economic profit. This is not the case for monopolistic competition or competitive markets.
- Why is advertising prevalent in monopolistic competition?
 - Advertising performs useful functions under monopolistic competition: it conveys information about the price of the goods offer for sale, the location of products, and new products. It also signals differences in quality, however, advertising also encourages brand loyalty, which makes it harder for other businesses to successfully enter the market.
Advertising can be manipulative and misleading.

What is Monopolistic Competition?

Each fast-food establishment has a unique set of menu items; the different products in fast-food restaurants give each seller a small degree of market power. This combination of market power and competition is typical of the market structure known as monopolistic competition. Indeed, **monopolistic competition** is characterized by low barriers to entry, many different firms, and product differentiation. **Product differentiation** is the process firms use to make a product more attractive to potential customers; firms use product differentiation to contrast their product's unique qualities with competing products. The differences, which we will examine in detail, can be minor and can involve subtle changes in packaging, quality, availability, and promotion; or the difference can be very significant.

How does monopolistic competition compare with other market structures we have studied? As **Table 12.1** shows, monopolistic competition falls between competitive markets and monopoly in terms of the number of sellers, the types of products sold, and competing firms' ability to enter and exit the market.

We have seen that firms in competitive markets do not have any market power, as a result, buyers can expect to find consistently low prices and wide availability; and we have seen that monopolies charge more and restrict the availability of a good or service. In markets that are monopolistically competitive, firms sell differentiated

products; this differentiation gives the monopolistic competitor some market power, though not as much as a monopolist, which controls the entire market.

Monopolistically competitive firms have a small amount of market power that enables them to search for the most profitable price.

Table 12.1

Competitive Markets, Monopolistic Competition, and Monopoly		
Competitive Markets	Monopolistic Competition	Monopoly
Many sellers	Many sellers	One seller
Similar products	Differentiated products	A unique product without close substitute
Free entry and exit	Low barriers to entry and exit	Significant barriers to entry and exit

To understand how monopolistic competition works, we begin with a closer look at product differentiation.

Product Differentiation

Monopolistically competitive firms create some market power through product differentiation. Differentiation can occur in a variety of ways, including style or type, location, and quality.

Style or Type

When you're ready for lunch at the mall, you can go to the food court, where many different place to eat offer a wide variety of choices. Where you decide to eat is a matter of your personal preference and the price you're willing to pay. Like most consumers, you will select the place that gives you what you want while providing the best value for your money. Consumers' differing tastes make it possible for a wide range of food vendors to compete side by side with rivals who provide many good substitutes.

Location

When consumers prefer to save time and to avoid the inconvenience of shopping for a better deal, a firm with a more convenient location will have some pricing power. As a result, producers who sell very similar products can generate some market power by locating their businesses

along routes most customers use to go to and from work or in other areas where customers frequently travel.

Quality

Firms also compete on the basis of quality, for example, if you want Mexican food you can go to Taco Bell, which is inexpensive and offers food cooked in advance. In contrast, at Moe's Southwestern Grill the food is freshly prepared and, as a result, more expensive; this form of product differentiation serves consumers quite well. Budget-conscious consumers can feast at Taco Bell, while those with a larger budget and a taste of higher-quality Mexican food can consider Moe's as another option.

What are the differences among Monopolistic Competition, Competitive Markets, and Monopoly?

Monopolistic competition occupies a place between competitive markets, which produces an efficient output at low prices, and monopoly, which produces an inefficient output at high prices. To help explain whether monopolistic competition is desirable or not, we consider the outcomes that individual firms can achieve when facing monopolistic competition in the short run and in the long run. Once you understand how monopolistic competition works, we will be able to compare the long-run equilibrium result with that of competitive markets and then determine if monopolistic competition is efficient.

Monopolistic Competition in the Short Run and the Long Run

A monopolistically competitive firm sells a differentiated product and for this reason has some market power. Recall that in perfect competition, each firm sells the same product, so competitors' products are perfect substitutes, which means that demand is perfectly elastic (flat). In monopolistic competition, each competitor provides a differentiated product, so competitors' products are imperfect substitutes for one another, which means that demand is relatively elastic (less flat), but still flatter (more elastic) than monopoly. Like a monopolist, the monopolistic competitor uses the profit-maximizing rule, $MR = MC$, and locates the corresponding point on its demand curve to determine the best price to charge and the best quantity to produce. Whether the firm earns a profit, experiences a loss, or breaks even is a function of other firms entering and exiting the market. Recall that entry and exit do not take place in the short run; in the long run, however, firms are free to enter an industry when they see a

potential for profits or leave if they are making losses. Therefore, entry and exit regulate how much profit a firm can make in the long run.

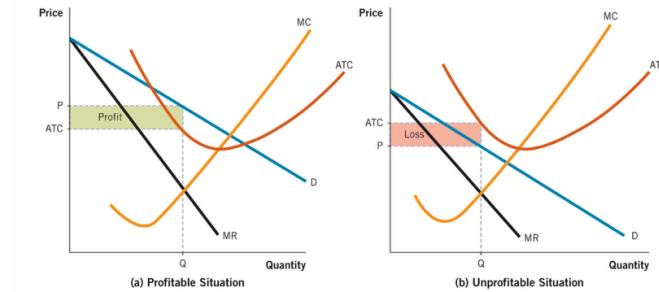
Monopolistic Competition in the short run

Figure 12.1 depicts a firm in a monopolistically competitive environment. In panel (a), the firm makes a profit; Panel (b) shows the same firm incurring a loss after a new competitor opens nearby. In each case, the firm uses the profit-maximizing rule to determine the best price to charge by locating the point at which marginal revenue equals marginal cost. This calculation establishes the profit-maximizing output (Q) along the vertical dashed line; the firm determines the best price to charge (P) by following the dashed horizontal line from the demand curve to the vertical axis.

FIGURE 12.1

The Monopolistically Competitive Firm in the Short Run

In this figure, we see how a single monopolistically competitive firm may make a profit or incur a loss depending on the demand conditions it faces. Notice that the marginal cost curve (MC) and average total cost curve (ATC) are identical in both panels because we are considering the same firm. The only functional difference is the location of the demand curve (D) and marginal revenue curve (MR). The demand in (a) is high enough for the firm to make a profit. In (b), however, there is not enough demand, so the firm experiences a loss.



In panel (a), we see that because price is greater than average total cost ($P > ATC$), the firm makes a short-run economic profit; the situation in panel (b) is different. Because $P < ATC$, the firm experiences a short-run economic loss; what accounts for the difference? Because we are considering the same firm, the marginal cost (MC) and average total cost (ATC) curves are identical in both panels. The only functional difference is the location of the demand (D) and marginal revenue (MR) curves. The demand in panel (a) is high enough for the firm to make profit. In panel (b), however, there is not enough demand; perhaps too many customers have switched to the new firm, so even though the monopolistic competitor

has some market power, if demand is too low, the firm may not be able to price its product high enough to make a profit.

Monopolistic Competition in the long run

In the long run, when firms can easily enter and exit a market, competition will drive economic profit to zero. This dynamic should be familiar to you from our previous discussions of competitive markets. If a firm is making economic profit, that profit attracts new entrants to the business; then the larger supply of competing firms will cause the demand for an individual firm's product to contract. Eventually, as more firms enter the market, it is no longer possible for existing firms to make an economic profit. A reverse process unfolds in the case of a market experiencing a loss. In this case, some firms exit the industry; then consumers have fewer options to choose from, and the remaining firms experience an increase in demand. Eventually, demand increases to the point at which firms no longer experience a loss.

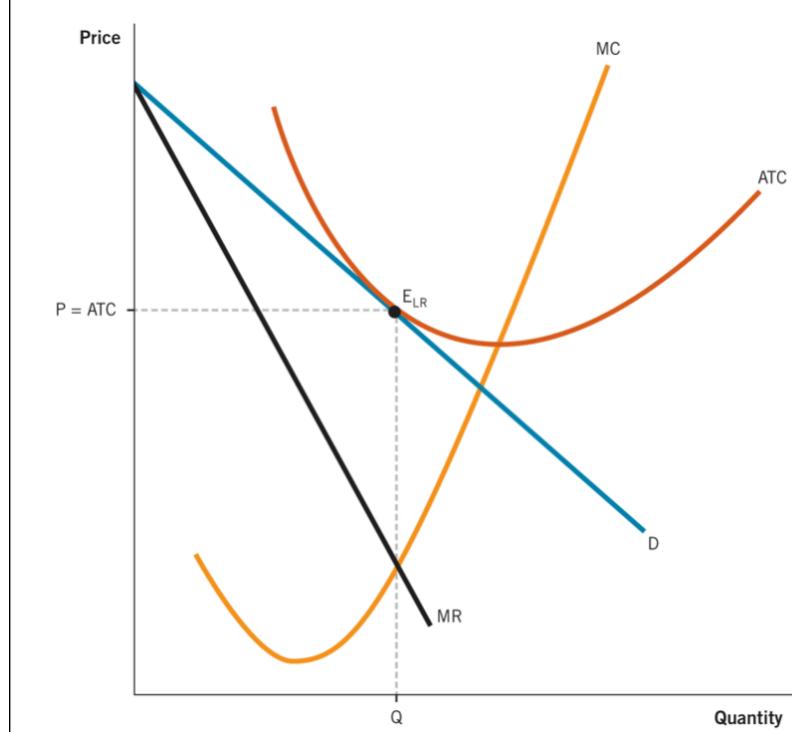
Figure 12.2 shows the market after the long-run adjustment process takes place. Price (P) is just equal to the average total cost of production (ATC) at the profit-maximizing rate of output (Q). At this point, firms are earning zero economic profit, as noted by $P = ATC$ along the vertical axis; the market reaches a long-run equilibrium at the point where there is no reason for firms to enter or exit the industry. Note that the demand curve is drawn **tangent** to the average total cost curve (touching at one place). If demand were any larger, the result would look like panel (a) in **Figure 12.1** and firms would experience an economic profit. Conversely, if demand were any lower, the result would look like panel (b) in **Figure 12.1** and firms would experience an economic loss. Where entry and exit exist, profits and losses are not possible in the long run. In this way, monopolistic competition resembles a competitive market.

