BUS 400 Business Economics

BUS 400 Business Economics

Adapted for Seneca College by Sandra Wellman

[AUTHOR REMOVED AT REQUEST OF ORIGINAL PUBLISHER]

SANDRA WELLMAN

UNIVERSITY OF MINNESOTA LIBRARIES PUBLISHING EDITION, 2016. THIS EDITION ADAPTED FROM A WORK ORIGINALLY PRODUCED IN 2012 BY A PUBLISHER WHO HAS REQUESTED THAT IT NOT RECEIVE ATTRIBUTION.

MINNEAPOLIS, MN

Contents

Publisher Information

This version has been adapted by Sandra Wellman for use at Seneca College. Content has been reorganized, revised, and some new content added.

The version this was based on *Principles of Economics* adapted from a work produced and distributed under a Creative Commons license (CC BY-NC-SA) in 2011 by a publisher who has requested that they and the original author not receive attribution. That edition was produced by the University of Minnesota Libraries Publishing through the eLearning Support Initiative and reformatted the original text, and replaced some images and figures to make the resulting whole more shareable. It was not significantly altered or updated from the original 2011 text.

CHAPTER 1: ECONOMICS: THE STUDY OF CHOICE

Start Up: Economics in the News

2008 seemed to be the year of economic news. From the worst financial crisis since the Great Depression to the possibility of a global recession, to gyrating gasoline and food prices, and to plunging housing prices, economic questions were the primary factors in the presidential campaign of 2008 and dominated the news generally.

What causes the prices of some good to rise while the prices of some other goods fall? Price determination is one of the things that we will study in this book. We will also consider factors that lead an economy to fall into a recession—and the attempts to limit it.

While the investigation of these problems surely falls within the province of economics, economics encompasses a far broader range of issues. Ultimately, economics is the study of choice. Because choices range over every imaginable aspect of human experience, so does economics. Economists have investigated the nature of family life, the arts, education, crime, sports, job creation—the list is virtually endless because so much of our lives involves making choices.

How do individuals make choices: Would you like better grades? More time to relax? More time watching movies? Getting better grades probably requires more time studying, and perhaps less relaxation and entertainment. Not only must we make choices as individuals, we must make choices as a society. Do we want a cleaner environment? Faster economic growth? Both may be desirable, but efforts to clean up the environment may conflict with faster economic growth. Society must make choices.

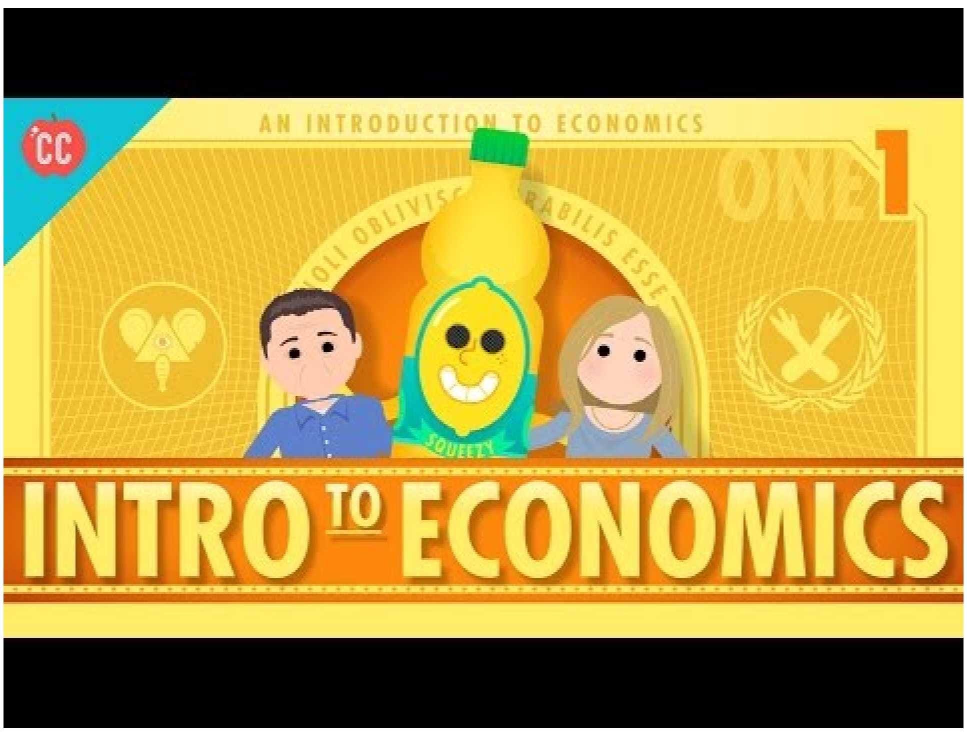
Economics is defined less by the subjects economists investigate than by the way in which economists investigate them. Economists have a way of looking at the world that differs from the way scholars in other disciplines look at the world. It is the economic way of thinking; this chapter introduces that way of thinking.

1.1 Defining Economics

Learning Objectives

1. Define economics.
2. Explain the concepts of scarcity and opportunity cost and how they relate to the definition of economics.
3. Understand the three fundamental economic questions: What should be produced? How should goods and services be produced? For whom should goods and services be produced?

Economics is a social science that examines how people choose among the alternatives available to them. It is social because it involves people and their behavior. It is a science because it uses, as much as possible, a scientific approach in its investigation of choices.



A YouTube element has been excluded from this version of the text. You can view it online here: https://pressbooks.senecacollege.ca/macroeconomics/?p=20

Scarcity, Choice, and Cost

All choices mean that one alternative is selected over another. Selecting among alternatives involves three ideas central to economics: scarcity, choice, and opportunity cost.

Scarcity

Our resources are limited. At any one time, we have only so much land, so many factories, so much oil, so many people. But our wants, our desires for the things that we can produce with those resources, are unlimited. We would always like more and better housing, more and better education—more and better of practically everything.

If our resources were also unlimited, we could say yes to each of our wants—and there would be no economics. Because our resources are limited, we cannot say yes to everything. To say yes to one thing requires that we say no to another. Whether we like it or not, we must make choices.

Our unlimited wants are continually colliding with the limits of our resources, forcing us to pick some activities and to reject others. Scarcity is the condition of having to choose among alternatives. A scarce good is one for which the choice of one alternative requires that another be given up.

Consider a parcel of land. The parcel presents us with several alternative uses. We could build a house on it. We could put a gas station on it. We could create a small park on it. We could leave the land undeveloped in order to be able to make a decision later as to how it should be used.

Suppose we have decided the land should be used for housing. Should it be a large and expensive house or several modest ones? Suppose it is to be a large and expensive house. Who should live in the house? If the Lees live in it, the Nguyens cannot. There are alternative uses of the land both in the sense of the type of use and also in the sense of who gets to use it. The fact that land is scarce means that society must make choices concerning its use.

Virtually everything is scarce. Consider the air we breathe, which is available in huge quantity at no charge to us. Could it possibly be scarce?

The test of whether air is scarce is whether it has alternative uses. What uses can we make of the air? We breathe it. We pollute it when we drive our cars, heat our houses, or operate our factories. In effect, one use of the air is as a garbage dump. We certainly need the air to breathe. But just as certainly, we choose to dump garbage in it. Those two uses are clearly alternatives to each other. The more garbage we dump in the air, the less desirable—and healthy—it will be to breathe. If we decide we want to breathe cleaner air, we must limit the activities that generate pollution. Air is a scarce good because it has alternative uses.

Not all goods, however, confront us with such choices. A free good is one for which the choice of one use does not require that we give up another. One example of a free good is gravity. The fact that gravity is holding you to the earth does not mean that your neighbor is forced to drift up into space! One person's use of gravity is not an alternative to another person's use.

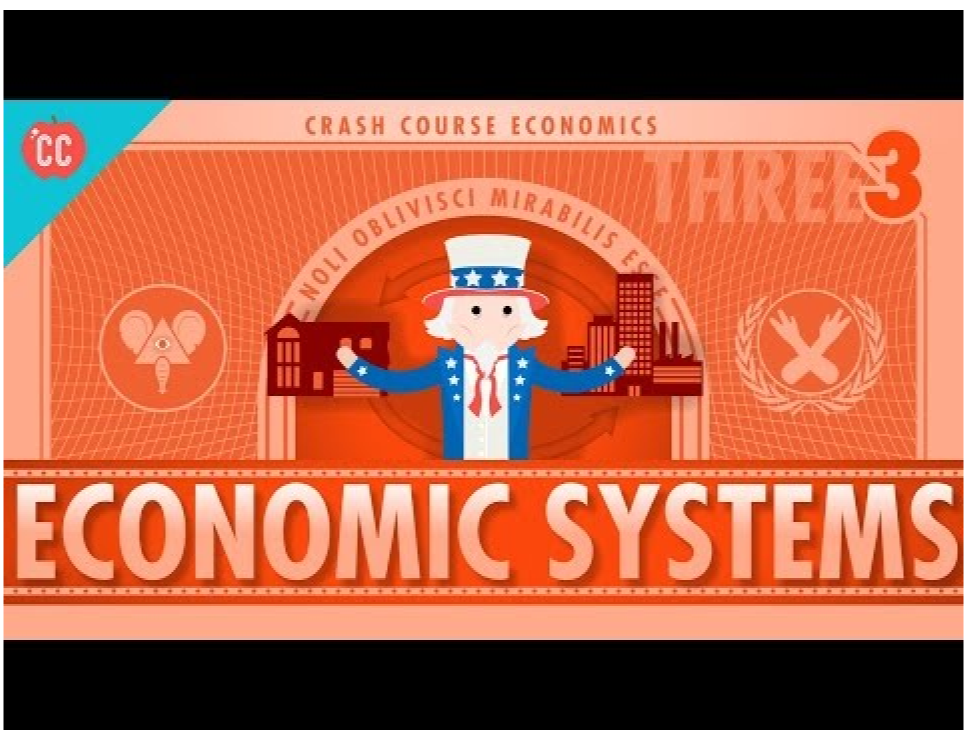
There are not many free goods. Outer space, for example, was a free good when the only use we made of it was to gaze at it. But now, our use of space has reached the point where one use can be an alternative to another. Conflicts have already arisen over the allocation of orbital slots for communications satellites. Thus, even parts of outer space are scarce. Space will surely become more scarce as we find new ways to use it. Scarcity characterizes virtually everything. Consequently, the scope of economics is wide indeed.

Scarcity and the Fundamental Economic Questions

The choices we confront as a result of scarcity raise three sets of issues. Every economy must answer the following questions:

1. **What should be produced?** Using the economy's scarce resources to produce one thing requires giving up another. Producing better education, for example, may require cutting back on other services, such as health care. A decision to preserve a wilderness area requires giving up other uses of the land. Every society must decide what it will produce with its scarce resources.
2. **How should goods and services be produced?** There are all sorts of choices to be made in determining how goods and services should be produced. Should a firm employ a few skilled or a lot of unskilled workers? Should it produce in its own country or should it use foreign plants? Should manufacturing firms use new or recycled raw materials to make their products?
3. **For whom should goods and services be produced?** If a good or service is produced, a decision must be made about who will get it. A decision to have one person or group receive a good or service usually means it will not be available to someone else. For example, representatives of the poorest nations on earth often complain that energy consumption per person in the United States is 17 *times* greater than energy consumption per person in the world's 62 poorest countries. Critics argue that the world's energy should be more evenly allocated. Should it? That is a "for whom" question.

Every economy must determine what should be produced, how it should be produced, and for whom it should be produced. We shall return to these questions again and again.



A YouTube element has been excluded from this version of the text. You can view it online here: https://pressbooks.senecacollege.ca/macroeconomics/?p=20

Opportunity Cost

It is within the context of scarcity that economists define what is perhaps the most important concept in all of economics, the concept of opportunity cost. Opportunity cost is the value of the best alternative forgone in making any choice.

The opportunity cost to you of reading the remainder of this chapter will be the value of the best other use to which you could have put your time. If you choose to spend $20 on a potted plant, you have simultaneously chosen to give up the benefits of spending the $20 on pizzas or a paperback book or a night at the movies. If the book is the most valuable of those alternatives, then the opportunity cost of the plant is the value of the enjoyment you otherwise expected to receive from the book.

The concept of opportunity cost must not be confused with the purchase price of an item. Consider the cost of a college or university education. That includes the value of the best alternative use of money spent for tuition, fees, and books. But the most important cost of a college education is the value of the forgone alternative uses of time spent studying and attending class instead of using the time in some other endeavor. Students sacrifice that time in hopes of even greater earnings in the future or because they place a value on the opportunity to learn. Or consider the cost of going to the doctor. Part of that cost is the value of the best alternative use of the money required to see the doctor. But, the cost also includes the value of the best alternative use of the time required to see the doctor. The essential thing to see in the concept of opportunity cost is found in the name of the concept. Opportunity cost is the value of the best opportunity forgone in a particular choice. It is not simply the amount spent on that choice.

The concepts of scarcity, choice, and opportunity cost are at the heart of economics. A good is scarce if the choice of one alternative requires that another be given up. The existence of alternative uses forces us to make choices. The opportunity cost of any choice is the value of the best alternative forgone in making it.

Key Takeaways

- Economics is a social science that examines how people choose among the alternatives available to them.
- Scarcity implies that we must give up one alternative in selecting another. A good that is not scarce is a free good.
- The three fundamental economic questions are: What should be produced? How should goods and services be produced? For whom should goods and services be produced?
- Every choice has an opportunity cost and opportunity costs affect the choices people make. The opportunity cost of any choice is the value of the best alternative that had to be forgone in making that choice.

Try It!

Identify the elements of scarcity, choice, and opportunity cost in each of the following:

1. The Environmental Protection Agency is considering an order that a 500-acre area on the outskirts of a large city be preserved in its natural state, because the area is home to a rodent that is considered an endangered species. Developers had planned to build a housing development on the land.
2. The manager of an automobile assembly plant is considering whether to produce cars or sport utility vehicles (SUVs) next month. Assume that the quantities of labour and other materials required would be the same for either type of production.
3. A young man who went to work as a nurses' aide after graduating from high school leaves his job to go to college, where he will obtain training as a registered nurse.

Case in Point: The Rising Cost of Energy

Figure 1.1

Oil is an exhaustible resource. The oil we burn today will not be available for use in the future. Part of the opportunity cost of our consumption of goods such as gasoline that are produced from oil includes the value people in the future might have placed on oil we use today.

It appears that the cost of our use of oil may be rising. We have been using "light crude," the oil found in the ground in deposits that can be readily tapped. As light crude becomes more scarce, the world may need to turn to so-called “heavy crude,” the crude oil that is found in the sandy soil of places such as Canada and Venezuela. That oil exists in such abundance that it propels Venezuela to the top of the world list of available oil. Saudi Arabia moves to the second position; Canada is third.

The difficulty with the oil mixed in the sand is that extracting it is far more costly than light crude, both in terms of the expenditures required and in terms of the environmental damage that mining it creates. Northern Alberta, in Canada, boasts a Florida-sized area whose sandy soils are rich in crude oil. Some of that oil is 1,200 feet underground. Extracting it requires pumping steam into the oily sand and then pumping up the resultant oily syrup. That syrup is then placed into huge, industrial-sized washing machines that separate crude oil. What is left over is toxic and will be placed in huge lakes that are being created by digging pits in the ground 200 feet deep. The oil produced from these sands has become important—Alberta is the largest foreign supplier of oil to the United States.

Sands that are closer to the surface are removed by bulldozers and giant cranes; the forest over it is cleared away. The oily sand is then hauled off in two-story dump trucks which, when filled, weigh more than a Boeing 747. Total SA, a French company, is leading the race to develop Canada’s oil. Jean Luc-Guiziou, the president of Total SA’s Canadian operations, says that the extraordinarily costly process of extracting heavy crude is something the world is going to have to get used to. “The light crude undiscovered today is getting scarcer and scarcer,” he told *The Wall Street Journal*. “We have to accept the reality of geoscience, which is that the next generation of oil resources will be heavier.”

Already, Total SA has clear-cut thousands of acres of forest land in order to gain access to the oily sand below. The process of extracting heavy crude oil costs the company $25 a barrel—compared to the $6 per barrel cost of extracting and refining light crude. Extracting heavy crude generates three times as much greenhouse gas per barrel as does light crude. By 2015, Fort McMurray, the small (population 61,000) town that has become the headquarters of Northern Alberta’s crude oil boom, will emit more greenhouse gas than the entire country of Denmark (population 5.4 million). Canada will exceed its greenhouse gas quota set by the Kyoto Accords—an international treaty aimed at limiting global warming—largely as a result of developing its heavy crude deposits.

No one even considered the extraction of heavy crude when light crude was cheap. In the late 1990s, oil cost just $12 per barrel, and deposits of heavy crude such as those in Canada attracted little attention. By mid-2006, oil sold for more than $70 per barrel, and Canada’s heavy crude was suddenly a hot commodity. “It moved from being just an interesting experiment in northern Canada to really this is the future source of oil supply,” Greg Stringham of the Canadian Association of Petroleum Producers told Al Jazeera.

Alberta’s energy minister, Greg Melchin, defends the province’s decision to proceed with the exploitation of its oily sand. “There is a cost to it, but the benefits are substantially greater,” he insists.

Not everyone agrees. George Poitras, a member of the Mikisew Cree tribe, lives downstream from the oil sands development. “You see a lot of the land dug up, a lot of the boreal forest struck down and it’s upsetting, it fills me with rage,” he says. Diana Gibson of the Parkland Institute, an environmental advocacy group, says that you can see the environmental damage generated by the extraction of oil sands around Fort McMurray from the moon. “What we are going to be having is destruction of very, very valuable ecosystems, and permanent pollution,” she says.

Sources: “Alberta’s Heavy Oil Burden,” Al Jazeera English, March 17, 2008 (see english.aljazeera.net); and Russell Gold, “As Prices Surge, Oil Giants Turn Sludge into Gold,” *The Wall Street Journal Online*, March 27, 2006, A1.

Answers to Try It! Problems

1. The 500-acre area is scarce because it has alternative uses: preservation in its natural state or a site for homes. A choice must be made between these uses. The opportunity cost of preserving the land in its natural state is the forgone value of the land as a housing development. The opportunity cost of using the land as a housing development is the forgone value of preserving the land.
2. The scarce resources are the plant and the labor at the plant. The manager must choose between producing cars and producing SUVs. The opportunity cost of producing cars is the profit that could be earned from producing SUVs; the opportunity cost of producing SUVs is the profit that could be earned from producing cars.
3. The man can devote his time to his current career or to an education; his time is a scarce resource. He must choose between these alternatives. The opportunity cost of continuing as a nurses' aide is the forgone benefit he expects from training as a registered nurse; the opportunity cost of going to college is the forgone income he could have earned working full-time as a nurses' aide.

1.2 The Field of Economics

Learning Objectives

1. Explain the distinguishing characteristics of the economic way of thinking.
2. Distinguish between microeconomics and macroeconomics.

We have examined the basic concepts of scarcity, choice, and opportunity cost in economics. In this section, we will look at economics as a field of study. We begin with the characteristics that distinguish economics from other social sciences.

The Economic Way of Thinking

Economists study choices that scarcity requires us to make. This fact is not what distinguishes economics from other social sciences; all social scientists are interested in choices. An anthropologist might study the choices of ancient peoples; a political scientist might study the choices of legislatures; a psychologist might study how people choose a mate; a sociologist might study the factors that have led to a rise in single-parent households. Economists study such questions as well. What is it about the study of choices by economists that makes economics different from these other social sciences?

Three features distinguish the economic approach to choice from the approaches taken in other social sciences:

1. Economists give special emphasis to the role of opportunity costs in their analysis of choices.
2. Economists assume that individuals make choices that seek to maximize the value of some objective, and that they define their objectives in terms of their own self-interest.
3. Individuals maximize by deciding whether to do a little more or a little less of something. Economists argue that individuals pay attention to the consequences of small changes in the levels of the activities they pursue.

The emphasis economists place on opportunity cost, the idea that people make choices that maximize the value of objectives that serve their self-interest, and a focus on the effects of small changes are ideas of great power. They constitute the core of economic thinking. The next three sections examine these ideas in greater detail.

Opportunity Costs Are Important

If doing one thing requires giving up another, then the expected benefits of the alternatives we face will affect the ones we choose. Economists argue that an understanding of opportunity cost is crucial to the examination of choices.

As the set of available alternatives changes, we expect that the choices individuals make will change. A rainy day could change the opportunity cost of reading a good book; we might expect more reading to get done in bad than in good weather. A high income can make it very costly to take a day off; we might expect highly paid individuals to work more hours than those who are not paid as well. If individuals are maximizing their level of satisfaction and firms are maximizing profits, then a change in the set of alternatives they face may affect their choices in a predictable way.

The emphasis on opportunity costs is an emphasis on the examination of alternatives. One benefit of the economic way of thinking is that it pushes us to think about the value of alternatives in each problem involving choice.

Individuals Maximize in Pursuing Self-Interest

What motivates people as they make choices? Perhaps more than anything else, it is the economist's answer to this question that distinguishes economics from other fields.

Economists assume that individuals make choices that they expect will create the maximum value of some objective, given the constraints they face. Furthermore, economists assume that people's objectives will be those that serve their own self-interest.

Economists assume, for example, that the owners of business firms seek to maximize profit. Given the assumed goal of profit maximization, economists can predict how firms in an industry will respond to changes in the markets in which they operate. As labor costs in the United States rise, for example, economists are not surprised to see firms moving some of their manufacturing operations overseas.

Similarly, economists assume that maximizing behavior is at work when they examine the behavior of consumers. In studying consumers, economists assume that individual consumers make choices aimed at maximizing their level of satisfaction. In the next chapter, we will look at the results of the shift from skiing to snowboarding; that is a shift that reflects the pursuit of self-interest by consumers and by manufacturers.

In assuming that people pursue their self-interest, economists are not assuming people are selfish. People clearly gain satisfaction by helping others, as suggested by the large charitable contributions people make. Pursuing one's own self-interest means pursuing the things that give one satisfaction. It need not imply greed or selfishness.

Choices Are Made at the Margin

Economists argue that most choices are made "at the margin." The margin is the current level of an activity. Think of it as the edge from which a choice is to be made. A choice at the margin is a decision to do a little more or a little less of something.

Assessing choices at the margin can lead to extremely useful insights. Consider, for example, the problem of curtailing water consumption when the amount of water available falls short of the amount people now use. Economists argue that one way to induce people to conserve water is to raise its price. A common response to this recommendation is that a higher price would have no effect on water consumption, because water is a necessity. Many people assert that prices do not affect water consumption because people "need" water.

But choices in water consumption, like virtually all choices, are made at the margin. Individuals do not make choices about whether they should or should not consume water. Rather, they decide whether to consume a little more or a little less water. Household water consumption in the United States totals about 105 gallons per person per day. Think of that starting point as the edge from which a choice at the margin in water consumption is made. Could a higher price cause you to use less water brushing your teeth, take shorter showers, or water your lawn less? Could a higher price cause people to reduce their use, say, to 104 gallons per person per day? To 103? When we examine the choice to consume water at the margin, the notion that a higher price would reduce consumption seems much more plausible. Prices affect our consumption of water because choices in water consumption, like other choices, are made at the margin.

The elements of opportunity cost, maximization, and choices at the margin can be found in each of two broad areas of economic analysis: microeconomics and macroeconomics. Your economics course, for example, may be designated as a “micro” or as a “macro” course. We will look at these two areas of economic thought in the next section.

Microeconomics and Macroeconomics

The field of economics is typically divided into two broad realms: microeconomics and macroeconomics. It is important to see the distinctions between these broad areas of study.

Microeconomics is the branch of economics that focuses on the choices made by individual decision-making units in the economy—typically consumers and firms—and the impacts those choices have on individual markets. Macroeconomics is the branch of economics that focuses on the impact of choices on the total, or aggregate, level of economic activity.

Why do tickets to the best concerts cost so much? How does the threat of global warming affect real estate prices in coastal areas? Why do women end up doing most of the housework? Why do senior citizens get discounts on public transit systems? These questions are generally regarded as microeconomic because they focus on individual units or markets in the economy.

Is the total level of economic activity rising or falling? Is the rate of inflation increasing or decreasing? What is happening to the unemployment rate? These are questions that deal with aggregates, or totals, in the economy; they are problems of macroeconomics. The question about the level of economic activity, for example, refers to the total value of all goods and services produced in the economy. Inflation is a measure of the rate of change in the average price level for the entire economy; it is a macroeconomic problem. The total levels of employment and unemployment in the economy represent the aggregate of all labor markets; unemployment is also a topic of macroeconomics.

Both microeconomics and macroeconomics give attention to individual markets. But in microeconomics that attention is an end in itself; in macroeconomics it is aimed at explaining the movement of major economic aggregates—the level of total output, the level of employment, and the price level.

We have now examined the characteristics that define the economic way of thinking and the two branches of this way of thinking: microeconomics and macroeconomics. In the next section, we will have a look at what one can do with training in economics.

Putting Economics to Work

Economics is one way of looking at the world. Because the economic way of thinking has proven quite useful, training in economics can be put to work in a wide range of fields. One, of course, is in work as an economist. Undergraduate work in economics can be applied to other careers as well.

Careers in Economics

Economists work in three types of organizations. About 58% of economists work for government agencies (Statistics Canada, Bank of Canada, Department of Finance are some examples). The remainder work for business firms or in colleges and universities.

Economists working for business firms and government agencies sometimes forecast economic activity to assist their employers in planning. They also apply economic analysis to the activities of the firms or agencies for which they work or consult. Economists employed at colleges and universities teach and conduct research.

Peruse the website of your college or university's economics department. Chances are the department will discuss the wide variety of occupations that their economics majors enter. Unlike engineering and accounting majors, economics and other social science majors tend to be distributed over a broad range of occupations.

Applying Economics to Other Fields

Suppose that you are considering something other than a career in economics. Would choosing to study economics help you?

The evidence suggests it may. Suppose, for example, that you are considering law school. The study of law requires keen analytical skills; studying economics sharpens such skills. Economists have traditionally argued that undergraduate work in economics serves as excellent preparation for law school. Economist Michael Nieswiadomy of the University of North Texas collected data on Law School Admittance Test (LSAT) scores for undergraduate majors listed by 2,200 or more students taking the test in 2003. Table 1.1 "LSAT Scores and Undergraduate Majors" gives the scores, as well as the ranking for each of these majors, in 2003 and in two previous years in which the rankings were compiled. In rankings for all three years, economics majors recorded the highest scores.

Table 1.1 LSAT Scores and Undergraduate Majors

Major field	LSAT average 2003–2004	2003–2004 Rank	1994–1995 Rank	1991–1992 Rank
Economics	156.6	1	1	1
Engineering	155.4	2	4	2
History	155.0	3	2	3
English	154.3	4	3	4
Finance	152.6	5	6	5
Political science	152.1	6	9	9
Psychology	152.1	7	7	8
Accounting	151.1	8	8	6
Communications	150.5	9	10	10
Sociology	150.2	10	12	13
Bus. Administration	149.6	11	13	12
Criminal Justice	144.7	12	14	14

Here are the average LSAT scores and rankings for the 12 undergraduate majors with more than 2200 students taking the test to enter law school in the 2003–2004 academic year.

Source: Michael Nieswiadomy, “LSAT Scores of Economics Majors: 2003–2004 Class Update,” *Journal of Economic Education*, 37(2) (Spring 2006): 244–247 and Michael Nieswiadomy, “LSAT Scores of Economics Majors” *Journal of Economic Education*, 29(4) (Fall 1998): 377–379.

Did the strong performance by economics, engineering, and history majors mean that training in those fields sharpens analytical skills tested in the LSAT, or that students with good analytical skills are more likely to major in them? Both factors were probably at work. Economics clearly attracts students with good analytical skills—and studying economics helps develop those skills.

Economics majors shine in other areas as well. According to the Bureau of Labor Statistics *Occupational Outlook Handbook*, a strong background in economic theory, mathematics, and statistics provides the basis for competing for the best job opportunities, particularly research assistant positions, in a broad range of fields. Many graduates with bachelor’s degrees will find good jobs in industry and business as management or sales trainees or as administrative assistants. Because economists are concerned with understanding and interpreting financial matters, among other subjects, they will also be attracted to and qualified for jobs as financial managers, financial analysts, underwriters, actuaries, securities and financial services sales workers, credit analysts, loan and budget officers, and urban and regional planners.

Table 1.2 “Average Yearly Salary Offers, May 2006 and Occupational Outlook 2004–2014, Selected Majors/Occupations” shows average yearly salary offers for bachelor degree candidates for May 2006 and the outlook for related occupations to 2014.

Table 1.2 Average Yearly Salary Offers, May 2006 and Occupational Outlook 2004–2014, Selected Majors/Occupations

Undergraduate major	Average $ Offer May, 2006	Projected % Change in Total Employment in Occupation 2004–2014
Computer Engineering	$54,200	10.1
Electrical/Electronic Engineering	54,053	11.8
Computer Science	50,892	25.6
Accounting	46,188	22.4
Economics and Finance	45,058	12.4
Management Information Systems	44,755	25.9
Logistics and Materials Management	43,426	13.2
Business Administration	40,976	17.0
Environmental Sciences (including forestry and conservation science)	39,750	6.3
Other Business Majors (e.g., Marketing)	37,446	20.8
Human Resources (incl. Labor Relations)	36,256	15.9
Geology and Geological Sciences	35,034	8.3
Sociology	33,752	4.7
Political Science/Government	33,151	7.3
Liberal Arts & Sciences (general studies)	32,627	na
Public Relations	32,623	21.7
Special Education	31,817	23.3
Elementary Education	31,778	18.2
Foreign Languages	31,364	na
Letters (incl. English)	31,204	20.4
Other Social Sciences (Including Criminal Justice and History)	30,788	12.3
Psychology	30,308	9.9
Pre-elementary Education	27,550	22.4
Social Work	25,865	19.6
Visual and Performing Arts	21,726	15.2

Sources: National Association of Colleges and Employers, *Salary Survey*, Spring 2006 http://naceweb.org; Bureau of Labor Statistics, 2006–2007 edition of the *Occupational Outlook Handbook*; *Occupational Employment, Training, and Earnings: Educational Level Report* (May, 2006) URL: http://data.bls.gov/oep/noeted/empoptd.jsp (note: na = not reported; that is, no specific occupation was reported in BLS report; Other business majors, Other social sciences, Social work (including Sociology), and Environmental Sciences are weighted averages of various disciplines, calculated by authors.)

One's choice of a major, or minor, is not likely to be based solely on considerations of potential earnings or the prospect of landing a spot in law school. You will also consider your interests and abilities in making a decision about whether to pursue further study in economics. And, of course, you will consider the expected benefits of alternative courses of study. What is *your* opportunity cost of pursuing study of economics? Does studying more economics serve your interests and will doing so maximize your satisfaction level? These considerations may be on your mind as you begin to study economics at the college level and obviously students will make many different choices. But, should you decide to pursue a major or minor in economics, you should know that a background in this field is likely to serve you well in a wide range of careers.

Key Takeaways

- Economists focus on the opportunity costs of choices, they assume that individuals make choices in a way that maximizes the value of an objective defined in terms of their own self-interest, and they assume that individuals make those choices at the margin.
- Economics is divided into two broad areas: microeconomics and macroeconomics.
- A wide range of career opportunities is open to economics majors. Empirical evidence suggests that students who enter the job market with a major in economics tend to earn more than do students in most other majors. Further, economics majors do particularly well on the LSAT.

Try It!

The Department of Agriculture estimated that the expenditures a middle-income, husband–wife family of three would incur to raise one additional child from birth in 2005 to age 17 would be $250,530. In what way does this estimate illustrate the economic way of thinking? Would the Department's estimate be an example of microeconomic or of macroeconomic analysis? Why?

Case in Point: The Financial Payoff to Studying Economics

Figure 1.2

Jeremy Wilburn – [Students in Classrooms at UIS](#) – CC BY-NC-ND 2.0.

College economics professors have long argued that studying economics is good preparation for a variety of careers. A recent study suggests they are right and that studying economics is even likely to make students more prosperous. Students who major in economics but did not pursue graduate work are likely to earn more than students in virtually every other college major. Students who major in economics and then go on to law school or an MBA program are likely to earn more than students who approach those areas of study having majored in most other areas.

Economists Dan A. Black, Seth Sanders, and Lowell Taylor used the 1993 National Survey of College Graduates, which included more than 86,000 college-educated workers between the ages of 25 and 55 that asked what field they had majored in. They then controlled for variables such as gender, race, and ethnicity. They found that students who had not done graduate work and had majored in economics earned more than students in any other major except engineering. Specifically, economics majors earned about 13% more than other social sciences majors, 11% more than business administration majors, and about the same as natural science and accounting majors. The economics majors in their survey, like those who majored in other social sciences and business administration and unlike those who majored in engineering or accounting, were spread out over a wide range of occupations but with many in management positions.

Based on the survey they used, over 40% of economics majors went on to earn graduate degrees, many in law and business. Economics majors ranked first in terms of wages, as compared to other law school graduates with the 12 most common pre-law majors (including such majors as business administration, finance, English, history, psychology, and political science). MBA graduates who had majored in economics earned more than those who had majored in any other field except chemical engineering. Specifically, undergraduate economics majors with MBAs earned about 15% more than those who had majored in other disciplines represented in the survey, including business-related majors.

It is remarkable that all of the business-related majors generated salaries much lower than those earned by economics majors with an MBA. One could argue that this reflects self-selection; that students who major in economics are simply brighter. But, students who major in physics have high SAT scores, yet they, too, earned wages that were about 20% lower than MBA students who had majored in economics. This finding lends some credence to the notion that the marketplace rewards training in the economic way of thinking.

Source: Dan A. Black, Seth Sanders, and Lowell Taylor, "The Economic Reward for Studying Economics," *Economic Inquiry*, 41(3), July 2003, 365–377.

Answer to Try It! Problem

The information given suggests one element of the economic way of thinking: assessing the choice at the margin. The estimate reflects the cost of one more child for a family that already has one. It is not clear from the information given how close the estimate of cost comes to the economic concept of opportunity cost. The Department of Agriculture's estimate included such costs as housing, food, transportation, clothing, health care, child care, and education. An economist would add the value of the best alternative use of the additional time that will be required for the child. If the couple is looking far ahead, it may want to consider the opportunity cost of sending a child to college. And, if it is looking *very* far ahead, it may want to consider the fact that nearly half of all parents over the age of 50 support at least one child over the age of 21. This is a problem in microeconomic analysis, because it focuses on the choices of individual households.

References

Bureau of Labor Statistics *Occupational Outlook* at http://www.bls.gov/oco/.

1.3 The Economists' Tool Kit

Learning Objectives

1. Explain how economists test hypotheses, develop economic theories, and use models in their analyses.
2. Explain how the all-other-things unchanged (ceteris paribus) problem and the fallacy of false cause affect the testing of economic hypotheses and how economists try to overcome these problems.
3. Distinguish between normative and positive statements.

Economics differs from other social sciences because of its emphasis on opportunity cost, the assumption of maximization in terms of one's own self-interest, and the analysis of choices at the margin. But certainly much of the basic methodology of economics and many of its difficulties are common to every social science—indeed, to every science. This section explores the application of the scientific method to economics.

Researchers often examine relationships between variables. A variable is something whose value can change. By contrast, a constant is something whose value does not change. The speed at which a car is traveling is an example of a variable. The number of minutes in an hour is an example of a constant.

Research is generally conducted within a framework called the scientific method, a systematic set of procedures through which knowledge is created. In the scientific method, hypotheses are suggested and then tested. A hypothesis is an assertion of a relationship between two or more variables that could be proven to be false. A statement is not a hypothesis if no conceivable test could show it to be false. The statement "Plants like sunshine" is not a hypothesis; there is no way to test whether plants like sunshine or not, so it is impossible to prove the statement false. The statement "Increased solar radiation increases the rate of plant growth" is a hypothesis; experiments could be done to show the relationship between solar radiation and plant growth. If solar radiation were shown to be unrelated to plant growth or to retard plant growth, then the hypothesis would be demonstrated to be false.

If a test reveals that a particular hypothesis is false, then the hypothesis is rejected or modified. In the case of the hypothesis about solar radiation and plant growth, we would probably find that more sunlight increases plant growth over some range but that too much can actually retard plant growth. Such results would lead us to modify our hypothesis about the relationship between solar radiation and plant growth.

If the tests of a hypothesis yield results consistent with it, then further tests are conducted. A hypothesis that has not been rejected after widespread testing and that wins general acceptance is commonly called a theory. A theory that has been subjected to even more testing and that has won virtually universal acceptance becomes a law. We will examine two economic laws in the next two chapters.

Even a hypothesis that has achieved the status of a law cannot be proven true. There is always a possibility that someone may find a case that invalidates the hypothesis. That possibility means that nothing in economics, or in any other social science, or in any science, can ever be *proven* true. We can have great confidence in a particular proposition, but it is always a mistake to assert that it is "proven."

Models in Economics

All scientific thought involves simplifications of reality. The real world is far too complex for the human mind—or the most powerful computer—to consider. Scientists use models instead. A model is a set of simplifying assumptions about some aspect of the real world. Models are always based on assumed conditions that are simpler than those of the real world, assumptions that are necessarily false. A model of the real world cannot *be* the real world.

We will encounter our first economic model in Chapter 35 "Appendix A: Graphs in Economics". For that model, we will assume that an economy can produce only two goods. Then we will explore the model of demand and supply. One of the assumptions we will make there is that all the goods produced by firms in a particular market are identical. Of course, real economies and real markets are not that simple. Reality is never as simple as a model; one point of a model is to simplify the world to improve our understanding of it.

Economists often use graphs to represent economic models. The appendix to this chapter provides a quick, refresher course, if you think you need one, on understanding, building, and using graphs.

Models in economics also help us to generate hypotheses about the real world. In the next section, we will examine some of the problems we encounter in testing those hypotheses.

Testing Hypotheses in Economics

Here is a hypothesis suggested by the model of demand and supply: an increase in the price of gasoline will reduce the quantity of gasoline consumers demand. How might we test such a hypothesis?

Economists try to test hypotheses such as this one by observing actual behavior and using empirical (that is, real-world) data. The average retail price of gasoline in the United States rose from an average of $2.12 per gallon on May 22, 2005 to $2.88 per gallon on May 22, 2006. The number of gallons of gasoline consumed by U.S. motorists rose 0.3% during that period.

The small increase in the quantity of gasoline consumed by motorists as its price rose is inconsistent with the hypothesis that an increased price will lead to an reduction in the quantity demanded. Does that mean that we should dismiss the original hypothesis? On the contrary, we must be cautious in assessing this evidence. Several problems exist in interpreting any set of economic data. One problem is that several things may be changing at once; another is that the initial event may be unrelated to the event that follows. The next two sections examine these problems in detail.

The All-Other-Things-Unchanged Problem

The hypothesis that an increase in the price of gasoline produces a reduction in the quantity demanded by consumers carries with it the assumption that there are no other changes that might also affect consumer demand. A better statement of the hypothesis would be: An increase in the price of gasoline will reduce the quantity consumers demand, ceteris paribus. Ceteris paribus is a Latin phrase that means "all other things unchanged."

But things changed between May 2005 and May 2006. Economic activity and incomes rose both in the United States and in many other countries, particularly China, and people with higher incomes are likely to buy more gasoline. Employment rose as well, and people with jobs use more gasoline as they drive to work. Population in the United States grew during the period. In short, many things happened during the period, all of which tended to increase the quantity of gasoline people purchased.

Our observation of the gasoline market between May 2005 and May 2006 did not offer a conclusive test of the hypothesis that an increase in the price of gasoline would lead to a reduction in the quantity demanded by consumers. Other things changed and affected gasoline consumption. Such problems are likely to affect any analysis of economic events. We cannot ask the world to stand still while we conduct experiments in economic phenomena. Economists employ a variety of statistical methods to allow them to isolate the impact of single events such as price changes, but they can never be certain that they have accurately isolated the impact of a single event in a world in which virtually everything is changing all the time.

In laboratory sciences such as chemistry and biology, it is relatively easy to conduct experiments in which only selected things change and all other factors are held constant. The economists' laboratory is the real world; thus, economists do not generally have the luxury of conducting controlled experiments.

The Fallacy of False Cause

Hypotheses in economics typically specify a relationship in which a change in one variable causes another to change. We call the variable that responds to the change the dependent variable; the variable that induces a change is called the independent variable. Sometimes the fact that two variables move together can suggest the false conclusion that one of the variables has acted as an independent variable that has caused the change we observe in the dependent variable.

Consider the following hypothesis: People wearing shorts cause warm weather. Certainly, we observe that more people wear shorts when the weather is warm. Presumably, though, it is the warm weather that causes people to wear shorts rather than the wearing of shorts that causes warm weather; it would be incorrect to infer from this that people cause warm weather by wearing shorts.