The firm's success will attract attention and encourage rivals to enter the market. As a result, the short-run profits that the firm enjoys will erode; as long as profits occur in the short run, other competitors will be encouraged to enter, while short-run losses will prompt some existing firms to close. The dynamic nature of competition guarantees that long-run profits and losses are not possible.

FIGURE 12.2

The Monopolistically Competitive Firm in the Long Run

Entry and exit cause short-run profits and losses to disappear in the long run, which means that the price charged (P) must be equal to the average total cost (ATC) of production. At this point, firms are earning zero economic profit, as noted by $P = ATC$ along the vertical axis. The market reaches a long-run equilibrium (E_{LR}) at the point where there is no reason for firms to enter or exit the industry.



Monopolistic Competition and Competitive Markets

We have seen that monopolistic competition and competitive markets similar; both market structures drive economic profit to zero in the long run, but monopolistic competitors enjoy some market power, which is a crucial difference. In this section, we compare pricing and output decisions in these two market structures, then we look at issues of scale and output.

The relationship between price, marginal cost, and long-run average cost

Monopolistically competitive firms have some market power, which enables them to charge slightly more than firms in competitive markets.

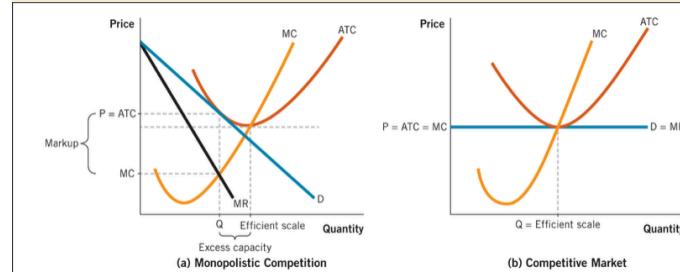
Figure 12.3 compares the long-run equilibrium between monopolistic competition and a competitive market; turning first to the firm in a market characterized by

monopolistic competition, show in panel (a), notice that the price (P) is greater than the marginal cost (MC) of making one more unit. The difference between P and MC is known as the markup. **Markup** is the difference between the price the firm charges and the marginal cost of production.

FIGURE 12.3

The Long-Run Equilibrium in Monopolistic Competition and Competitive Markets

There are two primary differences between the long-run equilibrium in monopolistic competition (a) and a competitive market (b). First, monopolistic competition produces markup, because P is greater than MC. In a competitive market, $P = MC$. Second, the output in monopolistic competition is smaller than the efficient scale. In a competitive market, the firm's output is equal to the most efficient scale.



A markup is possible when a firm enjoys some market power; products such as bottled water, cosmetics, prescription medicine, eyeglass frame, brand-name clothing, restaurant drinks, and greeting cards all have hefty markups...other companies use special packaging, while the marketing of the item is unquestionably a successful business strategy, the markup means consumers pay more. You can observe this result in panel (a) of Figure 12.3, where the price under monopolistic competition is higher than the price in a competitive market, shown in panel (b).

Next, look at the ATC curves in both panels because a monopolistic competitor has a downward-sloping demand curve, the point of tangency between the demand curve and the ATC curve is different from the point of tangency in a competitive market. The point where $P = ATC$ is higher

under monopolistic competition, panel (b) shows the demand curve just tangent to the ATC curve at the ATC's lowest point in a competitive market. Consequently, we can say that monopolistic competition produces higher prices than a competitive market does. If this result seems odd to you, recall that entry and exit do not ensure the lowest possible price, only that the price is equal to the average total cost of production; in a competitive market, where the demand curve is horizontal, the price is always the lowest possible average total cost of production. This is not the case under monopolistic competition, however, the price in monopolistic competition often reflects quality; cheap food is cheap for a reason. Firms may charge more for higher-quality food, but there will still be zero economic profit.

Scale and Output

When a firm produces at an output level smaller than the output level needed to minimize average total cost, we say it has **excessive capacity**. Turning back to panel (a) of **Figure 12.3**, we see excessive capacity in the difference between Q and the efficient scale.

This result differs from what we see in panel (b) of **Figure 12.3** for a competitive market. In a competitive market, the profit-maximizing output is equal to the most efficient scale of operation. This result is guaranteed because each firm sells an identical product and must therefore set its price equal to the minimum point on the average total cost curve. Under monopolistic competition, the profit-maximizing output is less than the minimum efficient scale; monopolistically competitive firms have the capacity to produce more output at a lower cost, but if they produced more, they would have to lower their price because a lower price decreases the firm's marginal revenue, it is more profitable for the monopolistic competitor to operate with excess capacity.

Monopolistic Competition, Inefficiency, and Social Welfare

Monopolistic competition produces a higher price and a lower level of output than a competitive market does. Recall that we looked at efficiency as a way to determine whether a firm's decisions are consistent with an output level beneficial to society. Does monopolistic competition display efficiency?

In **Figure 12.3**, panel (a), we observed that a monopolistic competitor has costs slightly above the lowest possible cost, so the average total costs of a monopolistically competitive firm are higher than those of a firm in a competitive market. This result is not efficient; to achieve efficiency, the monopolistically competitive firm could lower its price to what we would find in competitive markets. However, because a monopolistic competitor's goal is to make a profit, there is no incentive for the firm to lower its price. Every monopolistic competitor has a downward-sloping demand curve, so the demand curve cannot be tangent to the minimum point along the average total cost curve, as seen in panel (a).

Markup is a second source of inefficiency; we have seen that, for a monopolistically competitive firm at the profit-maximizing output level, $P > MC$ by an amount equal to the markup. The price reflects the consumer's willingness to pay, and this amount exceeds the marginal cost of production. A reduced markup would benefit consumers by lowering the price and decreasing the spread between the price and the marginal cost. If the firm did away with the markup entirely and set $P = MC$, the output level would benefit the greatest number of consumers. However, this result would not be practical, at the point where the greatest efficiency occurs, the demand curve would be below average total cost curve and the firm would lose money. It is unreasonable to expect a profit-seeking firm to pursue a pricing strategy that would benefit its customers at the expense of its own profit.

What if the government intervened on behalf of the consumer? Increased efficiency could be achieved through government regulation. After all, the government regulates monopolists to reduce market power and restore social welfare. Couldn't the government do the same in monopolistically competitive markets? Yes and no! It is certainly possible, but not desirable; monopolistically competitive markets have a limited amount of market power, so they cannot make a long-run economic profit like monopolists do. In addition, regulating the prices that firms in a monopolistically competitive market can charge would put many of them out of business, bear in mind we are talking about firms in markets like the fast-food industry...doing away with a significant percentage of these firms would mean fewer places for consumers to grab a quick bite, etc.

Regulating monopolistic competition through marginal cost pricing, or setting $P = MC$, would also create a host of problems like those we discussed for monopoly. A good proportion of the economy consists of monopolistically

competitive firms, so the scale of the regulatory effort would be enormous; and because implementing marginal cost pricing would result in widespread losses, the government would need to find a way to subsidize the regulated firms to keep them in business. The only way to fund these subsidies would be higher taxes, the inefficiencies present in monopolistic competition do not warrant government action.

Varying degrees of product differentiation

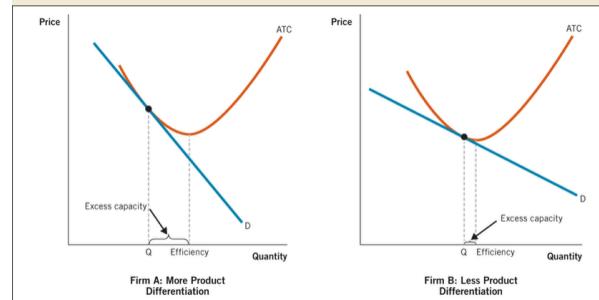
We have seen that products sold under monopolistic competition are more differentiated than those sold in a competitive market and less differentiated than those sold under monopoly. At one end of these two extremes we have competitive markets where firms sell identical products, have no market power, and face a perfectly elastic demand curve. At the other end we have a monopolist that sells a unique product without good substitutes and faces a steep downward-sloping demand curve indicative of highly inelastic demand. What about the firm that operates under monopolistic competition?

Figure 12.4 illustrates two monopolistic competitors with varying degree of product differentiation. High levels of differentiation occur when the firm has an especially attractive location, style, type, or quality of product that is in high demand among consumers and that competitors cannot easily replicate. Consumers have strong brand loyalty for the clothes these firms sell, so the demand curve is quite inelastic, the relatively steep slope of the demand curve means that the point of tangency between demand curve (D) and the total average at a high price; which produces a large amount of excess capacity. In contrast, firm B sells a product only slightly different from its' competitors.

FIGURE 12.4

Product Differentiation, Excess Capacity, and Efficiency

The difference in product differentiation is represented by the steepness (elasticity) of the demand curve, since the demand curve for firm A enjoys more product differentiation. As a result, it has more excess capacity and is less efficient. Firm B sells a product that is only slightly different from its competitors'. In this case, consumers have only weak preferences about which firm to buy from, and consumer demand is elastic. The results are a small amount of excess capacity and a more efficient result.



Why is advertising prevalent in monopolistic competition?

Advertising is a fact of daily life; it is also a means by which companies compete and therefore a cost of doing business in many industries. In the United States, advertising expenditures account for approximately 2% of all economic output annually.

Worldwide, advertising expenses are a little less—about 1% of global economic activity ; while the percentages are small in relative terms, in absolute terms worldwide advertising costs are over half a trillion dollars each year. Is this money well spent? Or is it a counterproductive contest that increases cost without adding value for the consumer? In this section, we will find that the answer is a little of both. Let's start by seeing who advertises.

Why firms advertise

No matter the company or slogan, the goal of advertising is to drive additional demand for the product being sold. Advertising campaigns use a variety of techniques to simulate demand.

Advertising and Demand		
Company/Product	Advertising slogan	How it increases demand
Convention and Visitors Authority/Las Vegas	“What happens here, stays here.”	The slogan attempts to convince travelers that they will have a better vacation in Las Vegas than anywhere else
Red Bull/energy drink	“Red Bull gives you wings”	The tagline promises energy and concentration to get things done and meet the deadlines.
Frito-Lay/potato chips	“Betcha can’t eat just one.”	The message that one potato chip is not enough to satisfy your craving appeals to chip buyers who choose better taste over lower-priced generics.
Apple Computers Inc/Electronics	“Think different”	The slogan subtly hints that competitors' products are all pretty much the same—and boring.
Twitter/Social Media	“It's What's happening.”	The suggestion is that not using the service means being out of touch with the world.
Visa/Debit Card	“It's everywhere you want to be...”	Widespread acceptance and usability are two of the major reasons for carrying a credit card.
Skittles/candy	“Taste the rainbow”	The emphasis is on taste and a variety of flavors.

A successful advertising campaign will change the demand curve in two dimensions: it will shift the demand curve to the right and alter its shape. Turning to **Figure 12.5**, we see this change; **first**, the demand curve shifts to the right in response to the additional demand created by the advertising. **Second**, the demand curve becomes more inelastic, or slightly more vertical; this change in shape happens because advertising has highlighted features that make the product attractive to specific customers who are now more likely to want it. Because demand is more elastic after advertising, the firm increases its market power and can raise its price.

In addition to increasing demand, advertising conveys information that consumers may find helpful in matching their preferences. Advertising tells us about the price of the goods offered, the location of products, and the introduction of new products. Firms also use advertising as a competitive mechanism to underprice one another. Finally, an advertising campaign signals quality; firms that run expensive advertising campaigns are making significant investment in their product, it is highly unlikely that a firm would spend a great deal on advertising if it did not think the process would yield a positive return. So a rational consumer can infer that firms spending a great deal on advertising are likely to have a higher-quality product than a competitor who does not advertise.

Advertising in different markets

Many firms engage in advertising, but advertising is not equally productive in all market structures. In our continuum from competitive markets to monopoly, markets that function under monopolistic competition invest the most in advertising.

Advertising in competitive markets

As you know by now, competitive firms sell nearly identical products at an identical price. For this reason, advertising raises a firm's costs without directly influencing its sales. Advertising for an undifferentiated good functions like a public good for the industry as a whole: the benefits flow to every firm in the market through increased market demand for the product. Each firm sells essentially the same good, so consumers can find the product at many competing locations at the same price. An individual firm that advertises in this market is at a competitive disadvantage because it will have higher costs that it cannot pass on to the consumer.

This does not mean that we never see advertising in competitive markets, although individual firms do not benefit from advertising, competitive industries as a whole can.

Advertising under monopolistic competition

Advertising is widespread under monopolistic competition because firms have differentiated products. Each firm's advertising increases the demand for its product and changes the slope of the demand curve. In short, the gains from advertising go directly to the firm spending the money; these benefits generate a strong incentive to advertise to gain new customers or to keep customers from switching to other products because each firm feels the same way. Advertising becomes the norm among monopolistically competitive firms.

Advertising as a monopolist

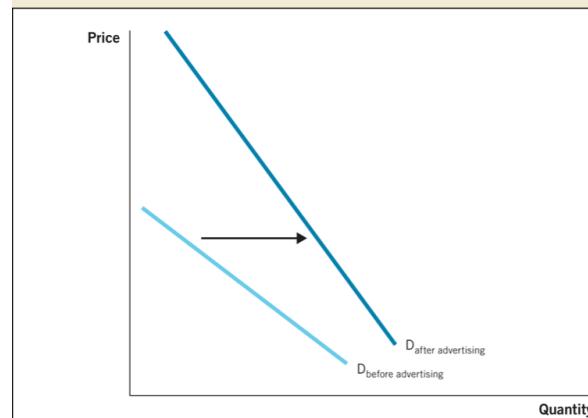
The monopolist sells a unique product without close substitutes. The fact that consumers have few, if any, good alternatives when deciding to buy the good makes the monopolist less likely to advertise than a monopolistic competitor. When consumer choice is limited, the firm does not have to advertise to get business. In addition, the competitive aspect is missing, so there is no need to advertise to prevent consumers from switching to rival products. However, that does not mean that the monopolist never advertises.

The monopolist may wish to advertise to inform the consumer about its product and stimulate demand. This strategy can be beneficial as long as the gains from advertising are enough to cover the cost of advertising.

FIGURE 12.5

Advertising and the Demand Curve

A successful advertising campaign increases demand. Advertising also makes the demand curve more inelastic, or vertical, by informing consumers about differences they care about. After advertising, consumers desire the good more intensely, which makes the demand curve for the firm's product somewhat more vertical.



The Negative Effects of Advertising

We have seen the benefits of advertising, but there are also drawbacks. Two of the most significant drawbacks are that advertising raises costs and can be deceitful.

Advertising and costs

Advertising costs are reflected in the firm's average total cost curve.

Figure 12.6 shows the paradox of advertising for most firms. When a firm advertises, it hopes to increase demand for the product and sell more units—say from point 1 at Q_1 to point 2 at the higher quantity Q_2 . If the firm can sell enough additional units, it will enjoy economies of scale, and the average total cost will fall from ATC_1 to ATC_2 . This return on the advertising investment looks like a good business decision.

However, the reality of advertising is much more complex; under monopolistic competition, each firm is competing with many other firms selling somewhat different products. Rival firms will respond with advertising of their own; this dynamic makes advertising the norm in monopolistic competition...each firm engages in competitive advertising to win new customers and keep the old ones, as a result, the impact on each individual firm's demand largely cancels out. This result is evident in the movement from point 1 to point 3 in **Figure 12.6**. Costs rise from ATC_1 to ATC_3 on the higher LRATC curve, but the quantity demanded may remain at Q_1 . The net result is that advertising creates higher costs but no change in quantity produced and a decrease in profit, in this case, we can think of advertising as causing negative **business-stealing externality** whereby no individual firm can easily gain market share but feels compelled to advertise to protect its customer base.

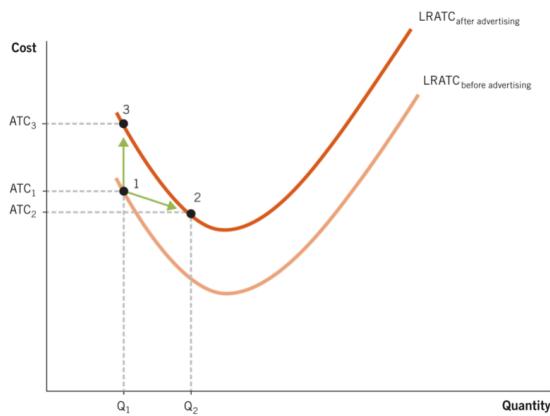
We have seen that advertising raises costs for the producer. It also raises prices for consumers, in fact, consumers who consistently favor a particular brand of a product have more inelastic demand than those who are willing to switch from one product to another. Therefore, brand loyalty often means higher prices.

Truth in advertising

Finally, many advertising campaigns are not just informative—they are designed to produce a psychological response. When an ad moves you

FIGURE 12.6**Advertising Increases Cost**

By advertising, the firm hopes to increase demand (or quantity) from point 1 to point 2. In this scenario, the increase in demand from Q_1 to Q_2 is large enough to create economies of scale even though advertising causes the long-run average total cost curve (LRATC) to rise. Because monopolistically competitive firms each advertise, the advertising efforts often cancel one another out. As a result, long-run average total costs rise without demand increasing much, so the firm may move from point 1 to point 3 instead.



to buy or act in a particular way, it becomes manipulative; because advertising can be such a powerful way to reach customers, there is a temptation to lie about a product. To prevent firms from spreading misinformation about their products, the Federal Trade Commission (FTC) regulates advertising and promotes economic efficiency.

Of course, even with regulatory oversight, consumers must still be vigilant, at best, the FTC can remove products from the market and levy fines against companies that make unsubstantial claims. However, the damage is often already done—the Latin phrase **caveat emptor**, or “buyer beware”, sums up the dangers of false information.

Sometimes the way a product is advertised is not illegal but is still borderline unethical. Firms often engage in price deception, or tricks to make you think a price is lower than it really is.