Reaching the incorrect conclusion that one event causes another because the two events tend to occur together is called the fallacy of false cause. The accompanying essay on baldness and heart disease suggests an example of this fallacy.

Because of the danger of the fallacy of false cause, economists use special statistical tests that are designed to determine whether changes in one thing actually do cause changes observed in another. Given the inability to perform controlled experiments, however, these tests do not always offer convincing evidence that persuades all economists that one thing does, in fact, cause changes in another.

In the case of gasoline prices and consumption between May 2005 and May 2006, there is good theoretical reason to believe the price increase should lead to a reduction in the quantity consumers demand. And economists have tested the hypothesis about price and the quantity demanded quite extensively. They have developed elaborate statistical tests aimed at ruling out problems of the fallacy of false cause. While we cannot prove that an increase in price will, ceteris paribus, lead to a reduction in the quantity consumers demand, we can have considerable confidence in the proposition.

Normative and Positive Statements

Two kinds of assertions in economics can be subjected to testing. We have already examined one, the hypothesis. Another testable assertion is a statement of fact, such as "It is raining outside" or "Microsoft is the largest producer of operating systems for personal computers in the world." Like hypotheses, such assertions can be demonstrated to be false. Unlike hypotheses, they can also be shown to be correct. A statement of fact or a hypothesis is a positive statement.

Although people often disagree about positive statements, such disagreements can ultimately be resolved through investigation. There is another category of assertions, however, for which investigation can never resolve differences. A normative statement is one that makes a value judgment. Such a judgment is the opinion of the speaker; no one can "prove" that the statement is or is not correct. Here are some examples of normative statements in economics: "We ought to do more to help the poor." "People in the United States should save more." "Corporate profits are too high." The statements are based on the values of the person who makes them. They cannot be proven false.

Because people have different values, normative statements often provoke disagreement. An economist whose values lead him or her to conclude that we should provide more help for the poor will disagree with one whose values lead to a conclusion that we should not. Because no test exists for these values, these two economists will continue to disagree, unless one persuades the other to adopt a different set of values. Many of the disagreements among economists are based on such differences in values and therefore are unlikely to be resolved.

Key Takeaways

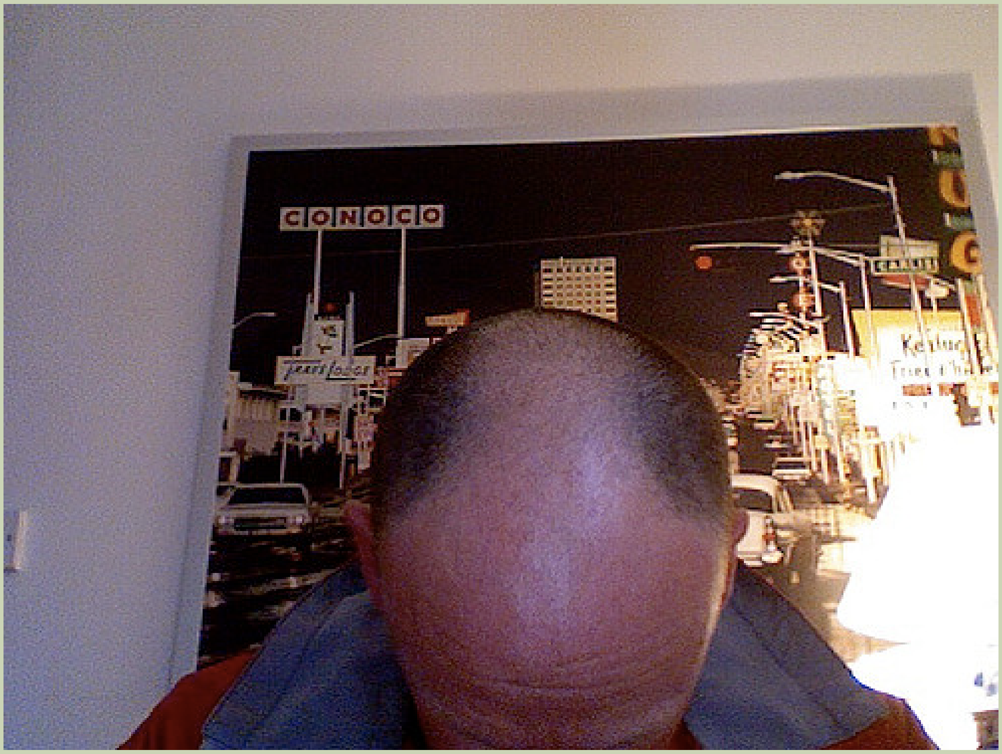
- Economists try to employ the scientific method in their research.
- Scientists cannot prove a hypothesis to be true; they can only fail to prove it false.
- Economists, like other social scientists and scientists, use models to assist them in their analyses.
- Two problems inherent in tests of hypotheses in economics are the all-other-things-unchanged problem and the fallacy of false cause.
- Positive statements are factual and can be tested. Normative statements are value judgments that cannot be tested. Many of the disagreements among economists stem from differences in values.

Try It!

Look again at the data in Table 1.1 "LSAT Scores and Undergraduate Majors". Now consider the hypothesis: "Majoring in economics will result in a higher LSAT score." Are the data given consistent with this hypothesis? Do the data prove that this hypothesis is correct? What fallacy might be involved in accepting the hypothesis?

Case in Point: Does Baldness Cause Heart Disease?

Figure 1.3



Mark Hunter – [bald](#) – CC BY-NC-ND 2.0.

A website called [embarrassingproblems.com](#) received the following email:

> "Dear Dr. Margaret,
> "I seem to be going bald. According to your website, this means I'm more likely to have a heart attack. If I take a drug to prevent hair loss, will it reduce my risk of a heart attack?"
> What did Dr. Margaret answer? Most importantly, she did not recommend that the questioner take drugs to treat his baldness, because doctors do not think that the baldness causes the heart disease. A more likely explanation for the association between baldness and heart disease is that both conditions are affected by an underlying factor. While noting that more research needs to be done, one hypothesis that Dr. Margaret offers is that higher testosterone levels might be triggering both the hair loss and the heart disease. The good news for people with early balding (which is really where the association with increased risk of heart disease has been observed) is that they have a signal that might lead them to be checked early on for heart disease.
>
> Source: http://www.embarrassingproblems.com/problems/problempage230701.htm.

Answer to Try It! Problem

The data are consistent with the hypothesis, but it is never possible to prove that a hypothesis is correct.

Accepting the hypothesis could involve the fallacy of false cause; students who major in economics may already have the analytical skills needed to do well on the exam.

CHAPTER 2: CONFRONTING SCARCITY: CHOICES IN PRODUCTION

Start Up: Tightening Security at the World's Airports

Do you want safer air travel or not? While that question is seldom asked so bluntly, any person who travels by air can tell you that our collective answer has been "yes," and it has been accompanied by increases in security and its associated costs at airports all over the world. Why? In short, "9/11." Terrorists hijacked four U.S. commercial airliners on September 11, 2001, and the tragic results that followed led to a sharp tightening in airport security.

In an effort to prevent similar disasters, airport security officials scrutinize luggage and passengers more carefully than ever before. In the months following 9/11, delays of as much as three hours were common as agents tried to assure that no weapons or bombs could be smuggled onto another plane.

"What to produce?" is a fundamental economic question. Every economy must answer this question. Should it produce more education, better health care, improved transportation, a cleaner environment? There are limits to what a nation can produce; deciding to produce more of one thing inevitably means producing less of something else. Individuals in much of the world, after the tragedy of 9/11, clearly were willing to give up time, and a fair amount of individual privacy, in an effort to obtain greater security. Nations and individual cities also devoted additional resources to police and other forms of protection in an effort to prevent tragedies such as 9/11. People all over the world chose to produce less of other goods in order to devote more resources to the production of greater security. And, as of early 2009, the choice to devote more resources to security had paid off; there had been no similar hijackings in the United States.

In this chapter we use our first model, the production possibilities model, to examine the nature of choices to produce more of some goods and less of others. As its name suggests, the **production possibilities model** shows the goods and services that an economy is capable of producing—its possibilities—given the factors of production and the technology it has available. The model specifies what it means to use resources fully and efficiently and suggests some important implications for international trade. We can also use the model to illustrate economic growth, a process that expands the set of production possibilities available to an economy.

We then turn to an examination of the type of economic system in which choices are made. An **economic system** is the set of rules that define how an economy's resources are to be owned and how decisions about their use are to be made. We will see that economic systems differ in terms of how they answer the fundamental economic questions. Many of the world's economic systems, including the systems that prevail in North America, Europe, and much of Asia and Central and South America, rely on individuals operating in a market economy to make those choices. Other economic systems, including those of Cuba and North Korea today and historically those of the former Soviet Union, Soviet bloc countries, and China, rely—or relied—on government to make these choices. Different economic systems result in different sets of choices and thus different outcomes; the fact that market economies generally outperform the others when it comes to providing more of the things that people want helps to explain the dramatic shift from government-dominated toward market-dominated economic systems that has occurred throughout the world in the past 25 years. The chapter concludes with an examination of the role of government in an economy that relies chiefly on markets to allocate goods and services.

2.1 Factors of Production

Learning Objectives

1. Define the three factors of production—labour, capital, and natural resources.
2. Explain the role of technology and entrepreneurs in the utilization of the economy's factors of production.

Choices concerning what goods and services to produce are choices about an economy's use of its factors of production, the resources available to it for the production of goods and services. The value, or satisfaction, that people derive from the goods and services they consume and the activities they pursue is called utility. Ultimately, then, an economy's factors of production create utility; they serve the interests of people.

The factors of production in an economy are its labor, capital, and natural resources. Labour is the human effort that can be applied to the production of goods and services. People who are employed or would like to be are considered part of the labour available to the economy. Capital is a factor of production that has been produced for use in the production of other goods and services. Office buildings, machinery, and tools are examples of capital. Natural resources are the resources of nature that can be used for the production of goods and services.

In the next three sections, we will take a closer look at the factors of production we use to produce the goods and services we consume. The three basic building blocks of labour, capital, and natural resources may be used in different ways to produce different goods and services, but they still lie at the core of production. We will then look at the roles played by technology and entrepreneurs in putting these factors of production to work. As economists began to grapple with the problems of scarcity, choice, and opportunity cost two centuries ago, they focused on these concepts, just as they are likely to do two centuries hence.

Labour

Laboor is human effort that can be applied to production. People who work to repair tires, pilot airplanes, teach children, or enforce laws are all part of the economy's labour. People who would like to work but have not found employment—who are unemployed—are also considered part of the labour available to the economy.

In some contexts, it is useful to distinguish two forms of labour. The first is the human equivalent of a natural resource. It is the natural ability an untrained, uneducated person brings to a particular production process. But most workers bring far more. The skills a worker has as a result of education, training, or experience that can be used in production are called human capital. Students who are attending a college or university are acquiring human capital. Workers who are gaining skills through experience or through training are acquiring human capital. Children who are learning to read are acquiring human capital.

The amount of labour available to an economy can be increased in two ways. One is to increase the total quantity of labor, either by increasing the number of people available to work or by increasing the average number of hours of work per week. The other is to increase the amount of human capital possessed by workers.

Capital

Long ago, when the first human beings walked the earth, they produced food by picking leaves or fruit off a plant or by catching an animal and eating it. We know that very early on, however, they began shaping stones into tools, apparently for use in butchering animals. Those tools were the first capital because they were produced for use in producing other goods—food and clothing.

Modern versions of the first stone tools include saws, meat cleavers, hooks, and grinders; all are used in butchering animals. Tools such as hammers, screwdrivers, and wrenches are also capital. Transportation equipment, such as cars and trucks, is capital. Facilities such as roads, bridges, ports, and airports are capital. Buildings, too, are capital; they help us to produce goods and services.

Capital does not consist solely of physical objects. The score for a new symphony is capital because it will be used to produce concerts. Computer software used by business firms or government agencies to produce goods and services is capital. Capital may thus include physical goods and intellectual discoveries. Any resource is capital if it satisfies two criteria:

1. The resource must have been produced.
2. The resource can be used to produce other goods and services.

One thing that is not considered capital is money. A firm cannot use money directly to produce other goods, so money does not satisfy the second criterion for capital. Firms can, however, use money to acquire capital. Money is a form of financial capital. Financial capital includes money and other "paper" assets (such as stocks and bonds) that represent claims on future payments. These financial assets are not capital, but they can be used directly or indirectly to purchase factors of production or goods and services.

Natural Resources

There are two essential characteristics of natural resources. The first is that they are found in nature—that no human effort has been used to make or alter them. The second is that they can be used for the production of goods and services. That requires knowledge; we must know how to use the things we find in nature before they become resources.

Consider oil. Oil in the ground is a natural resource because it is found (not manufactured) and can be used to produce goods and services. However, 250 years ago oil was a nuisance, not a natural resource. Pennsylvania farmers in the eighteenth century who found oil oozing up through their soil were dismayed, not delighted. No one knew what could be done with the oil. It was not until the mid-nineteenth century that a method was found for refining oil into kerosene that could be used to generate energy, transforming oil into a natural resource. Oil is now used to make all sorts of things, including clothing, drugs, gasoline, and plastic. It became a natural resource because people discovered and implemented a way to use it.

Defining something as a natural resource only if it can be used to produce goods and services does not mean that a tree has value only for its wood or that a mountain has value only for its minerals. If people gain utility from the existence of a beautiful wilderness area, then that wilderness provides a service. The wilderness is thus a natural resource.

The natural resources available to us can be expanded in three ways. One is the discovery of new natural resources, such as the discovery of a deposit of ore containing titanium. The second is the discovery of new uses for resources, as happened when new techniques allowed oil to be put to productive use or sand to be used in manufacturing computer chips. The third is the discovery of new ways to extract natural resources in order to use them. New methods of discovering and mapping oil deposits have increased the world's supply of this important natural resource.

Technology and the Entrepreneur

Goods and services are produced using the factors of production available to the economy. Two things play a crucial role in putting these factors of production to work. The first is technology, the knowledge that can be applied to the production of goods and services. The second is an individual who plays a key role in a market economy: the entrepreneur. An entrepreneur is a person who, operating within the context of a market economy, seeks to earn profits by finding new ways to organize factors of production. In non-market economies the role of the entrepreneur is played by bureaucrats and other decision makers who respond to incentives other than profit to guide their choices about resource allocation decisions.

The interplay of entrepreneurs and technology affects all our lives. Entrepreneurs put new technologies to work every day, changing the way factors of production are used. Farmers and factory workers, engineers and electricians, technicians and teachers all work differently than they did just a few years ago, using new technologies introduced by entrepreneurs. The music you enjoy, the books you read, the athletic equipment with which you play are produced differently than they were five years ago. The book you are reading was written and manufactured using technologies that did not exist ten years ago. We can dispute whether all the changes have made our lives better. What we cannot dispute is that they have made our lives different.

Key Takeaways

- Factors of production are the resources the economy has available to produce goods and services.
- Labour is the human effort that can be applied to the production of goods and services. Labour's contribution to an economy's output of goods and services can be increased either by increasing the quantity of labor or by increasing human capital.
- Capital is a factor of production that has been produced for use in the production of other goods and services.
- Natural resources are those things found in nature that can be used for the production of goods and services.
- Two keys to the utilization of an economy's factors of production are technology and, in the case of a market economic system, the efforts of entrepreneurs.

Try It!

Explain whether each of the following is labour, capital, or a natural resource.

1. An unemployed factory worker

2. A college professor
3. The library building on your campus
4. Yellowstone National Park
5. An untapped deposit of natural gas
6. The White House
7. The local power plant

Case in Point: Technology Cuts Costs, Boosts Productivity and Profits

Figure 2.1

Selbe Lynn – Oil Platform – CC BY-NC-ND 2.0.

Technology can seem an abstract force in the economy—important, but invisible.

It is not invisible to the 130 people who work on a Shell Oil Company oil rig called Mars, located in the deep waters of the Gulf of Mexico, about 160 miles southwest of Pensacola, Florida. The name Mars reflects its otherworld appearance—it extends 300 feet above the water's surface and has steel tendons that reach 3,000 feet to the floor of the gulf. This facility would not exist if it were not for the development of better oil discovery methods that include three-dimensional seismic mapping techniques, satellites that locate oil from space, and drills that can make turns as drilling foremen steer them by monitoring them on computer screens from the comfort of Mars. "We don't hit as many dry holes," commented Shell manager Miles Barrett. As a result of these new technologies, over the past two decades, the cost of discovering a barrel of oil dropped from $20 to under $5. And the technologies continue to improve. Three-dimensional surveys are being replaced with four-dimensional ones that allow geologists to see how the oil fields change over time.

The Mars project was destroyed by Hurricane Katrina in 2005. Royal Dutch Shell completed repairs in 2006—at a cost of $200 million. But, the facility is again pumping 130,000 barrels of oil per day and 150 million cubic feet of natural gas—the energy equivalent of an additional 26,000 barrels of oil.

Technology is doing more than helping energy companies track oil deposits. It is changing the way soft drinks and other grocery items are delivered to retail stores. For example, when a PepsiCo delivery driver arrives at a 7-Eleven, the driver keys into a handheld computer the inventory of soft drinks, chips, and other PepsiCo products. The information is transmitted to a main computer at the warehouse that begins processing the next order for that store. The result is that the driver can visit more stores in a day and PepsiCo can cover a given territory with fewer drivers and trucks.

New technology is even helping to produce more milk from cows. Ed Larsen, who owns a 1,200-cow dairy farm in Wisconsin, never gets up before dawn to milk the cows, the way he did as a boy. Rather, the cows are hooked up to electronic milkers. Computers measure each cow's output, and cows producing little milk are sent to a "hospital wing" for treatment. With the help of such technology, as well as better feed, today's dairy cows produce 50% more milk than did cows 20 years ago. Even though the number of dairy cows in the United States in the last 20 years has fallen 17%, milk output has increased 25%.

Who benefits from technological progress? Consumers gain from lower prices and better service. Workers gain: Their greater ability to produce goods and services translates into higher wages. And firms gain: Lower production costs mean higher profits. Of course, some people lose as technology advances. Some jobs are eliminated, and some firms find their services are no longer needed. One can argue about whether particular technological changes have improved our lives, but they have clearly made—and will continue to make—them far different.

Sources: David Ballingrud, "Drilling in the Gulf: Life on Mars," *St. Petersburg Times* (Florida), August 5, 2001, p. 1A; Barbara Hagenbaugh, "Dairy Farms Evolve to Survive," USA *Today*, August 7, 2003, p. 1B; Del Jones and Barbara Hansen, "Special Report: A Who's Who of Productivity," USA *Today*, August 30, 2001, p. 1B; and Christopher Helman, Shell Shocked, *Forbes Online*, July 27, 2006.

Answers to Try It! Problems

1. An unemployed factory worker could be put to work; he or she counts as labour.
2. A college professor is labour.
3. The library building on your campus is part of capital.
4. Yellowstone National Park. Those areas of the park left in their natural state are a natural resource. Facilities such as visitors' centers, roads, and campgrounds are capital.

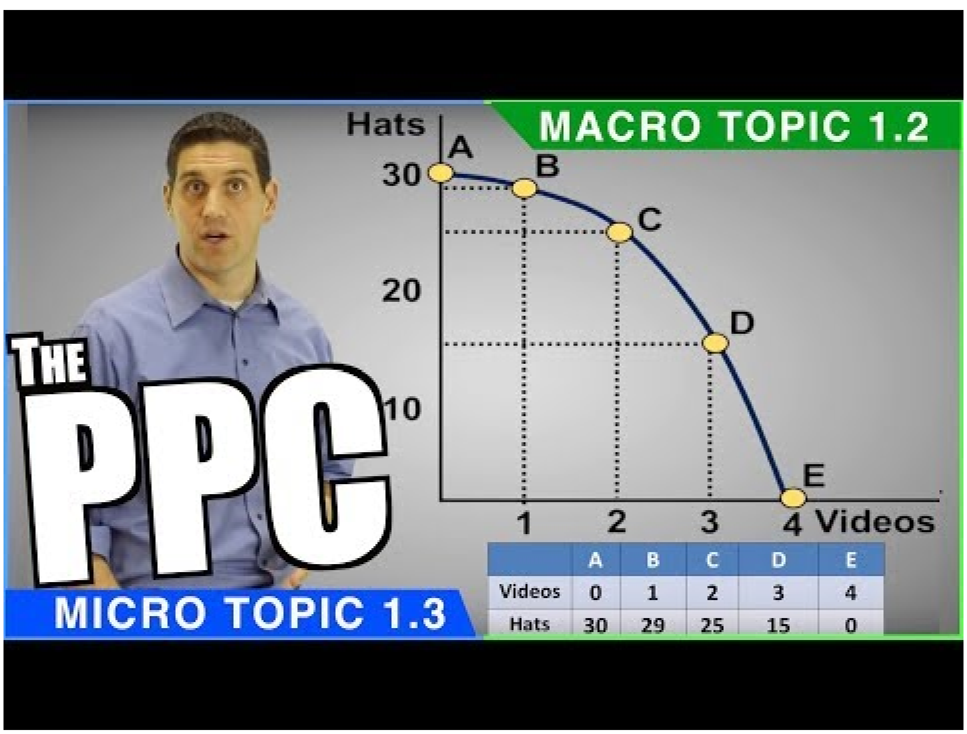
5. An untapped deposit of natural gas is a natural resource. Once extracted and put in a storage tank, natural gas is capital.
6. The White House is capital.
7. The local power plant is capital.

2.2 The Production Possibilities Curve

Learning Objectives

1. Explain the concept of the production possibilities curve and understand the implications of its downward slope and bowed-out shape.
2. Use the production possibilities model to distinguish between full employment and situations of idle factors of production and between efficient and inefficient production.
3. Understand specialization and its relationship to the production possibilities model and comparative advantage.

An economy's factors of production are scarce; they cannot produce an unlimited quantity of goods and services. A production possibilities curve is a graphical representation of the alternative combinations of goods and services an economy can produce. It illustrates the production possibilities model. In drawing the production possibilities curve, we shall assume that the economy can produce only two goods and that the quantities of factors of production and the technology available to the economy are fixed.



A YouTube element has been excluded from this version of the text. You can view it online here: https://pressbooks.senecacollege.ca/macroeconomics/?p=44

Constructing a Production Possibilities Curve

To construct a production possibilities curve, we will begin with the case of a hypothetical firm, Alpine Sports, Inc., a specialized sports equipment manufacturer. Christie Ryder began the business 15 years ago with a single ski production facility near Killington ski resort in central Vermont. Ski sales grew, and she also saw demand for snowboards rising—particularly after snowboard competition events were included in the 2002 Winter Olympics in Salt Lake City. She added a second plant in a nearby town. The second plant, while smaller than the first, was designed to produce snowboards as well as skis. She also modified the first plant so that it could produce both snowboards and skis. Two years later she added a third plant in another town. While even smaller than the second plant, the third was primarily designed for snowboard production but could also produce skis.

We can think of each of Ms. Ryder's three plants as a miniature economy and analyze them using the production possibilities model. We assume that the factors of production and technology available to each of the plants operated by Alpine Sports are unchanged.

Suppose the first plant, Plant 1, can produce 200 pairs of skis per month when it produces only skis. When devoted solely to snowboards, it produces 100 snowboards per month. It can produce skis and snowboards simultaneously as well.

The table in Figure 2.2 "A Production Possibilities Curve" gives three combinations of skis and snowboards that Plant 1 can produce each month. Combination A involves devoting the plant entirely to ski production; combination C means shifting all of the plant's resources to snowboard production; combination B involves the production of both goods. These values are plotted in a production possibilities curve for Plant 1. The curve is a downward-sloping straight line, indicating that there is a linear, negative relationship between the production of the two goods.

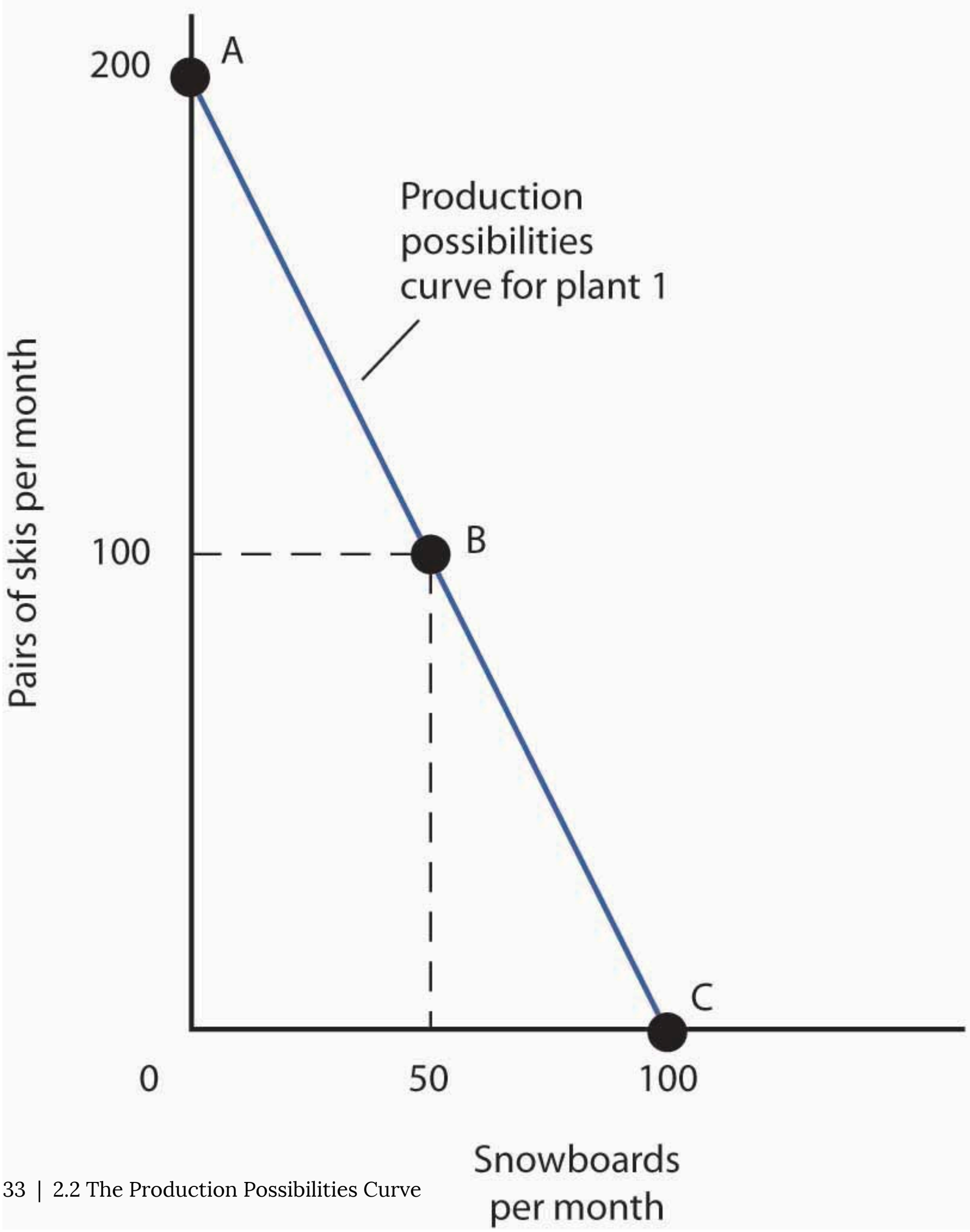
Neither skis nor snowboards is an independent or a dependent variable in the production possibilities model; we can assign either one to the vertical or to the horizontal axis. Here, we have placed the number of pairs of skis produced per month on the vertical axis and the number of snowboards produced per month on the horizontal axis.

The negative slope of the production possibilities curve reflects the scarcity of the plant's capital and labor. Producing more snowboards requires shifting resources out of ski production and thus producing fewer skis. Producing more skis requires shifting resources out of snowboard production and thus producing fewer snowboards.

The slope of Plant 1's production possibilities curve measures the rate at which Alpine Sports must give up ski production to produce additional snowboards. Because the production possibilities curve for Plant 1 is linear, we can compute the slope between any two points on the curve and get the same result. Between points A and B, for example, the slope equals −2 pairs of skis/snowboard (equals −100 pairs of skis/50 snowboards). (Many students are helped when told to read this result as "−2 pairs of skis *per* snowboard.") We get the same value between points B and C, and between points A and C.

Figure 2.2 A Production Possibilities Curve

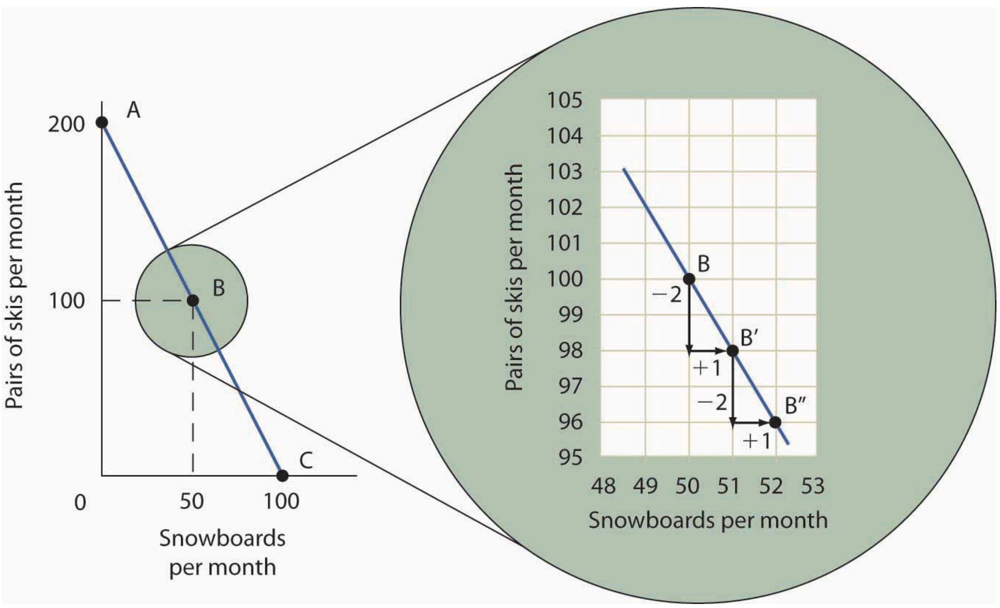
	Pairs of skis per month	Snowboards per month
A	200	0
B	100	50
C	0	100



The table shows the combinations of pairs of skis and snowboards that Plant 1 is capable of producing each month. These are also illustrated with a production possibilities curve. Notice that this curve is linear.

To see this relationship more clearly, examine [Figure 2.3 "The Slope of a Production Possibilities Curve"](). Suppose Plant 1 is producing 100 pairs of skis and 50 snowboards per month at point B. Now consider what would happen if Ms. Ryder decided to produce 1 more snowboard per month. The segment of the curve around point B is magnified in [Figure 2.3 "The Slope of a Production Possibilities Curve"](). The slope between points B and B′ is −2 pairs of skis/snowboard. Producing 1 additional snowboard at point B′ requires giving up 2 pairs of skis. We can think of this as the opportunity cost of producing an additional snowboard at Plant 1. This opportunity cost equals the absolute value of the slope of the production possibilities curve.

Figure 2.3 The Slope of a Production Possibilities Curve

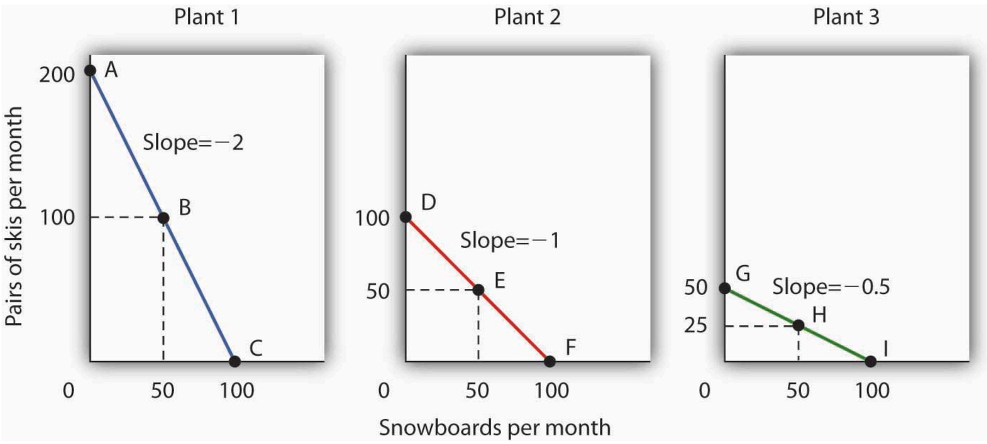


The slope of the linear production possibilities curve in [Figure 2.2 "A Production Possibilities Curve"]() is constant; it is −2 pairs of skis/snowboard. In the section of the curve shown here, the slope can be calculated between points B and B′. Expanding snowboard production to 51 snowboards per month from 50 snowboards per month requires a reduction in ski production to 98 pairs of skis per month from 100 pairs. The slope equals −2 pairs of skis/snowboard (that is, it must give up two pairs of skis to free up the resources necessary to produce one additional snowboard). To shift from B′ to B″, Alpine Sports must give up two more pairs of skis per snowboard. The absolute value of the slope of a production possibilities curve measures the opportunity cost of an additional unit of the good on the horizontal axis measured in terms of the quantity of the good on the vertical axis that must be forgone.

The absolute value of the slope of any production possibilities curve equals the opportunity cost of an additional unit of the good on the horizontal axis. It is the amount of the good on the vertical axis that must be given up in order to free up the resources required to produce one more unit of the good on the horizontal axis. We will make use of this important fact as we continue our investigation of the production possibilities curve.

[Figure 2.4 "Production Possibilities at Three Plants"]() shows production possibilities curves for each of the firm's three plants. Each of the plants, if devoted entirely to snowboards, could produce 100 snowboards. Plants 2 and 3, if devoted exclusively to ski production, can produce 100 and 50 pairs of skis per month, respectively. The exhibit gives the slopes of the production possibilities curves for each plant. The opportunity cost of an additional snowboard at each plant equals the absolute values of these slopes (that is, the number of pairs of skis that must be given up per snowboard).

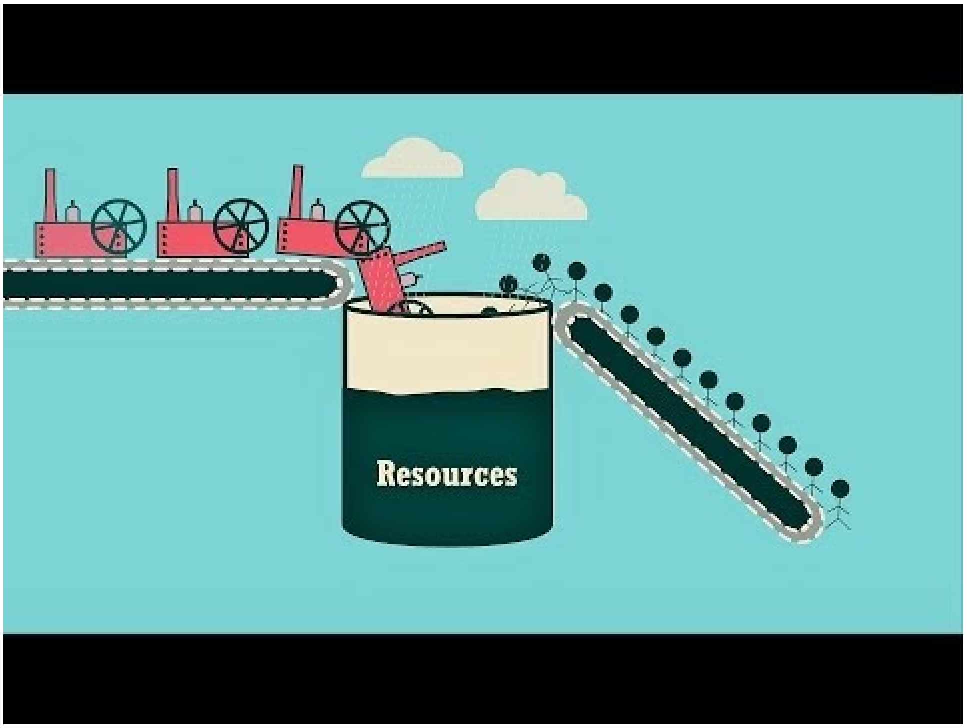
Figure 2.4 Production Possibilities at Three Plants



The slopes of the production possibilities curves for each plant differ. The steeper the curve, the greater the opportunity cost of an additional snowboard. Here, the opportunity cost is lowest at Plant 3 and greatest at Plant 1.

The exhibit gives the slopes of the production possibilities curves for each of the firm's three plants. The opportunity cost of an additional snowboard at each plant equals the absolute values of these slopes. More generally, the absolute value of the slope of any production possibilities curve at any point gives the opportunity cost of an additional unit of the good on the horizontal axis, measured in terms of the number of units of the good on the vertical axis that must be forgone.

The greater the absolute value of the slope of the production possibilities curve, the greater the opportunity cost will be. The plant for which the opportunity cost of an additional snowboard is greatest is the plant with the steepest production possibilities curve; the plant for which the opportunity cost is lowest is the plant with the flattest production possibilities curve. The plant with the lowest opportunity cost of producing snowboards is Plant 3; its slope of −0.5 means that Ms. Ryder must give up half a pair of skis in that plant to produce an additional snowboard. In Plant 2, she must give up one pair of skis to gain one more snowboard. We have already seen that an additional snowboard requires giving up two pairs of skis in Plant 1.



A YouTube element has been excluded from this version of the text. You can view it online here: https://pressbooks.senecacollege.ca/macroeconomics/?p=44

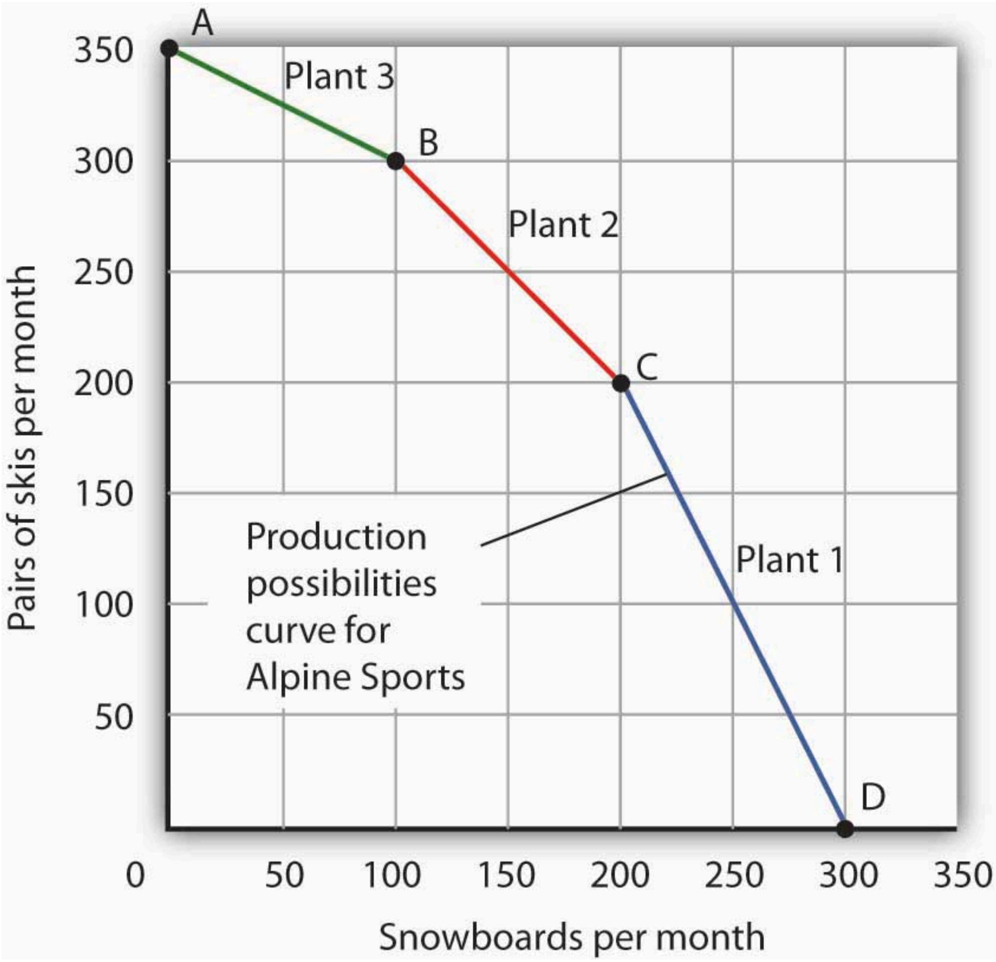
Comparative Advantage and the Production Possibilities Curve

To construct a combined production possibilities curve for all three plants, we can begin by asking how many pairs of skis Alpine Sports could produce if it were producing only skis. To find this quantity, we add up the values at the vertical intercepts of each of the production possibilities curves in Figure 2.4 "Production Possibilities at Three Plants". These intercepts tell us the maximum number of pairs of skis each plant can produce. Plant 1 can produce 200 pairs of skis per month, Plant 2 can produce 100 pairs of skis at per month, and Plant 3 can produce 50 pairs. Alpine Sports can thus produce 350 pairs of skis per month if it devotes its resources exclusively to ski production. In that case, it produces no snowboards.