Conclusion

Firms willingly spend on advertising because it can increase demand, build loyalty, and provide consumers with useful information about differences in products. Monopolistic competitors advertise and mark up their products like monopolists, but, like firms in a

competitive market, they cannot earn long-run profits. While an economic profit is possible in the short run in all three types of market structure (perfect competition, monopolistic competition, and monopoly), only the monopolist, whose business has significant barriers to entry can, earn an economic profit in the long run; entry and exit cause long-run profits to equal zero in competitive and monopolistically competitive markets.

Monopolistic competitors are price makers who fail to achieve the most efficient welfare-maximizing output for society, but this observation does not tell the entire story. Monopolistic competitors do not have as much market power or create as much excess capacity or markup as monopolists. Consequently, the monopolistic competitor lacks the ability to exploit consumers. The result is not perfect, but widespread monopolistic competition generally serves consumers and society well.

In the next chapter, we continue our exploration of market structure with oligopoly, which produces results that are much closer to monopoly than monopolistic competition.

Oligopoly and Strategic Behavior

ECO 304K: Introduction to Microeconomics

Cell Phone companies are competitive

The cell phone industry has features of both competition and monopoly: competition is fierce, but smaller firms and potential entrants into the market find it difficult to enter and compete. This mixture of characteristics represents another form of market structure—**oligopoly**. In this chapter, we examine oligopoly by comparing it with other market structures already familiar; we then look at some of the strategic behaviors firms in an oligopoly employ, an examination that leads us into the fascinating topic of game theory.

Big Questions

- What is oligopoly?
 - Oligopoly is a type of market structure that exists when a small number of firms sell a differentiated product in a market with significant barriers to entry. An oligopolist is like a monopolistic competitor in that it sells differentiated products; it is like a monopolist in that it enjoys significant barriers to entry. The small number of sellers in oligopoly leads to mutual interdependence.
 - Oligopolists have a tendency to collude and to form cartels in the hope of achieving monopoly-like profits.
 - Oligopolistic markets are socially inefficient because price and marginal cost are not equal. The result under oligopoly falls somewhere between the competitive market and monopoly outcomes.
- How does game theory explain strategic behavior?
 - Game theory helps to determine when cooperation among oligopolists is most likely to occur; in many cases, cooperation fails to occur because decision-makers have dominant strategies that lead them to be uncooperative. As a result, firms compete with price or advertising when they could potentially earn more profit by curtailing these activities.

- Games become more complicated when they are played multiple times, so short-run dominant strategies often disappear. Whenever repeated interaction occurs, decision-makers fare better under tit-for-tat, an approach that maximizes the long-run profit.
- How do government policies affect oligopoly behavior?
 - Antitrust law is complex, and cases are hard to persecute; nevertheless, these laws are essential in providing oligopolistic firms an incentive to compete rather than collude.
 - Antitrust policy limits price discrimination, exclusive dealings, tying arrangements, mergers and acquisitions that limit competition, and predatory pricing.
- What are network externalities?
 - A network externality occurs when the number of customers who purchase or use a good influences the quantity demanded. The presence of significant positive network externalities can cause small firms to go out of business.

What is Oligopoly?

Oligopoly is a form of market structure that exists when a small number of firms sell a product in a market with significant barriers to entry. An oligopolist is like a monopolistic competitor in that it often sells a differentiated product; but like pure monopolists, oligopolists enjoy significant barriers to entry. **Table 13.1** compares the differences and similarities between the four market structures.

Table 13.1

Comparing Oligopoly to other market structures			
Competitive Market	Monopolistic competition	Oligopoly	Monopoly
Many sellers	Many sellers	A few sellers	One seller
Similar products	Differentiated product	Differentiated product (most of the time)	Unique product without close substitutes
Free entry and exit	Easy entry and exit	Barriers to entry	Significant barriers to entry

We have seen that firms monopolistically competitive markets usually have a limited amount of market power. As a result, buyers often find low prices (but not as low as competitive markets) and wide availability. In contrast, an oligopolist sells in a market with significant barriers to entry and fewer rivals. Thus, the oligopolist has more market

power than a firm operating under monopolistic competition. However, because an oligopolist has as much market power as a monopolist.

Our study of oligopoly begins with a look at how economists measure market power in an industry; we then work through a simplified model of oligopoly to explore the choice that oligopolists make.

Measuring the concentration of industries

In markets with only a few sellers, industry output is highly concentrated among a few large firms. Economists use **concentration ratios** as a measure of the oligopoly power present in an industry; the most common measure, known as the four-firm concentration ratio, expresses the sales of the four largest firms in an industry as a percentage of that industry's total sales. The ratio is determined by taking the output of the four largest firms in an industry and dividing that output by the total production in the entire industry.

Collusion and Cartels in a Simple Duopoly Example

In this section, we explore the two conflicting tendencies found in oligopoly: oligopolists would like to act like monopolists, but they often end up competing like monopolistic competitors. To help us understand oligopoly behavior, we start with a simplified example: an industry consisting of only two firms, known as a **duopoly**. Duopolies (such as Boeing and Airbus in the wide-body jet) are rare in national and international markets, but not that uncommon in small, local markets. With **Table 13.3** as our guide, we will examine the output in this market under three different scenarios: competition, monopoly, and duopoly.

Duopoly sits between the two extremes; competition still exists, but is not as extensive as you would see in competitive markets, which ruthlessly drive the price down to cost. Nor does the result always mirror that of monopoly, where competitive pressures are completely absent. In an oligopoly, a small number of firms feel competitive pressures and also enjoy some of the advantages of monopoly.

Competitive Outcome

Recall that competitive markets drive the prices down to the point at which marginal revenue is equal to the marginal cost; because these firms are in business to make money, they will not provide something for nothing.

Monopoly Outcome

At the other extreme of the market structure continuum, a monopolist faces no competition, and price decisions do not depend on the activity of other firms. A

monopolist can search for the price that brings in the most profit. Compared with a competitive market, the monopoly price is higher and the quantity sold is lower; the result is a loss of efficiency.

Duopoly Outcome

In a duopoly, the two firms can decide to cooperate—even though this practice is illegal in the United States, as we will discuss shortly. If the duopolist cooperate, we say that they collude. **Collusion** is an agreement between rival firms that specifies the price each firm charges and the quantity it produces; the firms that collude can act like a single monopolist to maximize their profits.

When two or more firms act in unison, economists refer to them as a **cartel**. Many countries prohibit cartels; in the United States, **antitrust laws** prohibit collusion. However, even if collusion were legal, it would probably fail more often than not.

From what we know about competitive markets, we might expect the competition between the two firms to cause a price war in which price eventually falls to zero. Duopolist are unlikely to participate in an all-out price war because both firms would no longer be making any profit, but we cannot know to what extent competitive pressures will determine each firm's decision. We see that a market with a small number of sellers is characterized by **mutual interdependence**, which is a market situation in which the actions of one firm have an impact on the price and output of its competitors. As a result, a firm's market share is determined by the products it offers, the price it charges, and the actions of its rivals.

Oligopolists want to emulate the monopoly outcome, but the push to compete with their rivals often makes it difficult to maintain a cartel. Yet the idea that cartels are unstable is not guaranteed. When a stable cartel is not achieved, firms in oligopoly fall short of fully maximizing profits; but they also do not compete to the same degree as firms in competitive markets. Therefore, when a market is an oligopoly, output is likely to be higher than under a monopoly and lower than within a competitive market. As you would expect, the amount of output affect the prices; the higher output (compared with monopoly) makes oligopoly prices generally lower than monopoly prices, and lower output (compared with a competitive market) makes oligopoly prices higher than those found in competitive markets.

In many industries, smaller firms may take a cue from the decisions made by the price leader. **Price leadership** occurs when a single firm, known as the price leader, produces a larger share of the total output in the industry. The price leader sets the price and the output level that maximizes its own profits. Smaller firms then set their prices to match the price leader; because the impact on the price is small to begin with, it makes sense that smaller rivals tend to follow the price leader.

Price leadership is not illegal because it does not involve collusion, rather, it relies on an understanding that an effort to resist changes implemented by the price leader will lead to both increased price competition and lower profits for every firm in the industry; because the firms act in accordance with one another, this practice is commonly known as **tacit collusion**.

Oligopoly with more than two firms

We have seen how firms behave in a duopoly; what happens when more firms enter the market? The additional of a third firm complicates efforts to maintain a cartel and increases the possibility of a more competitive result.

This a **price effect**, and it reflects how a change in price affects the firm's revenue. When the firm sells an additional unit, it generates additional revenues for the firm; this **output effect** occurs when a change in price affects the number of customers in a market.

The price effect and output effect make it difficult to maintain a cartel when there are more than two firms. Generally, as the number of firms grow, each individual firm becomes less concerned about its impact on the overall price, because any price above marginal cost creates a profit. Therefore, individual firms are more willing to lower prices because doing so creates a large output effect for the individual firm and only a small price effect in the market.

Of course, not all firms are the same size, therefore, smaller and larger firms in an oligopolistic market react differently to the price and output effects. Increased output at smaller firms will have a negligible impact on overall prices because smaller firms represent only a tiny fraction of the market supply; but the same is not true for firms with a large market share. Decisions at these firms will have a substantial impact on price and output because the overall amount supplied in the market will change appreciably...in other words, in an oligopoly, the decisions of one firm directly affect other firms.

How does Game Theory explain strategic behavior?

Decision-making under oligopoly can be complex; **game theory** is a branch of mathematics that economists use to analyze the strategic behavior of decision-makers who have to consider the behavior of others around them. In particular, game theory can help us determine what level of cooperation is most likely to occur; a game consists of a set of players, a set of strategies available to those players, and specification of the payoffs for each combination of strategies. The game is usually represented by a payoff matrix that shows the players, strategies, and payoffs; it is presumed that each player acts simultaneously or without knowing the actions of the other.

In this section, we will learn about the prisoner's dilemma, an example from game theory that helps us understand how dominant strategies often frame short-run decisions. (In its simplest form, the prisoner's dilemma is a game played just once, not

repeatedly over time.) We will use the idea of the dominant strategy to explain why oligopolists often choose to advertise. Finally, we will come full circle and argue that the dominant strategy in a game may be overcome in the long run, through repeated interactions.

Strategic Behavior and Dominant Strategy

We have seen that in oligopoly there is mutual interdependence: rival's business choices affect the earnings the other rivals can expect to make. To learn more about the decisions firms make, we will explore a fundamental problem in game theory known as **prisoner's dilemma**.

The scenario goes like this: two prisoners are being interrogated separately about a crime they both participated in, and each is offered a plea bargain to cooperate with the authorities by testifying against each other. If both suspects refuse to cooperate with the authorities, neither can be convicted of a more serious crime, though they will both have to spend more time in jail; but the police have offered full immunity if one cooperates and the other does not. This means that each suspect has an incentive to betray the other. The problem is that if they both confess, they will spend more time in jail than if they had both stayed quiet; when decision-makers face incentives that make it difficult to achieve mutually beneficial outcomes, we say they are in a **prisoner's dilemma**. This situation makes the payoff for cooperating with the authorities more attractive than the result of keeping quiet.

We can understand the outcomes of the prisoner's dilemma by looking at the payoff matrix in Figure 13.1.

When a player always prefers one strategy, regardless of what his opponent chooses, we say it is a **dominant strategy**. We can see a dominant strategy at work in the case of our two suspects; they know that if they both keep quiet, they will spend one year in jail. The dilemma occurs because both suspects are more likely to testify and get 10 years in jail. The choice to testify is obvious for two reasons; **first**, neither suspect can monitor the actions of the other after they are separated. **Second**, once each suspect understands that his partner will save jail time if he testifies, he realizes that the incentives are not in favor of keeping quiet.

The dominant strategy in our example is also a Nash equilibrium, named for mathematician John Nash. A **Nash equilibrium** occurs when all economic decision-makers have no incentive to change their current decision; if each

suspect reasons that the other will testify, the best response is also to testify. Each suspect may wish that he and his partner could coordinate their actions and agree to keep quiet. However, without possibility of coordination, neither has an incentive to withhold testimony; so they both think strategically and decide to testify.

Duopoly and the Prisoner's Dilemma

The prisoner's dilemma example suggests that cooperation can be difficult to achieve; what this means for oligopoly is that it is not natural or easy for firms to collude. This is the prisoner's dilemma: each firm has an incentive to lower its price to generate more revenue than if they had colluded and kept prices high, but acting on this incentive causes them both to earn less revenue than if they had kept prices high.

Advertising and Game Theory

We have seen that oligopolists function like monopolistic competitors in that they sell differentiated products. We know that advertising is commonplace in markets with a differentiated product; in the case of an oligopoly, mutual interdependence means that advertising can create a contest between firms trying to gain customers. The result may be skyrocketing advertising budgets and little, or no, net gain of customers. Therefore, oligopolists have an incentive to scale back advertising, but only if their rivals also agree to scale back. Like all cooperative action among competitors, this is easier said than done.

This dilemma is that each firm needs to advertise to market its product and retain its customer base, but most advertising expenditures end up canceling each other out and costing the companies millions of dollars.

Escaping the Prisoner's Dilemma in the long run

We have seen how game theory can be a useful tool for understanding strategic decision-making in noncooperative environments. However, the dominant strategy does not consider the possible long-run benefits of cooperation.

Game theorist Robert Axelrod decided to examine the choices participants make in the long-run setting; he ran a sophisticated computer simulation in which he invited scholars to submit strategies for securing points in a prisoner's dilemma tournament over many rounds. All the submissions were collected and paired, and the results were scored; After each simulation, Axelrod eliminated the weakest strategy and reran the tournament with the remaining strategies.

This evolutionary approach continued until the best strategy remained; among all strategies, including those that were solely cooperative or noncooperative, tit-for-tat dominated. **Tit-for-tat** is a long-run strategy that promotes cooperation among participants by mimicking the opponent's more recent decision with repayment in kind. As the name implies, a tit-for-tat strategy is one in which you do whatever your opponent does. If your opponent breaks the agreement, you break the agreement, too. If the opponent behave properly, then you behave properly, too.

Because the joint payoffs for cooperation are high in a prisoner's dilemma, tit-for-tat begins with the players cooperating. In subsequent rounds, the tit-for-tat strategy mimics whatever the other player did in the previous round. The genius behind tit-for-tat is that it changes the incentives and encourages cooperation.

The prisoner's dilemma nicely captures why cooperation is so difficult in the short run, but most interactions in life occur over the long run. For example, scam artists and sketchy companies take advantage of short-run opportunities that cannot last because relationships in the long run—with businesses and with people—involve mutual trust. Cooperation is the default because you know that the other side is invested in the relationship; under these circumstances, the tit-for-tat strategy works well.

Sequential Games

Not all games involve simultaneous decisions; sometimes one player must move first and then the other player responds to the first move. In this case it is possible for the first player to utilize **backward induction** to get the best possible result. Backward

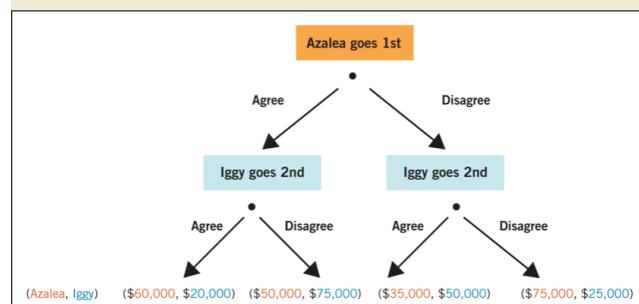
induction in game theory is the process of deducing backward from the end of a scenario to infer a sequence of optimal actions.

However, if we let one player go first, the game has a predictable conclusion; to see this process at work, look at Figure 13.5. This type of diagram is known as **decision tree**; it illustrates all of the possible outcomes.

FIGURE 13.5

Decision Tree for a Sequential Game

Azalea will agree, knowing that Iggy will disagree. This game guarantees Azalea \$50,000 and Iggy \$75,000.



There are many examples of sequential games in life; chess and checkers are two popular board games that utilize backward induction. Likewise, many businesses decisions are also sequential in nature, and once a particular path is taken, it becomes easier to predict how future decisions will unfold. For instance, when a firm decide to launch a new advertising campaign, it is easier for the firm to predict how a rival will react by examining the remaining choices along a decision tree.

A caution about game theory

Game theory is a decision-making tool, but not all games have dominant strategies that make players decisions easy to predict. The preferred choice is strictly a function of what the other player selects; winning at business in the long run often occurs because you are one step ahead of the competition, not because you deploy strategy that attempts to take advantage of a short-run opportunity.

How do government policies affect oligopoly behavior?

When oligopolists in an industry form a cooperative alliance, they function like a monopoly. Competition disappears, which is not good for society; one way to improve the social welfare of society is to restore competition and limit monopoly practice through policy legislation.

Antitrust policy

Efforts to curtail the adverse consequences of oligopolistic cooperation began in the **Sherman Antitrust Act of 1890**, the first federal law to place limits on cartels and monopolies. The Sherman Act was created in response to the increase in concentration ratios in many leading U.S industries, including steel, railroads, mining, textiles, and oil. Prior to the passage of the Sherman Act, firms were free to pursue contracts that created mutually beneficial outcomes; once the act took effect, however, certain cooperative actions became criminal.

Section 2 of the Sherman Act reads, “Every person who shall monopolize, or attempt to monopolize, or combine or conspire with any person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a felony.”

The **Clayton Act of 1914** targets corporate behaviors that reduce competition. Large corporations had been vilified during the presidential election of 1912, and

the Sherman Act was seen as largely ineffective in curbing monopoly power. To strengthen antitrust policy, the Clayton Act added to the list of activities deemed socially detrimental, including:

1. *Price discrimination*

1. If it lessens competition or creates monopoly

2. *Exclusive dealings*

1. That restrict a buyer's ability to deal with competitors

3. *Tying arrangements*

1. That requires the buyer to purchase an additional product in order to purchase it the first.

4. *Mergers and acquisitions*

1. That lessen competition, or situations in which a person serves as a director on more than one board in the same industry.

As the Clayton Act makes clear, there are many ways to reduce competition

Over the past hundred years, lawmakers have continued to refine antitrust policy; additional legislation along with court interpretations of existing antitrust law have made it difficult to determine whether a company has violated the law. The U.S Justice Department is charged with oversight, but it often lacks the resources to fully investigate every case; antitrust law is complex, and cases are hard to prosecute, but these laws are essential to maintaining a competitive business environment. Without effective restraints on excessive market power, firms would organize into cartels more often or would find other ways to restrict competition.

Predatory Pricing

While firms have a strong incentive to cooperate in order to keep prices high, they also want to keep potential rivals out of the market. **Predatory pricing** is the practice of setting prices deliberately below average variable costs with the intent of driving rivals out of the market; the short-run loss in order to prevent rivals from entering the market or to drive rival firms out of business in the long run. Once the rivals are gone, the firm should be able to act like a monopolist.