Now suppose the firm decides to produce 100 snowboards. That will require shifting one of its plants out of ski production. Which one will it choose to shift? The sensible thing for it to do is to choose the plant in which snowboards have the lowest opportunity cost—Plant 3. It has an advantage not because it can produce more snowboards than the other plants (all the plants in this example are capable of producing up to 100 snowboards per month) but because it is the least productive plant for making skis. Producing a snowboard in Plant 3 requires giving up just half a pair of skis.

Economists say that an economy has a comparative advantage in producing a good or service if the opportunity cost of producing that good or service is lower for that economy than for any other. Plant 3 has a comparative advantage in snowboard production because it is the plant for which the opportunity cost of additional snowboards is lowest. To put this in terms of the production possibilities curve, Plant 3 has a comparative advantage in snowboard production (the good on the horizontal axis) because its production possibilities curve is the flattest of the three curves.

Figure 2.5 The Combined Production Possibilities Curve for Alpine Sports



The curve shown combines the production possibilities curves for each plant. At point A, Alpine Sports produces 350 pairs of skis per month and no snowboards. If the firm wishes to increase snowboard production, it will first use Plant 3, which has a comparative advantage in snowboards.

Plant 3's comparative advantage in snowboard production makes a crucial point about the nature of comparative advantage. It need not imply that a particular plant is especially good at an activity. In our example, all three plants are equally good at snowboard production. Plant 3, though, is the least efficient of the three in ski production. Alpine thus gives up fewer skis when it produces snowboards in Plant 3. Comparative advantage thus can stem from a lack of efficiency in the production of an alternative good rather than a special proficiency in the production of the first good.

The combined production possibilities curve for the firm's three plants is shown in Figure 2.5 "The Combined Production Possibilities Curve for Alpine Sports". We begin at point A, with all three plants producing only skis. Production totals 350 pairs of skis per month and zero snowboards. If the firm were to produce 100 snowboards at Plant 3, ski production would fall by 50 pairs per month (recall that the opportunity cost per snowboard at Plant 3 is half a pair of skis). That would bring ski production to 300 pairs, at point B. If Alpine Sports were to produce still more snowboards in a single month, it would shift production to Plant 2, the facility with the next-lowest opportunity cost. Producing 100 snowboards at Plant 2 would leave Alpine Sports producing 200 snowboards and 200 pairs of skis per month, at point

C. If the firm were to switch entirely to snowboard production, Plant 1 would be the last to switch because the cost of each snowboard there is 2 pairs of skis. With all three plants producing only snowboards, the firm is at point D on the combined production possibilities curve, producing 300 snowboards per month and no skis.

Notice that this production possibilities curve, which is made up of linear segments from each assembly plant, has a bowed-out shape; the absolute value of its slope increases as Alpine Sports produces more and more snowboards. This is a result of transferring resources from the production of one good to another according to comparative advantage. We shall examine the significance of the bowed-out shape of the curve in the next section.

The Law of Increasing Opportunity Cost

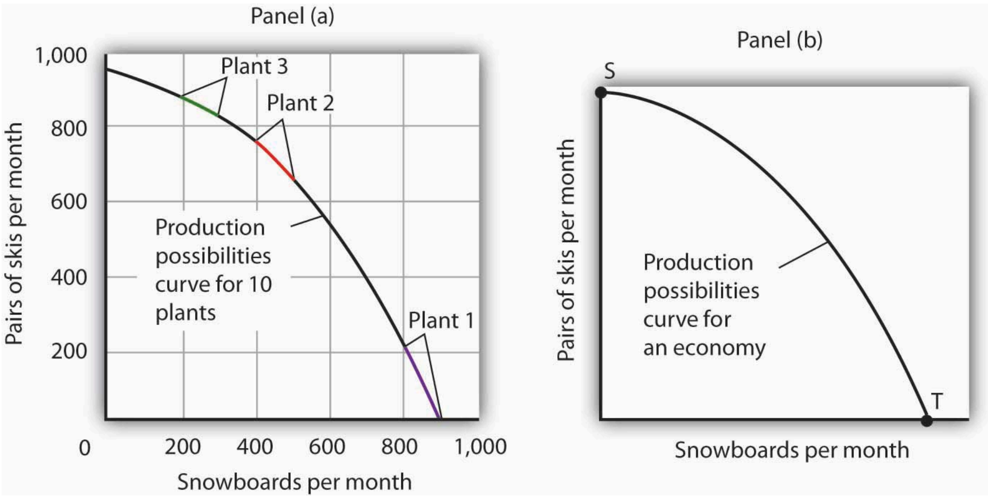
We see in Figure 2.5 "The Combined Production Possibilities Curve for Alpine Sports" that, beginning at point A and producing only skis, Alpine Sports experiences higher and higher opportunity costs as it produces more snowboards. The fact that the opportunity cost of additional snowboards increases as the firm produces more of them is a reflection of an important economic law. The law of increasing opportunity cost holds that as an economy moves along its production possibilities curve in the direction of producing more of a particular good, the opportunity cost of additional units of that good will increase.

We have seen the law of increasing opportunity cost at work traveling from point A toward point D on the production possibilities curve in Figure 2.5 "The Combined Production Possibilities Curve for Alpine Sports". The opportunity cost of each of the first 100 snowboards equals half a pair of skis; each of the next 100 snowboards has an opportunity cost of 1 pair of skis, and each of the last 100 snowboards has an opportunity cost of 2 pairs of skis. The law also applies as the firm shifts from snowboards to skis. Suppose it begins at point D, producing 300 snowboards per month and no skis. It can shift to ski production at a relatively low cost at first. The opportunity cost of the first 200 pairs of skis is just 100 snowboards at Plant 1, a movement from point D to point C, or 0.5 snowboards per pair of skis. We would say that Plant 1 has a comparative advantage in ski production. The next 100 pairs of skis would be produced at Plant 2, where snowboard production would fall by 100 snowboards per month. The opportunity cost of skis at Plant 2 is 1 snowboard per pair of skis. Plant 3 would be the last plant converted to ski production. There, 50 pairs of skis could be produced per month at a cost of 100 snowboards, or an opportunity cost of 2 snowboards per pair of skis.

The bowed-out production possibilities curve for Alpine Sports illustrates the law of increasing opportunity cost. Scarcity implies that a production possibilities curve is downward sloping; the law of increasing opportunity cost implies that it will be bowed out, or concave, in shape.

The bowed-out curve of Figure 2.5 "The Combined Production Possibilities Curve for Alpine Sports" becomes smoother as we include more production facilities. Suppose Alpine Sports expands to 10 plants, each with a linear production possibilities curve. Panel (a) of Figure 2.6 "Production Possibilities for the Economy" shows the combined curve for the expanded firm, constructed as we did in Figure 2.5 "The Combined Production Possibilities Curve for Alpine Sports". This production possibilities curve includes 10 linear segments and is almost a smooth curve. As we include more and more production units, the curve will become smoother and smoother. In an actual economy, with a tremendous number of firms and workers, it is easy to see that the production possibilities curve will be smooth. We will generally draw production possibilities curves for the economy as smooth, bowed-out curves, like the one in Panel (b). This production possibilities curve shows an economy that produces only skis and snowboards. Notice the curve still has a bowed-out shape; it still has a negative slope. Notice also that this curve has no numbers. Economists often use models such as the production possibilities model with graphs that show the general shapes of curves but that do not include specific numbers.

Figure 2.6 Production Possibilities for the Economy



As we combine the production possibilities curves for more and more units, the curve becomes smoother. It retains its negative slope and bowed-out shape. In Panel (a) we have a combined production possibilities curve for Alpine Sports, assuming that it now has 10 plants producing skis and snowboards. Even though each of the plants has a linear curve, combining them according to comparative advantage, as we did with 3 plants in Figure 2.5 "The Combined Production Possibilities Curve for Alpine Sports", produces what appears to be a smooth, nonlinear curve, even though it is made up of linear segments. In drawing production possibilities curves for the economy, we shall generally assume they are smooth and "bowed out," as in Panel (b). This curve depicts an entire economy that produces only skis and snowboards.

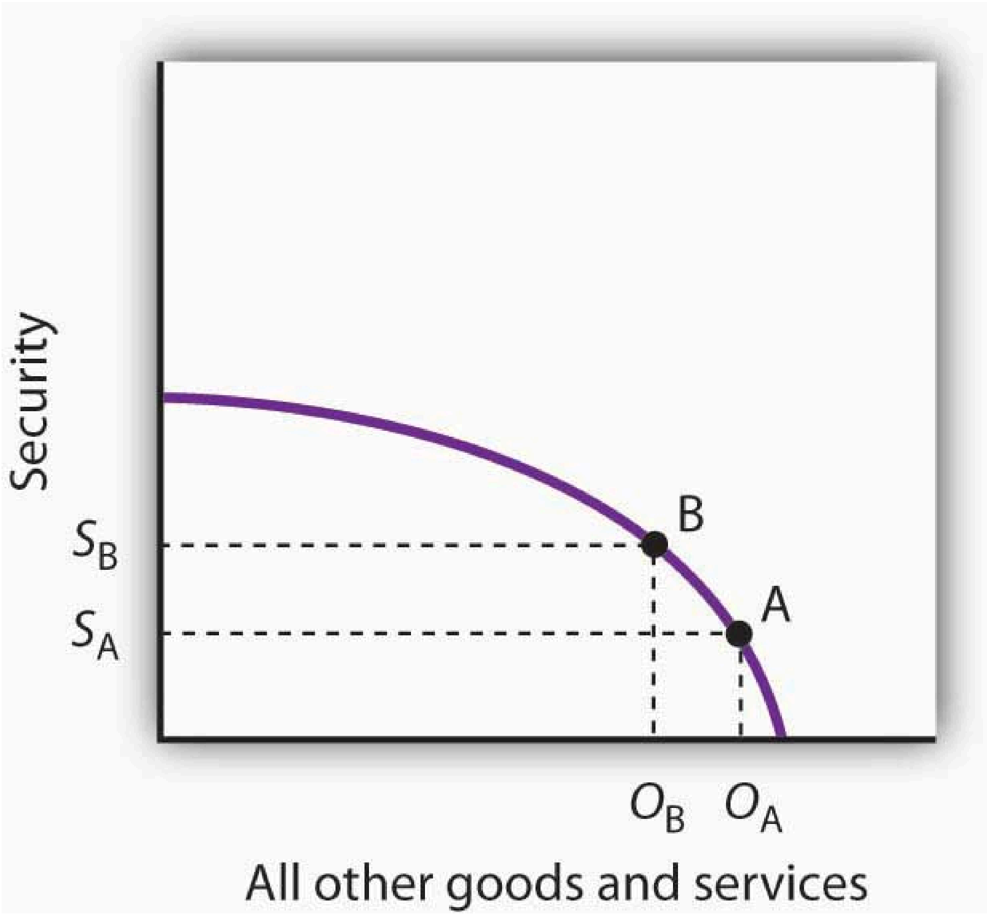
Movements Along the Production Possibilities Curve

We can use the production possibilities model to examine choices in the production of goods and services. In applying the model, we assume that the economy can produce two goods, and we assume that technology and the factors of production available to the economy remain unchanged. In this section, we shall assume that the economy operates on its production possibilities curve so that an increase in the production of one good in the model implies a reduction in the production of the other.

We shall consider two goods and services: national security and a category we shall call "all other goods and services." This second category includes the entire range of goods and services the economy can produce, aside from national defense and security. Clearly, the transfer of resources to the effort to enhance national security reduces the quantity of other goods and services that can be produced. In the wake of the 9/11 attacks in 2001, nations throughout the world increased their spending for national security. This spending took a variety of forms. One, of course, was increased defense spending. Local and state governments also increased spending in an effort to prevent terrorist attacks. Airports around the world hired additional agents to inspect luggage and passengers.

The increase in resources devoted to security meant fewer "other goods and services" could be produced. In terms of the production possibilities curve in Figure 2.7 "Spending More for Security", the choice to produce more security and less of other goods and services means a movement from A to B. Of course, an economy cannot really *produce* security; it can only attempt to provide it. The attempt to provide it requires resources; it is in that sense that we shall speak of the economy as "producing" security.

Figure 2.7 Spending More for Security



Here, an economy that can produce two categories of goods, security and "all other goods and services," begins at point A on its production possibilities curve. The economy produces S_A units of security and O_A units of all other goods and services per period. A movement from A to B requires shifting resources out of the production of all other goods and services and into spending on security. The increase in spending on security, to S_A units of security per period, has an opportunity cost of reduced production of all other goods and services. Production of all other goods and services falls by $O_A - O_B$ units per period.

At point A, the economy was producing S_A units of security on the vertical axis—defense services and various forms of police protection—and O_A units of other goods and services on the horizontal axis. The decision to devote more resources to security and less to other goods and services represents the choice we discussed in the chapter introduction. In this case we have categories of goods rather than specific goods. Thus, the economy chose to increase spending on security in the effort to defeat terrorism. Since we have assumed that the economy has a fixed quantity of available resources, the increased use of resources for security and national defense necessarily reduces the number of resources available for the production of other goods and services.

The law of increasing opportunity cost tells us that, as the economy moves along the production possibilities curve in the direction of more of one good, its opportunity cost will increase. We may conclude that, as the economy moved along this curve in the direction of greater production of security, the opportunity cost of the additional security began to increase. That is because the resources transferred from the production of other goods and services to the production of security had a greater and greater comparative advantage in producing things other than security.

The production possibilities model does not tell us where on the curve a particular economy will operate. Instead, it lays out the possibilities facing the economy. Many countries, for example, chose to move along their respective production possibilities curves to produce more security and national defense and less of all other goods in the wake of 9/11. We will see in the chapter on demand and supply how choices about what to produce are made in the marketplace.

Producing on Versus Producing Inside the Production Possibilities Curve

An economy that is operating inside its production possibilities curve could, by moving onto it, produce more of all the goods and services that people value, such as food, housing, education, medical care, and music. Increasing the availability of these goods would improve the standard of living. Economists conclude that it is better to be on the production possibilities curve than inside it.

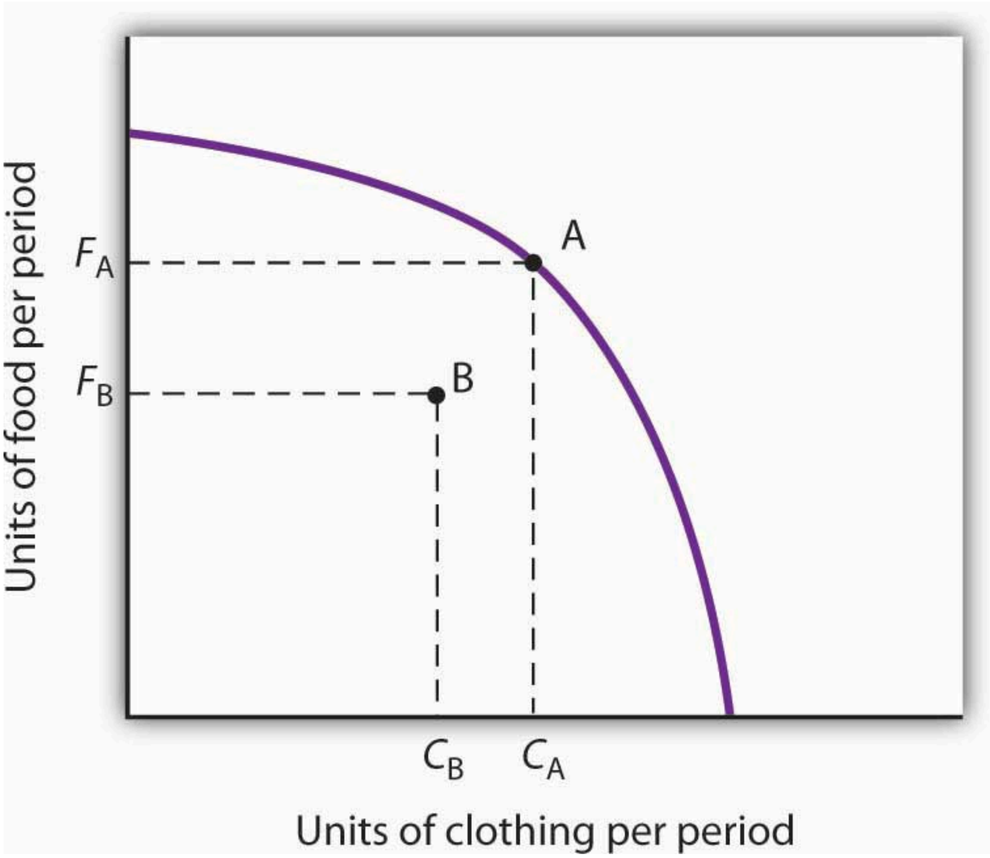
Two things could leave an economy operating at a point inside its production possibilities curve. First, the economy might fail to use fully the resources available to it. Second, it might not allocate resources on the basis of comparative advantage. In either case, production within the production possibilities curve implies the economy could improve its performance.

Idle Factors of Production

Suppose an economy fails to put all its factors of production to work. Some workers are without jobs, some buildings are without occupants, some fields are without crops. Because an economy's production possibilities curve assumes the full use of the factors of production available to it, the failure to use some factors results in a level of production that lies inside the production possibilities curve.

If all the factors of production that are available for use under current market conditions are being utilized, the economy has achieved full employment. An economy cannot operate on its production possibilities curve unless it has full employment.

Figure 2.8 Idle Factors and Production



The production possibilities curve shown suggests an economy that can produce two goods, food and clothing. As a result of a failure to achieve full employment, the economy operates at a point such as B, producing F_B units of food and C_B units of clothing per period. Putting its factors of production to work allows a move to the production possibilities curve, to a point such as A. The production of both goods rises.

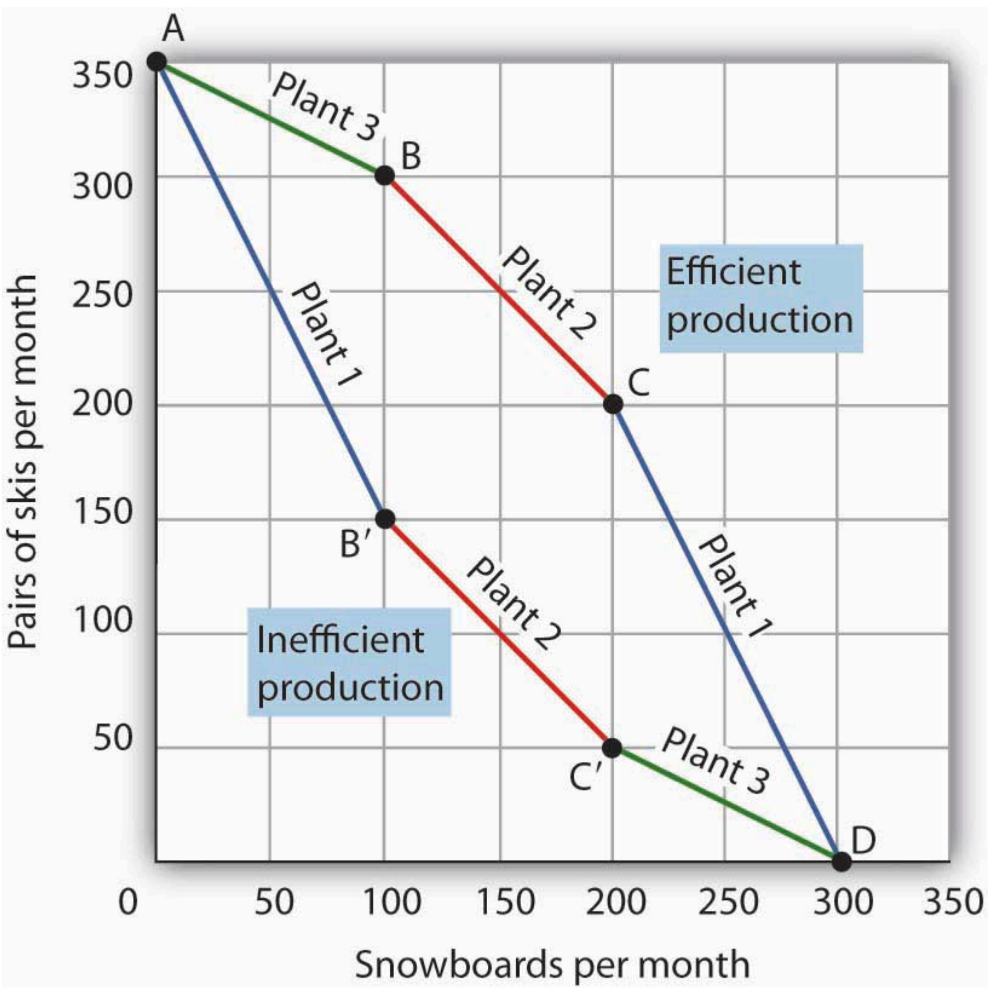
Figure 2.8 "Idle Factors and Production" shows an economy that can produce food and clothing. If it chooses to produce at point A, for example, it can produce F_A units of food and C_A units of clothing. Now suppose that a large fraction of the economy's workers lose their jobs, so the economy no longer makes full use of one factor of production: labor. In this example, production moves to point B, where the economy produces less food (F_B) and less clothing (C_B) than at point A. We often think of the loss of jobs in terms of the workers; they have lost a chance to work and to earn income. But the production possibilities model points to another loss: goods and services the economy could have produced that are not being produced.

Inefficient Production

Now suppose Alpine Sports is fully employing its factors of production. Could it still operate inside its production possibilities curve? Could an economy that is using all its factors of production still produce less than it could? The answer is "Yes," and the key lies in comparative advantage. An economy achieves a point on its production possibilities curve only if it allocates its factors of production on the basis of comparative advantage. If it fails to do that, it will operate inside the curve.

Suppose that, as before, Alpine Sports has been producing only skis. With all three of its plants producing skis, it can produce 350 pairs of skis per month (and no snowboards). The firm then starts producing snowboards. This time, however, imagine that Alpine Sports switches plants from skis to snowboards in numerical order: Plant 1 first, Plant 2 second, and then Plant 3. Figure 2.9 "Efficient Versus Inefficient Production" illustrates the result. Instead of the bowed-out production possibilities curve ABCD, we get a bowed-in curve, AB′C′D. Suppose that Alpine Sports is producing 100 snowboards and 150 pairs of skis at point B′. Had the firm based its production choices on comparative advantage, it would have switched Plant 3 to snowboards and then Plant 2, so it could have operated at a point such as C. It would be producing more snowboards and more pairs of skis—and using the same quantities of factors of production it was using at B′. Had the firm based its production choices on comparative advantage, it would have switched Plant 3 to snowboards and then Plant 2, so it would have operated at point C. It would be producing more snowboards and more pairs of skis—and using the same quantities of factors of production it was using at B′. When an economy is operating on its production possibilities curve, we say that it is engaging in efficient production. If it is using the same quantities of factors of production but is operating inside its production possibilities curve, it is engaging in inefficient production. Inefficient production implies that the economy could be producing more goods without using any additional labor, capital, or natural resources.

Figure 2.9 Efficient Versus Inefficient Production



When factors of production are allocated on a basis other than comparative advantage, the result is inefficient production. Suppose Alpine Sports operates the three plants we examined in Figure 2.4 "Production Possibilities at Three Plants". Suppose further that all three plants are devoted exclusively to ski production; the firm operates at A. Now suppose that, to increase snowboard production, it transfers plants in numerical order: Plant 1 first, then Plant 2, and finally Plant 3. The result is the bowed-in curve AB′C′D. Production on the production possibilities curve ABCD requires that factors of production be transferred according to comparative advantage.

Points on the production possibilities curve thus satisfy two conditions: the economy is making full use of its factors of production, and it is making efficient use of its factors of production. If there are idle or inefficiently allocated factors of production, the economy will operate inside the production possibilities curve. Thus, the production possibilities curve not only shows what can be produced; it provides insight into how goods and services should be produced. It suggests that to obtain efficiency in production, factors of production should be allocated on the basis of comparative advantage. Further, the economy must make full use of its factors of production if it is to produce the goods and services it is capable of producing.

Specialization

The production possibilities model suggests that specialization will occur. Specialization implies that an economy is producing the goods and services in which it has a comparative advantage. If Alpine Sports selects point C in Figure 2.9 "Efficient Versus Inefficient Production", for example, it will assign Plant 1 exclusively to ski production and Plants 2 and 3 exclusively to snowboard production.

Such specialization is typical in an economic system. Workers, for example, specialize in particular fields in which they have a comparative advantage. People work and use the income they earn to buy—perhaps import—goods and services from people who have a comparative advantage in doing other things. The result is a far greater quantity of goods and services than would be available without this specialization.

Think about what life would be like without specialization. Imagine that you are suddenly completely cut off from the rest of the economy. You must produce everything you consume; you obtain nothing from anyone else. Would you be able to consume what you consume now? Clearly not. It is hard to imagine that most of us could even survive in such a setting. The gains we achieve through specialization are enormous.

Nations specialize as well. Much of the land in the United States has a comparative advantage in agricultural production and is devoted to that activity. Hong Kong, with its huge population and tiny endowment of land, allocates virtually none of its land to agricultural use; that option would be too costly. Its land is devoted largely to nonagricultural use.

A YouTube element has been excluded from this version of the text. You can view it online here: https://pressbooks.senecacollege.ca/macroeconomics/?p=44

Key Takeaways

- A production possibilities curve shows the combinations of two goods an economy is capable of producing.
- The downward slope of the production possibilities curve is an implication of scarcity.
- The bowed-out shape of the production possibilities curve results from allocating resources based on comparative advantage. Such an allocation implies that the law of increasing opportunity cost will hold.
- An economy that fails to make full and efficient use of its factors of production will operate inside its production possibilities curve.
- Specialization means that an economy is producing the goods and services in which it has a comparative advantage.

Try It!

Suppose a manufacturing firm is equipped to produce radios or calculators. It has two plants, Plant R and Plant S, at which it can produce these goods. Given the labor and the capital available at both plants, it can produce the combinations of the two goods at the two plants shown.

Output per day, Plant S		
Combination	**Calculators**	**Radios**
D	50	0
E	25	50
F	0	100

Put calculators on the vertical axis and radios on the horizontal axis. Draw the production possibilities curve for Plant R. On a separate graph, draw the production possibilities curve for Plant S. Which plant has a comparative advantage in calculators? In radios? Now draw the combined curves for the two plants. Suppose the firm decides to produce 100 radios. Where will it produce them? How many calculators will it be able to produce? Where will it produce the calculators?

Case in Point: The Cost of the Great Depression

Figure 2.10

Wikimedia Commons – public domain.

The U.S. economy looked very healthy in the beginning of 1929. It had enjoyed seven years of dramatic growth and unprecedented prosperity. Its resources were fully employed; it was operating quite close to its production possibilities curve.

In the summer of 1929, however, things started going wrong. Production and employment fell. They continued to fall for several years. By 1933, more than 25% of the nation's workers had lost their jobs. Production had plummeted by almost 30%. The economy had moved well within its production possibilities curve.

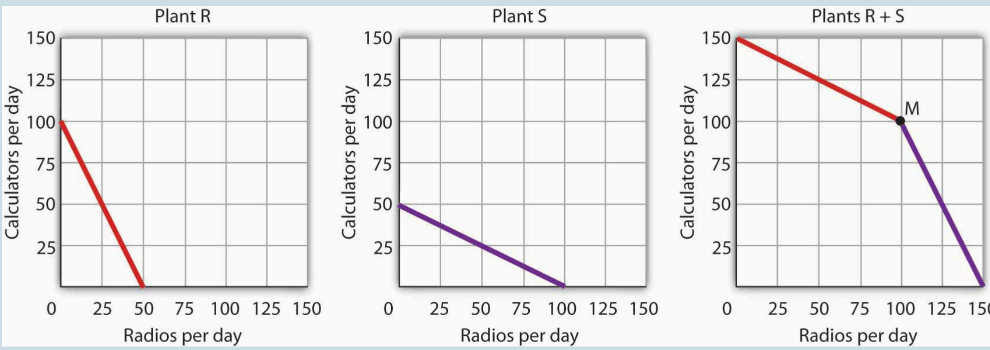
Output began to grow after 1933, but the economy continued to have vast numbers of idle workers, idle factories, and idle farms. These resources were not put back to work fully until 1942, after the U.S. entry into World War II demanded mobilization of the economy's factors of production.

Between 1929 and 1942, the economy produced 25% fewer goods and services than it would have if its resources had been fully employed. That was a loss, measured in today's dollars, of well over $3 trillion. In material terms, the forgone output represented a greater cost than the United States would ultimately spend in World War II. The Great Depression was a costly experience indeed.

Answer to Try It! Problem

The production possibilities curves for the two plants are shown, along with the combined curve for both plants. Plant R has a comparative advantage in producing calculators. Plant S has a comparative advantage in producing radios, so, if the firm goes from producing 150 calculators and no radios to producing 100 radios, it will produce them at Plant S. In the production possibilities curve for both plants, the firm would be at M, producing 100 calculators at Plant R.

Figure 2.11



2.3 Applications of the Production Possibilities Model

Learning Objectives

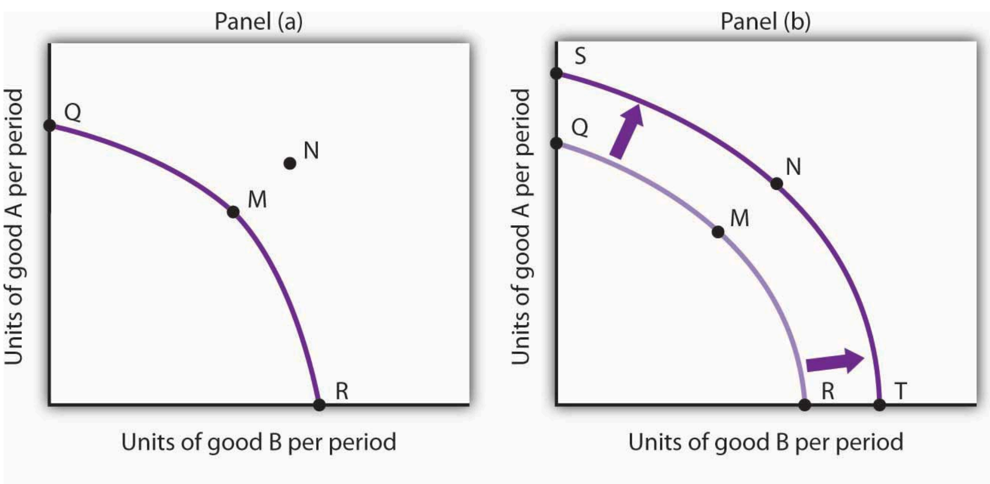
1. Define economic growth in terms of the production possibilities model and discuss factors that make such growth possible.
2. Explain the classification of economic systems, the role of government in different economic systems, and the strengths and weaknesses of different systems.

The production possibilities curve gives us a model of an economy. The model provides powerful insights about the real world, insights that help us to answer some important questions: What determines the rate at which production will increase over time? What is the role of economic freedom in the economy? In this section we explore applications of the model to questions of economic growth and the choice of an economic system.

Economic Growth

An increase in the physical quantity or in the quality of factors of production available to an economy or a technological gain will allow the economy to produce more goods and services; it will shift the economy's production possibilities curve outward. The process through which an economy achieves an outward shift in its production possibilities curve is called economic growth. An outward shift in a production possibilities curve is illustrated in [Figure 2.13 "Economic Growth and the Production Possibilities Curve"](). In Panel (a), a point such as N is not attainable; it lies outside the production possibilities curve. Growth shifts the curve outward, as in Panel (b), making previously unattainable levels of production possible.

Figure 2.13 Economic Growth and the Production Possibilities Curve



An economy capable of producing two goods, A and B, is initially operating at point M on production possibilities curve OMR in Panel (a). Given this production possibilities curve, the economy could not produce a combination such as shown by point N, which lies outside the curve. An increase in the factors of production available to the economy would shift the curve outward to SNT, allowing the choice of a point such as N, at which more of both goods will be produced.

The Sources of Economic Growth

Economic growth implies an outward shift in an economy's production possibilities curve. Recall that when we draw such a curve, we assume that the quantity and quality of the economy's factors of production and its technology are unchanged. Changing these will shift the curve. Anything that increases the quantity or quality of the factors of production available to the economy or that improves the technology available to the economy contributes to economic growth.

Consider, for example, the dramatic gains in human capital that have occurred in the United States since the beginning of the past century. In 1900, about 3.5% of U.S. workers had completed a high school education. By 2006, that percentage rose almost to 92. Fewer than 1% of the workers in 1900 had graduated from college; as late as 1940 only 3.5% had graduated from college. By 2006, nearly 32% had graduated from college. In addition to being better educated, today's workers have received more and better training on the job. They bring far more economically useful knowledge and skills to their work than did workers a century ago.

Moreover, the technological changes that have occurred within the past 100 years have greatly reduced the time and effort required to produce most goods and services. Automated production has become commonplace. Innovations in transportation (automobiles, trucks, and airplanes) have made the movement of goods and people cheaper and faster. A dizzying array of new materials is available for manufacturing. And the development of modern information technology—including computers, software, and communications equipment—that seemed to proceed at breathtaking pace especially during the final years of the last century and continuing to the present has transformed the way we live and work.

Look again at the technological changes of the last few years described in the Case in Point on advances in technology. Those examples of technological progress through applications of computer technology—from new ways of mapping oil deposits to new methods of milking cows—helped propel the United States and other economies to dramatic gains in the ability to produce goods and services. They have helped shift the countries' production possibilities curve outward. They have helped fuel economic growth.

Table 2.1 "Sources of U.S. Economic Growth, 1948–2002" summarizes the factors that have contributed to U.S. economic growth in the past half century. When looking at the period of 1948–2002 as a whole we see that about 60% of economic growth stems from increases in the quantities of capital and labour and 40% from increases in the qualities of the factors of production and improvements in technology. In the most recent period, 1995–2002, however, these percentages are essentially reversed, with a little less than 30% explained by increases in quantities of the factors of production and a whopping 70% explained by improvements in factor quality and technology.

Table 2.1 Sources of U.S. Economic Growth, 1948–2002

Year	Percentage contribution to growth	Period growth rate
Years 1948–2002		**3.46%**
Increase in quantity of labour	21%	
Increase in quantity of capital	41%	
Increase in quality of labour	10%	
Increase in quality of capital	20%	
Improved technology	25%	
Years 1948–1973		**3.99%**
Increase in quantity of labour	15%	
Increase in quantity of capital	44%	
Increase in quality of labour	11%	
Increase in quality of capital	5%	
Improved technology	25%	
Years 1973–1989		**2.97%**
Increase in quantity of labour	31%	
Increase in quantity of capital	39%	
Increase in quality of labour	7%	
Increase in quality of capital	12%	
Improved technology	10%	
Years 1989–1995		**2.43%**
Increase in quantity of labour	26%	
Increase in quantity of capital	33%	
Increase in quality of labour	15%	
Increase in quality of capital	17%	
Improved technology	11%	
Years 1995–2002		**3.59%**
Increase in quantity of labour	19%	
Increase in quantity of capital	8%	
Increase in quality of labour	5%	
Increase in quality of capital	47%	
Improved technology	20%	

Total output during the period shown increased sixfold. The chart shows the percentage of this increase accounted for by increases in the quantity of labor and of capital and by increases in the quality of labour and of capital and improvements in technology. In the 1995–2002 period, the incorporation of information technology led to improvements in the quality of capital and technology that greatly contributed to growth.

Source: Based on Dale W. Jorgenson, "Accounting for Growth in the Information Age," *Handbook of Economic Growth*, Phillipe Aghion and Steven Durlauf, eds. Amsterdam: North Holland, 2005.

Another way of looking at these data for the most recent period is to notice that the increase in the rate of economic growth between the 1989 to 1995 period and the 1995 to 2002 period of more than one percentage point per year is largely explained by better-quality capital and better technology. The study by economist Dale Jorgenson on which the data shown in Table 2.1 "Sources of U.S. Economic Growth, 1948–2002" are derived notes that these two main contributors to higher economic growth can be largely attributed to the development of information technology and its incorporation in the workplace.

Waiting for Growth

One key to growth is, in effect, the willingness to wait, to postpone current consumption in order to enhance future productive capability. When Stone Age people fashioned the first tools, they were spending time building capital rather than engaging in consumption. They delayed current consumption to enhance their future consumption; the tools they made would make them more productive in the future.

Resources society could have used to produce consumer goods are being used to produce new capital goods and new knowledge for production instead—all to enhance future production. An even more important source of growth in many nations has been increased human capital. Increases in human capital often require the postponement of consumption. If you are a college student, you are engaged in precisely this effort. You are devoting time to study that could have been spent working, earning income, and thus engaging in a higher level of consumption. If you are like most students, you are making this choice to postpone consumption because you expect it will allow you to earn more income, and thus enjoy greater consumption, in the future.

Think of an economy as being able to produce two goods, capital and consumer goods (those destined for immediate use by consumers). By focusing on the production of consumer goods, the people in the economy will be able to enjoy a higher standard of living today. If they reduce their consumption—and their standard of living—today to enhance their ability to produce goods and services in the future, they will be able to shift their production possibilities curve outward. That may allow them to produce even more consumer goods. A decision for greater growth typically involves the sacrifice of present consumption.

Arenas for Choice: A Comparison of Economic Systems

Under what circumstances will a nation achieve efficiency in the use of its factors of production? The discussion above suggested that Christie Ryder would have an incentive to allocate her plants efficiently because by doing so she could achieve greater output of skis and snowboards than would be possible from inefficient production. But why would she want to produce more of these two goods—or of any goods? Why would decision makers throughout the economy want to achieve such efficiency?

Economists assume that privately owned firms seek to maximize their profits. The drive to maximize profits will lead firms such as Alpine Sports to allocate resources efficiently to gain as much production as possible from their factors of production. But whether firms will seek to maximize profits depends on the nature of the economic system within which they operate.

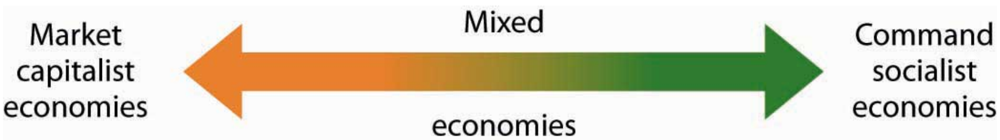
Classifying Economic Systems

Each of the world's economies can be viewed as operating somewhere on a spectrum between market capitalism and command socialism. In a market capitalist economy, resources are generally owned by private individuals who have the power to make decisions about their use. A market capitalist system is often referred to as a free enterprise economic system. In a command socialist economy, the government is the primary owner of capital and natural resources and has broad power to allocate the use of factors of production. Between these two categories lie mixed economies that combine elements of market capitalist and of command socialist economic systems.

No economy represents a pure case of either market capitalism or command socialism. To determine where an economy lies between these two types of systems, we evaluate the extent of government ownership of capital and natural resources and the degree to which government is involved in decisions about the use of factors of production.

Figure 2.14 "Economic Systems" suggests the spectrum of economic systems. Market capitalist economies lie toward the left end of this spectrum; command socialist economies appear toward the right. Mixed economies lie in between. The market capitalist end of the spectrum includes countries such as the United States, the United Kingdom, and Canada. Hong Kong, though now part of China, has a long history as a market capitalist economy and is generally regarded as operating at the market capitalist end of the spectrum. Countries at the command socialist end of the spectrum include North Korea and Cuba.

Figure 2.14 Economic Systems

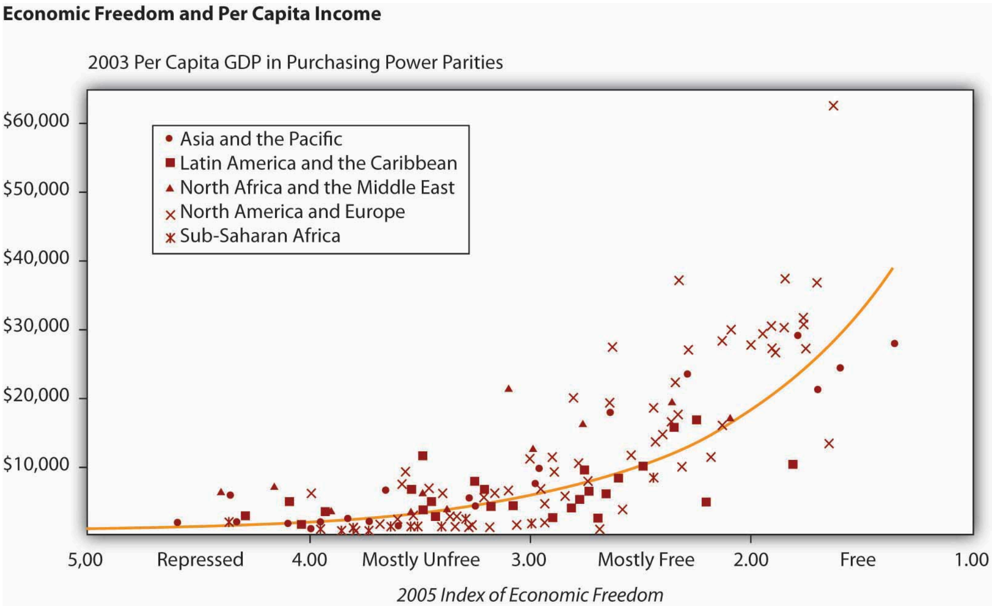


Some European economies, such as France, Germany, and Sweden, have a sufficiently high degree of regulation that we consider them as operating more toward the center of the spectrum. Russia and China, which long operated at the command socialist end of the spectrum, can now be considered mixed economies. Most economies in Latin America once operated toward the right end of the spectrum. While their governments did not exercise the extensive ownership of capital and natural resources that are one characteristic of command socialist systems, their governments did impose extensive regulations. Many of these nations are in the process of carrying out economic reforms that will move them further in the direction of market capitalism.

The global shift toward market capitalist economic systems that occurred in the 1980s and 1990s was in large part the result of three important features of such economies. First, the emphasis on individual ownership and decision-making power has generally yielded greater individual freedom than has been available under command socialist or some more heavily regulated mixed economic systems that lie toward the command socialist end of the spectrum. People seeking political, religious, and economic freedom have thus gravitated toward market capitalism. Second, market economies are more likely than other systems to allocate resources on the basis of comparative advantage. They thus tend to generate higher levels of production and income than do other economic systems. Third, market capitalist-type systems appear to be the most conducive to entrepreneurial activity.

Suppose Christie Ryder had the same three plants we considered earlier in this chapter but was operating in a mixed economic system with extensive government regulation. In such a system, she might be prohibited from transferring resources from one use to another to achieve the gains possible from comparative advantage. If she were operating under a command socialist system, she would not be the owner of the plants and thus would be unlikely to profit from their efficient use. If that were the case, there is no reason to believe she would make any effort to assure the efficient use of the three plants. Generally speaking, it is economies toward the market capitalist end of the spectrum that offer the greatest inducement to allocate resources on the basis of comparative advantage. They tend to be more productive and to deliver higher material standards of living than do economies that operate at or near the command socialist end of the spectrum.

Figure 2.15 Economic Freedom and Income



The horizontal axis shows the degree of economic freedom—"free," "mostly free," "mostly unfree," and "repressed"—according to the measures used by the Heritage Foundation and *The Wall Street Journal*. The graph shows the relationship between economic freedom and per capita income. Countries with higher degrees of economic freedom tended to have higher per capita incomes.

Source: World Bank, *World Development Indicators Online*, available by subscription at www.worldbank.org/data; Central Intelligence Agency, *The World Factbook* 2004, available at http://www.cia.gov/cia/publications/factbook/index.html for the following countries: Bahamas, Burma, Cuba, Cyprus, Equatorial Guinea, North Korea, Libya, Qatar, Suriname, Taiwan, Zimbabwe; Marc A. Miles, Edwin J. Feulner, and Mary Anastasia O'Grady, 2005 *Index of Economic Freedom* (Washington, D.C.: The Heritage Foundation and Dow Jones & Company, Inc., 2005), at www.heritage.org/index.

Market capitalist economies rely on economic freedom. Indeed, one way we can assess the degree to which a country can be considered market capitalist is by the degree of economic freedom it permits. Several organizations have attempted to compare economic freedom in various countries. One of the most extensive comparisons is a joint annual effort by the Heritage Foundation and *The Wall Street Journal*. The 2008 rating was based on policies in effect in 162 nations early that year. The report ranks these nations on the basis of such things as the degree of regulation of firms, tax levels, and restrictions on international trade. Hong Kong ranked as the freest economy in the world. North Korea received the dubious distinction of being the least free.

It seems reasonable to expect that the greater the degree of economic freedom a country permits, the greater the amount of income per person it will generate. This proposition is illustrated in Figure 2.15 "Economic Freedom and Income". The group of countries categorized as "free" generated the highest incomes in the Heritage Foundation/*Wall Street Journal* study; those rated as "repressed" had the lowest. The study also found that countries that over the last decade have done the most to improve their positions in the economic freedom rankings have also had the highest rates of growth. We must be wary of slipping into the fallacy of false cause by concluding from this evidence that economic freedom generates higher incomes. It could be that higher incomes lead nations to opt for greater economic freedom. But in this case, it seems reasonable to conclude that, in general, economic freedom does lead to higher incomes.

Government in a Market Economy

The production possibilities model provides a menu of choices among alternative combinations of goods and services. Given those choices, which combinations will be produced?

In a market economy, this question is answered in large part through the interaction of individual buyers and sellers. As we have already seen, government plays a role as well. It may seek to encourage greater consumption of some goods and discourage consumption of others. In the United States, for example, taxes imposed on cigarettes discourage smoking, while special treatment of property taxes and mortgage interest in the federal income tax encourages home ownership. Government may try to stop the production and consumption of some goods altogether, as many governments do with drugs such as heroin and cocaine. Government may supplement the private consumption of some goods by producing more of them itself, as many U.S. cities do with golf courses and tennis courts. In other cases, there may be no private market for a good or service at all. In the choice between security and defense versus all other goods and services outlined at the beginning of this chapter, government agencies are virtually the sole providers of security and national defense.

All nations also rely on government to provide defense, enforce laws, and redistribute income. Even market economies rely on government to regulate the activities of private firms, to protect the environment, to provide education, and to produce a wide range of other goods and services. Government's role may be limited in a market economy, but it remains fundamentally important.

Key Takeaways

- The ideas of comparative advantage and specialization suggest that restrictions on international trade are likely to reduce production of goods and services.
- Economic growth is the result of increasing the quantity or quality of an economy's factors of production and of advances in technology.
- Policies to encourage growth generally involve postponing consumption to increase capital and human capital.
- Market capitalist economies have generally proved more productive than mixed or command socialist economies.
- Government plays a crucial role in any market economy.

Try It!

Draw a production possibilities curve for an economy that can produce two goods, CD players and jackets. You do not have numbers for this one—just draw a curve with the usual bowed-out shape. Put the quantity of CD players per period on the vertical axis and the quantity of jackets per period on the horizontal axis. Now mark a point A on the curve you have drawn; extend dotted lines from this point to the horizontal and vertical axes. Mark the initial quantities of the two goods as CD_A and J_A, respectively. Explain why, in the absence of economic growth, an increase in jacket production requires a reduction in the production of CD players. Now show how economic growth could lead to an increase in the production of both goods.

Case in Point: The European Union and the Production Possibilities Curve

Figure 2.16

Michael Simmons – Euros – CC BY-NC-ND 2.0.

Formed by the Maastricht Treaty of 1993, The European Union represents one of the boldest efforts of our time to exploit the theory of comparative advantage. The Treaty sought to eliminate all trade barriers between the European Union's members. It established a European Parliament and a European Central Bank. The Bank introduced the euro in 1999, a currency that replaced national currencies such as the German deutsche mark and the French franc. At first, the euro was used only for transactions between banks. 320 million people in 15 EU nations (Austria, Belgium, Cyprus, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, the Netherlands, Portugal, Slovenia, and Spain) used the euro by 2008. While the dollar continues to be more widely used, the total value of euros in circulation exceeds that of dollars.

The movement toward European integration can be dated back more than half a century. In 1950, just five years after a war that had devastated much of the world, Robert Schuman, the French Minister of Foreign Affairs, proposed a union between France and Germany to cooperate in the production of iron and steel. In the context of the time, Schuman's proposal was a radical one. World War II had begun with Germany's attempt to seize control of Europe—and ultimately the world. Japan and Italy joined Germany in this effort. Germany had captured France; France had been liberated in 1944 by the Allied invasion in Normandy. The proposal for cooperation between two countries that had been the most bitter of enemies was a revolutionary one. Schuman's speech, delivered on May 9, 1950, is celebrated throughout Europe as "Europe Day."

In effect, the European Union has created an entity very much like the United States. Countries within the European Union retain their own languages and cultural differences, but they have ceded a remarkable degree of sovereignty to the Union. Members of the European Union can trade as freely with each other as can states within the United States. Just as the U.S. Constitution prohibits states from restricting trade with other states, the European Union has dismantled all forms of restrictions that countries within the Union used to impose on one another. Just as restrictions on specialization among Ms. Ryder's plants in Alpine Sports would have forced it to operate inside its production possibilities curve, restrictions that had existed among members of the European Union once put the members of the Union inside their collective production possibilities curve.

The experiment appears to have been a success. Trade among member nations has expanded sharply. A study by Carmen Diaz Mora, an economist at the University of Castilla-La Mancha in Spain, found that the bulk of the expanded trade within the Union was trade within industries and that it was driven by comparative advantage. In particular, she found that countries in the northern part of the Union, such as France and Germany, tended to specialize in relatively high-valued goods—office equipment and electrical goods—while countries in the southern part of the Union specialized in relatively low-valued goods such as food and textile products. In trade within the clothing industry, countries such as Italy tend to specialize in the production of higher-valued clothing, while lower-income countries such as Portugal specialize in the production of cheaper clothing. In sparkling wines, France specializes in the higher-quality end of the spectrum, while Spain specializes in the low-quality end. Similarly, Germany specializes in the production of higher-quality cars while Spain specializes in lower-quality vehicles. Similar exchanges occur across a wide range of goods and services.

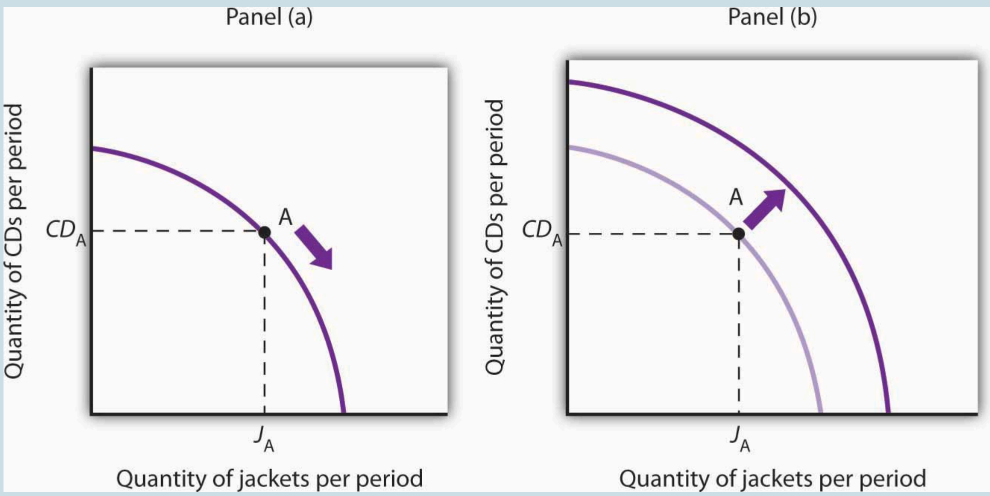
Diaz Mora found that comparative advantage tended to correspond to income levels. Countries in the northern part of the European Union tend to have high per capita incomes and high levels of human capital and technology—these countries gained by specializing in the production of high-valued goods. Countries in the southern part of the Union also gained by specialization—in the production of low-valued goods. This specialization has increased the welfare of people throughout the Union.

Sources: Carmen Diaz Mora, "The Role of Comparative Advantage in Trade Within Industries: A Panel Data Approach for the European Union," *Weltwirtschaftliches Archiv* 138:2 (2002), 291–316.

Answer to Try It! Problem

Your first production possibilities curve should resemble the one in Panel (a). Starting at point A, an increase in jacket production requires a move down and to the right along the curve, as shown by the arrow, and thus a reduction in the production of CD players. Alternatively, if there is economic growth, it shifts the production possibilities curve outward, as in Panel (b). This shift allows an increase in production of both goods, as suggested by the arrow.

Figure 2.17



2.4 Review and Practice

Summary

Economics deals with choices. In this chapter we have examined more carefully the range of choices in production that must be made in any economy. In particular, we looked at choices involving the allocation of an economy's factors of production: labour, capital, and natural resources.

In addition, in any economy, the level of technology plays a key role in determining how productive the factors of production will be. In a market economy, entrepreneurs organize factors of production and act to introduce technological change.

The production possibilities model is a device that assists us in thinking about many of the choices about resource allocation in an economy. The model assumes that the economy has factors of production that are fixed in both quantity and quality. When illustrated graphically, the production possibilities model typically limits our analysis to two goods. Given the economy's factors of production and technology, the economy can produce various combinations of the two goods. If it uses its factors of production efficiently and has full employment, it will be operating on the production possibilities curve.

Two characteristics of the production possibilities curve are particularly important. First, it is downward sloping. This reflects the scarcity of the factors of production available to the economy; producing more of one good requires giving up some of the other. Second, the curve is bowed out. Another way of saying this is to say that the curve gets steeper as we move from left to right; the absolute value of its slope is increasing. Producing each additional unit of the good on the horizontal axis requires a greater sacrifice of the good on the vertical axis than did the previous units produced. This fact, called the law of increasing opportunity cost, is the inevitable result of efficient choices in production—choices based on comparative advantage.

The production possibilities model has important implications for international trade. It suggests that free trade will allow countries to specialize in the production of goods and services in which they have a comparative advantage. This specialization increases the production of all goods and services.

Increasing the quantity or quality of factors of production and/or improving technology will shift the production possibilities curve outward. This process is called economic growth. In the last 50 years, economic growth in the United States has resulted chiefly from increases in human capital and from technological advance.

Choices concerning the use of scarce resources take place within the context of a set of institutional arrangements that define an economic system. The principal distinctions between systems lie in the degree to which ownership of capital and natural resources and decision making authority over scarce resources are held by government or by private individuals. Economic systems include market capitalist, mixed, and command socialist economies. An increasing body of evidence suggests that market capitalist economies tend to be most productive; many command socialist and mixed economies are moving in the direction of market capitalist systems.

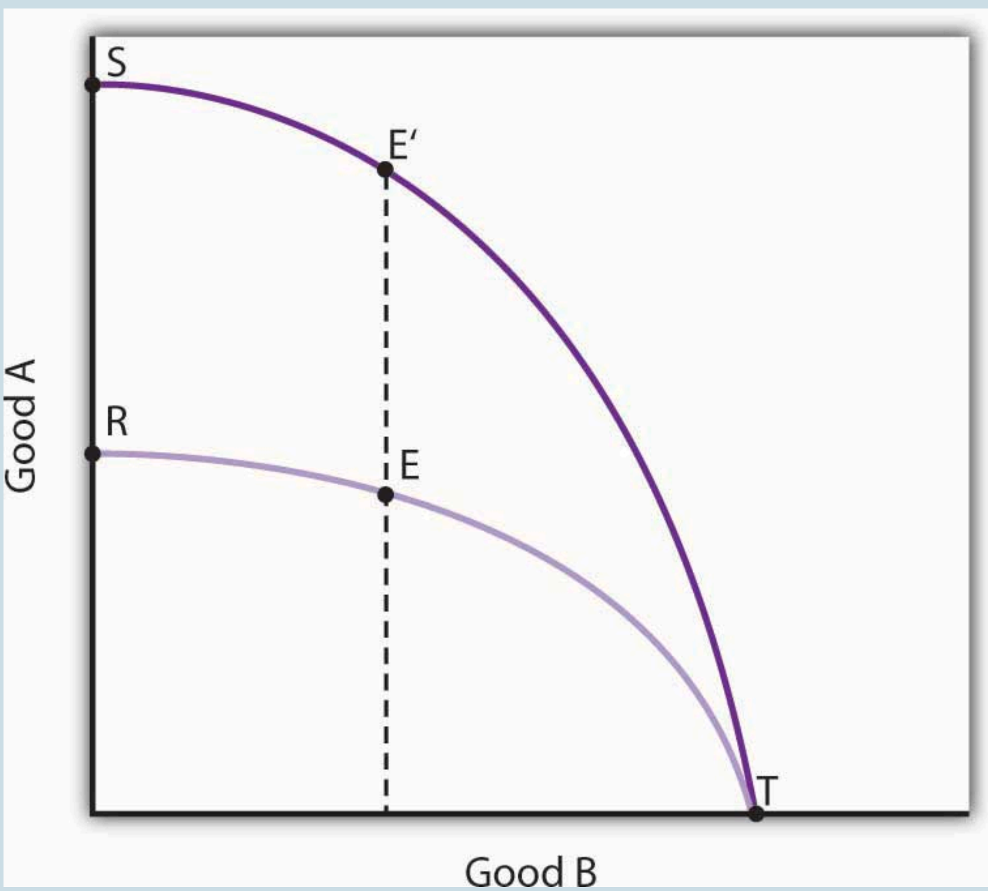
The presumption in favour of market-based systems does not preclude a role for government. Government is

necessary to provide the system of laws on which market systems are founded. It may also be used to provide certain goods and services, to help individuals in need, and to regulate the actions of individuals and firms.

Concept Problems

1. How does a college education increase one's human capital?
2. Why does the downward-sloping production possibilities curve imply that factors of production are scarce?
3. In what ways are the bowed-out shape of the production possibilities curve and the law of increasing opportunity cost related?
4. Suppose an economy can produce two goods, A and B. It is now operating at point E on production possibilities curve RT. An improvement in the technology available to produce good A shifts the curve to ST, and the economy selects point E′. How does this change affect the opportunity cost of producing an additional unit of good B?

Figure 2.18



6. Could a nation's production possibilities curve ever shift inward? Explain what such a shift would mean, and discuss events that might cause such a shift to occur.
7. Suppose blue-eyed people were banned from working. How would this affect a nation's production possibilities curve?
8. Evaluate this statement: "The Canadian economy could achieve greater growth by devoting fewer resources to consumption and more to investment; it follows that such a shift would be desirable."
9. Two countries, Sportsland and Foodland, have similar total quantities of labour, capital, and natural resources. Both can produce two goods, figs and footballs. Sportsland's resources are particularly well suited to the production of footballs but are not very productive in producing figs. Foodland's resources are very productive when used for figs but are not capable of producing many footballs. In which country is the cost of additional footballs generally greater? Explain.
10. Suppose a country is committed to using its resources based on the reverse of comparative advantage doctrine: it first transfers those resources for which the cost is greatest, not lowest. Describe this country's production possibilities curve.
11. The U.S. Constitution bans states from restricting imports of goods and services from other states. Suppose this restriction did not exist and that states were allowed to limit imports of goods and services produced in other states. How do you think this would affect U.S. output? Explain.
12. By 1993, nations in the European Union (EU) had eliminated all barriers to the flow of goods, services, labor, and capital across their borders. Even such things as consumer protection laws and the types of plugs required to plug in appliances have been standardized to ensure that there will be no barriers to trade. How do you think this elimination of trade barriers affected EU output?
13. How did the technological changes described in the Case in Point "Technology Cuts Costs, Boosts Productivity and Profits" affect the production possibilities curve for the United States?

Numerical Problems

1. Nathan can mow four lawns in a day or plant 20 trees in a day.

    1. Draw Nathan's production possibilities curve for mowing lawns and planting trees. Assume the production possibilities curve is linear and put the quantity of lawns mowed per day on the horizontal axis and the quantity of trees planted per day on the vertical axis.
    2. What is Nathan's opportunity cost of planting trees?
    3. What is Nathan's opportunity cost of mowing lawns?

2. David can mow four lawns in a day or plant four trees in a day.

    1. Draw David's production possibilities curve for mowing lawns and planting trees. Again, assume a linear production possibilities curve and put the quantity of lawns mowed per day on the horizontal axis.
    2. What is David's opportunity cost of planting trees?

3. What is David's opportunity cost of mowing lawns?

3. Given the production information in problems 1 and 2 above, who has the comparative advantage in planting trees? Mowing lawns?
4. The exhibits below describe the production possibilities for Germany and Turkey.

Figure 2.19

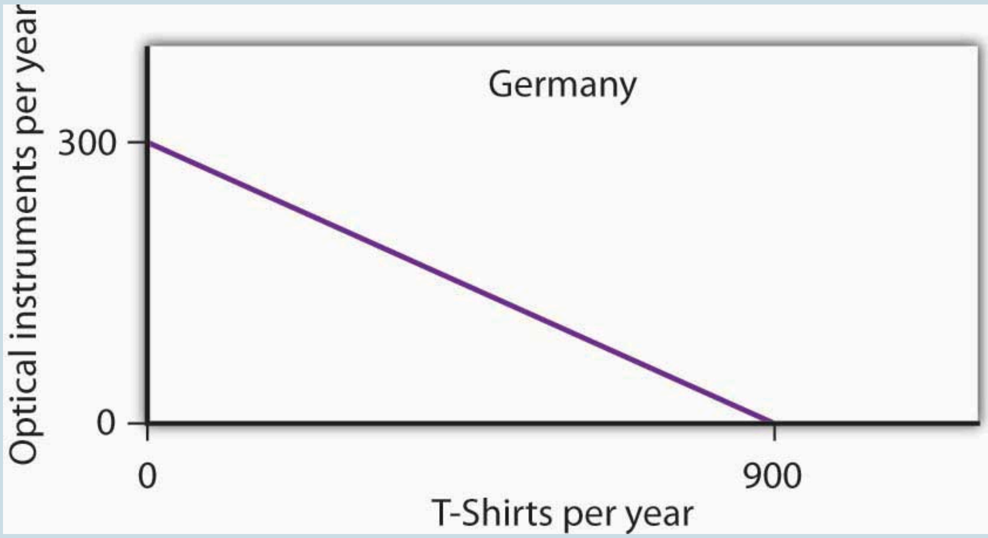
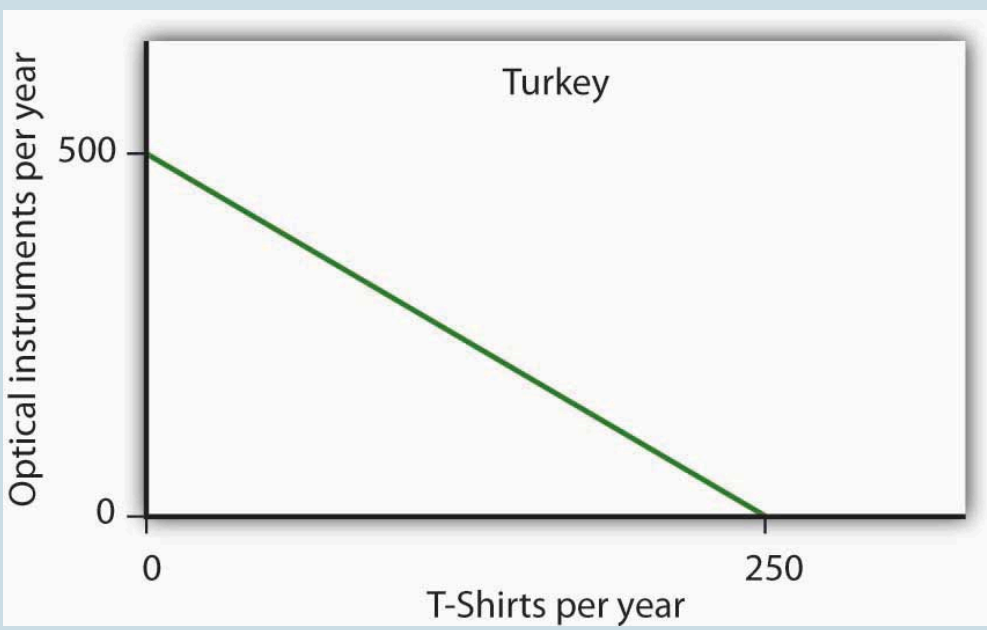


Figure 2.20



1. What is the slope of Germany's production possibilities curve?
2. What is the slope of Turkey's production possibilities curve?

3. What is the opportunity cost of producing T-shirts in Germany?
4. What is the opportunity cost of producing T-shirts in Turkey?
5. What is the opportunity cost of producing optical instruments in Germany?
6. What is the opportunity cost of producing optical instruments in Turkey?
7. In which good does Germany have a comparative advantage?
8. In which good does Turkey have a comparative advantage?

5. The nation of Leisureland can produce two goods, bicycles and bowling balls. The western region of Leisureland can, if it devotes all its resources to bicycle production, produce 100 bicycles per month. Alternatively, it could devote all its resources to bowling balls and produce 400 per month—or it could produce any combination of bicycles and bowling balls lying on a straight line between these two extremes.

    1. Draw a production possibilities curve for western Leisureland (with bicycles on the vertical axis).
    2. What it is the opportunity cost of producing an additional bowling ball measured in terms of forgone bicycles in western Leisureland?
    3. Suppose that eastern Leisureland can, if it devotes all its resources to the production of bicycles, produce 400. If it devotes all its resources to bowling ball production, though, it can produce only 100. Draw the production possibilities curve for eastern Leisureland (again, assume it is linear and put bicycles on the vertical axis).
    4. What is the opportunity cost of producing an additional bowling ball measured in terms of forgone bicycles in eastern Leisureland?
    5. Explain the difference in opportunity cost between western and eastern Leisureland. Which region has a comparative advantage in producing bowling balls? Bicycles?
    6. Draw the production possibilities curve for Leisureland, one that combines the curves for western and eastern Leisureland.
    7. Suppose it is determined that 400 bicycles must be produced. How many bowling balls can be produced?
    8. Where will these goods be produced?

6. The table below shows the production possibilities schedule for an economy.

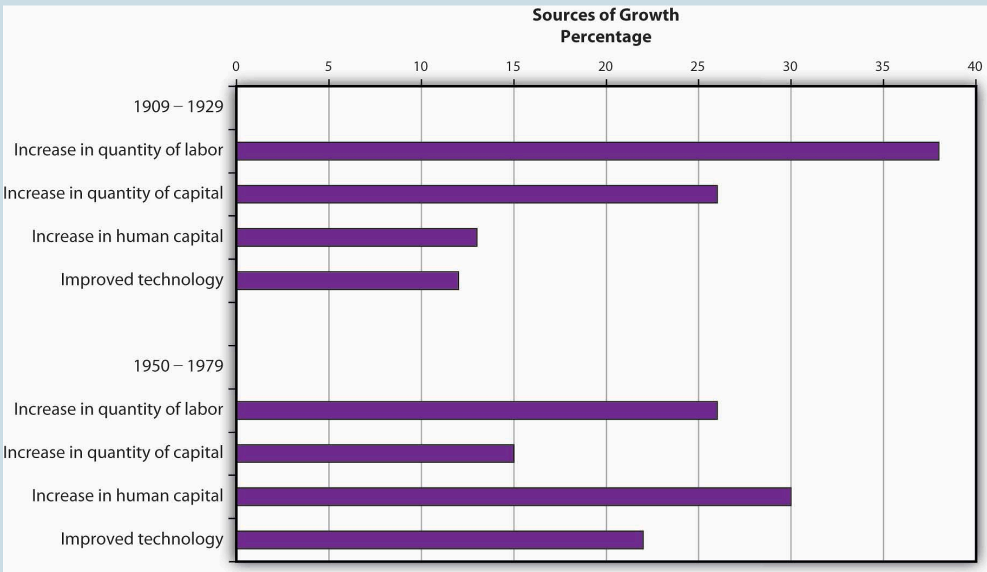
Production Alternatives	Capital goods per period	Consumer goods per period
A	0	40
B	1	36
C	2	28
D	3	16
E	4	0

    1. Putting capital goods per period on the horizontal axis and consumer goods per period on the vertical axis, graph the production possibilities curve for the economy.
    2. If the economy is producing at alternative B, what is the opportunity cost to it of producing at alternative C instead?
    3. If the economy is producing at alternative C, what is the opportunity cost to it of producing at alternative D instead?

4. Is it possible for this economy to produce 30 units of consumer goods per period while producing 1 unit of capital goods? Would this combination of goods represent efficient or inefficient production? Explain.
5. Which point, B or C, would lead to higher economic growth? Explain your answer.

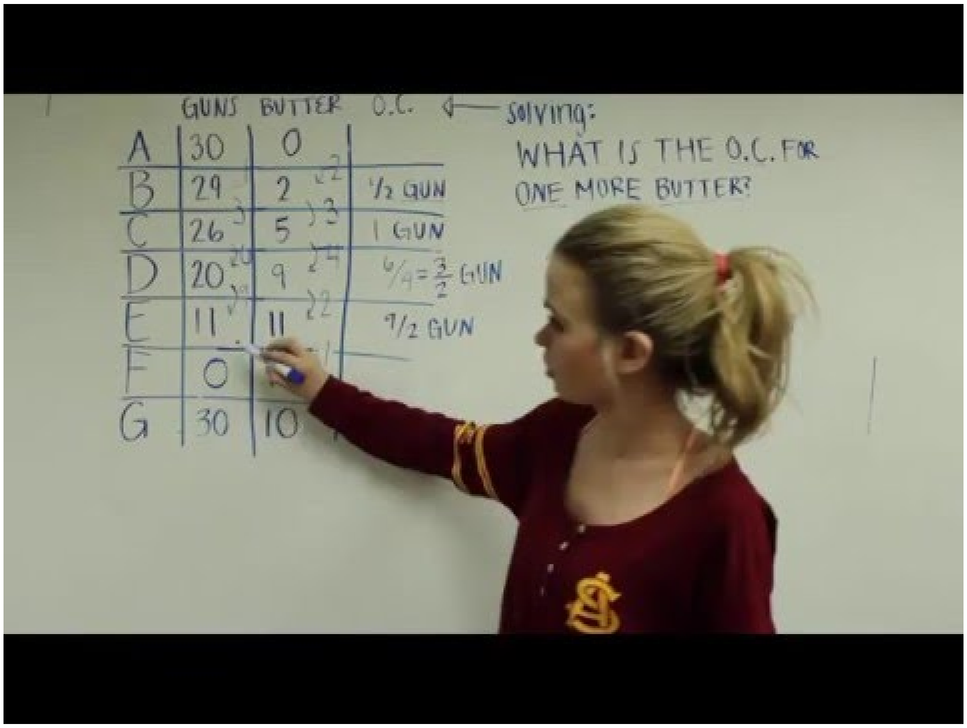
7. The exhibit below shows the sources of growth in the United States between 1909 and 1929 and between 1950 and 1979, according to a study by Edward Denison (Denison, E., 2962). (Note: The sources of economic growth are cumulative and, taken collectively, explain 100% of total growth over the period.)

Figure 2.21



1. Approximately what percentage of U.S. growth between 1909 and 1929 was due to increases in quantities of factors of production?
2. Approximately what percentage of U.S. growth between 1909 and 1929 was due to increases in quality of factors of production and technological improvement?
3. Approximately what percentage of U.S. growth between 1950 and 1979 was due to increases in quantities of factors of production?
4. Approximately what percentage of U.S. growth between 1950 and 1979 was due to increases in quality of factors of production and technological improvement?

More Practice



A YouTube element has been excluded from this version of the text. You can view it online here: https://pressbooks.senecacollege.ca/macroeconomics/?p=58

References

Denison, E., *The Sources of Economic Growth in the United States* (New York: Committee for Economic Development, 1962) and Edward Denison, *Trends in American Growth* 1929–1982 (Washington, D.C.: Brookings Institutions, 1985).

CHAPTER 3: DEMAND AND SUPPLY

Start Up: Crazy for Coffee

Starbucks Coffee Company revolutionized the coffee-drinking habits of millions of people. Starbucks, whose bright green-and-white logo is almost as familiar as the golden arches of McDonald's, began in Seattle in 1971. Fifteen years later it had grown into a chain of four stores in the Seattle area. Then in 1987 Howard Schultz, a former Starbucks employee, who had become enamored with the culture of Italian coffee bars during a trip to Italy, bought the company from its founders for $3.8 million. In 2008, Americans were willingly paying $3 or more for a cappuccino or a latté, and Starbuck's had grown to become an international chain, with over 16,000 stores around the world.

The change in consumers' taste for coffee and the profits raked in by Starbucks lured other companies to get into the game. Retailers such as Seattle's Best Coffee and Gloria Jean's Coffees entered the US market, and today there are thousands of coffee bars, carts, drive-throughs, and kiosks in downtowns, malls, and airports all around the country. Even McDonald's began selling specialty coffees.

But over the last decade the price of coffee beans has been quite volatile. Just as consumers were growing accustomed to their cappuccinos and lattés, in 1997, the price of coffee beans shot up. Excessive rain and labor strikes in coffee-growing areas of South America had reduced the supply of coffee, leading to a rise in its price. In the early 2000s, Vietnam flooded the market with coffee, and the price of coffee beans plummeted. More recently, weather conditions in various coffee-growing countries reduced supply, and the price of coffee beans went back up.

Markets, the institutions that bring together buyers and sellers, are always responding to events, such as bad harvests and changing consumer tastes that affect the prices and quantities of particular goods. The demand for some goods increases, while the demand for others decreases. The supply of some goods rises, while the supply of others falls. As such events unfold, prices adjust to keep markets in balance. This chapter explains how the market forces of demand and supply interact to determine equilibrium prices and equilibrium quantities of goods and services. We will see how prices and quantities adjust to changes in demand and supply and how changes in prices serve as signals to buyers and sellers.

The model of demand and supply that we shall develop in this chapter is one of the most powerful tools in all of economic analysis. You will be using it throughout your study of economics. We will first look at the variables that influence demand. Then we will turn to supply, and finally we will put demand and supply together to explore how the model of demand and supply operates. As we examine the model, bear in mind that demand is a representation of the behavior of buyers and that supply is a representation of the behavior of sellers. Buyers may be consumers purchasing groceries or producers purchasing iron ore to make steel. Sellers may be firms selling cars or households selling their labor services. We shall see that the ideas of demand and supply apply, whatever the identity of the buyers or sellers and whatever the good or service being exchanged in the market. In this chapter, we shall focus on buyers and sellers of goods and services.

3.1 Demand

Learning Objectives

1. Define the quantity demanded of a good or service and illustrate it using a demand schedule and a demand curve.
2. Distinguish between the following pairs of concepts: demand and quantity demanded, demand schedule and demand curve, movement along and shift in a demand curve.
3. Identify demand shifters and determine whether a change in a demand shifter causes the demand curve to shift to the right or to the left.

How many pizzas will people eat this year? How many doctor visits will people make? How many houses will people buy?

Each good or service has its own special characteristics that determine the quantity people are willing and able to consume. One is the price of the good or service itself. Other independent variables that are important determinants of demand include consumer preferences, prices of related goods and services, income, demographic characteristics such as population size, and buyer expectations. The number of pizzas people will purchase, for example, depends very much on whether they like pizza. It also depends on the prices for alternatives such as hamburgers or spaghetti. The number of doctor visits is likely to vary with income—people with higher incomes are likely to see a doctor more often than people with lower incomes. The demands for pizza, for doctor visits, and for housing are certainly affected by the age distribution of the population and its size.

While different variables play different roles in influencing the demands for different goods and services, economists pay special attention to one: the price of the good or service. Given the values of all the other variables that affect demand, a higher price tends to reduce the quantity people demand, and a lower price tends to increase it. A medium pizza typically sells for $5 to $10. Suppose the price were $30. Chances are, you would buy fewer pizzas at that price than you do now. Suppose pizzas typically sold for $2 each. At that price, people would be likely to buy more pizzas than they do now.

We will discuss first how price affects the quantity demanded of a good or service and then how other variables affect demand.

Price and the Demand Curve

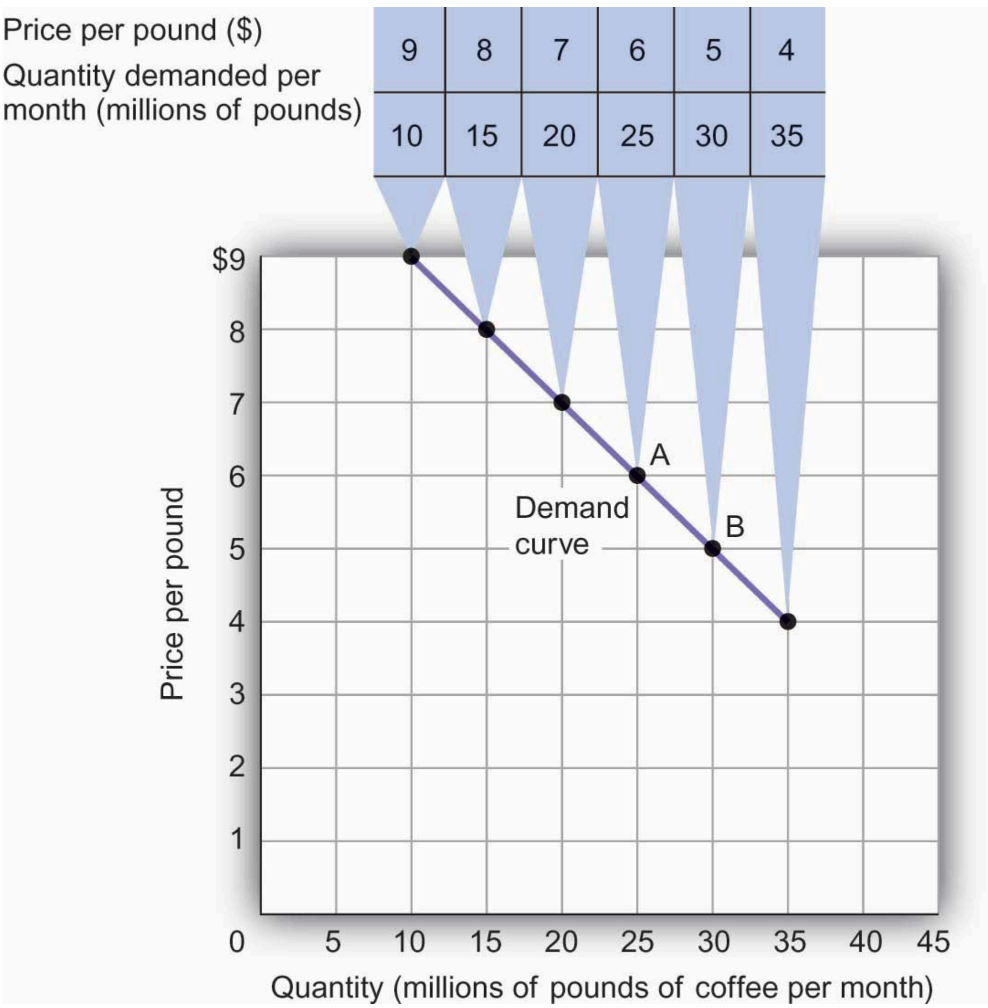
Because people will purchase different quantities of a good or service at different prices, economists must be careful when speaking of the "demand" for something. They have therefore developed some specific terms for expressing the general concept of demand.

The quantity demanded of a good or service is the quantity buyers are willing and able to buy at a particular price during a particular period, all other things unchanged. (As we learned, we can substitute the Latin phrase "ceteris paribus" for "all other things unchanged.") Suppose, for example, that 100,000 movie tickets are sold each month in a particular town at a price of $8 per ticket. That quantity—100,000—is the quantity of movie admissions demanded per month at a price of

$8. If the price were $12, we would expect the quantity demanded to be less. If it were $4, we would expect the quantity demanded to be greater. The quantity demanded at each price would be different if other things that might affect it, such as the population of the town, were to change. That is why we add the qualifier that other things have not changed to the definition of quantity demanded.

A demand schedule is a table that shows the quantities of a good or service demanded at different prices during a particular period, all other things unchanged. To introduce the concept of a demand schedule, let us consider the demand for coffee in the United States. We will ignore differences among types of coffee beans and roasts, and speak simply of coffee. The table in Figure 3.1 "A Demand Schedule and a Demand Curve" shows quantities of coffee that will be demanded each month at prices ranging from $9 to $4 per pound; the table is a demand schedule. We see that the higher the price, the lower the quantity demanded.