Predatory pricing is illegal, but difficult to prosecute; neither the court system or economists have a simple rule that helps to determine when a firm steps over the line; predatory pricing can look and feel like spirited competition. Moreover, the concern is not the competitive aspect or lower prices, but the effect on the market when all rivals fail. To prove that predatory pricing has occurred, the courts need evidence that firm's prices increased significantly after its rivals failed.

What are network externalities?

We end this chapter by considering a special kind of externality that often occurs in oligopoly. A **network externality** occurs when the number of customers who purchase or use a good influence the demand. When a network externality exists, firms with many customers often find it easier to attract new customers and to keep their regular customers from switching to other rivals.

Most network externalities involve the introduction of new technologies. For instance, some technologies need to reach a critical mass before consumers can effectively use them.

In addition to the advantages of forming a larger network, customers may face significant switching costs if they leave. **Switching costs** are the costs incurred when a consumer changes from one supplier to another. When consumers face switching costs, the demand for the existing product becomes more inelastic; as a result, oligopolists not only leverage the number of customers they maintain in their network, but also try to make switching to another network more difficult. For instance, firms promote customer loyalty through frequent flier benefits, hotel reward points, and credit card reward programs to create higher switching costs.

Oligopolists are keenly aware of the power of network externalities; as new markets develop, the first firm into an industry often gains a large customer base. When there are positive network externalities, the customer base enables the firm to grow quickly. In addition, consumers are often more comfortable purchasing from an established firm; these two factors favor the formation of large firms and make it difficult for smaller competitors to gain customers. As a result, the presence of significant positive network externalities causes small firms to be driven out of business or force them to merge with larger competitors.

Conclusion

Firms in oligopoly markets can compete or collude to create monopoly conditions; the result us often hard to predict. In many cases, the presence of a dominant short-run

strategy causes firms to compete on price and advertising even though doing so yields a lower economic profit. In contrast, the potential success for a tit-for-tat strategy suggests that oligopolistic firms are capable of cooperating to jointly maximize their long-run profits. Whether oligopoly mirror the result found in monopolistic competition or monopoly matters a great deal because society's welfare is higher when more competition is present; because oligopoly is not a market structure with a predictable outcome, each oligopolistic industry must be assessed on a case-by-case basis by examining data and utilizing game theory. For these reasons, the study of oligopoly is one of the most fascinating parts of the theory of the firm.

Consumer Choice

ECO 304K: Introduction to Microeconomics

The more money you have, the happier you'll be?... Not so fast

Some lottery winners do just fine, of course; but it's common for winners to deal badly with the sudden wealth, and all the attention from "friends" who suddenly emerge from the woodwork, that there's a name for this phenomenon: the lottery curse. The root problem is that winners suddenly face a whole ton of choices they hadn't faced before and aren't well-equipped to make. In this chapter, we use our understanding of income constraints, price, and personal satisfaction to determine which economic choices yield the greatest benefits.

Big Questions

- How do economists model consumer satisfaction?
 - Economists model consumer satisfaction by examining utility, which is a measure of the level of satisfaction that a consumer enjoys from the measure of the level of satisfaction that a consumer enjoys from the consumption of goods and services
 - Utility diminishes with additional consumption; this property limits the amount of any particular good or service that a person will consume
- How do consumers optimize their purchasing decisions?
 - Consumers optimize their purchasing decisions by finding the combination of goods and services that maximizes the level of satisfaction from a given income or budget. The consumer optimum occurs when a consumer maximizes the utility from his or her income or budget, so that the marginal utility per dollar spent on every item purchased is equal to that of every other item purchased.
 - Changes in price have two distinct effects on consumer behavior: if the price falls, the marginal utility per dollar spent will be higher, as a result, consumers will substitute the product that has become relatively less expensive. This is the substitution effect, if the lower price also results in substantial savings, it causes an increase in purchasing power, known as the real-income effect.

- What is the diamond-water paradox?
 - The diamond-water paradox explains why water, which is essential to life, is inexpensive, while diamonds, which do not sustain life are expensive; Many people of Adam Smith's era, in the 18th century, found the paradox perplexing. We can save the diamond-water paradox by recognizing that the price of water is low because its supply is abundant, and the price of diamonds, and the price of diamonds is high because their supply is small; if water is rare as diamonds, there is no doubt that the price of water exceed the price of diamonds.

How do economists model consumer satisfaction?

Economists typically feel the ideal solution to this problem achieves the highest **utility** summed across all people—what is often called “the greatest good for the greatest number”. In this chapter, we show how economists put numbers to happiness and calamity; that may seem coldhearted, but the goal is to benefit society by using scarce resources to do the most good. In doing so, economists are often forced to consider the opportunity costs of taking one path and forgoing another; when economists make these choices, they use particular economic tools.

To better understand the decisions that consumers make, economists attempt to measure the satisfaction that consumers get when they make purchases. **Utility** is defined as a measure of satisfaction that a consumer enjoys from the consumption of goods and services. Utility theory seeks to measure contentment, or satisfaction; to understand why people buy the goods and services they do, we need to recognize that some products produce more utility than others and that everyone receives the different levels of satisfaction from the same good or service. In other words, utility varies from individual to individual; to quantify this idea of relative satisfaction, economists measure utility with a unit they refer to as a **util**.

There is tremendous value in modeling decisions this way; when we understand utility, we can predict what people are likely to purchase. We model consumer behavior in a manner similar to the way we model how a firm makes decisions or how the labor market works. We expect the firm to maximize profits, the laborer to accept the best offer, and the consumer to find the combination of goods that gives the most utility.

Utility, or what most of us think of as satisfaction, is a balance between economic and personal factors. Even though there is an inherent problem with equating money and satisfaction, this has not stopped researchers from exploring the connection.

In the next section, we explore the connection between total utility and marginal utility; this connection will help us understand why more money does not necessarily bring more satisfaction.

Total Utility and Marginal Utility

Thinking about the choices that consumers make can help us understand how to increase total utility. **Marginal utility** is the additional satisfaction derived from consuming one more unit of a good or service.

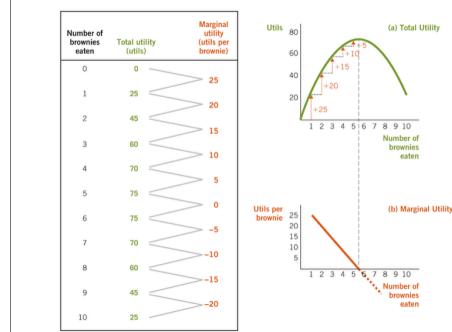
The graph in panel (a) of Figure 16.1 reveals that while the total utility (*the green curve*) rises until it reaches 75, the rate of increase (that is, the increase in marginal utility) falls from 25 utils for the first brownie down to 5 additional utils for the fifth. The marginal utility values from the table are plotted in panel (b), which shows that marginal utility declines steadily as consumption rises.

The relationship between total utility and marginal utility is evident when we observe the dashed line that connects panel (a) and (b); because the marginal utility becomes negative after five brownies are consumed, the total utility eventually falls. To the left of the dashed line, the marginal utility is positive in panel (b) and the total utility is rising in panel (a). Conversely, to the right of the dashed line, the marginal utility is negative, and the total utility is falling.

FIGURE 16.1

Total Utility and Marginal Utility

The relationship between total utility and marginal utility can be seen by observing the dashed line that connects panels (a) and (b). Because the marginal utility becomes negative after five brownies are consumed, the total utility eventually falls after a certain number of brownies are eaten. To the left of the dashed line, the marginal utility is positive in panel (b) and the total utility is rising in panel (a). Conversely, to the right of the dashed line, the marginal utility is negative and the total utility is falling.



Diminishing Marginal Utility

As you can see in panel (b) of Figure 16.1, the satisfaction that a consumer derives from consuming a good or service declines with each additional unit is consumed. Consider what happens when you participate in a favorite activity for an hour and then decide to do something else. **Diminishing marginal utility** occurs when marginal utility declines as consumption increases; the concept of diminishing marginal utility is so universal that it is one of the most widely held ideas in all of economics.

Your own intuition should confirm this theory; if increasing marginal utility were possible, you would find that with every passing second you would enjoy what

you were doing more and never want to stop, because economists do not observe this behavior among rational consumers, we can be highly confident that diminishing marginal utility has tremendous explanatory power.

How do consumers optimize their purchasing decisions?

Maximizing utility requires that consumers get the most satisfaction out of every dollar they spend, or what is commonly called “getting the biggest bang for the buck.” When a consumer gets the most bang for the buck, we say the consumer has optimized his or her purchasing decisions; however, optimization is easier said than done. Over the course of the year, each of us will make thousand of purchases of different amounts; our budgets are generally not unlimited, and we try to spend in a way that enables us to meet both short-run and our long-run needs. The combination of goods and services that maximizes the satisfaction, or utility, we get from income or budget is the **consumer optimum**.

In this section, we examine the decision process that leads to the consumer optimum. We start with two goods and then generalize those findings across a consumer’s entire income or budget.

Consumer Purchasing Decisions

To reach your consumer optimum, you must allocate your available money by choosing goods that give you the most utility per dollar spent; by getting the biggest bang for your buck, you will end up optimizing your choices. This relationship, shown below in terms of marginal utility (MU), helps quantify the decision. If we divide the marginal utility of a good by its price, we get the utility per dollar spent.

By thinking at the margin about which good provides the highest marginal utility, you also maximize your total utility; of course, most people rarely think this way, but as consumers we make marginal choices all the time. Instead of adding up utils, we think “that isn’t worth it” or “that’s a steal”. Consumer choice is not such a conscious calculation as an instinct to seek the most satisfaction; next we extend our analysis by generalizing two-good example.