Figure 3.1 A Demand Schedule and a Demand Curve



The table is a demand schedule; it shows quantities of coffee demanded per month in the United States at particular prices, all other things unchanged. These data are then plotted on the demand curve. At point A on the curve, 25 million pounds of coffee per month are demanded at a price of $6 per pound. At point B, 30 million pounds of coffee per month are demanded at a price of $5 per pound.

The information given in a demand schedule can be presented with a demand curve, which is a graphical representation of a demand schedule. A demand curve thus shows the relationship between the price and quantity demanded of a good or service during a particular period, all other things unchanged. The demand curve in Figure 3.1 "A Demand Schedule and a Demand Curve" shows the prices and quantities of coffee demanded that are given in the demand schedule. At point A, for example, we see that 25 million pounds of coffee per month are demanded at a price of $6 per pound. By convention, economists graph price on the vertical axis and quantity on the horizontal axis.

Price alone does not determine the quantity of coffee or any other good that people buy. To isolate the effect of changes in price on the quantity of a good or service demanded, however, we show the quantity demanded at each price, assuming that those other variables remain unchanged. We do the same thing in drawing a graph of the relationship between any two variables; we assume that the values of other variables that may affect the variables shown in the graph (such as income or population) remain unchanged for the period under consideration.

A change in price, with no change in any of the other variables that affect demand, results in a movement *along* the demand curve. For example, if the price of coffee falls from $6 to $5 per pound, consumption rises from 25 million pounds to 30 million pounds per month. That is a movement from point A to point B along the demand curve in Figure 3.1 "A Demand Schedule and a Demand Curve". A movement along a demand curve that results from a change in price is called a change in quantity demanded. Note that a change in quantity demanded is not a change or shift in the demand curve; it is a movement *along* the demand curve.

The negative slope of the demand curve in Figure 3.1 "A Demand Schedule and a Demand Curve" suggests a key behavioral relationship of economics. All other things unchanged, the law of demand holds that, for virtually all goods and services, a higher price leads to a reduction in quantity demanded and a lower price leads to an increase in quantity demanded.

The law of demand is called a law because the results of countless studies are consistent with it. Undoubtedly, you have observed one manifestation of the law. When a store finds itself with an overstock of some item, such as running shoes or tomatoes, and needs to sell these items quickly, what does it do? It typically has a sale, expecting that a lower price will increase the quantity demanded. In general, we expect the law of demand to hold. Given the values of other variables that influence demand, a higher price reduces the quantity demanded. A lower price increases the quantity demanded. Demand curves, in short, slope downward.

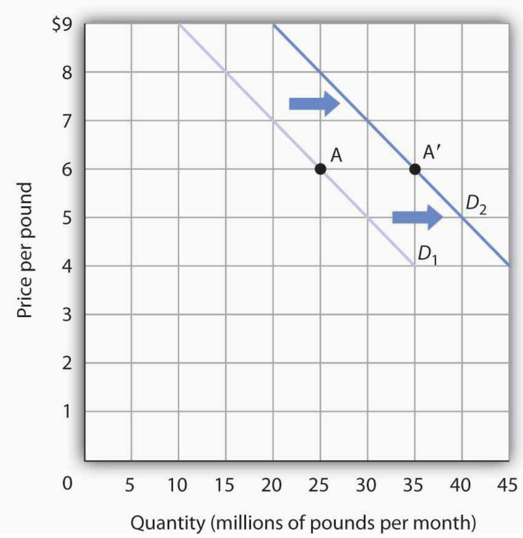
Changes in Demand

Of course, price alone does not determine the quantity of a good or service that people consume. Coffee consumption, for example, will be affected by such variables as income and population. Preferences also play a role. The story at the beginning of the chapter illustrates as much. Starbucks "turned people on" to coffee. We also expect other prices to affect coffee consumption. People often eat doughnuts or bagels with their coffee, so a reduction in the price of doughnuts or bagels might induce people to drink more coffee. An alternative to coffee is tea, so a reduction in the price of tea might result in the consumption of more tea and less coffee. Thus, a change in any one of the variables held constant in constructing a demand schedule will change the quantities demanded at each price. The result will be a *shift* in the entire demand curve rather than a movement along the demand curve. A *shift* in a demand curve is called a change in demand.

Suppose, for example, that something happens to increase the quantity of coffee demanded at each price. Several events could produce such a change: an increase in incomes, an increase in population, or an increase in the price of tea would each be likely to increase the quantity of coffee demanded at each price. Any such change produces a new demand schedule. Figure 3.2 "An Increase in Demand" shows such a change in the demand schedule for coffee. We see that the quantity of coffee demanded per month is greater at each price than before. We show that graphically as a shift in the demand curve. The original curve, labeled D_1, shifts to the right to D_2. At a price of $6 per pound, for example, the quantity demanded rises from 25 million pounds per month (point A) to 35 million pounds per month (point A′).

Figure 3.2 An Increase in Demand

Price	Old quantity demanded	New quantity demanded
$9	10	20
8	15	25
7	20	30
6	25	35
5	30	40
4	35	45



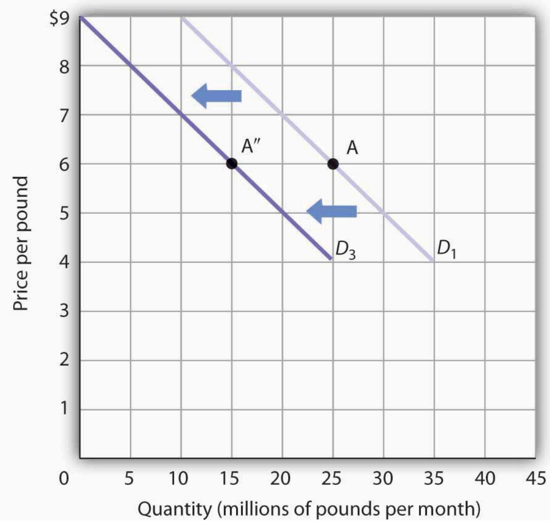
An increase in the quantity of a good or service demanded at each price is shown as an increase in demand. Here, the original demand curve D_1 shifts to D_2. Point A on D_1 corresponds to a price of $6 per pound and a quantity demanded of 25 million pounds of coffee per month. On the new demand curve D_2, the quantity demanded at this price rises to 35 million pounds of coffee per month (point A′).

Just as demand can increase, it can decrease. In the case of coffee, demand might fall as a result of events such as a reduction in population, a reduction in the price of tea, or a change in preferences. For example, a definitive finding that the caffeine in coffee contributes to heart disease, which is currently being debated in the scientific community, could change preferences and reduce the demand for coffee.

A reduction in the demand for coffee is illustrated in Figure 3.3 "A Reduction in Demand". The demand schedule shows that less coffee is demanded at each price than in Figure 3.1 "A Demand Schedule and a Demand Curve". The result is a shift in demand from the original curve D_1 to D_3. The quantity of coffee demanded at a price of $6 per pound falls from 25 million pounds per month (point A) to 15 million pounds per month (point A″). Note, again, that a change in quantity demanded, ceteris paribus, refers to a movement *along* the demand curve, while a change in demand refers to a *shift* in the demand curve.

Figure 3.3 A Reduction in Demand

Price	Old quantity demanded	New quantity demanded
$9	10	0
8	15	5
7	20	10
6	25	15
5	30	20
4	35	25



A reduction in demand occurs when the quantities of a good or service demanded fall at each price. Here, the demand schedule shows a lower quantity of coffee demanded at each price than we had in Figure 3.1 "A Demand Schedule and a Demand Curve". The reduction shifts the demand curve for coffee to D_3 from D_1. The quantity demanded at a price of $6 per pound, for example, falls from 25 million pounds per month (point A) to 15 million pounds of coffee per month (point A″).

A variable that can change the quantity of a good or service demanded at each price is called a demand shifter. When these other variables change, the all-other-things-unchanged conditions behind the original demand curve no longer hold. Although different goods and services will have different demand shifters, the demand shifters are likely to include (1) consumer preferences, (2) the prices of related goods and services, (3) income, (4) demographic characteristics, and (5) buyer expectations. Next we look at each of these.

Preferences

Changes in preferences of buyers can have important consequences for demand. We have already seen how Starbucks supposedly increased the demand for coffee. Another example is reduced demand for cigarettes caused by concern about the effect of smoking on health. A change in preferences that makes one good or service more popular will shift the demand curve to the right. A change that makes it less popular will shift the demand curve to the left.

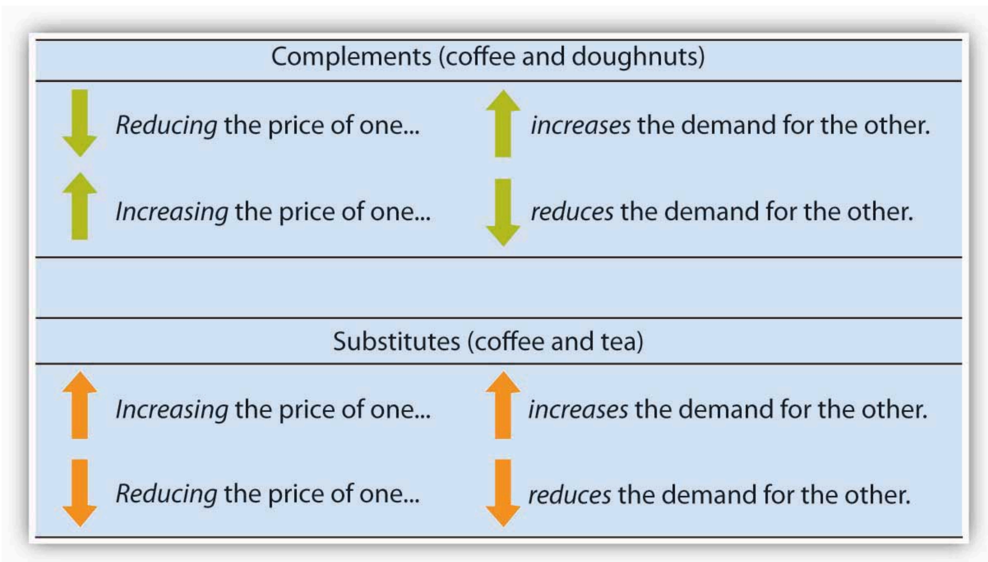
Prices of Related Goods and Services

Suppose the price of doughnuts were to fall. Many people who drink coffee enjoy dunking doughnuts in their coffee; the lower price of doughnuts might therefore increase the demand for coffee, shifting the demand curve for coffee to the right. A lower price for tea, however, would be likely to reduce coffee demand, shifting the demand curve for coffee to the left.

In general, if a reduction in the price of one good increases the demand for another, the two goods are called complements. If a reduction in the price of one good reduces the demand for another, the two goods are called substitutes. These definitions hold in reverse as well: two goods are complements if an increase in the price of one reduces the demand for the other, and they are substitutes if an increase in the price of one increases the demand for the other. Doughnuts and coffee are complements; tea and coffee are substitutes.

Complementary goods are goods used in conjunction with one another. Tennis rackets and tennis balls, eggs and bacon, and stationery and postage stamps are complementary goods. Substitute goods are goods used instead of one another. iPODs, for example, are likely to be substitutes for CD players. Breakfast cereal is a substitute for eggs. A file attachment to an e-mail is a substitute for both a fax machine and postage stamps.

Figure 3.4



Income

As incomes rise, people increase their consumption of many goods and services, and as incomes fall, their consumption of these goods and services falls. For example, an increase in income is likely to raise the demand for gasoline, ski trips, new cars, and jewelry. There are, however, goods and services for which consumption falls as income rises—and rises as income falls. As incomes rise, for example, people tend to consume more fresh fruit but less canned fruit.

A good for which demand increases when income increases is called a normal good. A good for which demand decreases when income increases is called an inferior good. An increase in income shifts the demand curve for fresh fruit (a normal good) to the right; it shifts the demand curve for canned fruit (an inferior good) to the left.

Demographic Characteristics

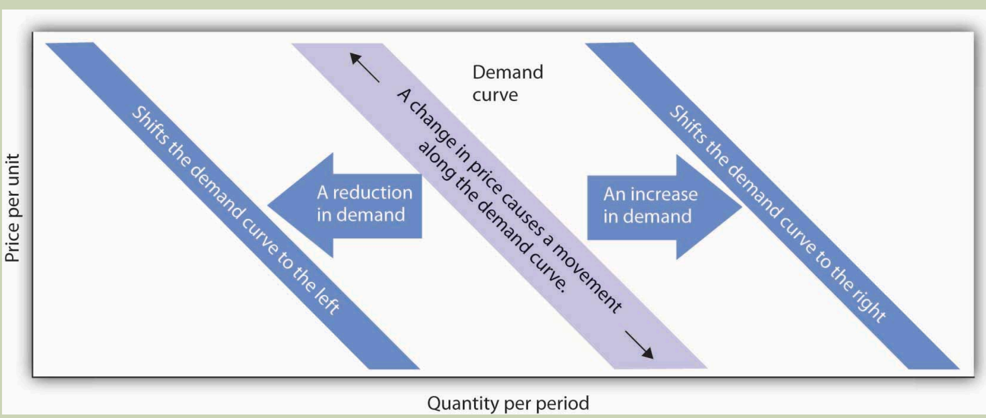
The number of buyers affects the total quantity of a good or service that will be bought; in general, the greater the population, the greater the demand. Other demographic characteristics can affect demand as well. As the share of the population over age 65 increases, the demand for medical services, ocean cruises, and motor homes increases. The birth rate in the United States fell sharply between 1955 and 1975 but has gradually increased since then. That increase has raised the demand for such things as infant supplies, elementary school teachers, soccer coaches, in-line skates, and college education. Demand can thus shift as a result of changes in both the number and characteristics of buyers.

Buyer Expectations

The consumption of goods that can be easily stored, or whose consumption can be postponed, is strongly affected by buyer expectations. The expectation of newer TV technologies, such as high-definition TV, could slow down sales of regular TVs. If people expect gasoline prices to rise tomorrow, they will fill up their tanks today to try to beat the price increase. The same will be true for goods such as automobiles and washing machines: an expectation of higher prices in the future will lead to more purchases today. If the price of a good is expected to fall, however, people are likely to reduce their purchases today and await tomorrow's lower prices. The expectation that computer prices will fall, for example, can reduce current demand.

Heads Up!

Figure 3.5



It is crucial to distinguish between a change in quantity demanded, which is a movement along the demand curve caused by a change in price, and a change in demand, which implies a shift of the demand curve itself. A change in demand is caused by a change in a demand shifter. An increase in demand is a shift of the demand curve to the right. A decrease in demand is a shift in the demand curve to the left. This drawing of a demand curve highlights the difference.

Key Takeaways

- The quantity demanded of a good or service is the quantity buyers are willing and able to buy at a particular price during a particular period, all other things unchanged.
- A demand schedule is a table that shows the quantities of a good or service demanded at different prices during a particular period, all other things unchanged.
- A demand curve shows graphically the quantities of a good or service demanded at different prices during a particular period, all other things unchanged.
- All other things unchanged, the law of demand holds that, for virtually all goods and services, a higher price induces a reduction in quantity demanded and a lower price induces an increase in quantity demanded.
- A change in the price of a good or service causes a change in the quantity demanded—a movement *along* the demand curve.
- A change in a demand shifter causes a change in demand, which is shown as a *shift* of the demand curve. Demand shifters include preferences, the prices of related goods and services, income, demographic characteristics, and buyer expectations.
- Two goods are substitutes if an increase in the price of one causes an increase in the demand for the other. Two goods are complements if an increase in the price of one causes a decrease in the demand for the other.
- A good is a normal good if an increase in income causes an increase in demand. A good is an inferior good if an increase in income causes a decrease in demand.

Try It!

All other things unchanged, what happens to the demand curve for DVD rentals if there is (a) an increase in the price of movie theater tickets, (b) a decrease in family income, or (c) an increase in the price of DVD rentals? In answering this and other "Try It!" problems in this chapter, draw and carefully label a set of axes. On the horizontal axis of your graph, show the quantity of DVD rentals. It is necessary to specify the time period to which your quantity pertains (e.g., "per period," "per week," or "per year"). On the vertical axis show the price per DVD rental. Since you do not have specific data on prices and quantities demanded, make a "free-hand" drawing of the curve or curves you are asked to examine. Focus on the general shape and position of the curve(s) before and after events occur. Draw new curve(s) to show what happens in each of the circumstances given. The curves could shift to the left or to the right, or stay where they are.

Case in Point: Solving Campus Parking Problems Without Adding More Parking Spaces

Figure 3.6

Alden Jewell – The Parking Lot – CC BY 2.0.

Unless you attend a "virtual" campus, chances are you have engaged in more than one conversation about how hard it is to find a place to park on campus. Indeed, according to Clark Kerr, a former president of the University of California system, a university is best understood as a group of people "held together by a common grievance over parking."

Clearly, the demand for campus parking spaces has grown substantially over the past few decades. In surveys conducted by Daniel Kenney, Ricardo Dumont, and Ginger Kenney, who work for the campus design company Sasaki and Associates, it was found that 7 out of 10 students own their own cars. They have interviewed "many students who confessed to driving from their dormitories to classes that were a five-minute walk away," and they argue that the deterioration of college environments is largely attributable to the increased use of cars on campus and that colleges could better service their missions by not adding more parking spaces.

Since few universities charge enough for parking to even cover the cost of building and maintaining parking lots, the rest is paid for by all students as part of tuition. Their research shows that "for every 1,000 parking spaces, the median institution loses almost $400,000 a year for surface parking, and more than $1,200,000 for structural parking." Fear of a backlash from students and their parents, as well as from faculty and staff, seems to explain why campus administrators do not simply raise the price of parking on campus.

While Kenney and his colleagues do advocate raising parking fees, if not all at once then over time, they also suggest some subtler, and perhaps politically more palatable, measures—in particular, shifting the demand for parking spaces to the left by lowering the prices of substitutes.

Two examples they noted were at the University of Washington and the University of Colorado at Boulder. At the University of Washington, car poolers may park for free. This innovation has reduced purchases of single-occupancy parking permits by 32% over a decade. According to University of Washington assistant director of transportation services Peter Dewey, "Without vigorously managing our parking and providing commuter alternatives, the university would have been faced with adding approximately 3,600 parking spaces, at a cost of over $100 million…The university has created opportunities to make capital investments in buildings supporting education instead of structures for cars." At the University of Colorado, free public transit has increased use of buses and light rail from 300,000 to 2 million trips per year over the last decade. The increased use of mass transit has allowed the university to avoid constructing nearly 2,000 parking spaces, which has saved about $3.6 million annually.

Sources: Daniel R. Kenney, "How to Solve Campus Parking Problems Without Adding More Parking," *The Chronicle of Higher Education*, March 26, 2004, Section B, pp. B22-B23.

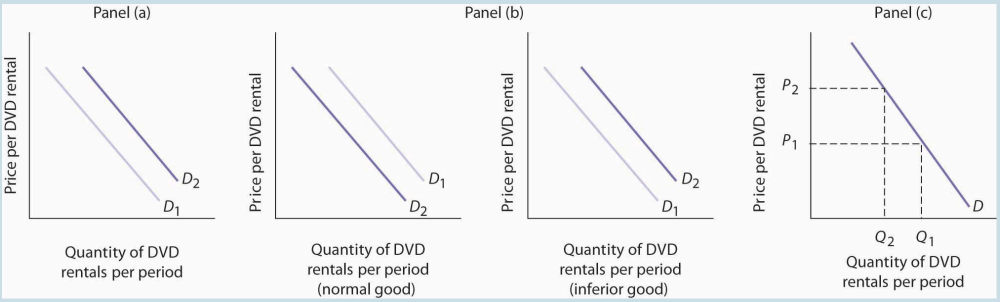
Answer to Try It! Problem

Since going to the movies is a substitute for watching a DVD at home, an increase in the price of going to the movies should cause more people to switch from going to the movies to staying at home and renting DVDs. Thus, the demand curve for DVD rentals will shift to the right when the price of movie theater tickets increases [Panel (a)].

A decrease in family income will cause the demand curve to shift to the left if DVD rentals are a normal good but to the right if DVD rentals are an inferior good. The latter may be the case for some families, since staying at home and watching DVDs is a cheaper form of entertainment than taking the family to the movies. For most others, however, DVD rentals are probably a normal good [Panel (b)].

An increase in the price of DVD rentals does not shift the demand curve for DVD rentals at all; rather, an increase in price, say from P_1 to P_2, is a movement upward to the left along the demand curve. At a higher price, people will rent fewer DVDs, say Q_2 instead of Q_1, ceteris paribus [Panel (c)].

Figure 3.7



Panel (a)
Price per DVD rental
D_2
D_1
Quantity of DVD rentals per period
Panel (b)
Price per DVD rental
D_1
D_2
Quantity of DVD rentals per period (normal good)
Price per DVD rental
D_2
D_1
Quantity of DVD rentals per period (inferior good)
Panel (c)
Price per DVD rental
P_2
P_1
D
Q_2 Q_1
Quantity of DVD rentals per period

3.2 Supply

Learning Objectives

1. Define the quantity supplied of a good or service and illustrate it using a supply schedule and a supply curve.
2. Distinguish between the following pairs of concepts: supply and quantity supplied, supply schedule and supply curve, movement along and shift in a supply curve.
3. Identify supply shifters and determine whether a change in a supply shifter causes the supply curve to shift to the right or to the left.

What determines the quantity of a good or service sellers are willing to offer for sale? Price is one factor; ceteris paribus, a higher price is likely to induce sellers to offer a greater quantity of a good or service. Production cost is another determinant of supply. Variables that affect production cost include the prices of factors used to produce the good or service, returns from alternative activities, technology, the expectations of sellers, and natural events such as weather changes. Still another factor affecting the quantity of a good that will be offered for sale is the number of sellers—the greater the number of sellers of a particular good or service, the greater will be the quantity offered at any price per time period.

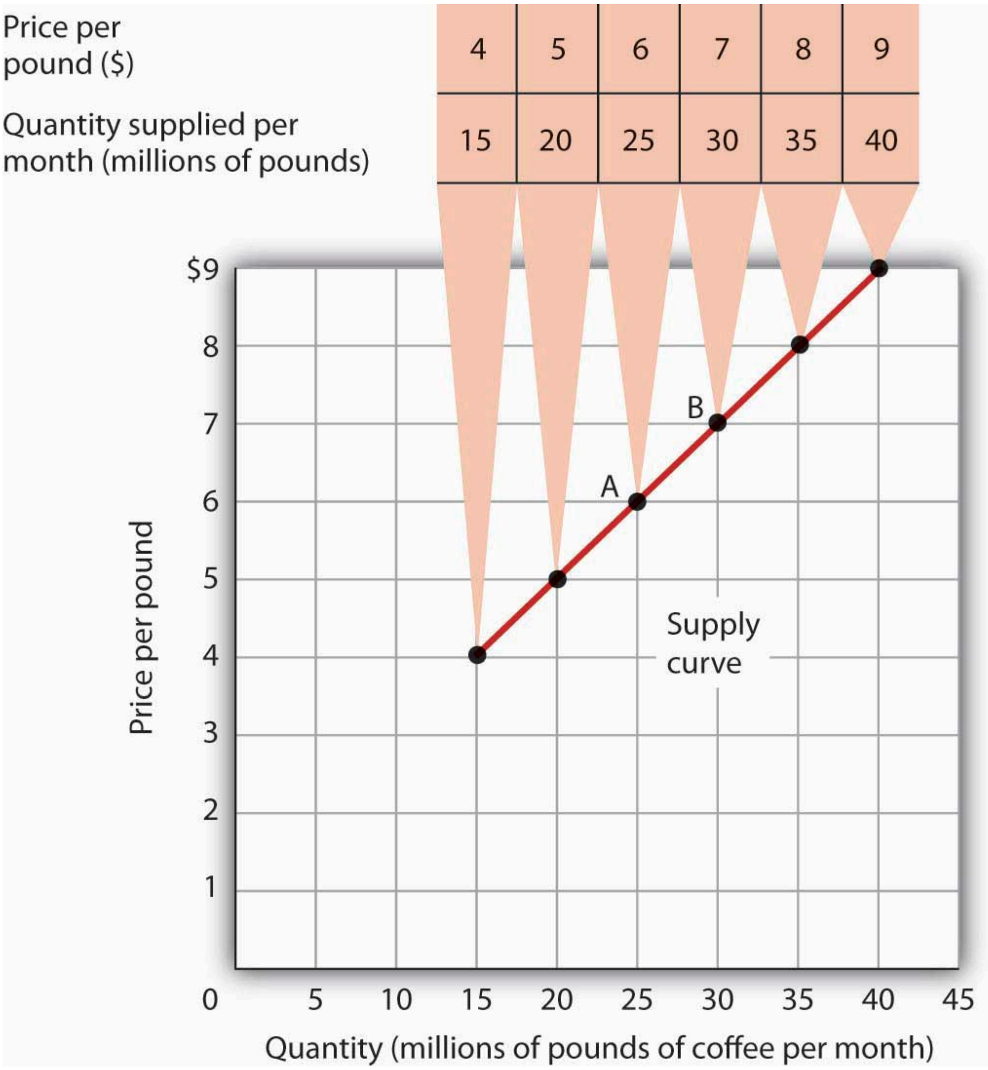
Price and the Supply Curve

The quantity supplied of a good or service is the quantity sellers are willing to sell at a particular price during a particular period, all other things unchanged. Ceteris paribus, the receipt of a higher price increases profits and induces sellers to increase the quantity they supply.

In general, when there are many sellers of a good, an increase in price results in an increase in quantity supplied, and this relationship is often referred to as the law of supply. We will see, though, through our exploration of microeconomics, that there are a number of exceptions to this relationship. There are cases in which a higher price will not induce an increase in quantity supplied. Goods that cannot be produced, such as additional land on the corner of Park Avenue and 56th Street in Manhattan, are fixed in supply—a higher price cannot induce an increase in the quantity supplied. There are even cases, which we investigate in microeconomic analysis, in which a higher price induces a reduction in the quantity supplied.

Generally speaking, however, when there are many sellers of a good, an increase in price results in a greater quantity supplied. The relationship between price and quantity supplied is suggested in a supply schedule, a table that shows quantities supplied at different prices during a particular period, all other things unchanged. Figure 3.8 "A Supply Schedule and a Supply Curve" gives a supply schedule for the quantities of coffee that will be supplied per month at various prices, ceteris paribus. At a price of $4 per pound, for example, producers are willing to supply 15 million pounds of coffee per month. A higher price, say $6 per pound, induces sellers to supply a greater quantity—25 million pounds of coffee per month.

Figure 3.8 A Supply Schedule and a Supply Curve



The supply schedule shows the quantity of coffee that will be supplied in the United States each month at particular prices, all other things unchanged. The same information is given graphically in the supply curve. The values given here suggest a positive relationship between price and quantity supplied.

A supply curve is a graphical representation of a supply schedule. It shows the relationship between price and quantity supplied during a particular period, all other things unchanged. Because the relationship between price and quantity supplied is generally positive, supply curves are generally upward sloping. The supply curve for coffee in Figure 3.8 "A Supply Schedule and a Supply Curve" shows graphically the values given in the supply schedule.

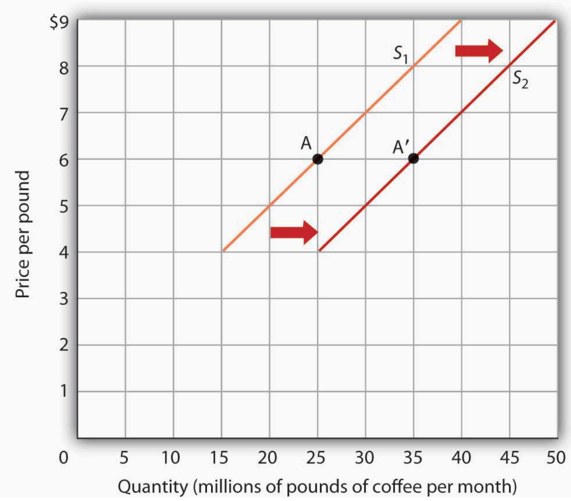
A change in price causes a movement *along* the supply curve; such a movement is called a change in quantity supplied. As is the case with a change in quantity demanded, a change in quantity supplied does not shift the supply curve. By definition, it is a movement along the supply curve. For example, if the price rises from $6 per pound to $7 per pound, the quantity supplied rises from 25 million pounds per month to 30 million pounds per month. That's a movement from point A to point B along the supply curve in Figure 3.8 "A Supply Schedule and a Supply Curve".

Changes in Supply

When we draw a supply curve, we assume that other variables that affect the willingness of sellers to supply a good or service are unchanged. It follows that a change in any of those variables will cause a change in supply, which is a shift in the supply curve. A change that increases the quantity of a good or service supplied at each price shifts the supply curve to the right. Suppose, for example, that the price of fertilizer falls. That will reduce the cost of producing coffee and thus increase the quantity of coffee producers will offer for sale at each price. The supply schedule in Figure 3.9 "An Increase in Supply" shows an increase in the quantity of coffee supplied at each price. We show that increase graphically as a shift in the supply curve from S_1 to S_2. We see that the quantity supplied at each price increases by 10 million pounds of coffee per month. At point A on the original supply curve S_1, for example, 25 million pounds of coffee per month are supplied at a price of $6 per pound. After the increase in supply, 35 million pounds per month are supplied at the same price (point A′ on curve S_2).

Figure 3.9 An Increase in Supply

Price	Old quantity supplied	New quantity supplied
$4	15	25
5	20	30
6	25	35
7	30	40
8	35	45
9	40	50

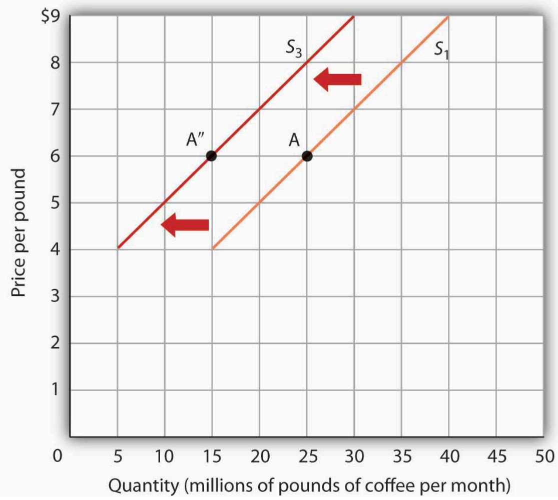


If there is a change in supply that increases the quantity supplied at each price, as is the case in the supply schedule here, the supply curve shifts to the right. At a price of $6 per pound, for example, the quantity supplied rises from the previous level of 25 million pounds per month on supply curve S_1 (point A) to 35 million pounds per month on supply curve S_2 (point A′).

An event that reduces the quantity supplied at each price shifts the supply curve to the left. An increase in production costs and excessive rain that reduces the yields from coffee plants are examples of events that might reduce supply. Figure 3.10 "A Reduction in Supply" shows a reduction in the supply of coffee. We see in the supply schedule that the quantity of coffee supplied falls by 10 million pounds of coffee per month at each price. The supply curve thus shifts from S_1 to S_3.

Figure 3.10 A Reduction in Supply

Price	Old quantity supplied	New quantity supplied
$4	15	5
5	20	10
6	25	15
7	30	20
8	35	25
9	40	30



A change in supply that reduces the quantity supplied at each price shifts the supply curve to the left. At a price of $6 per pound, for example, the original quantity supplied was 25 million pounds of coffee per month (point A). With a new supply curve S_3, the quantity supplied at that price falls to 15 million pounds of coffee per month (point A″).

A variable that can change the quantity of a good or service supplied at each price is called a supply shifter. Supply shifters include (1) prices of factors of production, (2) returns from alternative activities, (3) technology, (4) seller expectations, (5) natural events, and (6) the number of sellers. When these other variables change, the all-other-things-unchanged conditions behind the original supply curve no longer hold. Let us look at each of the supply shifters.

Prices of Factors of Production

A change in the price of labor or some other factor of production will change the cost of producing any given quantity of the good or service. This change in the cost of production will change the quantity that suppliers are willing to offer at any price. An increase in factor prices should decrease the quantity suppliers will offer at any price, shifting the supply curve to the left. A reduction in factor prices increases the quantity suppliers will offer at any price, shifting the supply curve to the right.

Suppose coffee growers must pay a higher wage to the workers they hire to harvest coffee or must pay more for fertilizer. Such increases in production cost will cause them to produce a smaller quantity at each price, shifting the supply curve for coffee to the left. A reduction in any of these costs increases supply, shifting the supply curve to the right.

Returns from Alternative Activities

To produce one good or service means forgoing the production of another. The concept of opportunity cost in economics suggests that the value of the activity forgone is the opportunity cost of the activity chosen; this cost should affect supply. For example, one opportunity cost of producing eggs is not selling chickens. An increase in the price people are willing to pay for fresh chicken would make it more profitable to sell chickens and would thus increase the opportunity cost of producing eggs. It would shift the supply curve for eggs to the left, reflecting a decrease in supply.

Technology

A change in technology alters the combinations of inputs or the types of inputs required in the production process. An improvement in technology usually means that fewer and/or less costly inputs are needed. If the cost of production is lower, the profits available at a given price will increase, and producers will produce more. With more produced at every price, the supply curve will shift to the right, meaning an increase in supply.

Impressive technological changes have occurred in the computer industry in recent years. Computers are much smaller and are far more powerful than they were only a few years ago—and they are much cheaper to produce. The result has been a huge increase in the supply of computers, shifting the supply curve to the right.

While we usually think of technology as enhancing production, declines in production due to problems in technology are also possible. Outlawing the use of certain equipment without pollution-control devices has increased the cost of production for many goods and services, thereby reducing profits available at any price and shifting these supply curves to the left.

Seller Expectations

All supply curves are based in part on seller expectations about future market conditions. Many decisions about production and selling are typically made long before a product is ready for sale. Those decisions necessarily depend on expectations. Changes in seller expectations can have important effects on price and quantity.

Consider, for example, the owners of oil deposits. Oil pumped out of the ground and used today will be unavailable in the future. If a change in the international political climate leads many owners to expect that oil prices will rise in the future, they may decide to leave their oil in the ground, planning to sell it later when the price is higher. Thus, there will be a decrease in supply; the supply curve for oil will shift to the left.

Natural Events

Storms, insect infestations, and drought affect agricultural production and thus the supply of agricultural goods. If something destroys a substantial part of an agricultural crop, the supply curve will shift to the left. The terrible cyclone that killed more than 50,000 people in Myanmar in 2008 also destroyed some of the country's prime rice growing land. That shifted the supply curve for rice to the left. If there is an unusually good harvest, the supply curve will shift to the right.

The Number of Sellers

The supply curve for an industry, such as coffee, includes all the sellers in the industry. A change in the number of sellers in an industry changes the quantity available at each price and thus changes supply. An increase in the number of sellers supplying a good or service shifts the supply curve to the right; a reduction in the number of sellers shifts the supply curve to the left.

The market for cellular phone service has been affected by an increase in the number of firms offering the service. Over the past decade, new cellular phone companies emerged, shifting the supply curve for cellular phone service to the right.

Heads Up!

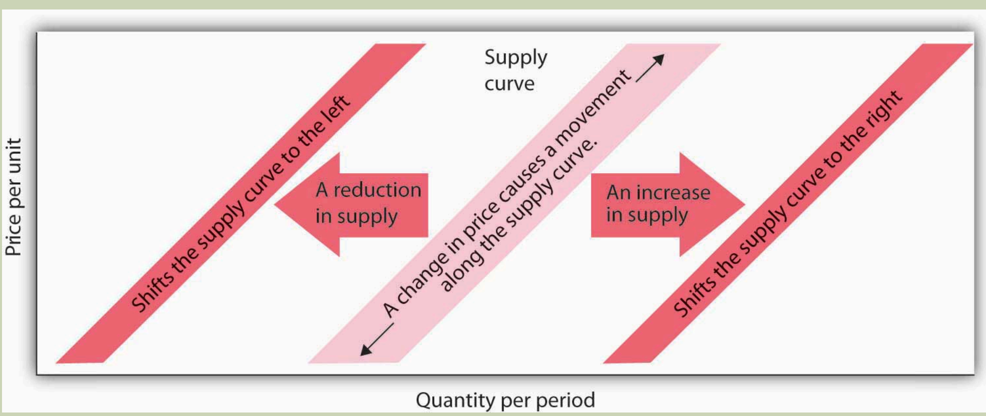
There are two special things to note about supply curves. The first is similar to the Heads Up! on demand curves: it is important to distinguish carefully between changes in supply and changes in quantity supplied. A change in supply results from a change in a supply shifter and implies a shift of the supply curve to the right or left. A change in price produces a change in quantity supplied and induces a movement along the supply curve. A change in price does not shift the supply curve.

The second caution relates to the interpretation of increases and decreases in supply. Notice that in [Figure 3.9 "An Increase in Supply"](#) an increase in supply is shown as a shift of the supply curve to the right; the curve shifts in the direction of increasing quantity with respect to the horizontal axis. In [Figure 3.10 "A Reduction in Supply"](#) a reduction in supply is shown as a shift of the supply curve to the left; the curve shifts in the direction of decreasing quantity with respect to the horizontal axis.

Because the supply curve is upward sloping, a shift to the right produces a new curve that in a sense lies "below" the original curve. Students sometimes make the mistake of thinking of such a shift as a shift "down" and therefore as a reduction in supply. Similarly, it is easy to make the mistake of showing an increase in supply with a new curve that lies "above" the original curve. But that is a reduction in supply!

To avoid such errors, focus on the fact that an increase in supply is an increase in the quantity supplied at each price and shifts the supply curve in the direction of increased quantity on the horizontal axis. Similarly, a reduction in supply is a reduction in the quantity supplied at each price and shifts the supply curve in the direction of a lower quantity on the horizontal axis.

Figure 3.11



Key Takeaways

- The quantity supplied of a good or service is the quantity sellers are willing to sell at a particular price during a particular period, all other things unchanged.
- A supply schedule shows the quantities supplied at different prices during a particular period, all other things unchanged. A supply curve shows this same information graphically.
- A change in the price of a good or service causes a change in the quantity supplied—a movement *along* the supply curve.
- A change in a supply shifter causes a change in supply, which is shown as a *shift* of the supply curve. Supply shifters include prices of factors of production, returns from alternative activities, technology, seller expectations, natural events, and the number of sellers.
- An increase in supply is shown as a shift to the right of a supply curve; a decrease in supply is shown as a shift to the left.

Try It!

If all other things are unchanged, what happens to the supply curve for DVD rentals if there is (a) an increase in wages paid to DVD rental store clerks, (b) an increase in the price of DVD rentals, or (c) an increase in the number of DVD rental stores? Draw a graph that shows what happens to the supply curve in each circumstance. The supply curve can shift to the left or to the right, or stay where it is. Remember to label the axes and curves, and remember to specify the time period (e.g., “DVDs rented per week”).

Case in Point: The Monks of St. Benedict’s Get Out of the Egg Business

Figure 3.12

Randy OHC – Untitled – CC BY 2.0.

It was cookies that lured the monks of St. Benedict's out of the egg business, and now private retreat sponsorship is luring them away from cookies.

St. Benedict's is a Benedictine monastery, nestled on a ranch high in the Colorado Rockies, about 20 miles down the road from Aspen. The monastery's 15 monks operate the ranch to support themselves and to provide help for poor people in the area. They lease out about 3,500 acres of their land to cattle and sheep grazers, produce cookies, and sponsor private retreats. They used to produce eggs.

Attracted by potential profits and the peaceful nature of the work, the monks went into the egg business in 1967. They had 10,000 chickens producing their Monastery Eggs brand. For a while, business was good. Very good. Then, in the late 1970s, the price of chicken feed started to rise rapidly.

"When we started in the business, we were paying $60 to $80 a ton for feed—delivered," recalls the monastery's abbot, Father Joseph Boyle. "By the late 1970s, our cost had more than doubled. We were paying $160 to $200 a ton. That really hurt, because feed represents a large part of the cost of producing eggs."

The monks adjusted to the blow. "When grain prices were lower, we'd pull a hen off for a few weeks to molt, then return her to laying. After grain prices went up, it was 12 months of laying and into the soup pot," Father Joseph says.

Grain prices continued to rise in the 1980s and increased the costs of production for all egg producers. It caused the supply of eggs to fall. Demand fell at the same time, as Americans worried about the cholesterol in eggs. Times got tougher in the egg business.

"We were still making money in the financial sense," Father Joseph says. "But we tried an experiment in 1985 producing cookies, and it was a success. We finally decided that devoting our time and energy to the cookies would pay off better than the egg business, so we quit the egg business in 1986."

The mail-order cookie business was good to the monks. They sold 200,000 ounces of Monastery Cookies in 1987.

By 1998, however, they had limited their production of cookies, selling only locally and to gift shops. Since 2000, they have switched to "providing private retreats for individuals and groups—about 40 people per month," according to Brother Charles.

The monks' calculation of their opportunity costs revealed that they would earn a higher return through sponsorship of private retreats than in either cookies or eggs. This projection has proved correct.

And there is another advantage as well.

"The chickens didn't stop laying eggs on Sunday," Father Joseph chuckles. "When we shifted to cookies we could take Sundays off. We weren't hemmed in the way we were with the chickens." The move to providing retreats is even better in this regard. Since guests provide their own meals, most of the monastery's effort goes into planning and scheduling, which frees up even more of their time for other worldly as well as spiritual pursuits.

Source: Personal interviews.

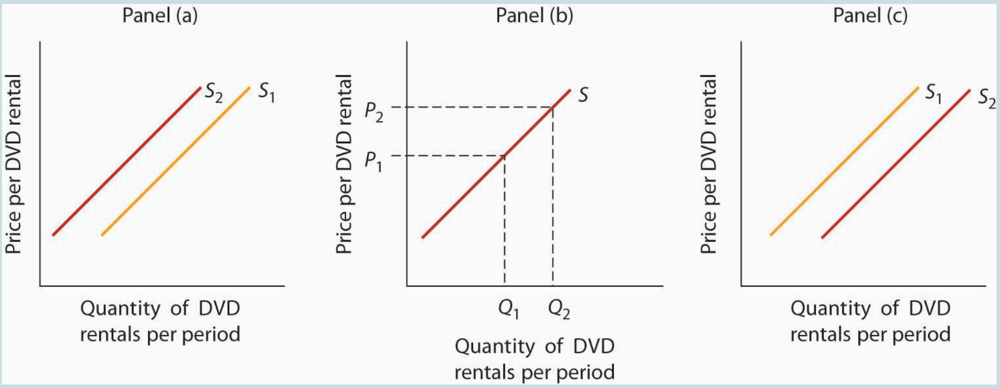
Answer to Try It! Problem

DVD rental store clerks are a factor of production in the DVD rental market. An increase in their wages raises the cost of production, thereby causing the supply curve of DVD rentals to shift to the left [Panel (a)]. (*Caution*: It is possible that you thought of the wage increase as an increase in income, a demand shifter, that would lead to an increase in demand, but this would be incorrect. The question refers only to wages of DVD rental store clerks. They may rent some DVD, but their impact on total demand would be negligible. Besides, we have no information on what has happened overall to incomes of people who rent DVDs. We do know, however, that the cost of a factor of production, which is a supply shifter, increased.)

An increase in the price of DVD rentals does not shift the supply curve at all; rather, it corresponds to a movement upward to the right along the supply curve. At a higher price of P_2 instead of P_1, a greater quantity of DVD rentals, say Q_2 instead of Q_1, will be supplied [Panel (b)].

An increase in the number of stores renting DVDs will cause the supply curve to shift to the right [Panel (c)].

Figure 3.13



Panel (a)
Price per DVD rental
S_2
S_1
Quantity of DVD rentals per period
Panel (b)
Price per DVD rental
P_2
P_1
S
Q_1
Q_2
Quantity of DVD rentals per period
Panel (c)
Price per DVD rental
S_1
S_2
Quantity of DVD rentals per period

3.3 Demand, Supply, and Equilibrium

Learning Objectives

1. Use demand and supply to explain how equilibrium price and quantity are determined in a market.
2. Understand the concepts of surpluses and shortages and the pressures on price they generate.
3. Explain the impact of a change in demand or supply on equilibrium price and quantity.
4. Explain how the circular flow model provides an overview of demand and supply in product and factor markets and how the model suggests ways in which these markets are linked.

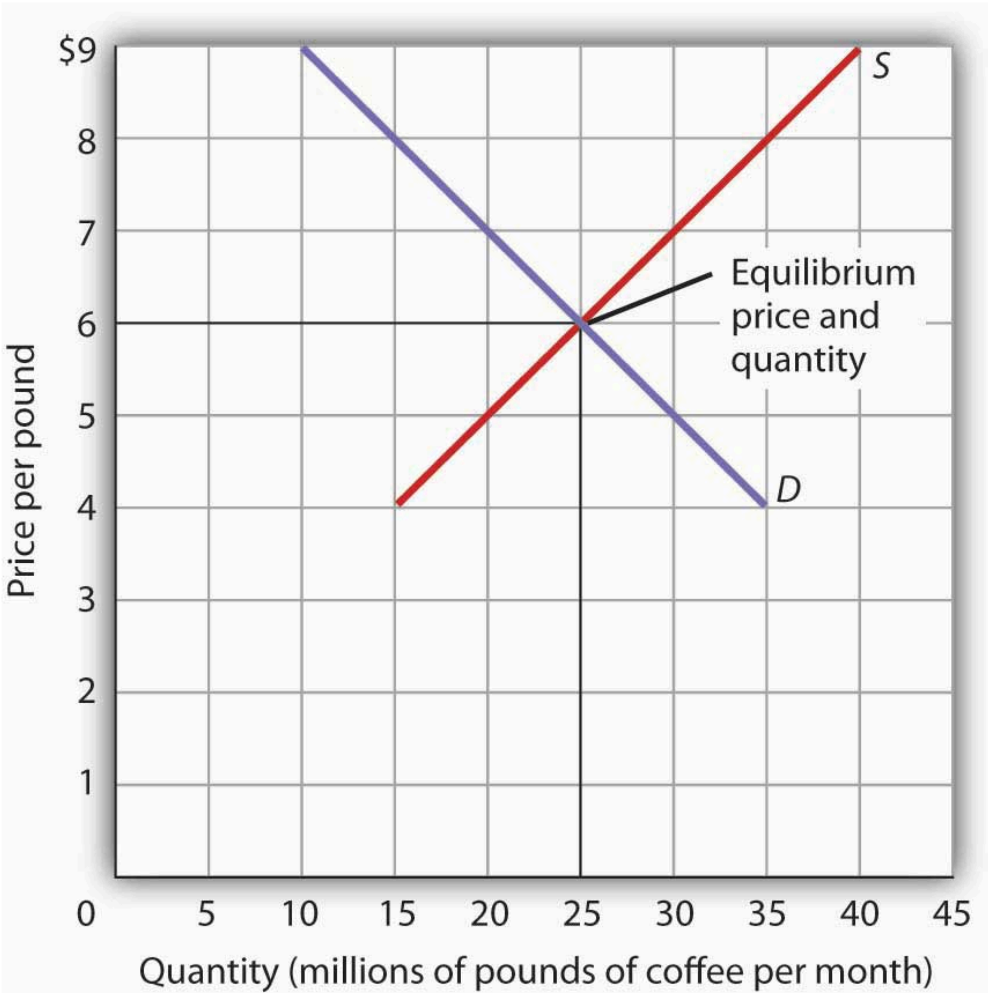
In this section we combine the demand and supply curves we have just studied into a new model. The model of demand and supply uses demand and supply curves to explain the determination of price and quantity in a market.

The Determination of Price and Quantity

The logic of the model of demand and supply is simple. The demand curve shows the quantities of a particular good or service that buyers will be willing and able to purchase at each price during a specified period. The supply curve shows the quantities that sellers will offer for sale at each price during that same period. By putting the two curves together, we should be able to find a price at which the quantity buyers are willing and able to purchase equals the quantity sellers will offer for sale.

Figure 3.14 “The Determination of Equilibrium Price and Quantity” combines the demand and supply data introduced in Figure 3.1 “A Demand Schedule and a Demand Curve” and Figure 3.8 “A Supply Schedule and a Supply Curve” Notice that the two curves intersect at a price of $6 per pound—at this price the quantities demanded and supplied are equal. Buyers want to purchase, and sellers are willing to offer for sale, 25 million pounds of coffee per month. The market for coffee is in equilibrium. Unless the demand or supply curve shifts, there will be no tendency for price to change. The equilibrium price in any market is the price at which quantity demanded equals quantity supplied. The equilibrium price in the market for coffee is thus $6 per pound. The equilibrium quantity is the quantity demanded and supplied at the equilibrium price.

Figure 3.14 The Determination of Equilibrium Price and Quantity



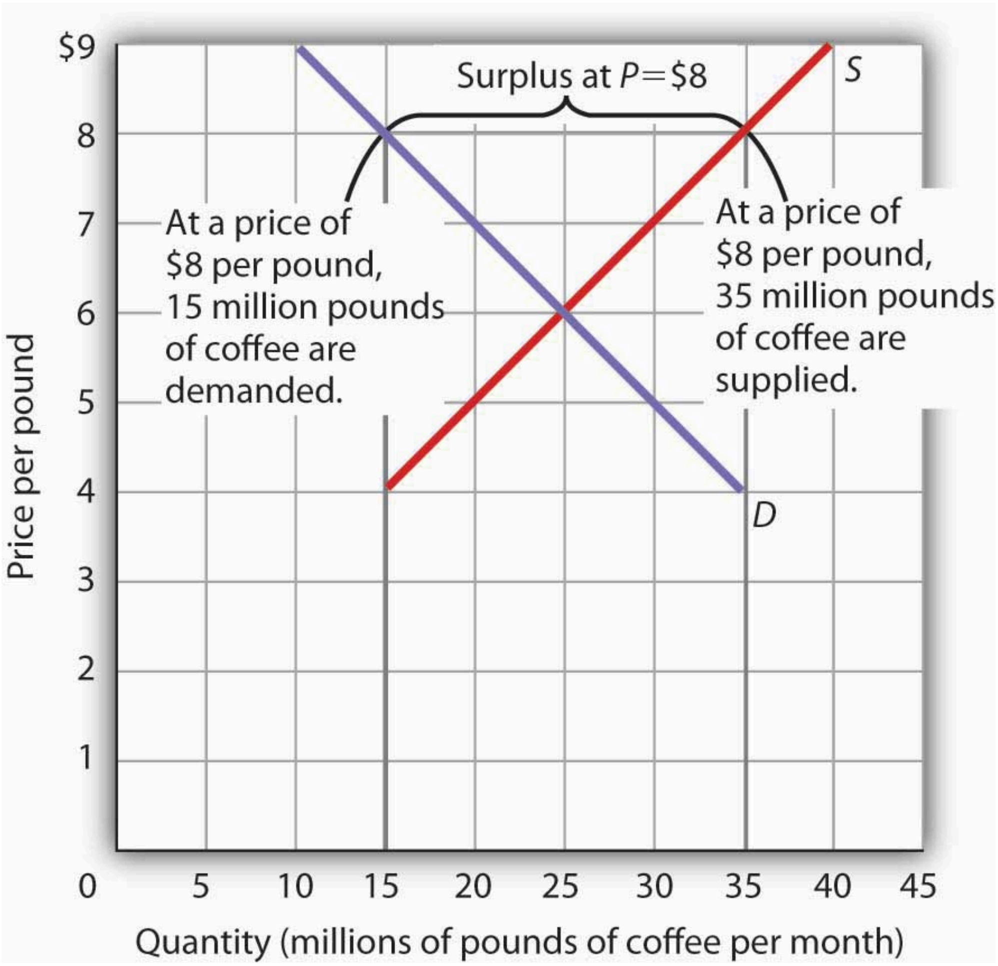
When we combine the demand and supply curves for a good in a single graph, the point at which they intersect identifies the equilibrium price and equilibrium quantity. Here, the equilibrium price is $6 per pound. Consumers demand, and suppliers supply, 25 million pounds of coffee per month at this price.

With an upward-sloping supply curve and a downward-sloping demand curve, there is only a single price at which the two curves intersect. This means there is only one price at which equilibrium is achieved. It follows that at any price other than the equilibrium price, the market will not be in equilibrium. We next examine what happens at prices other than the equilibrium price.

Surpluses

Figure 3.15 "A Surplus in the Market for Coffee" shows the same demand and supply curves we have just examined, but this time the initial price is $8 per pound of coffee. Because we no longer have a balance between quantity demanded and quantity supplied, this price is not the equilibrium price. At a price of $8, we read over to the demand curve to determine the quantity of coffee consumers will be willing to buy—15 million pounds per month. The supply curve tells us what sellers will offer for sale—35 million pounds per month. The difference, 20 million pounds of coffee per month, is called a surplus. More generally, a surplus is the amount by which the quantity supplied exceeds the quantity demanded at the current price. There is, of course, no surplus at the equilibrium price; a surplus occurs only if the current price exceeds the equilibrium price.

Figure 3.15 A Surplus in the Market for Coffee



At a price of $8, the quantity supplied is 35 million pounds of coffee per month and the quantity demanded is 15 million pounds per month; there is a surplus of 20 million pounds of coffee per month. Given a surplus, the price will fall quickly toward the equilibrium level of $6.

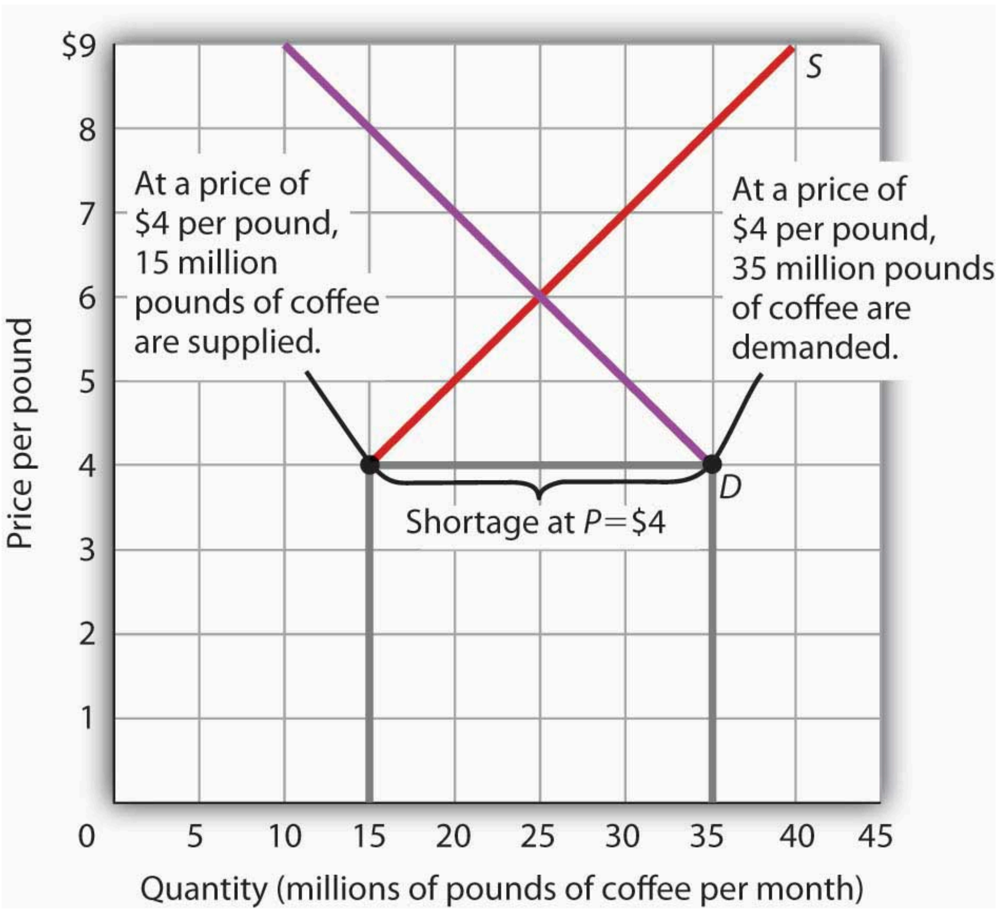
A surplus in the market for coffee will not last long. With unsold coffee on the market, sellers will begin to reduce their prices to clear out unsold coffee. As the price of coffee begins to fall, the quantity of coffee supplied begins to decline. At the same time, the quantity of coffee demanded begins to rise. Remember that the reduction in quantity supplied is a movement *along* the supply curve—the curve itself does not shift in response to a reduction in price. Similarly, the increase in quantity demanded is a movement *along* the demand curve—the demand curve does not shift in response to a reduction in price. Price will continue to fall until it reaches its equilibrium level, at which the demand and supply curves intersect. At that point, there will be no tendency for price to fall further. In general, surpluses in the marketplace are short-lived. The prices of most goods and services adjust quickly, eliminating the surplus. Later on, we will discuss some markets in which adjustment of price to equilibrium may occur only very slowly or not at all.

Shortages

Just as a price above the equilibrium price will cause a surplus, a price below equilibrium will cause a shortage. A shortage is the amount by which the quantity demanded exceeds the quantity supplied at the current price.

Figure 3.16 "A Shortage in the Market for Coffee" shows a shortage in the market for coffee. Suppose the price is $4 per pound. At that price, 15 million pounds of coffee would be supplied per month, and 35 million pounds would be demanded per month. When more coffee is demanded than supplied, there is a shortage.

Figure 3.16 A Shortage in the Market for Coffee

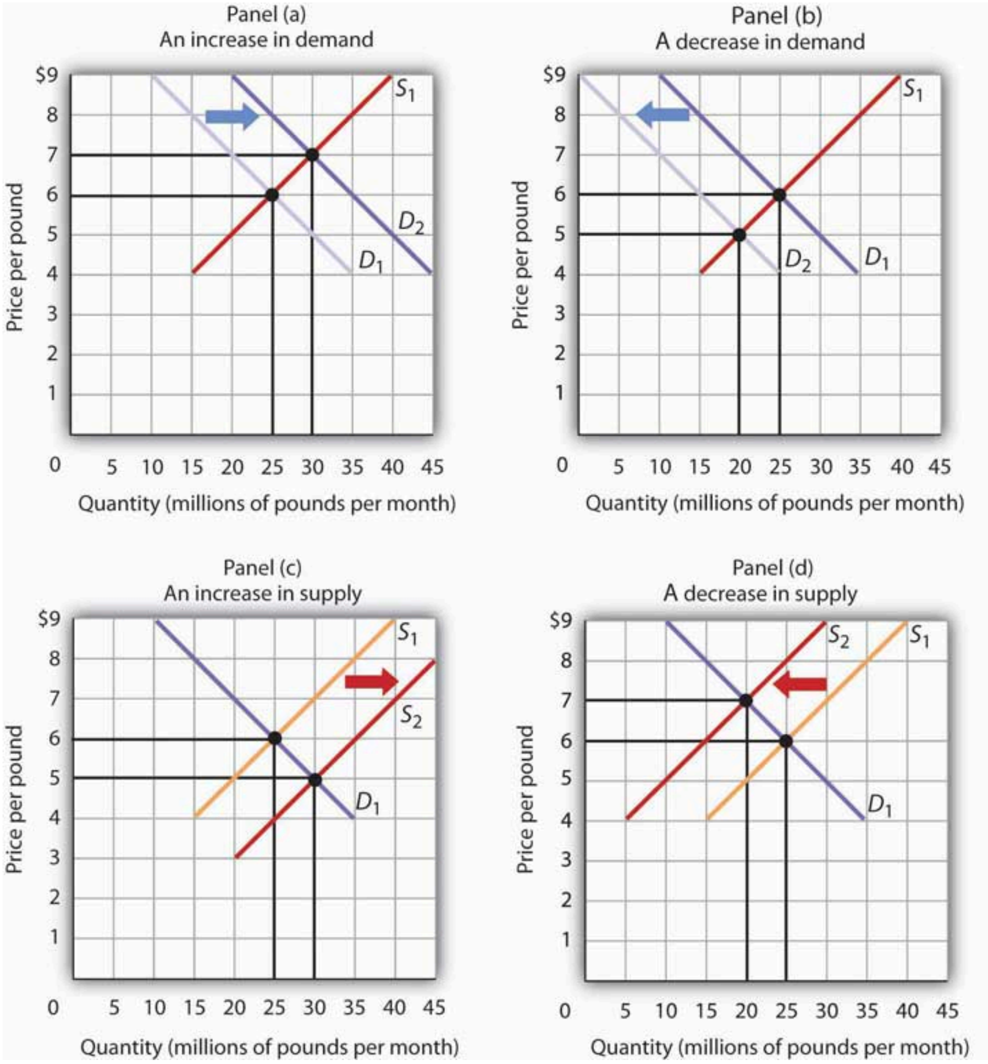


At a price of $4 per pound, the quantity of coffee demanded is 35 million pounds per month and the quantity supplied is 15 million pounds per month. The result is a shortage of 20 million pounds of coffee per month.

In the face of a shortage, sellers are likely to begin to raise their prices. As the price rises, there will be an increase in the quantity supplied (but not a change in supply) and a reduction in the quantity demanded (but not a change in demand) until the equilibrium price is achieved.

Shifts in Demand and Supply

Figure 3.17 Changes in Demand and Supply



A change in demand or in supply changes the equilibrium solution in the model. Panels (a) and (b) show an increase and a decrease in demand, respectively; Panels (c) and (d) show an increase and a decrease in supply, respectively.

A change in one of the variables (shifters) held constant in any model of demand and supply will create a change in demand or supply. A shift in a demand or supply curve changes the equilibrium price and equilibrium quantity for a good or service. Figure 3.17 "Changes in Demand and Supply" combines the information about changes in the demand and supply of coffee presented in Figure 3.2 "An Increase in Demand" Figure 3.3 "A Reduction in Demand" Figure 3.9 "An Increase in Supply" and Figure 3.10 "A Reduction in Supply" In each case, the original equilibrium price is $6 per pound, and the corresponding equilibrium quantity is 25 million pounds of coffee per month. Figure 3.17 "Changes in Demand and Supply" shows what happens with an increase in demand, a reduction in demand, an increase in supply, and a reduction in supply. We then look at what happens if both curves shift simultaneously. Each of these possibilities is discussed in turn below.

An Increase in Demand

An increase in demand for coffee shifts the demand curve to the right, as shown in Panel (a) of Figure 3.17 "Changes in

Demand and Supply”. The equilibrium price rises to $7 per pound. As the price rises to the new equilibrium level, the quantity supplied increases to 30 million pounds of coffee per month. Notice that the supply curve does not shift; rather, there is a movement along the supply curve.

Demand shifters that could cause an increase in demand include a shift in preferences that leads to greater coffee consumption; a lower price for a complement to coffee, such as doughnuts; a higher price for a substitute for coffee, such as tea; an increase in income; and an increase in population. A change in buyer expectations, perhaps due to predictions of bad weather lowering expected yields on coffee plants and increasing future coffee prices, could also increase current demand.

A Decrease in Demand

Panel (b) of Figure 3.17 “Changes in Demand and Supply” shows that a decrease in demand shifts the demand curve to the left. The equilibrium price falls to $5 per pound. As the price falls to the new equilibrium level, the quantity supplied decreases to 20 million pounds of coffee per month.

Demand shifters that could reduce the demand for coffee include a shift in preferences that makes people want to consume less coffee; an increase in the price of a complement, such as doughnuts; a reduction in the price of a substitute, such as tea; a reduction in income; a reduction in population; and a change in buyer expectations that leads people to expect lower prices for coffee in the future.

An Increase in Supply

An increase in the supply of coffee shifts the supply curve to the right, as shown in Panel (c) of Figure 3.17 “Changes in Demand and Supply”. The equilibrium price falls to $5 per pound. As the price falls to the new equilibrium level, the quantity of coffee demanded increases to 30 million pounds of coffee per month. Notice that the demand curve does not shift; rather, there is movement along the demand curve.

Possible supply shifters that could increase supply include a reduction in the price of an input such as labor, a decline in the returns available from alternative uses of the inputs that produce coffee, an improvement in the technology of coffee production, good weather, and an increase in the number of coffee-producing firms.

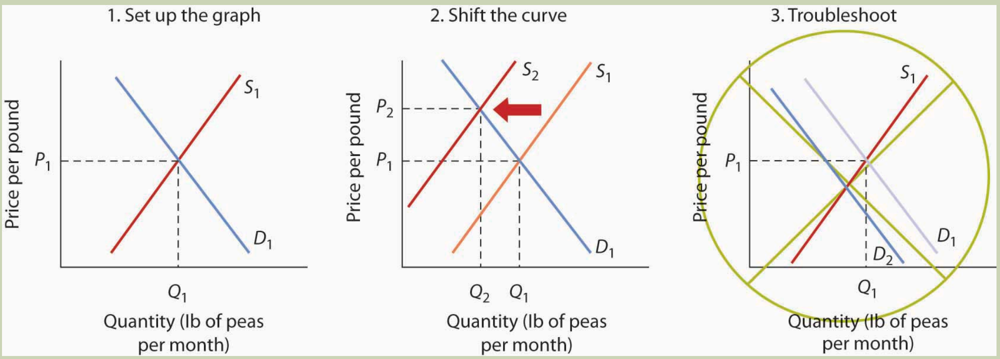
A Decrease in Supply

Panel (d) of Figure 3.17 “Changes in Demand and Supply” shows that a decrease in supply shifts the supply curve to the left. The equilibrium price rises to $7 per pound. As the price rises to the new equilibrium level, the quantity demanded decreases to 20 million pounds of coffee per month.

Possible supply shifters that could reduce supply include an increase in the prices of inputs used in the production of coffee, an increase in the returns available from alternative uses of these inputs, a decline in production because of problems in technology (perhaps caused by a restriction on pesticides used to protect coffee beans), a reduction in the number of coffee-producing firms, or a natural event, such as excessive rain.

Heads Up!

Figure 3.18



You are likely to be given problems in which you will have to shift a demand or supply curve.

Suppose you are told that an invasion of pod-crunching insects has gobbled up half the crop of fresh peas, and you are asked to use demand and supply analysis to predict what will happen to the price and quantity of peas demanded and supplied. Here are some suggestions.

Put the quantity of the good you are asked to analyze on the horizontal axis and its price on the vertical axis. Draw a downward-sloping line for demand and an upward-sloping line for supply. The initial equilibrium price is determined by the intersection of the two curves. Label the equilibrium solution. You may find it helpful to use a number for the equilibrium price instead of the letter "P." Pick a price that seems plausible, say, 79¢ per pound. Do not worry about the precise positions of the demand and supply curves; you cannot be expected to know what they are.

Step 2 can be the most difficult step; the problem is to decide which curve to shift. The key is to remember the difference between a change in demand or supply and a change in quantity demanded or supplied. At each price, ask yourself whether the given event would change the quantity demanded. Would the fact that a bug has attacked the pea crop change the quantity demanded at a price of, say, 79¢ per pound? Clearly not; none of the demand shifters have changed. The event would, however, reduce the quantity supplied at this price, and the supply curve would shift to the left. There is a change in supply and a reduction in the quantity demanded. There is no change in demand.

Next check to see whether the result you have obtained makes sense. The graph in Step 2 makes sense; it shows price rising and quantity demanded falling.

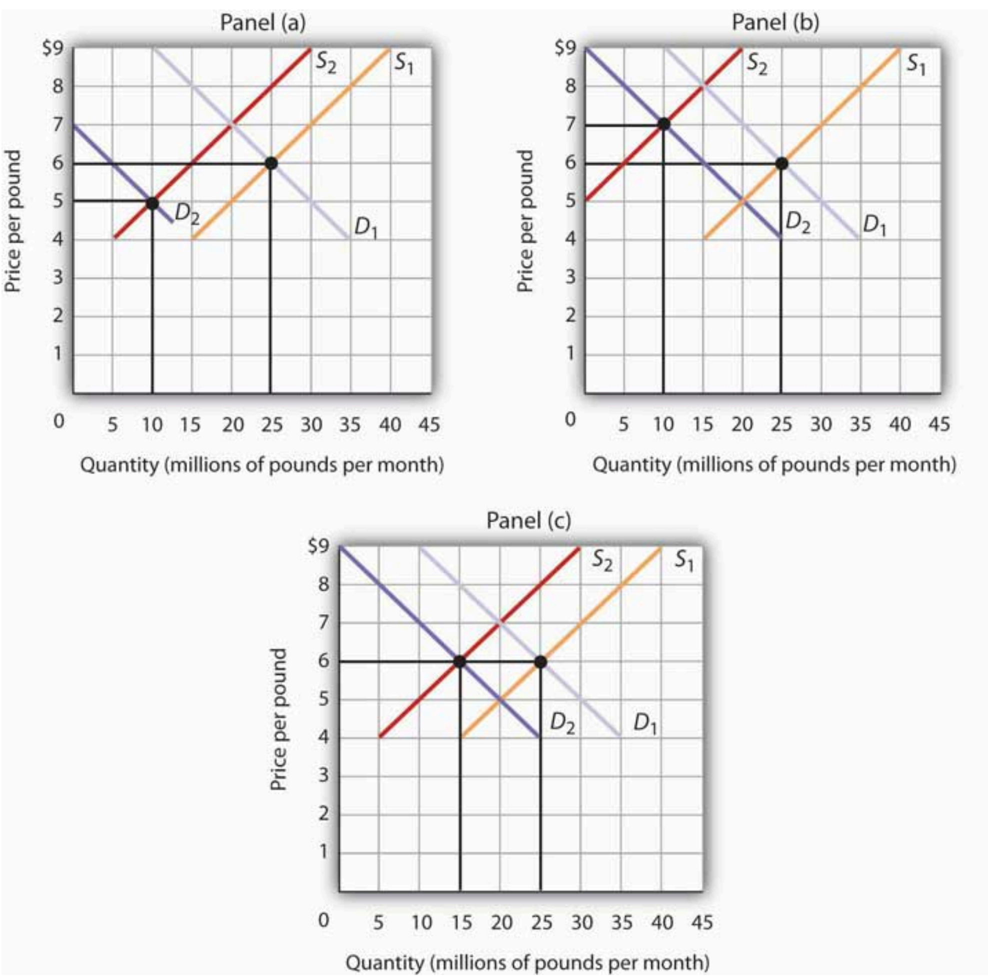
It is easy to make a mistake such as the one shown in the third figure of this Heads Up! One might, for example, reason that when fewer peas are available, fewer will be demanded, and therefore the demand curve will shift to the left. This suggests the price of peas will fall—but that does not make sense. If only half as many fresh peas were available, their price would surely rise. The error here lies in confusing a change in quantity demanded with a change in demand. Yes, buyers will end up buying fewer peas. But no, they will not demand fewer peas at each price than before; the demand curve does not shift.

Simultaneous Shifts

As we have seen, when *either* the demand or the supply curve shifts, the results are unambiguous; that is, we know what will happen to both equilibrium price and equilibrium quantity, so long as we know whether demand or supply increased or decreased. However, in practice, several events may occur at around the same time that cause *both* the demand and supply curves to shift. To figure out what happens to equilibrium price and equilibrium quantity, we must know not only in which direction the demand and supply curves have shifted but also the relative amount by which each curve shifts. Of course, the demand and supply curves could shift in the same direction or in opposite directions, depending on the specific events causing them to shift.

For example, all three panels of [Figure 3.19 "Simultaneous Decreases in Demand and Supply"] show a decrease in demand for coffee (caused perhaps by a decrease in the price of a substitute good, such as tea) and a simultaneous decrease in the supply of coffee (caused perhaps by bad weather). Since reductions in demand and supply, considered separately, each cause the equilibrium quantity to fall, the impact of both curves shifting simultaneously to the left means that the new equilibrium quantity of coffee is less than the old equilibrium quantity. The effect on the equilibrium price, though, is ambiguous. Whether the equilibrium price is higher, lower, or unchanged depends on the extent to which each curve shifts.

Figure 3.19 Simultaneous Decreases in Demand and Supply

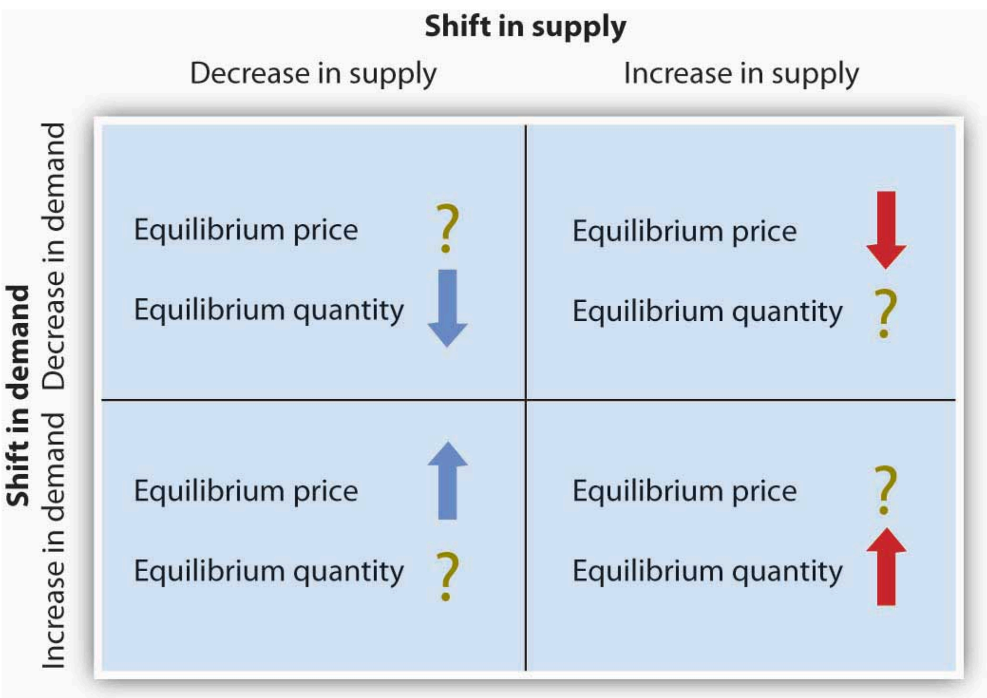


Both the demand and the supply of coffee decrease. Since decreases in demand and supply, considered separately, each cause equilibrium quantity to fall, the impact of both decreasing simultaneously means that a new equilibrium quantity of coffee must be less than the old equilibrium quantity. In Panel (a), the demand curve shifts farther to the left than does the supply curve, so equilibrium price falls. In Panel (b), the supply curve shifts farther to the left than does the demand curve, so the equilibrium price rises. In Panel (c), both curves shift to the left by the same amount, so equilibrium price stays the same.

If the demand curve shifts farther to the left than does the supply curve, as shown in Panel (a) of [Figure 3.19 "Simultaneous Decreases in Demand and Supply"](), then the equilibrium price will be lower than it was before the curves shifted. In this case the new equilibrium price falls from $6 per pound to $5 per pound. If the shift to the left of the supply curve is greater than that of the demand curve, the equilibrium price will be higher than it was before, as shown in Panel (b). In this case, the new equilibrium price rises to $7 per pound. In Panel (c), since both curves shift to the left by the same amount, equilibrium price does not change; it remains $6 per pound.

Regardless of the scenario, changes in equilibrium price and equilibrium quantity resulting from two different events need to be considered separately. If both events cause equilibrium price or quantity to move in the same direction, then clearly price or quantity can be expected to move in that direction. If one event causes price or quantity to rise while the other causes it to fall, the extent by which each curve shifts is critical to figuring out what happens. [Figure 3.20 "Simultaneous Shifts in Demand and Supply"]() summarizes what may happen to equilibrium price and quantity when demand and supply both shift.

Figure 3.20 Simultaneous Shifts in Demand and Supply



If simultaneous shifts in demand and supply cause equilibrium price or quantity to move in the same direction, then equilibrium price or quantity clearly moves in that direction. If the shift in one of the curves causes equilibrium price or quantity to rise while the shift in the other curve causes equilibrium price or quantity to fall, then the relative amount by which each curve shifts is critical to figuring out what happens to that variable.

As demand and supply curves shift, prices adjust to maintain a balance between the quantity of a good demanded and the quantity supplied. If prices did not adjust, this balance could not be maintained.

Notice that the demand and supply curves that we have examined in this chapter have all been drawn as linear. This simplification of the real world makes the graphs a bit easier to read without sacrificing the essential point: whether the curves are linear or nonlinear, demand curves are downward sloping and supply curves are generally upward sloping. As circumstances that shift the demand curve or the supply curve change, we can analyze what will happen to price and what will happen to quantity.

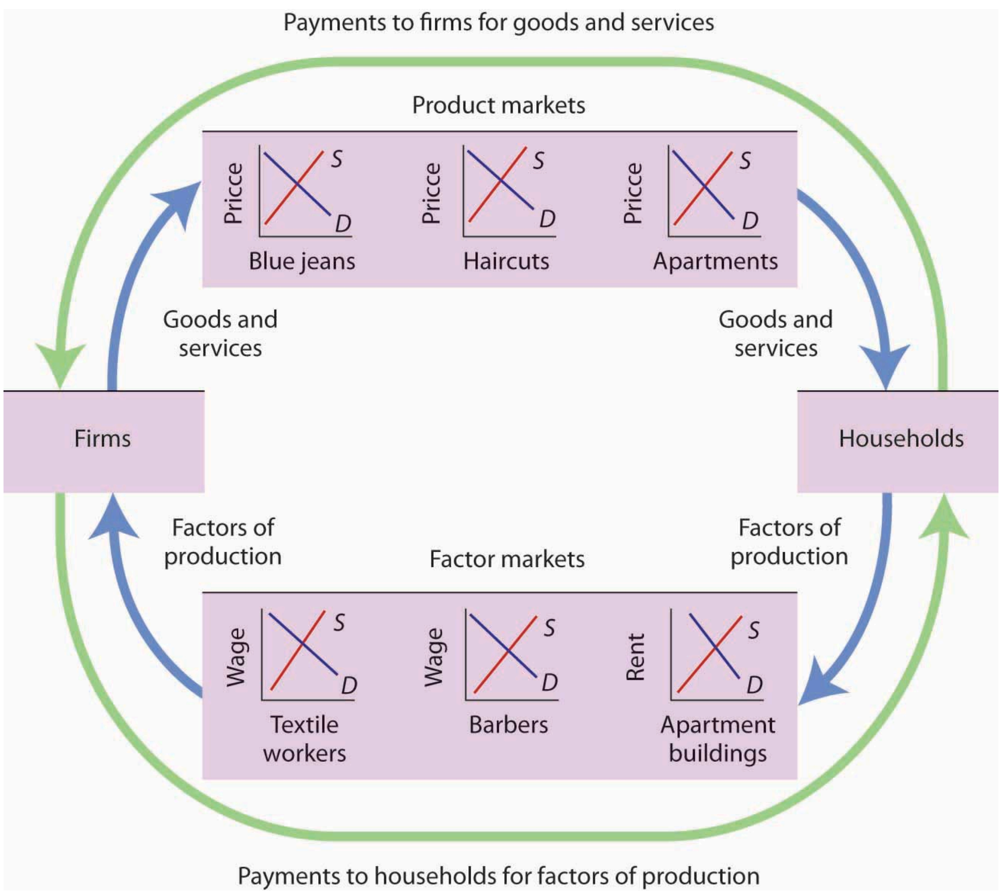
An Overview of Demand and Supply: The Circular Flow Model

Implicit in the concepts of demand and supply is a constant interaction and adjustment that economists illustrate with the circular flow model. The circular flow model provides a look at how markets work and how they are related to each other. It shows flows of spending and income through the economy.

A great deal of economic activity can be thought of as a process of exchange between households and firms. Firms supply goods and services to households. Households buy these goods and services from firms. Households supply factors of production—labor, capital, and natural resources—that firms require. The payments firms make in exchange for these factors represent the incomes households earn.

The flow of goods and services, factors of production, and the payments they generate is illustrated in Figure 3.21 "The Circular Flow of Economic Activity". This circular flow model of the economy shows the interaction of households and firms as they exchange goods and services and factors of production. For simplicity, the model here shows only the private domestic economy; it omits the government and foreign sectors.

Figure 3.21 The Circular Flow of Economic Activity



This simplified circular flow model shows flows of spending between households and firms through product and factor markets. The inner arrows show goods and services flowing from firms to households and factors of production flowing from households to firms. The outer flows show the payments for goods, services, and factors of production. These flows, in turn, represent millions of individual markets for products and factors of production.

The circular flow model shows that goods and services that households demand are supplied by firms in product markets. The exchange for goods and services is shown in the top half of Figure 3.21 "The Circular Flow of Economic Activity". The bottom half of the exhibit illustrates the exchanges that take place in factor markets. factor markets are markets in which households supply factors of production—labor, capital, and natural resources—demanded by firms.

Our model is called a circular flow model because households use the income they receive from their supply of factors of production to buy goods and services from firms. Firms, in turn, use the payments they receive from households to pay for their factors of production.