Marginal Thinking with more than two goods

The idea of measuring utility makes our instinctive sense more explicit and enables us to solve simple optimization problems. Consumers optimum captures this idea by comparing the utility gained with the price paid for every

item a consumer buys. In other words, a consumer's income or budget is balanced so that the ratio of the marginal utility (MU) per dollar spent on every item, from good A to good Z, is equal. In mathematical terms:

$$\frac{MU_A}{Price_A} = \frac{MU_B}{Price_B} = \dots = \frac{MU_Z}{Price_Z}$$

It should be noted that, because goods aren't infinitely divisible (we can't buy a fraction of a soda can), we can't always make the fraction come out exactly equal. Still, we maximize utility by buying the goods with the higher fractions, and in the end the marginal utilities per dollar spent for each good end up approximately equal.

In the next section, we explore the relationship between changes in price and changes in the consumer optimum.

Price Changes and the Consumer Optimum

We can say lower prices increase the marginal utility per dollar spent and cause consumers to buy more of a good. Higher prices have the opposite effect by lowering the marginal utility per dollar spent; if that conclusion sounds an awful lot like the law of demand, it is! We have just restated the law of demand in terms of marginal utility.

We know that according to the law of demand (see Chapter 3), the quantity demanded falls when the price rises, and the quantity demanded rises when price falls—all other things being equal. If we think of consumer desire for a particular product as demand, it makes sense to find a connection between the prices consumers pay, the quantity they buy, and the marginal utility they receive.

A lower price has two effects, ***first***, because the marginal utility per dollar spent is now higher, consumers substitute the product that has become relatively less expensive—that is the **substitution effect**. Second, at the same time, a lower price can also change the purchasing power of income—this is the **real-income effect**.

The real-income effect matters only when prices change enough to cause a measurable effect on the purchasing power of the consumer's income or budget; however, the real-income effect is negligible.

What is the Diamond-Water Paradox?

Now that you understand the connection between prices and utility we can tackle one of the most interesting puzzles in economics—the **diamond-water paradox**. First described by Adam Smith in 1776, the diamond-water paradox explains why water, which is essential to life, is inexpensive, while diamonds, which do not sustain life, are expensive. Many people of Smith’s era found that the paradox perplexing; today, we can use consumer choice theory to answer the question.

Essentially, the diamond-water paradox unfairly compares the amount of marginal utility a person receives from a small quantity of something rare (the diamond) with the marginal utility received from consuming a small amount of additional water after already consuming a large amount.

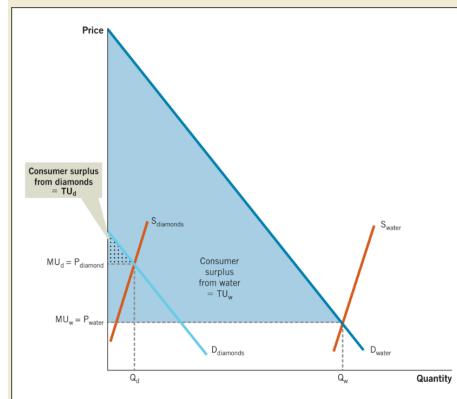
We know that marginal utility is captured in the law of demand and therefore by the price of a good. For example, when the price of diamonds increases, the quantity demanded declines; we learned in Chapter 5 that in graphical terms, the consumer surplus is the area under the demand curve and above the price, or the gains from trade that a consumer enjoys. Therefore, if the price of diamonds rises, consumers will enjoy less surplus when buying them.

Figure 16.2 contrasts the demand supply equilibrium in both the market for water and the market for diamonds. Notice that the consumer surplus is the area highlighted in blue for water and the triangular area highlighted with dots for diamonds; the blue area of total utility for water (TU_W) is much larger than the dotted area of total utility for diamonds (TU_D) because water is essential for life. Therefore, water creates significantly more total utility than diamonds do; however, in most places in the United States, water is very plentiful, so people take additional units of it for granted. In fact, it is so plentiful that if someone were to offer you a gallon of water, you would probably hesitate to take it; but what if someone offered a gallon-size bucket of diamonds? You bet you would take that, therefore, it should not surprise you that something quite as plentiful, water, would yield less marginal utility than

FIGURE 16.2

The Diamond-Water Paradox

The diamond-water paradox exists because people fail to recognize that demand and supply are equally important in determining the value of a good creates in society. The demand for water is large, while the demand for diamonds is small. If we look at the amount of consumer surplus, we observe that the blue area (TU_W), which represents the consumer surplus for water) is much larger than the dotted area (TU_D), which represents the consumer surplus for diamonds) because water is essential for life. As a result, water creates significantly more total utility (TU) than diamonds. However, because water is abundant in most places, the price, P_{water} , is low. In contrast, diamonds are rare and the price, P_{Diamond} , is high.



something rare, diamonds ($MU_W < MU_D$). However, if water were as rare as diamonds, there is no doubt that the price of water would exceed the price of diamonds.

Let's consider how we use water; we can bathe in it, cook with it, and drink it. Each of these uses has high value, so the marginal utility of water is high; but we also use it to water our lawns and fill our fish tanks, those uses are not nearly as essential, so the marginal utility of water for those uses is much lower. The reason we use water in both essential and nonessential ways is that its price is relatively low, so low-value uses, like filling fish tanks, yield enough utility to justify the cost; because water is abundant in most places, the price P_{water} is low, in contrast, diamonds are rare and their price P_{Diamond} is high. The cost of obtaining a diamond means that a consumer must get a great deal of marginal utility from the purchase of a diamond to justify the expense, which explains why diamonds are given as gifts for extremely special occasions.

Conclusion

Does having more money make people more satisfied? The answer is no. More money enables people to buy more goods, but because of diminishing marginal utility, the increase in satisfaction from being able to buy more goods or higher-quality goods become progressively smaller with rising income, so we could say that having more money makes people somewhat satisfied; but it seems appropriate to add that the relationship between quality of life and money is not direct. More money sometimes lead to more utility, and at other times more money means more problems.

As we have seen in this chapter, price plays a key role in determining utility because consumers face a budget and wish to maximize their utility, the prices they pay determine their marginal utility per dollar spent. Comparing the marginal utility per dollar spent across many goods helps us understand individuals' consumption patterns; diminishing marginal utility also helps to describe consumer choice because marginal utility declines with additional consumption, consumer do not exclusively purchase their favorite products. Instead, they diversify their choices in order to gain more utility; in addition, changes in prices have two different effects: one on real income and a separate substitution effect that together determine the composition of the bundle of goods purchased.

In the next chapter, we question how much individuals use consumer choice theory to make their decisions. The approach known as behavioral economics argue that decision-makers are not entirely rational about the choices they make.

Finally, in the appendix that follows, we refine consumer theory by discussing indifference curves; please read the appendix to get a glimpse into how economists model consumer choice in great detail.

Behavioral Economics and Risk Taking

ECO 304K: Introduction to Microeconomics

In this textbook, we have proceeded as if every person were ***Homo economicus***, a rationally self-interested decision-maker. This idealized individual is acutely aware of opportunities in the environment and strives to maximize the benefits received from each course of action while minimizing the costs.

Sometimes our instincts do steer us in the right direction; more commonly, however, they lead us astray in various subtle ways. In this chapter, we'll see how that happens; to fold the broadest possible set of human behavior into economic analysis, we must turn to the field of ***behavioral economics***. Which enables us to capture a wider range of human motivations than the rational-agent model alone affords.

Big Questions

- How do economists explain irrational behavior?
 - Economists use a number of concepts from behavioral economics to explain how people make choices that display irrational behavior. These concepts include bounded rationality, misperception of probabilities, framing effects and priming effects, the status quo bias, intertemporal decision-making, the endowment effect, judgement about fairness, preference reversals, and prospect theory.
 - Folding the behavioral approach into the standard model makes economists' predictions about human behavior much more robust.
- What is the role of risk in decision-making?
 - Risk influences decision-making because people can be risk-averse, risk-neutral, or risk-takers.
 - In the traditional economic model, risk tolerances are assumed to be constant; if an individual is a risk-taker by nature, he or she will takes risks in any circumstances. Likewise, if an individual does not like to take chances, he or she will avoid risk.
 - Maurice Allais proved that many people have inconsistent risk preferences, or what are known as preference reversals. Moreover, he showed that simply

because some people's preferences are not constant does not necessarily mean that their decisions are irrational.

- Prospect theory suggests that individuals weigh the utilities and risks of gains and losses differently and are therefore willing to take on additional risk to try recover losses caused by negative shocks.

How do economists explain irrational behavior?

Like economics, psychology endeavors to understand the choices people make; one key difference is that psychologists do not assume that people always behave in a fully rational way. As a result, psychologists have a much broader toolbox at their disposal to describe human behavior; **behavioral economics** is the field of economics that draws on insights from experimental psychology to explore how people make economic decisions.

Until relatively recently, economists have ignored many human behaviors that do not fit their models. For example, because traditional economic theory assumed that people make optimal decisions, economic theorists did not try to explain why people might make an impulse purchase. Behavioral economists, however, understand that many behaviors contradict standard assumptions about rationality; they employ the idea of **bounded rationality**, which proposes that although decision-makers want a good outcome, either they are not capable of performing the problem-solving that traditional economic theory assumes or they are not inclined to do so.

Bounded rationality, or limited reasoning, can be explained in three ways: **first**, the information the individual uses to make the decision may be limited or incomplete; **second**, the human brain has a limited capacity to process information; **third**, there is often a limited amount of time in which to make a decision. These limitations prevent the decision-maker from reaching the results predicted under perfect rationality.