The demand and supply model developed in this chapter gives us a basic tool for understanding what is happening in each of these product or factor markets and also allows us to see how these markets are interrelated. In Figure 3.21 "The Circular Flow of Economic Activity", markets for three goods and services that households want—blue jeans, haircuts, and apartments—create demands by firms for textile workers, barbers, and apartment buildings. The equilibrium of supply and demand in each market determines the price and quantity of that item. Moreover, a change in equilibrium in one market will affect equilibrium in related markets. For example, an increase in the demand for haircuts would lead to an increase in demand for barbers. Equilibrium price and quantity could rise in both markets. For some purposes, it will be adequate to simply look at a single market, whereas at other times we will want to look at what happens in related markets as well.

In either case, the model of demand and supply is one of the most widely used tools of economic analysis. That widespread use is no accident. The model yields results that are, in fact, broadly consistent with what we observe in the marketplace. Your mastery of this model will pay big dividends in your study of economics.

Key Takeaways

- The equilibrium price is the price at which the quantity demanded equals the quantity supplied. It is determined by the intersection of the demand and supply curves.
- A surplus exists if the quantity of a good or service supplied exceeds the quantity demanded at the current price; it causes downward pressure on price. A shortage exists if the quantity of a good or service demanded exceeds the quantity supplied at the current price; it causes upward pressure on price.
- An increase in demand, all other things unchanged, will cause the equilibrium price to rise; quantity supplied will increase. A decrease in demand will cause the equilibrium price to fall; quantity supplied will decrease.
- An increase in supply, all other things unchanged, will cause the equilibrium price to fall; quantity demanded will increase. A decrease in supply will cause the equilibrium price to rise; quantity demanded will decrease.
- To determine what happens to equilibrium price and equilibrium quantity when both the supply and demand curves shift, you must know in which direction each of the curves shifts and the extent to which each curve shifts.

- The circular flow model provides an overview of demand and supply in product and factor markets and suggests how these markets are linked to one another.

Try It!

What happens to the equilibrium price and the equilibrium quantity of DVD rentals if the price of movie theater tickets increases and wages paid to DVD rental store clerks increase, all other things unchanged? Be sure to show all possible scenarios, as was done in Figure 3.19 "Simultaneous Decreases in Demand and Supply". Again, you do not need actual numbers to arrive at an answer. Just focus on the general position of the curve(s) before and after events occurred.

Case in Point: Demand, Supply, and Obesity

Figure 3.22

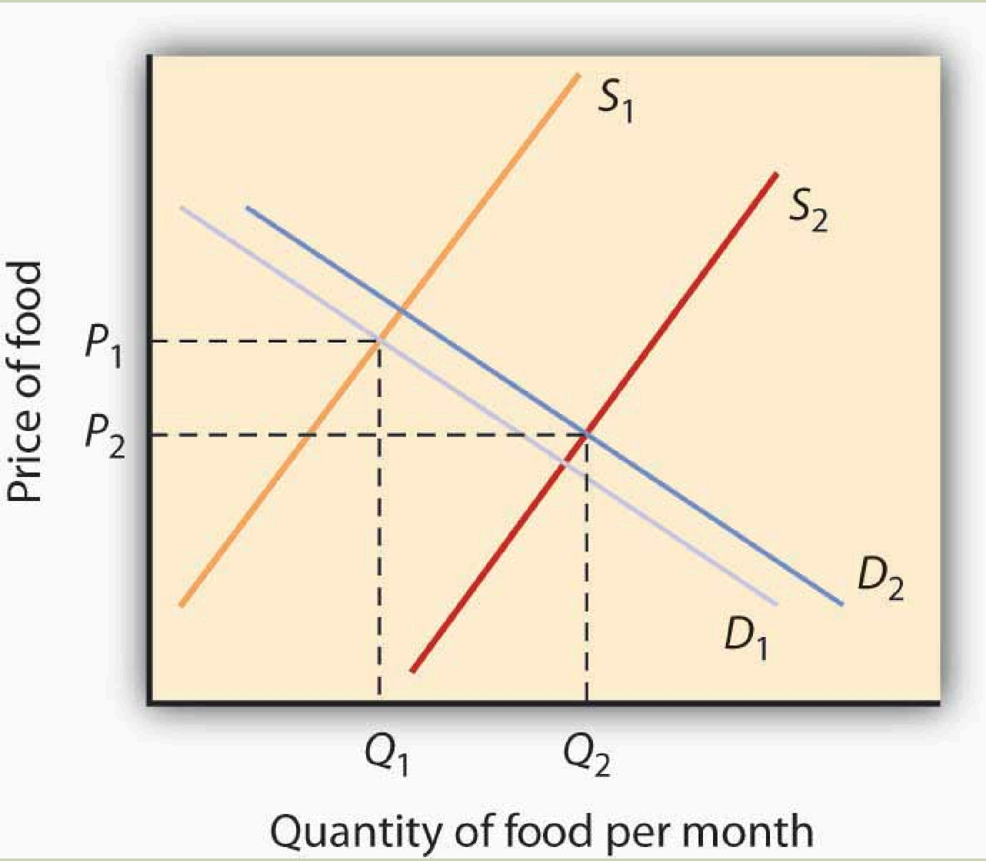
Rattopennugu – American fat beauty – CC BY-NC-ND 2.0.

Why are so many Americans fat? Put so crudely, the question may seem rude, but, indeed, the number of obese Americans has increased by more than 50% over the last generation, and obesity may now be the nation's number one health problem. According to Sturm Roland in a recent RAND Corporation study, "Obesity appears to have a stronger association with the occurrence of chronic medical conditions, reduced physical health-related quality of life and increased health care and medication expenditures than smoking or problem drinking."

Many explanations of rising obesity suggest higher demand for food. What more apt picture of our sedentary life style is there than spending the afternoon watching a ballgame on TV, while eating chips and salsa, followed by a dinner of a lavishly topped, take-out pizza? Higher income has also undoubtedly contributed to a rightward shift in the demand curve for food. Plus, any additional food intake translates into more weight increase because we spend so few calories preparing it, either directly or in the process of earning the income to buy it. A study by economists Darius Lakdawalla and Tomas Philipson suggests that about 60% of the recent growth in weight may be explained in this way—that is, demand has shifted to the right, leading to an increase in the equilibrium quantity of food consumed and, given our less strenuous life styles, even more weight gain than can be explained simply by the increased amount we are eating.

What accounts for the remaining 40% of the weight gain? Lakdawalla and Philipson further reason that a rightward shift in demand would by itself lead to an increase in the quantity of food as well as an increase in the price of food. The problem they have with this explanation is that over the post-World War II period, the relative price of food has declined by an average of 0.2 percentage points per year. They explain the fall in the price of food by arguing that agricultural innovation has led to a substantial rightward shift in the supply curve of food. As shown, lower food prices and a higher equilibrium quantity of food have resulted from simultaneous rightward shifts in demand and supply and that the rightward shift in the supply of food from S_1 to S_2 has been substantially larger than the rightward shift in the demand curve from D_1 to D_2.

Figure 3.23



Sources: Roland, Sturm, "The Effects of Obesity, Smoking, and Problem Drinking on Chronic Medical Problems and Health Care Costs," *Health Affairs*, 2002; 21(2): 245–253. Lakdawalla, Darius and Tomas Philipson, "The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination," National Bureau of Economic Research Working Paper no. w8946, May 2002.

Answer to Try It! Problem

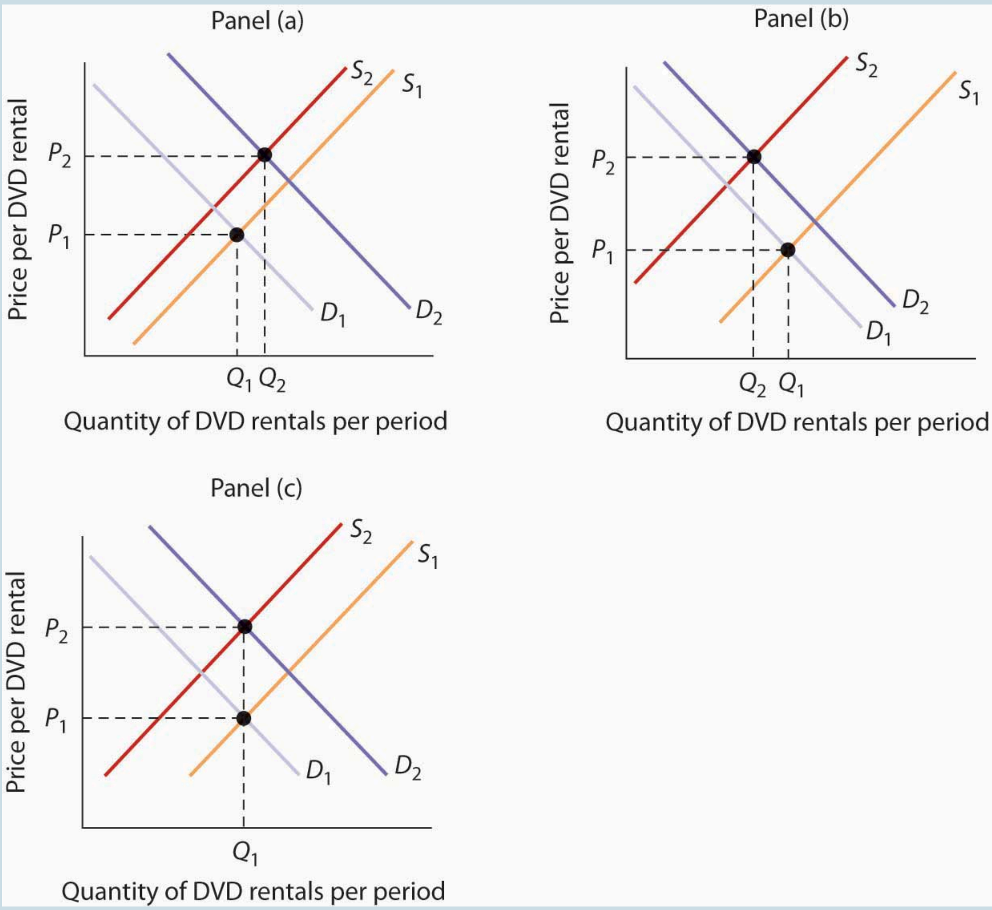
An increase in the price of movie theater tickets (a substitute for DVD rentals) will cause the demand curve for DVD rentals to shift to the right. An increase in the wages paid to DVD rental store clerks (an increase in the cost of a factor of production) shifts the supply curve to the left. Each event taken separately causes equilibrium price to rise. Whether equilibrium quantity will be higher or lower depends on which curve shifted more.

If the demand curve shifted more, then the equilibrium quantity of DVD rentals will rise [Panel (a)].

If the supply curve shifted more, then the equilibrium quantity of DVD rentals will fall [Panel (b)].

If the curves shifted by the same amount, then the equilibrium quantity of DVD rentals would not change [Panel (c)].

Figure 3.24



3.4 Review and Practice

Summary

In this chapter we have examined the model of demand and supply. We found that a demand curve shows the quantity demanded at each price, all other things unchanged. The law of demand asserts that an increase in price reduces the quantity demanded and a decrease in price increases the quantity demanded, all other things unchanged. The supply curve shows the quantity of a good or service that sellers will offer at various prices, all other things unchanged. Supply curves are generally upward sloping: an increase in price generally increases the quantity supplied, all other things unchanged.

The equilibrium price occurs where the demand and supply curves intersect. At this price, the quantity demanded equals the quantity supplied. A price higher than the equilibrium price increases the quantity supplied and reduces the quantity demanded, causing a surplus. A price lower than the equilibrium price increases the quantity demanded and reduces the quantity supplied, causing a shortage. Usually, market surpluses and shortages are short-lived. Changes in demand or supply, caused by changes in the determinants of demand and supply otherwise held constant in the analysis, change the equilibrium price and output. The circular flow model allows us to see how demand and supply in various markets are related to one another.

Concept Problems

1. What do you think happens to the demand for pizzas during the Super Bowl? Why?
2. Which of the following goods are likely to be classified as normal goods or services? Inferior? Defend your answer.
    1. Beans
    2. Tuxedos
    3. Used cars
    4. Used clothing
    5. Computers
    6. Books reviewed in *The New York Times*
    7. Macaroni and cheese
    8. Calculators
    9. Cigarettes
    10. Caviar
    11. Legal services

3. Which of the following pairs of goods are likely to be classified as substitutes? Complements? Defend your answer.

    1. Peanut butter and jelly
    2. Eggs and ham
    3. Nike brand and Reebok brand sneakers
    4. IBM and Apple Macintosh brand computers
    5. Dress shirts and ties
    6. Airline tickets and hotels
    7. Gasoline and tires
    8. Beer and wine
    9. Faxes and first-class mail
    10. Cereal and milk
    11. Cereal and eggs

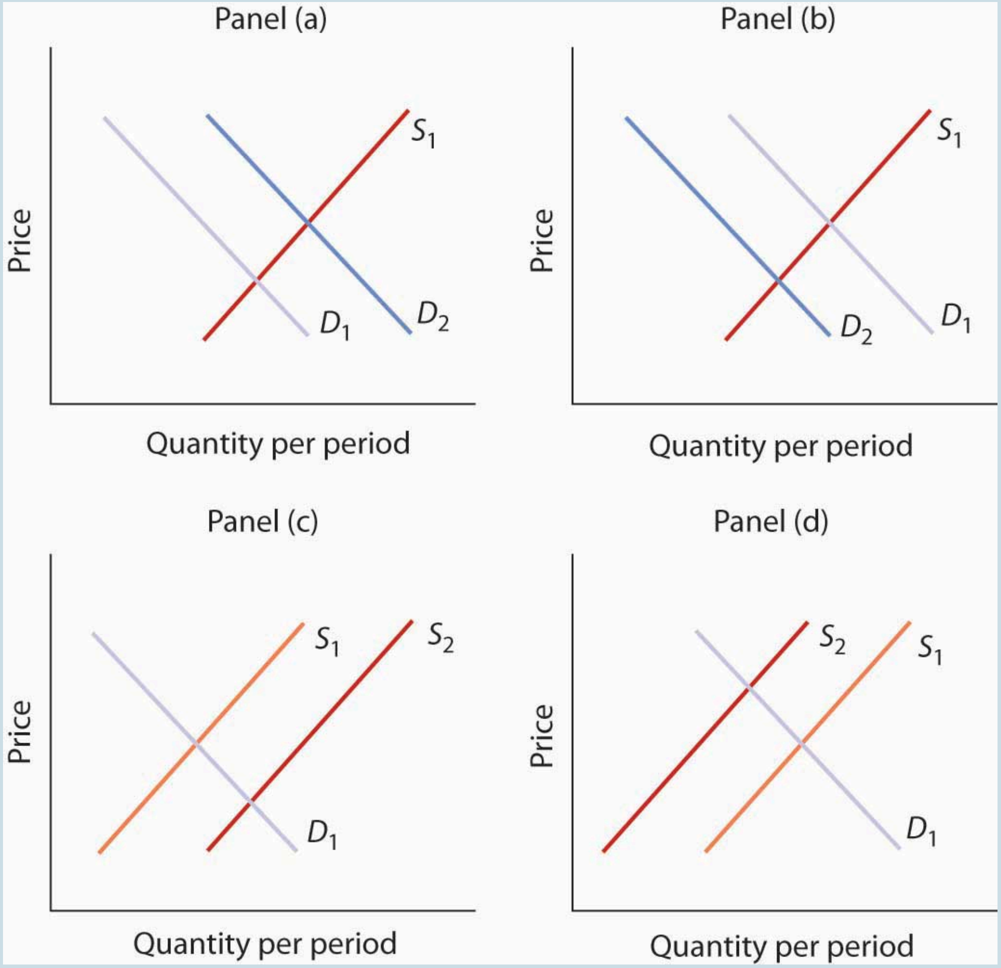
4. A study found that lower airfares led some people to substitute flying for driving to their vacation destinations. This reduced the demand for car travel and led to reduced traffic fatalities, since air travel is safer per passenger mile than car travel. Using the logic suggested by that study, suggest how each of the following events would affect the number of highway fatalities in any one year.

    1. An increase in the price of gasoline
    2. A large reduction in rental rates for passenger vans
    3. An increase in airfares

5. Children under age 2 are now allowed to fly free on U.S. airlines; they usually sit in their parents' laps. Some safety advocates have urged that they be required to be strapped in infant seats, which would mean their parents would have to purchase tickets for them. Some economists have argued that such a measure would actually increase infant fatalities. Can you say why?

6. The graphs below show four possible shifts in demand or in supply that could occur in particular markets. Relate each of the events described below to one of them.

    Figure 3.25



1. How did the heavy rains in South America in 1997 affect the market for coffee?
2. The Surgeon General decides french fries are not bad for your health after all and issues a report endorsing their use. What happens to the market for french fries?
3. How do you think rising incomes affect the market for ski vacations?
4. A new technique is discovered for manufacturing computers that greatly lowers their production cost. What happens to the market for computers?
5. How would a ban on smoking in public affect the market for cigarettes?

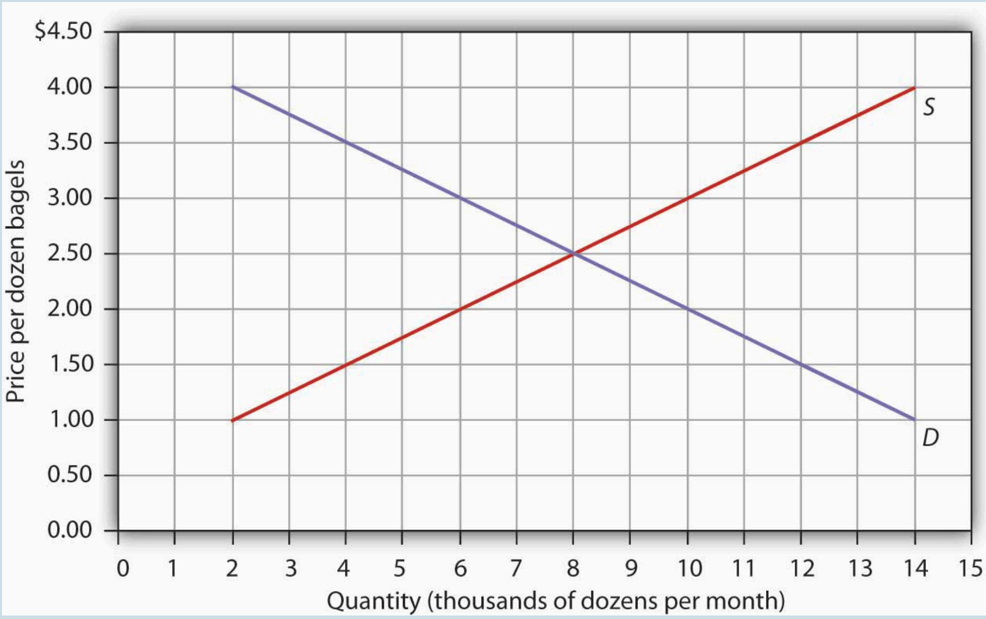
7. As low-carb diets increased in popularity, egg prices rose sharply. How might this affect the monks' supply of cookies or private retreats? (See the Case in Point on the Monks of St. Benedict's.)
8. Gasoline prices typically rise during the summer, a time of heavy tourist traffic. A "street talk" feature on a radio station sought tourist reaction to higher gasoline prices. Here was one response: "I don't like 'em [the higher prices] much. I think the gas companies just use any excuse to jack up prices, and they're doing it again now." How does this tourist's perspective differ from that of economists who use the model of demand and supply?
9. The introduction to the chapter argues that preferences for coffee changed in the 1990s and that excessive rain hurt yields from coffee plants. Show and explain the effects of these two circumstances on the coffee market.

10. With preferences for coffee remaining strong in the early part of the century, Vietnam entered the market as a major exporter of coffee. Show and explain the effects of these two circumstances on the coffee market.
11. The study on the economics of obesity discussed in the Case in Point in this chapter on that topic also noted that another factor behind rising obesity is the decline in cigarette smoking as the price of cigarettes has risen. Show and explain the effect of higher cigarette prices on the market for food. What does this finding imply about the relationship between cigarettes and food?
12. In 2004, *The New York Times* reported that India might be losing its outsourcing edge due to rising wages (Scheiber, N., 2004) The reporter noted that a recent report "projected that if India continued to produce college graduates at the current rate, demand would exceed supply by 20% in the main outsourcing markets by 2008." Using the terminology you learned in this chapter, explain what he meant to say was happening in the market for Indian workers in outsourcing jobs. In particular, is demand for Indian workers increasing or decreasing? Is the supply of Indian workers increasing or decreasing? Which is shifting faster? How do you know?
13. For more than a century, milk producers have produced skim milk, which contains virtually no fat, along with regular milk, which contains 4% fat. But a century ago, skim milk accounted for only about 1% of total production, and much of it was fed to hogs. Today, skim and other reduced-fat milks make up the bulk of milk sales. What curve shifted, and what factor shifted it?
14. Suppose firms in the economy were to produce fewer goods and services. How do you think this would affect household spending on goods and services? (*Hint*: Use the circular flow model to analyze this question.)

Numerical Problems

Problems 1–5 are based on the graph below.

Figure 3.26



1. At a price of $1.50 per dozen, how many bagels are demanded per month?
2. At a price of $1.50 per dozen, how many bagels are supplied per month?
3. At a price of $3.00 per dozen, how many bagels are demanded per month?
4. At a price of $3.00 per dozen, how many bagels are supplied per month?
5. What is the equilibrium price of bagels? What is the equilibrium quantity per month?

Problems 6–9 are based on the model of demand and supply for coffee as shown in Figure 3.17 "Changes in Demand and Supply" You can graph the initial demand and supply curves by using the following values, with all quantities in millions of pounds of coffee per month:

Price	Quantity demanded	Quantity supplied
$3	40	10
4	35	15
5	30	20
6	25	25
7	20	30
8	15	35
9	10	40

1. Suppose the quantity demanded rises by 20 million pounds of coffee per month at each price. Draw the initial demand and supply curves based on the values given in the table above. Then draw the new demand curve given by this change, and show the new equilibrium price and quantity.
2. Suppose the quantity demanded falls, relative to the values given in the above table, by 20 million pounds per month at prices between $4 and $6 per pound; at prices between $7 and $9 per pound, the quantity demanded becomes zero. Draw the new demand curve and show the new equilibrium price and quantity.
3. Suppose the quantity supplied rises by 20 million pounds per month at each price, while the quantities demanded retain the values shown in the table above. Draw the new supply curve and show the new

equilibrium price and quantity.

4. Suppose the quantity supplied falls, relative to the values given in the table above, by 20 million pounds per month at prices above $5; at a price of $5 or less per pound, the quantity supplied becomes zero. Draw the new supply curve and show the new equilibrium price and quantity.

Problems 10–15 are based on the demand and supply schedules for gasoline below (all quantities are in thousands of gallons per week):

Price per gallon	Quantity demanded	Quantity supplied
$1	8	0
2	7	1
3	6	2
4	5	3
5	4	4
6	3	5
7	2	6
8	1	7

1. Graph the demand and supply curves and show the equilibrium price and quantity.
2. At a price of $3 per gallon, would there be a surplus or shortage of gasoline? How much would the surplus or shortage be? Indicate the surplus or shortage on the graph.
3. At a price of $6 per gallon, would there be a surplus or shortage of gasoline? How much would the surplus or shortage be? Show the surplus or shortage on the graph.
4. Suppose the quantity demanded increased by 2,000 gallons per month at each price. At a price of $3 per gallon, how much would the surplus or shortage be? Graph the demand and supply curves and show the surplus or shortage.
5. Suppose the quantity supplied decreased by 2,000 gallons per month at each price for prices between $4 and $8 per gallon. At prices less than $4 per gallon the quantity supplied becomes zero, while the quantities demanded retain the values shown in the table. At a price of $4 per gallon, how much would the surplus or shortage be? Graph the demand and supply curves and show the surplus or shortage.
6. If the demand curve shifts as in problem 13 and the supply curve shifts as in problem 14, without drawing a graph or consulting the data, can you predict whether equilibrium price increases or decreases? What about equilibrium quantity? Now draw a graph that shows what the new equilibrium price and quantity are.

References

Scheiber, N., "As a Center for Outsourcing, India Could Be Losing Its Edge," *New York Times*, May 9, 2004, p. BU3.

CHAPTER 4: APPLICATIONS OF DEMAND AND SUPPLY

Start Up: A Composer Logs On

"Since the age of seven, I knew that I would be a musician. And from age fourteen, I knew that I would be a composer," says Israeli-born Ofer Ben-Amots. What he did not know was that he would use computers to carry out his work. He is now a professor of music at Colorado College, and Dr. Ben-Amots's compositions and operas have been performed in the United States, Europe, and Japan.

For over 15 years, he has used musical notation software to help in composing music. "The output is extremely elegant. Performers enjoy looking at such a clear and clean score. The creation of parts out of a full score is as easy as pressing the key on the keyboard." Changes can easily be inserted into the notation file, which eliminates the need for recopying. In addition, Dr. Ben-Amots uses computers for playback. "I can listen to a relatively accurate 'digital performance' of the score at any given point, with any tempo or instrumentation I choose. The sound quality has improved so much that digital files sound almost identical to real performance." He can also produce CDs on his own and create Podcasts so that anyone in the world can hear his music. He engages in self-publication of scores and self-marketing. "In my case, I get to keep the copyrights on all of my music. This would have been impossible ten to twelve years ago when composers transferred their rights to publishers. Home pages on the World Wide Web allow me to promote my own work." Professor Ben-Amots also changed the way he teaches music composition. New application software, such as GarageBand, has opened the way for anyone interested to try to compose music. Whereas his music composition classes used to have music theory prerequisites, today his classes are open to all.

Dr. Ben-Amots started out in 1989 with a Macintosh SE30 that had 4 megabytes of random access memory (RAM) and an 80-megabyte hard drive. It cost him about $3,000. Today, he uses a Macintosh Powerbook G4 laptop with 1.5 gigabytes of memory, built-in DVD/CD burner, and wireless Internet connections. His new computer cost about $2,000. How personal computers rose so dramatically in power as they fell so steeply in price is just one of the stories about markets we will tell in this chapter, which aims to help you understand how the model of demand and supply applies to the real world.

In the first section of this chapter, we will look at several markets that you are likely to have participated in or be familiar with—the market for personal computers, the markets for crude oil and for gasoline, and the stock market. You probably own or have access to a computer. Each of us was affected by the sharp rise in crude oil and gasoline prices from 2004 to mid-2008. The performance of the stock market is always a major news item and may affect you personally, if not now, then in the future. The concepts of demand and supply go a long way in explaining the behavior of equilibrium prices and quantities in all of these markets. The purpose of this section is to allow you to practice using the model of demand and supply and get you to start thinking about the myriad ways the model of demand and supply can be applied.

In the second part of the chapter we will look at markets in which the government has historically played a large role in regulating prices. By legislating maximum or minimum prices, the government has kept the prices of certain goods below or above equilibrium. We will look at the arguments for direct government intervention in controlling prices as well as the consequences of such policies. As we shall see, preventing the price of a good from finding its own equilibrium often has consequences that may be at odds with the intentions of the policy makers who put the regulations in place.

In the third section of the chapter we will look at the market for health care. This market is interesting because how well (or poorly) it works can be a matter of life and death and because it has special characteristics. In particular, markets in which participants do not pay for goods directly, but rather pay insurers who then pay the suppliers of the goods, operate somewhat differently from those in which participants pay directly for their purchases. This extension of demand and supply analysis reveals much about how such markets operate.

4.1 Putting Demand and Supply to Work

Learning Objectives

1. Learn how to apply the model of demand and supply to explaining the behavior of equilibrium prices and quantities in a variety of markets.
2. Explain how technological change can be represented using the model of demand and supply.
3. Explain how the model of demand and supply can be used to explain changes in prices of shares of stock.

A shift in either demand or supply, or in both, leads to a change in equilibrium price and equilibrium quantity. We begin this chapter by examining markets in which prices adjust quickly to changes in demand or supply: the market for personal computers, the markets for crude oil and gasoline, and the stock market. These markets are thus direct applications of the model of demand and supply.

The Personal Computer Market

In the 1960s, to speak of computers was to speak of IBM, the dominant maker of large mainframe computers used by business and government agencies. Then between 1976, when Apple Computer introduced its first desktop computer, and 1981, when IBM produced its first personal computers (PCs), the old world was turned upside down. In 1984, just 8.2% of U.S. households owned a personal computer. By 2007, Google estimates that 78% did. The tools of demand and supply tell the story from an economic perspective.

Technological change has been breathtakingly swift in the computer industry. Because personal computers have changed so dramatically in performance and in the range of the functions they perform, we shall speak of "quality-adjusted" personal computers. The price per unit of quality-adjusted desktop computers fell by about half every 50 months during the period 1976–1989. In the first half of the 1990s, those prices fell by half every 28 months. In the second half of the 1990s, the "halving time" fell to every 24 months (Tuomi, I).

Consider another indicator of the phenomenal change in computers. Between 1993 and 1998, the Bureau of Labor Statistics estimates that central processing unit (CPU) speed rose 1,263%, system memory increased 1,500%, hard drive capacity soared by 3,700%, and monitor size went up 13%. It seems safe to say that the dizzying pace of change recorded in the 1990s has increased in this century. A "computer" today is not the same good as a "computer" even five years ago. To make them comparable, we must adjust for these changes in quality.

Initially, most personal computers were manufactured by Apple or Compaq; both companies were very profitable. The potential for profits attracted IBM and other firms to the industry. Unlike large mainframe computers, personal computer clones turned out to be fairly easy things to manufacture. As shown in Table 4.1 "Personal Computer Shipments, Market Percentage Shares by Vendors, World and United States", the top five personal computer manufacturers produced only 48% of the personal computers sold in the world in 2005, and the largest manufacturer, Dell, sold only about 19% of the total in that year. This is a far cry from the more than 90% of the mainframe computer market that IBM once held. The market has become far more competitive.

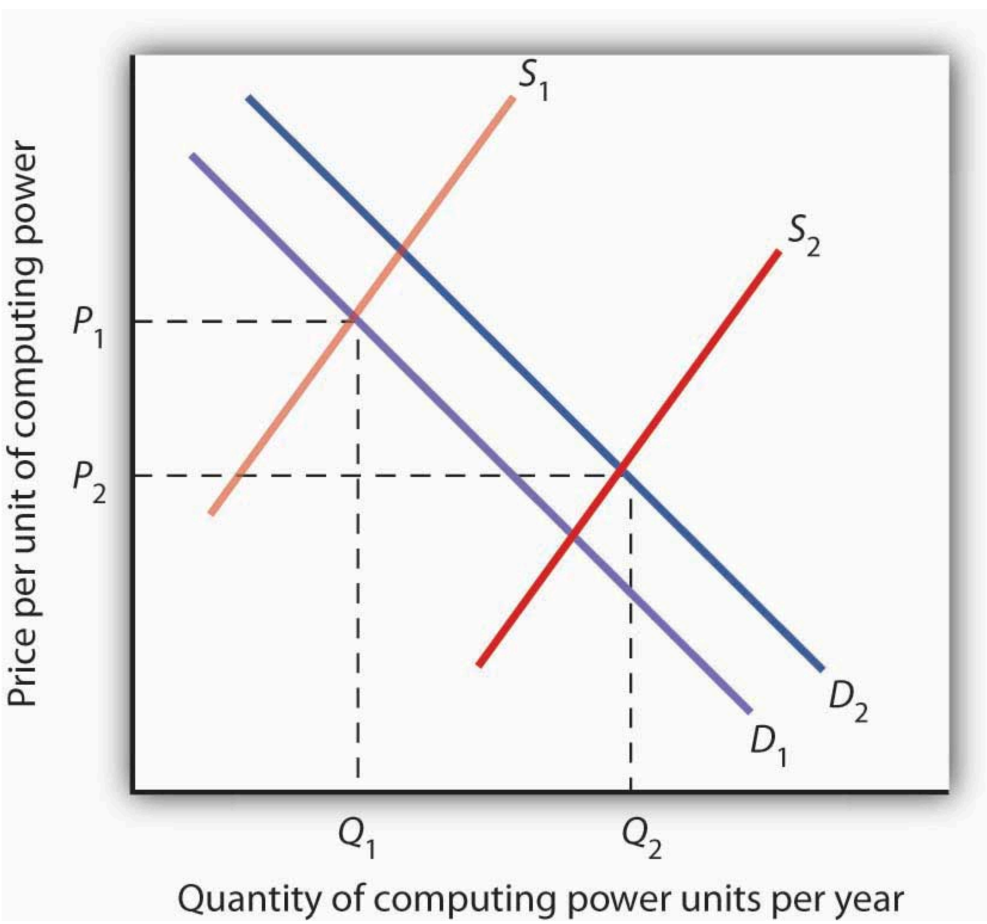
Table 4.1 Personal Computer Shipments, Market Percentage Shares by Vendors, World and United States

Company	% of World Shipments	Company	% of U.S. Shipments
Dell	18.9	Dell	34
Hewlett-Packard	15.4	Hewlett-Packard	18.2
IBM	5.1	Gateway	5.7
Fujitsu Seimens	4.6	IBM	4.3
Acer	4	Apple	3.9
Others	52	Others	34
Total	**100.0**	**Total**	**100.0**

Source: IDC—Press Release 15 Apr 2005 "PC Market Approaches 11% Growth as International Demand Remains Strong, According to IDC" (http://www.idc.com/getdoc.jsp?containerId=pr2005_04_14_17070722) (Totals may not add due to rounding)

Figure 4.1 "The Personal Computer Market" illustrates the changes that have occurred in the computer market. The horizontal axis shows the quantity of quality-adjusted personal computers. Thus, the quantity axis can be thought of as a unit of computing power. Similarly, the price axis shows the price per unit of computing power. The rapid increase in the number of firms, together with dramatic technological improvements, led to an increase in supply, shifting the supply curve in Figure 4.1 "The Personal Computer Market" to the right from S_1 to S_2.

Figure 4.1 The Personal Computer Market



The supply curve for quality-adjusted personal computers has shifted markedly to the right, reducing the equilibrium price from P_1 to P_2 and increasing the equilibrium quantity from Q_1 to Q_2 in 2005.

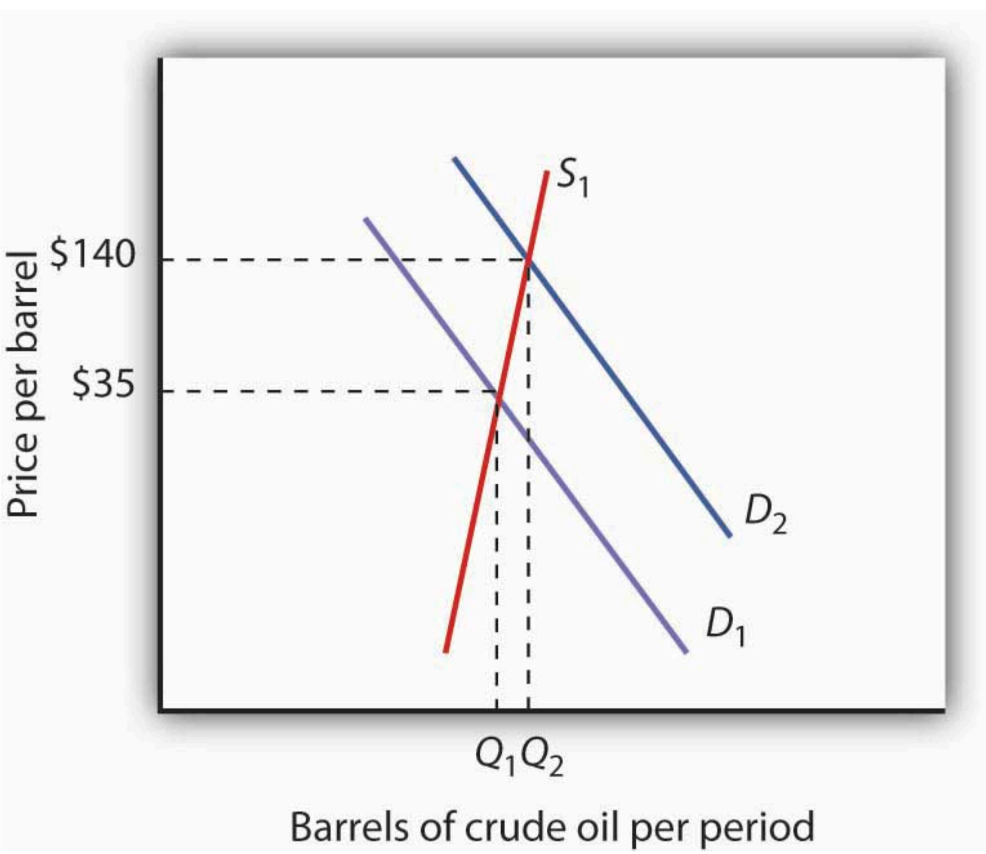
Demand also shifted to the right from D_1 to D_2, as incomes rose and new uses for computers, from e-mail and social networking to Voice over Internet Protocol (VoIP) and Radio Frequency ID (RFID) tags (which allow wireless tracking of commercial shipments via desktop computers), altered the preferences of consumer and business users. Because we observe a fall in equilibrium price and an increase in equilibrium quantity, we conclude that the rightward shift in supply has outweighed the rightward shift in demand. The power of market forces has profoundly affected the way we live and work.

The Markets for Crude Oil and for Gasoline

The market for crude oil took a radical turn in 1973. The price per barrel of crude oil quadrupled in 1973 and 1974. Price remained high until the early 1980s but then fell back drastically and remained low for about two decades. In 2004, the price of oil began to move upward and by 2008 had reached $147 per barrel.

What caused the dramatic increase in gasoline and oil prices in 2008? It appeared to be increasing worldwide demand outpacing producers' ability–or willingness–to increase production much. This increase in demand is illustrated in [Figure 4.2 "The Increasing Demand for Crude Oil"](#).

Figure 4.2 The Increasing Demand for Crude Oil

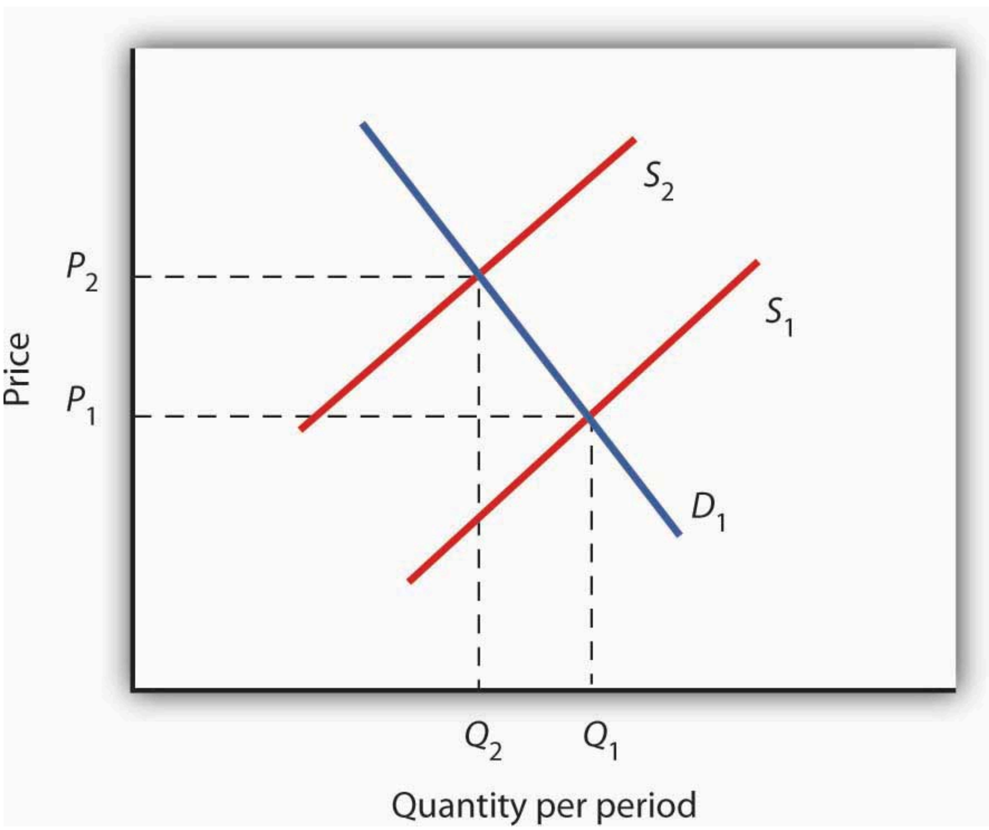


The price of oil was $35 per barrel at the beginning of 2004, as determined by the intersection of world demand, D_1, and world supply, S_1. Increasing world demand, prompted largely by increasing demand from China as well as from other countries, shifted world demand to D_2, pushing the price as high as $140 per barrel by the middle of 2008.