We will continue our discussion of behavioral economics by examining various behaviors that do not fit assumptions about fully rational behavior. These include misperceptions of probabilities, inconsistencies in decision-making, and judgements about fairness when making decisions. The goal in this section is to help you recognize and understand many of the behaviors that lead to contradictions between what economic models predict and what people actually do.

Misperceptions of Probabilities

Economic models that assume rationality in decision-making do not account for the way people perceive the probability of events. Low-probability events are

often overanticipated, and high-probability events are often underanticipated; to understand why, we consider several familiar examples, including games of chance, difficulties in assessing probabilities, and seeing patterns where none exist.

Games of chance

Playing games of chance is generally a losing proposition, yet even with great odds against winning, millions of people spend money to play games of chance. How can we explain this behavior?

For some people, the remote chance of winning a lottery offers hope that they will be able to purchase something they need but cannot afford or even to escape from poverty. In many cases, people have incomplete information about the probabilities and price structures; most lottery players do not calculate the exact odds of winning. Lottery agencies typically highlight winners, as if the game has a positive expected value, which gets people excited about playing. Imagine how sobering it would be if every headline trumpeting the newest lottery millionaire was followed by all the names of people who lost. In fact, almost all games of chance have **negative expected values** for the participants, meaning that players are not likely to succeed at the game.

Players often operate under the irrational belief that they have control over the outcome; they're sure that playing certain numbers or patterns will bring success. Many players also feel they must stick with their favorite numbers to avoid regret; everyone has heard stories about players who changed from their lucky pattern only to watch it win.

In contrast, some gaming behaviors are rational; when the expected value of a gamble is positive, we actually expect that the more you play, the more likely it is that your earnings will be higher than your losses.

Gambles can also make sense when you have very little to lose or no other options; and some people find the thrill of gambling enjoyable as entertainment, whether they win or lose. However, most gambling behaviors do not have rational motivations; gambling often creates addictions that lead players to make poor financial decisions.

The difficulties in assessing probabilities

In our discussion of games of chance, we saw that people who gamble do not usually evaluate probabilities in a rational way; but this irrational

decision-making also happens with many other behaviors besides gambling.

The difficulty in recognizing the true underlying probabilities, combined with an irrational fear of regret, leads to many poor decisions.

Understanding these tendencies help economists to evaluate why some decisions are difficult to get right.

Seeing patterns where none exist

Two fallacies, or false ways of thinking, help explain how some people make decisions: **the gambler's fallacy** and **the hot hand fallacy**.

The gambler's fallacy is the belief that recent outcomes are unlikely to be repeated and that outcomes that have not occurred recently are due to happen soon. In other words, someone who uses the gambler's fallacy believes that if many "heads" have occurred in a row, then "tails" is more likely to occur next.

The **hot hand fallacy** is the belief that random sequences exhibit a positive correlation (relationship).

Inconsistencies in decision-making

If people were entirely rational, they would always be consistent; so the way a question is asked should not alter our responses, but research has shown that it does. Likewise, rational decision-making requires the ability to take the long-run trade-offs into account: if the returns are large enough, people should be willing to sacrifice current employment for future benefits. Yet many of us make shortsighted decisions; in this section, we examine a variety of decision-making mistakes, including framing effects, priming effects, status quo bias, inter temporal decision-making, and the endowment effect.

Framing Effects and Priming Effects

We have seen a number of ways in which economic models do not entirely account for the behavior of real people; one common mistake that people make involves the **framing effect**, which occurs when an answer is influenced by the way a question is asked or a decision is influenced by the way alternatives are presented.

Another decision-making pitfall, known as the **priming effect**, occurs when the order of the questions influences the answers,

Status Quo Bias

When people want to keep things the way they are, they may exhibit what is known as the **status quo bias**; this bias leads decision-makers to try to protect what they have, even when an objective evaluation of their circumstances suggests that a change would be beneficial.

The status quo bias causes people to behave conservatively; the cost of this behavior is missed opportunities that could potentially enhance welfare.

Status quo bias also explains why new products and ideas have trouble gaining traction: many potential customers prefer to leave things the way they are, even if something new might make more sense.

Intertemporal Decision-making

Intertemporal decisions occur across time; **intertemporal decision-making**—that is, planning to do something over a period of time—requires the ability to value the present and the future consistently. The ability to resist temptation is illustrated by a classic research experiment conducted at a preschool at Stanford University in 1972.

The Endowment Effect

What belongs to us, and what we give away, always seems very precious, because once we own something, we value it more. That's the **endowment effect**, and it explains why the loss of value associated with selling an item, or giving it up, is typically greater than the financial or emotional gain associated with obtaining the item in the first place. The endowment effect lies outside traditional economic theory, which assumes that humans consistently make rational decisions.

There's extensive evidence that physically handling an object can increase our perceived ownership of it.

Judgement about fairness

The pursuit of fairness is another common behavior that is important in economic decisions but that standard economic theory cannot explain.

While fairness is not normally modeled in economics, behavioral economists have developed experiments to determine the role of fairness in personal decisions. The **ultimatum game** is an economic experiment in which two

players decide how to divide a sum of money; the game shows how fairness enters into the rational decision-making process. In the game, player 1 is given a sum of money and is asked to propose a way of splitting it with player 2; Player 2 can either accept or reject the proposal, if player 2 accept, the sum is split according to the proposal. However, if player 2 rejects the proposal, neither player gets anything; the game is played only once, so the 1st player does not have to worry about reciprocity.

Each of the ideas that we have presented in this section, including misperceptions of probability, inconsistency in decision-making, and judgements about fairness, represent a departure from the traditional economic model of rational maximization. In the next section, we focus on risk-taking; as you will soon learn, not everyone evaluates risk in the same way...this fact has led economists to reconsider their models of human behavior.

What is the role of risk in decision-making?

In this section, we examine the role that risk plays in decision-making; the standard economic model of consumer choice assumes that people are consistent in their risk-taking preference. However, people's risk tolerances actually vary widely and are subject to change; thus, risk-taking behavior is not nearly as simple or predictable as economists once believed. We begin with a phenomenon known as **preference reversal**; we then consider how negative surprises can cause people to take more risk, which is explained by **prospect theory**.

Preference Reversals

As you know, trying to predict human behavior is not easy; Maurice Allais, the recipient of the 1988 Nobel Prize in Economics, noticed that people's tolerance for risk appeared to change in different situations. This observation did not agree with the standard economic model, which assumes that an individual's risk tolerance is constant and places the individual into one of the three distinct groups: **risk-averse people**, who prefer a sure thing over a gamble with a higher expected value; **risk-neutral people**, who choose the highest expected value regardless of the risk; and **risk-takers**, who prefer gambles with lower expected values, and potentially higher winnings, over a sure thing.

Allais developed a means of assessing risk behavior by presenting the set of choices (known as the Allais paradox) depicted in Table 17.1. Individuals were asked to choose their preferred options between gambles A and B and then again between gambles C and D.

TABLE 17.1	
The Allais Paradox	
Choose gamble A or B	
Gamble A	Gamble B book content
No gamble—receive \$1 million in cash 100% of the time	A lottery ticket that pays \$5 million 10% of the time, \$1 million 89% of the time, and nothing 1% of the time
Choose gamble C or D	
Gamble C	Gamble D
A lottery ticket that pays \$5 million 10% of the time and nothing 90% of the time	A lottery ticket that pays \$1 million 11% of the time and nothing 89% of the time

Economic science predicts that people will choose constantly according to their risk preference. As a result, economists understood that risk-averse individuals would choose the pair A and D; likewise, the pair B and C makes sense if the participants are risk-neutral and wish to maximize the expected value of the gambles.

A **preference reversal** occurs when risk tolerance is not consistent; Allais argued that a person's risk tolerance depends on his or her financial circumstances.

It turns out that preference reversals are more common than economists once believed; for example, almost 80% of all income tax filers expect to get a refund because they overpaid in the previous year. This behavior is odd, since there is an opportunity cost of waiting to get money back from the government when it didn't need to be paid in the first place. Employees could have asked their employers to withhold less and enjoyed their money sooner; individuals who choose to wait to receive their money later are said to have a difference that is weakly positive. In most circumstances, people have strongly positive time preferences: they prefer to have what they want sooner rather than later; so what do these taxpayers do when they learn the amount of their refund? In many cases, they pay their tax preparers an additional fee to have their refunds sent to their bank accounts electronically so they can receive the money sooner! Traditional economic analysis is unable to explain behavior; but armed with Allais' insights, we now see this behavior as a preference reversal.

Prospect Theory

Prospect theory, developed by Daniel Kahneman and Amos Tversky, suggests that people weigh decisions according to subjective utilities of gains and losses. The theory implies that people evaluate the risks that lead to gains separately from the risks that lead to losses; this concept is useful because it explains why some investors try to make up for losses by taking more chances rather than by maximizing the utility they receive from money under a rigid calculation of expected value.

Conclusion

Behavioral economics challenges the traditional economic model and invites a deeper understanding of human behavior. Armed with the insights from behavioral economics, we can answer questions that span a wider range of behaviors; we have seen behavioral economics at work in the examples in this chapter, which include the “opt-in” or “opt-out” debate, the economies of risk-taking, the effects of question design, the status quo bias. These ideas do not fit squarely into traditional economic analysis; you have learned enough at this point to question the assumptions we have made throughout the book. In next chapter, we apply all tools we have acquired to examine one of the most important sectors of the economy—health care and health insurance.