Higher oil prices also increase the cost of producing virtually every good or service, as at a minimum, the production of most goods requires transportation. These costs inevitably translate into higher prices for nearly all goods and services. Supply curves of the goods and services thus affected shift to the left, putting downward pressure on output and upward pressure on prices.

Graphically, the impact of higher gasoline prices on businesses that use gasoline is illustrated in Figure 4.3 "The Impact of Higher Gasoline Prices". Because higher gasoline prices increase the cost of doing business, they shift the supply curves for nearly all businesses to the left, putting upward pressure on prices and downward pressure on output. In the case shown here, the supply curve in a typical industry shifts from S_1 to S_2. This increases the equilibrium price from P_1 to P_2 and reduces the equilibrium quantity from Q_1 to Q_2.

Figure 4.3 The Impact of Higher Gasoline Prices



Higher gasoline prices increase the cost of producing virtually every good or service. In the case shown here, the supply curve in a typical industry shifts from S_1 to S_2. This increases the equilibrium price from P_1 to P_2 and reduces equilibrium quantity from Q_1 to Q_2.

Then, as the world economy slowed dramatically in the second half of 2008, the demand curve for oil shifted back to the left. By November 2008, the price per barrel had dropped back to below $60 per barrel. As gas prices also subsided, so did the threat of higher prices in other industries.

The Stock Market

The circular flow model suggests that capital, like other factors of production, is supplied by households to firms. Firms, in turn, pay income to those households for the use of their capital. Generally speaking, however, capital is actually owned by firms themselves. General Motors owns its assembly plants, and Wal-Mart owns its stores; these firms therefore own their capital. But firms, in turn, are owned by people—and those people, of course, live in households. It is through their ownership of firms that households own capital.

A firm may be owned by one individual (a sole proprietorship), by several individuals (a partnership), or by shareholders who own stock in the firm (a corporation). Although most firms in the United States are sole proprietorships or partnerships, the bulk of the nation's total output (about 90%) is produced by corporations. Corporations also own most of the capital (machines, plants, buildings, and the like).

This section describes how the prices of shares of corporate stock, shares in the ownership of a corporation, are determined by the interaction of demand and supply. Ultimately, the same forces that determine the value of a firm's stock determine the value of a sole proprietorship or partnership.

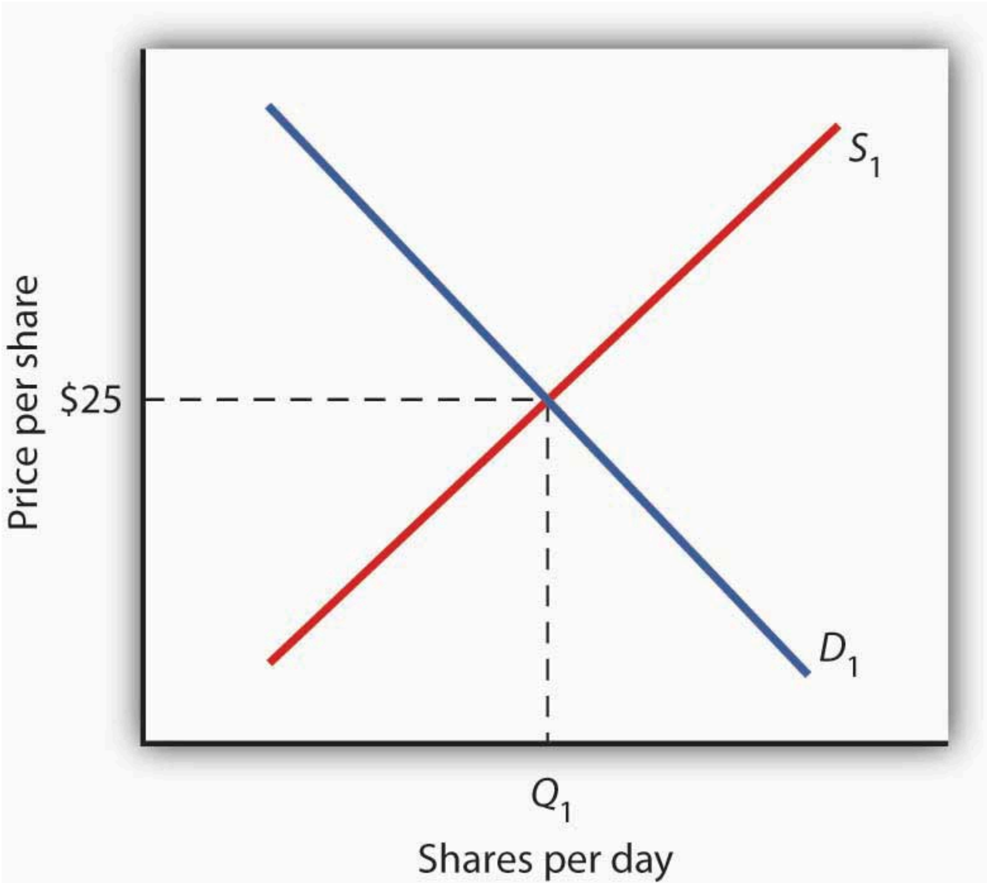
When a corporation needs funds to increase its capital or for other reasons, one means at its disposal is to issue new stock in the corporation. (Other means include borrowing funds or using past profits.) Once the new shares have been sold in what is called an initial public offering (IPO), the corporation receives no further funding as shares of its stock are bought and sold on the secondary market. The secondary market is the market for stocks that have been issued in the past, and the daily news reports about stock prices almost always refer to activity in the secondary market. Generally, the corporations whose shares are traded are not involved in these transactions.

The stock market is the set of institutions in which shares of stock are bought and sold. The New York Stock Exchange (NYSE) is one such institution. There are many others all over the world, such as the DAX in Germany and the Bolsa in Mexico. To buy or sell a share of stock, one places an order with a stockbroker who relays the order to one of the traders at the NYSE or at some other exchange.

The process through which shares of stock are bought and sold can seem chaotic. At many exchanges, traders with orders from customers who want to buy stock shout out the prices those customers are willing to pay. Traders with orders from customers who want to sell shout out offers of prices at which their customers are willing to sell. Some exchanges use electronic trading, but the principle is the same: if the price someone is willing to pay matches the price at which someone else is willing to sell, the trade is made. The most recent price at which a stock has traded is reported almost instantaneously throughout the world.

Figure 4.4 "Demand and Supply in the Stock Market" applies the model of demand and supply to the determination of stock prices. Suppose the demand curve for shares in Intel Corporation is given by D_1 and the supply by S_1. (Even though the total number of shares outstanding is fixed at any point in time, the supply curve is not vertical. Rather, the supply curve is upward sloping because it represents how many shares current owners are prepared to sell at each price, and that number will be greater at higher prices.) Suppose that these curves intersect at a price of $25, at which Q_1 shares are traded each day. If the price were higher, more shares would be offered for sale than would be demanded, and the price would quickly fall. If the price were lower, more shares would be demanded than would be supplied, and the price would quickly rise. In general, we can expect the prices of shares of stock to move quickly to their equilibrium levels.

Figure 4.4 Demand and Supply in the Stock Market



The equilibrium price of stock shares in Intel Corporation is initially $25, determined by the intersection of demand and supply curves D_1 and S_1, at which Q_1 million shares are traded each day.

The intersection of the demand and supply curves for shares of stock in a particular company determines the equilibrium price for a share of stock. But what determines the demand and supply for shares of a company's stock?

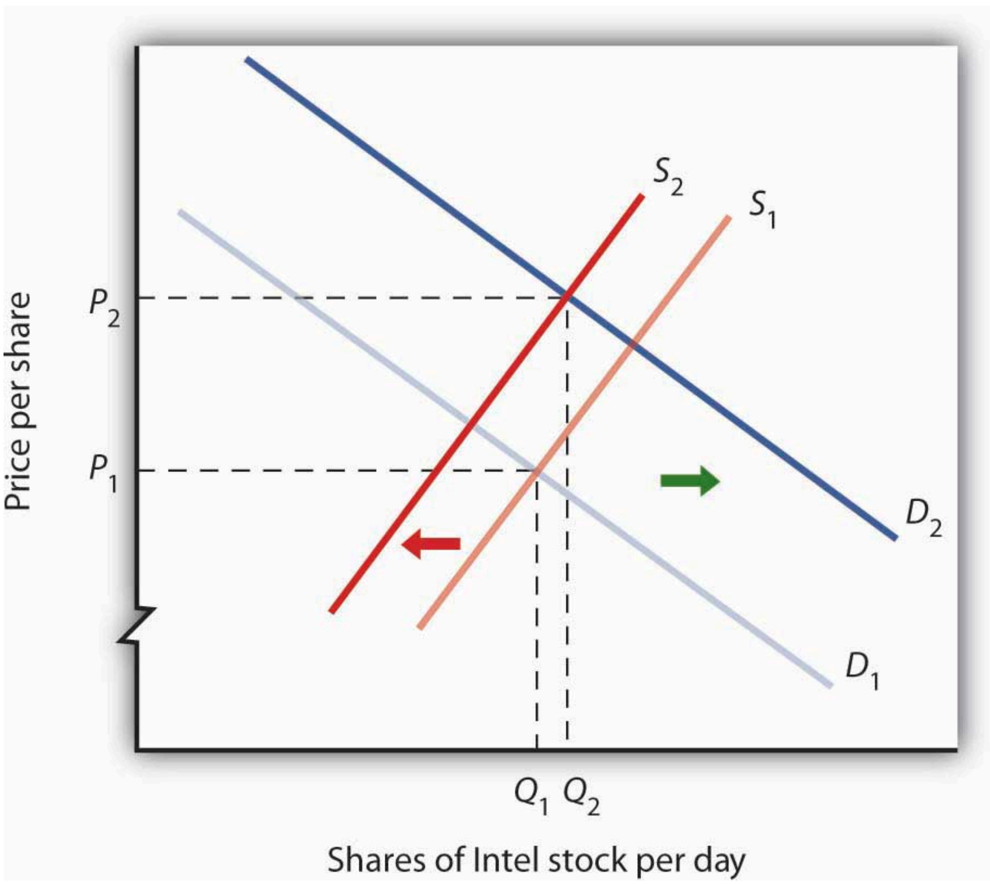
The owner of a share of a company's stock owns a share of the company, and, hence, a share of its profits; typically, a corporation will retain and reinvest some of its profits to increase its future profitability. The profits kept by a company are called retained earnings. Profits distributed to shareholders are called dividends. Because a share of stock gives its owner a claim on part of a company's future profits, it follows that the expected level of future profits plays a role in determining the value of its stock.

Of course, those future profits cannot be known with certainty; investors can only predict what they might be, based on information about future demand for the company's products, future costs of production, information about the soundness of a company's management, and so on. Stock prices in the real world thus reflect estimates of a company's profits projected into the future.

The downward slope of the demand curve suggests that at lower prices for the stock, more people calculate that the firm's future earnings will justify the stock's purchase. The upward slope of the supply curve tells us that as the price of the stock rises, more people conclude that the firm's future earnings do not justify holding the stock and therefore offer to sell it. At the equilibrium price, the number of shares supplied by people who think holding the stock no longer makes sense just balances the number of shares demanded by people who think it does.

What factors, then, cause the demand or supply curves for shares of stocks to shift? The most important factor is a change in the expectations of a company's future profits. Suppose Intel announces a new generation of computer chips that will lead to faster computers with larger memories. Current owners of Intel stock would adjust upward their estimates of what the value of a share of Intel stock should be. At the old equilibrium price of $25 fewer owners of Intel stock would be willing to sell. Since this would be true at every possible share price, the supply curve for Intel stock would shift to the left, as shown in Figure 4.5 "A Change in Expectations Affects the Price of Corporate Stock". Just as the expectation that a company will be more profitable shifts the supply curve for its stock to the left, that same change in expectations will cause more people to want to purchase the stock, shifting the demand curve to the right. In Figure 4.5 "A Change in Expectations Affects the Price of Corporate Stock", we see the supply curve shifting to the left, from S_1 to S_2, while the demand curve shifts to the right, from D_1 to D_2.

Figure 4.5 A Change in Expectations Affects the Price of Corporate Stock



If financial investors decide that a company is likely to be more profitable, then the supply of the stock shifts to the left (in this case, from S_1 to S_2), and the demand for the stock shifts to the right (in this case, from D_1 to D_2), resulting in an increase in price from P_1 to P_2.

Other factors may alter the price of an individual corporation's share of stock or the level of stock prices in general. For example, demographic change and rising incomes have affected the demand for stocks in recent years. For example, with a large proportion of the U.S. population nearing retirement age and beginning to think about and plan for their lives during retirement, the demand for stocks has risen.

Information on the economy as a whole is also likely to affect stock prices. If the economy overall is doing well and people expect that to continue, they may become more optimistic about how profitable companies will be in general, and thus the prices of stocks will rise. Conversely, expectations of a sluggish economy, as happened in the fall of 2008, could cause stock prices in general to fall.

The stock market is bombarded with new information every minute of every day. Firms announce their profits of the previous quarter. They announce that they plan to move into a new product line or sell their goods in another country. We learn that the price of Company A's good, which is a substitute for one sold by Company B, has risen. We learn that countries sign trade agreements, launch wars, or make peace. All of this information may affect stock prices because any information can affect how buyers and sellers value companies.

Key Takeaways

- Technological change, which has caused the supply curve for computing power to shift to the right, is the main reason for the rapid increase in equilibrium quantity and decrease in equilibrium price of personal computers.
- The increase in crude oil and gasoline prices in 2008 was driven primarily by increased demand for crude oil, an increase that was created by economic growth throughout the world. Crude oil and gas prices fell markedly as world economic growth subsided later in the year.
- Higher gasoline prices increased the cost of producing virtually every good and service, shifting supply curves for most goods and services to the left. This tended to push prices up and output down.
- Demand and supply determine prices of shares of corporate stock. The equilibrium price of a share of stock strikes a balance between those who think the stock is worth more and those who think it is worth less than the current price.
- If a company's profits are expected to increase, the demand curve for its stock shifts to the right and the supply curve shifts to the left, causing equilibrium price to rise. The opposite would occur if a company's profits were expected to decrease.
- Other factors that influence the price of corporate stock include demographic and income changes and the overall health of the economy.

Try It!

Suppose an airline announces that its earnings this year are lower than expected due to reduced ticket sales. The airline spokesperson gives no information on how the company plans to turn things around. Use the model of demand and supply to show and explain what is likely to happen to the price of the airline's stock.

Case in Point: 9/11 and the Stock Market

Figure 4.6

Randy Lemoine – New York Stock Exchange – CC BY 2.0.

The hijacking of four airplanes and the steering of them into buildings is perhaps the only disaster that has become universally known by its date: September 11, 2001—hence, 9/11. "9/11" will remain etched in our collective memory for a great many generations.

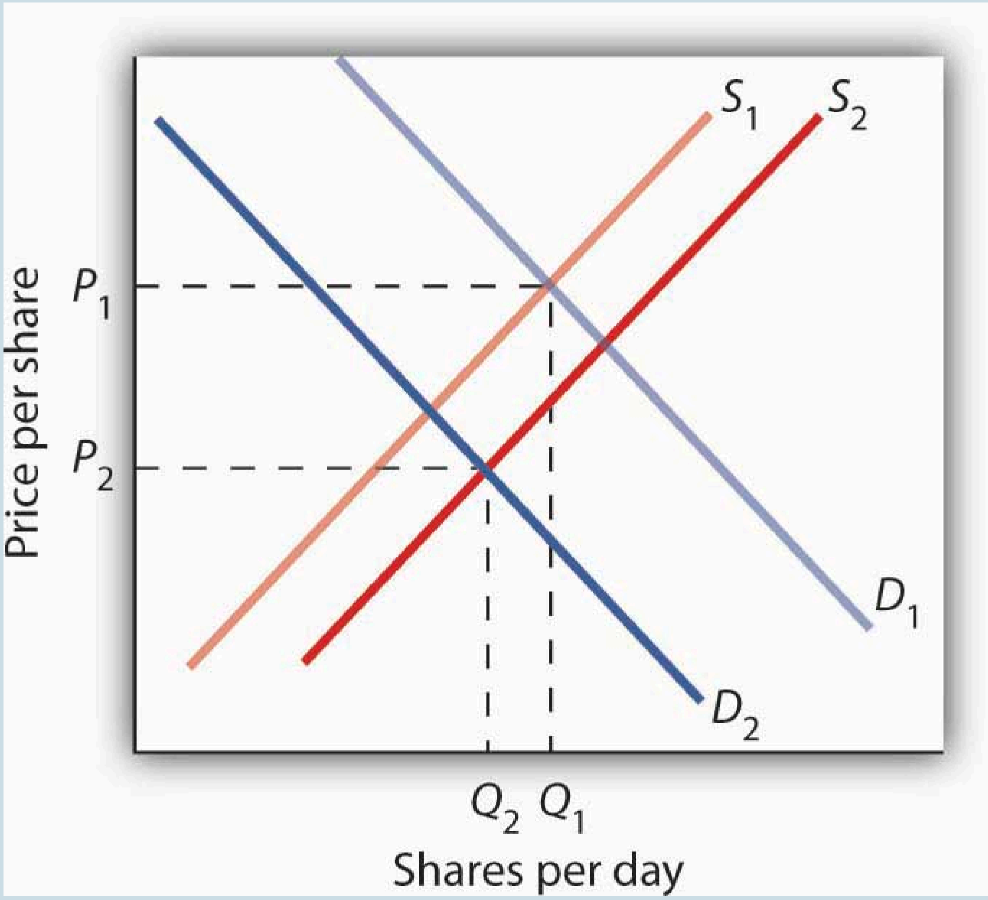
Disasters such as 9/11 represent the kind of complete surprises that dramatically affect stock prices, if only temporarily. The New York Stock Exchange was closed on the day of the attack and remained closed for six days. On the day the market opened, the Dow Jones Industrial Average (the "DOW", a widely used gauge of stock prices) fell nearly 685 points to 8,920. It was one of the biggest one-day decline in U.S. history.

Why did the attacks on September 11, 2001, have such a dramatic short-term impact on the stock market? The attacks of 9/11 plunged the United States and much of the rest of the world into a very frightening war against terrorism. The realization that terrorists could strike anytime and in any place sapped consumer and business confidence alike and affected both the demand and supply of most stocks. The attacks on 9/11 provoked fear and uncertainty—two things that are certain to bring stock prices down, at least until other events and more information cause expectations to change again in this very responsive market.

Answer to Try It! Problem

The information given in the problem suggests that the airline's profits are likely to fall below expectations. Current owners of the airline's stock and potential buyers of the stock would adjust downward their estimates of what the value of the corporation's stock should be. As a result the supply curve for the stock would increase, shifting it to the right, while the demand curve for the stock would decrease, shifting it to the left. As a result, equilibrium price of the stock falls from P_1 to P_2. What happens to equilibrium quantity depends on the extent to which each curve shifts. In the diagram, equilibrium quantity is shown to decrease from Q_1 to Q_2.

Figure 4.7



References

Tuomi, I., "The Lives and Death of Moore's Law." http://www.firstmonday.org/issues/issue7_11/tuomi/index. *First Monday* (http://www.firstmonday.org) is a peer-reviewed journal on the Internet.

4.2 Government Intervention in Market Prices: Price Floors and Price Ceilings

Learning Objectives

1. Use the model of demand and supply to explain what happens when the government imposes price floors or price ceilings.
2. Discuss the reasons why governments sometimes choose to control prices and the consequences of price control policies.

So far in this chapter and in the previous chapter, we have learned that markets tend to move toward their equilibrium prices and quantities. Surpluses and shortages of goods are short-lived as prices adjust to equate quantity demanded with quantity supplied.

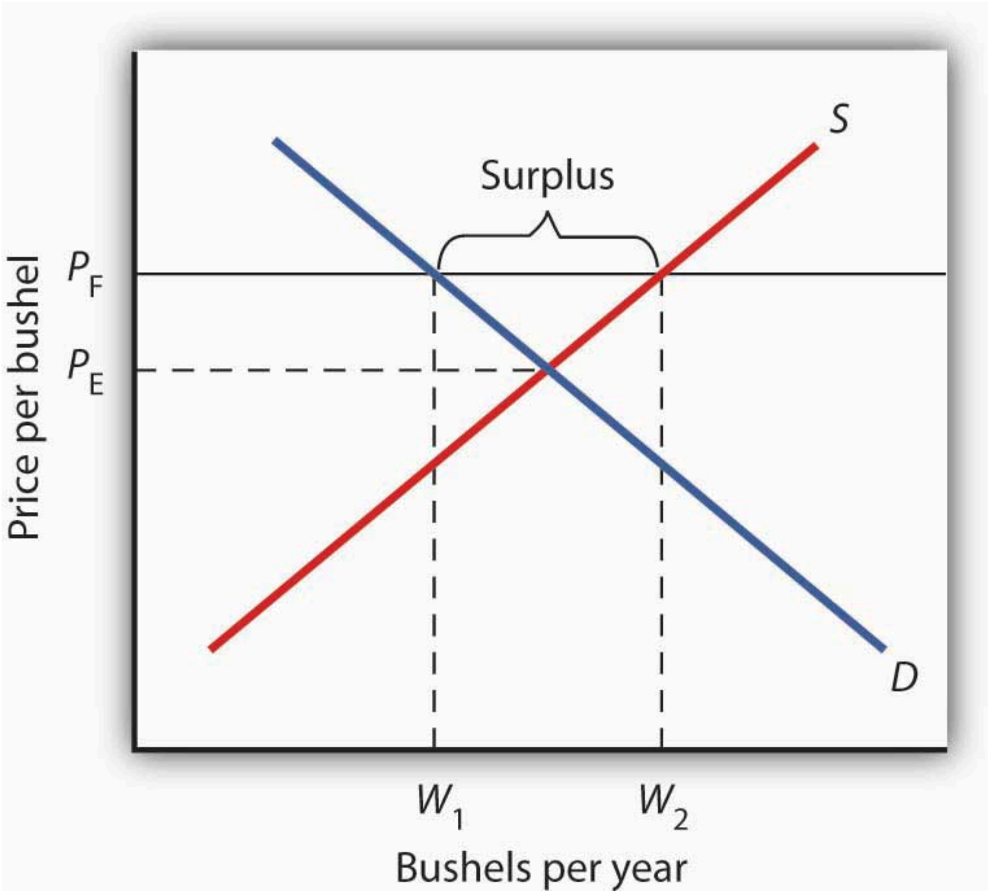
In some markets, however, governments have been called on by groups of citizens to intervene to keep prices of certain items higher or lower than what would result from the market finding its own equilibrium price. In this section we will examine agricultural markets and apartment rental markets—two markets that have often been subject to price controls. Through these examples, we will identify the effects of controlling prices. In each case, we will look at reasons why governments have chosen to control prices in these markets and the consequences of these policies.

Agricultural Price Floors

Governments often seek to assist farmers by setting price floors in agricultural markets. A minimum allowable price set above the equilibrium price is a price floor. With a price floor, the government forbids a price below the minimum. (Notice that, if the price floor were for whatever reason set below the equilibrium price, it would be irrelevant to the determination of the price in the market since nothing would prohibit the price from rising to equilibrium.) A price floor that is set above the equilibrium price creates a surplus.

Figure 4.8 "Price Floors in Wheat Markets" shows the market for wheat. Suppose the government sets the price of wheat at P_F. Notice that P_F is above the equilibrium price of P_E. At P_F, we read over to the demand curve to find that the quantity of wheat that buyers will be willing and able to purchase is W_1 bushels. Reading over to the supply curve, we find that sellers will offer W_2 bushels of wheat at the price floor of P_F. Because P_F is above the equilibrium price, there is a surplus of wheat equal to (W_2 – W_1) bushels. The surplus persists because the government does not allow the price to fall.

Figure 4.8 Price Floors in Wheat Markets



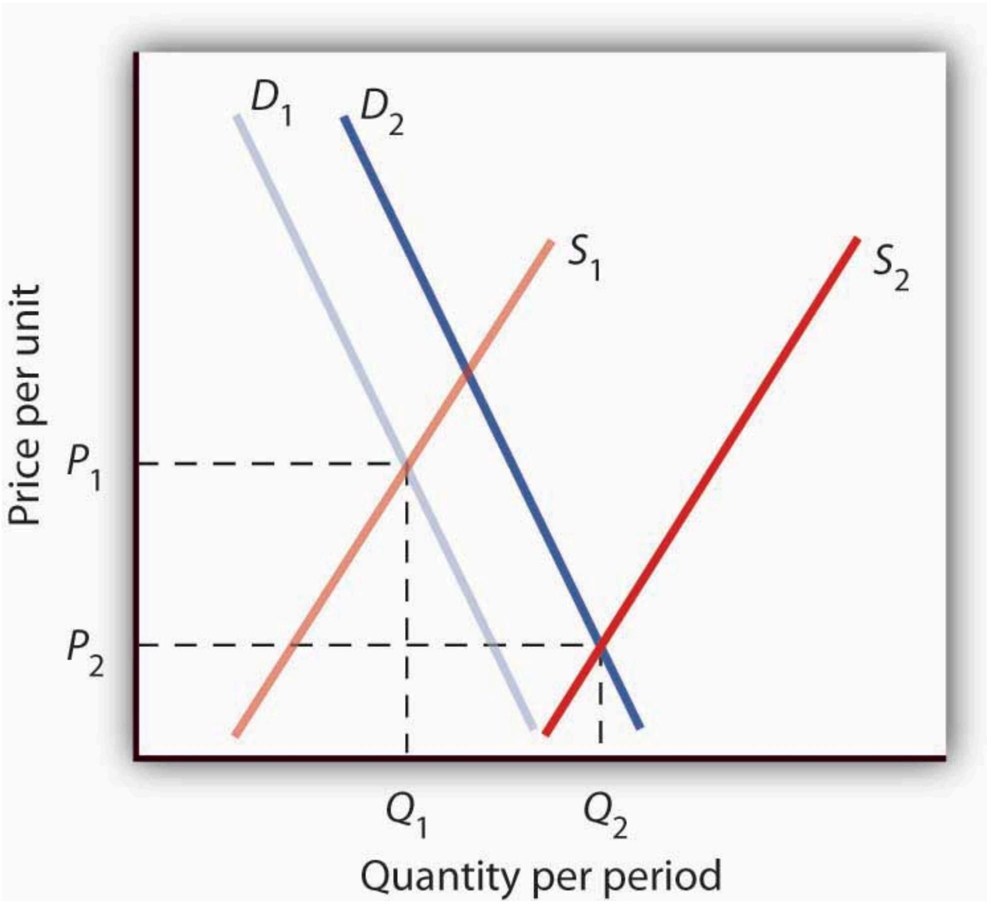
A price floor for wheat creates a surplus of wheat equal to (W_2 – W_1) bushels.

Why have many governments around the world set price floors in agricultural markets? Farming has changed dramatically over the past two centuries. Technological improvements in the form of new equipment, fertilizers, pesticides, and new varieties of crops have led to dramatic increases in crop output per acre. Worldwide production capacity has expanded markedly. As we have learned, technological improvements cause the supply curve to shift to the right, reducing the price of food. While such price reductions have been celebrated in computer markets, farmers have successfully lobbied for government programs aimed at keeping their prices from falling.

While the supply curve for agricultural goods has shifted to the right, the demand has increased with rising population and with rising income. But as incomes rise, people spend a smaller and smaller fraction of their incomes on food. While the demand for food has increased, that increase has not been nearly as great as the increase in supply. Figure 4.9 "Supply and Demand Shifts for Agricultural Products" shows that the supply curve has shifted much farther to the right, from S_1 to S_2, than the demand curve has, from D_1 to D_2. As a result, equilibrium quantity has risen dramatically, from Q_1 to Q_2, and equilibrium price has fallen, from P_1 to P_2.

On top of this long-term historical trend in agriculture, agricultural prices are subject to wide swings over shorter periods. Droughts or freezes can sharply reduce supplies of particular crops, causing sudden increases in prices. Demand for agricultural goods of one country can suddenly dry up if the government of another country imposes trade restrictions against its products, and prices can fall. Such dramatic shifts in prices and quantities make incomes of farmers unstable.

Figure 4.9 Supply and Demand Shifts for Agricultural Products



A relatively large increase in the supply of agricultural products, accompanied by a relatively small increase in demand, has reduced the price received by farmers and increased the quantity of agricultural goods.

The Great Depression of the 1930s led to a major federal role in agriculture. The Depression affected the entire economy, but it hit farmers particularly hard. Prices received by farmers plunged nearly two-thirds from 1930 to 1933. Many farmers had a tough time keeping up mortgage payments. By 1932, more than half of all farm loans were in default.

Farm legislation passed during the Great Depression has been modified many times, but the federal government has continued its direct involvement in agricultural markets. This has meant a variety of government programs that guarantee a minimum price for some types of agricultural products. These programs have been accompanied by government purchases of any surplus, by requirements to restrict acreage in order to limit those surpluses, by crop or production restrictions, and the like.

To see how such policies work, look back at [Figure 4.8 "Price Floors in Wheat Markets"](). At P_F, W_2 bushels of wheat will be supplied. With that much wheat on the market, there is market pressure on the price of wheat to fall. To prevent price from falling, the government buys the surplus of (W_2 – W_1) bushels of wheat, so that only W_1 bushels are actually available to private consumers for purchase on the market. The government can store the surpluses or find special uses for them. For example, surpluses generated in the United States have been shipped to developing countries as grants-in-aid or distributed to local school lunch programs. As a variation on this program, the government can require farmers who want to participate in the price support program to reduce acreage in order to limit the size of the surpluses.

After 1973, the government stopped buying the surpluses (with some exceptions) and simply guaranteed farmers a "target price." If the average market price for a crop fell below the crop's target price, the government paid the difference. If, for example, a crop had a market price of $3 per unit and a target price of $4 per unit, the government would give farmers a payment of $1 for each unit sold. Farmers would thus receive the market price of $3 plus a government payment of $1 per unit. For farmers to receive these payments, they had to agree to remove acres from production and to comply with certain conservation provisions. These restrictions sought to reduce the size of the surplus generated by the target price, which acted as a kind of price floor.

What are the effects of such farm support programs? The intention is to boost and stabilize farm incomes. But, with price floors, consumers pay more for food than they would otherwise, and governments spend heavily to finance the programs. With the target price approach, consumers pay less, but government financing of the program continues. U.S. federal spending for agriculture averaged well over $22 billion per year between 2003 and 2007, roughly $70 per person.

Help to farmers has sometimes been justified on the grounds that it boosts incomes of "small" farmers. However, since farm aid has generally been allotted on the basis of how much farms produce rather than on a per-farm basis, most federal farm support has gone to the largest farms. If the goal is to eliminate poverty among farmers, farm aid could be redesigned to supplement the incomes of small or poor farmers rather than to undermine the functioning of agricultural markets.

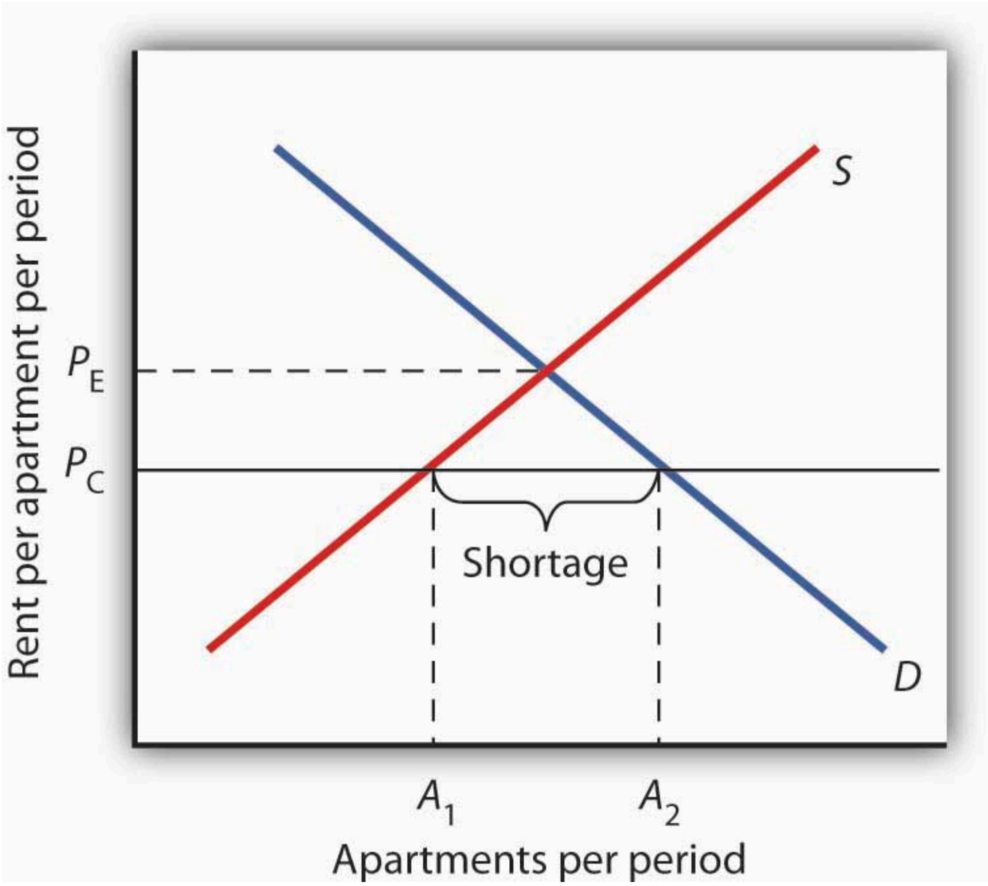
In 1996, the U.S. Congress passed the Federal Agriculture Improvement and Reform Act of 1996, or FAIR. The thrust of the new legislation was to do away with the various programs of price support for most crops and hence provide incentives for farmers to respond to market price signals. To protect farmers through a transition period, the act provided for continued payments that were scheduled to decline over a seven-year period. However, with prices for many crops falling in 1998, the U.S. Congress passed an emergency aid package that increased payments to farmers. In 2008, as farm prices reached record highs, Congress passed a farm bill that increased subsidy payments to $40 billion. It did, however, for the first time limit payments to the wealthiest farmers. Individual farmers whose farm incomes exceed $750,000 (or $1.5 million for couples) would be ineligible for some subsidy programs.

Rental Price Ceilings

The purpose of rent control is to make rental units cheaper for tenants than they would otherwise be. Unlike agricultural price controls, rent control in the United States has been largely a local phenomenon, although there were national rent controls in effect during World War II. Currently, about 200 cities and counties have some type of rent control provisions, and about 10% of rental units in the United States are now subject to price controls. New York City's rent control program, which began in 1943, is among the oldest in the country. Many other cities in the United States adopted some form of rent control in the 1970s. Rent controls have been pervasive in Europe since World War I, and many large cities in poorer countries have also adopted rent controls.

Rent controls in different cities differ in terms of their flexibility. Some cities allow rent increases for specified reasons, such as to make improvements in apartments or to allow rents to keep pace with price increases elsewhere in the economy. Often, rental housing constructed after the imposition of the rent control ordinances is exempted. Apartments that are vacated may also be decontrolled. For simplicity, the model presented here assumes that apartment rents are controlled at a price that does not change.

Figure 4.10 Effect of a Price Ceiling on the Market for Apartments



A price ceiling on apartment rents that is set below the equilibrium rent creates a shortage of apartments equal to (A_2 – A_1) apartments.

Figure 4.10 "Effect of a Price Ceiling on the Market for Apartments" shows the market for rental apartments. Notice that the demand and supply curves are drawn to look like all the other demand and supply curves you have encountered so far in this text: the demand curve is downward-sloping and the supply curve is upward-sloping.

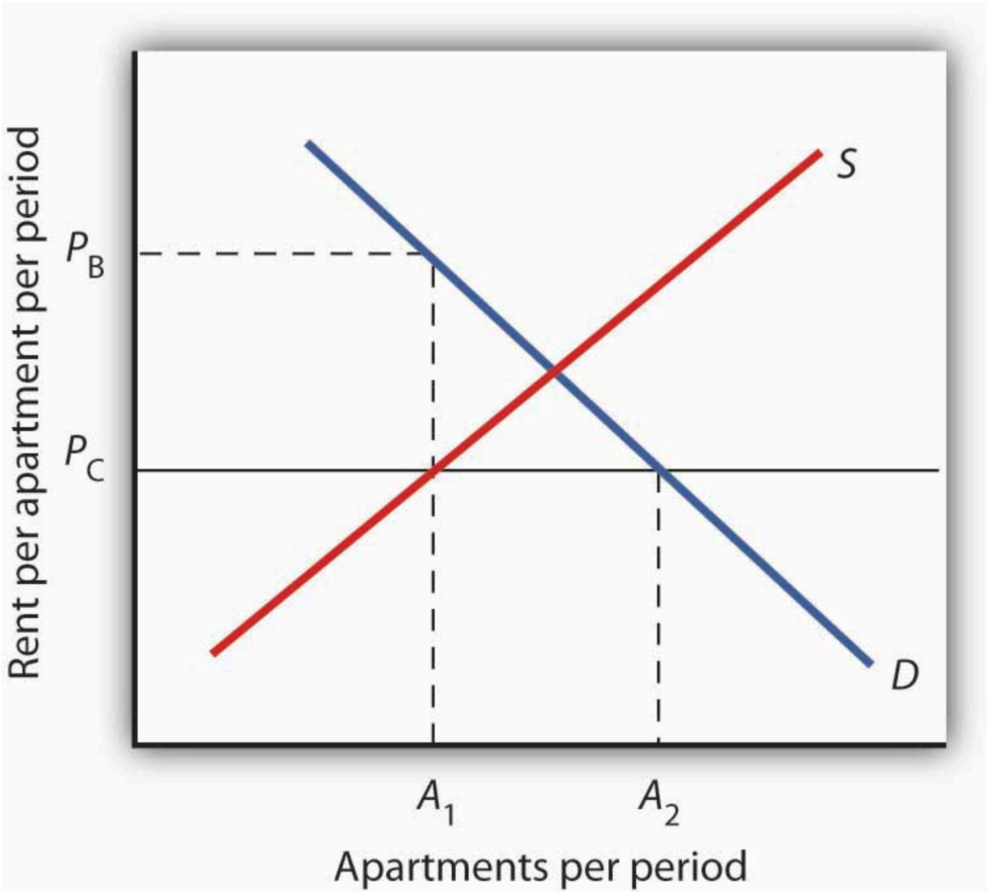
The demand curve shows that a higher price (rent) reduces the quantity of apartments demanded. For example, with higher rents, more young people will choose to live at home with their parents. With lower rents, more will choose to live in apartments. Higher rents may encourage more apartment sharing; lower rents would induce more people to live alone.

The supply curve is drawn to show that as rent increases, property owners will be encouraged to offer more apartments to rent. Even though an aerial photograph of a city would show apartments to be fixed at a point in time, owners of those properties will decide how many to rent depending on the amount of rent they anticipate. Higher rents may also induce some homeowners to rent out apartment space. In addition, renting out apartments implies a certain level of service to renters, so that low rents may lead some property owners to keep some apartments vacant.

Rent control is an example of a price ceiling, a maximum allowable price. With a price ceiling, the government forbids a price above the maximum. A price ceiling that is set below the equilibrium price creates a shortage that will persist.

Suppose the government sets the price of an apartment at P_C in Figure 4.10 "Effect of a Price Ceiling on the Market for Apartments". Notice that P_C is below the equilibrium price of P_E. At P_C, we read over to the supply curve to find that sellers are willing to offer A_1 apartments. Reading over to the demand curve, we find that consumers would like to rent A_2 apartments at the price ceiling of P_C. Because P_C is below the equilibrium price, there is a shortage of apartments equal to (A_2 – A_1). (Notice that if the price ceiling were set above the equilibrium price it would have no effect on the market since the law would not prohibit the price from settling at an equilibrium price that is lower than the price ceiling.)

Figure 4.11 The Unintended Consequences of Rent Control



Controlling apartment rents at P_C creates a shortage of (A_2 – A_1) apartments. For A_1 apartments, consumers are willing and able to pay P_B, which leads to various "backdoor" payments to apartment owners.

If rent control creates a shortage of apartments, why do some citizens nonetheless clamor for rent control and why do governments often give in to the demands? The reason generally given for rent control is to keep apartments affordable for low- and middle-income tenants.

But the reduced quantity of apartments supplied must be rationed in some way, since, at the price ceiling, the quantity demanded would exceed the quantity supplied. Current occupants may be reluctant to leave their dwellings because finding other apartments will be difficult. As apartments do become available, there will be a line of potential renters waiting to fill them, any of whom is willing to pay the controlled price of P_C or more. In fact, reading up to the demand curve in Figure 4.11 "The Unintended Consequences of Rent Control" from A_1 apartments, the quantity available at P_C, you can see that for A_1 apartments, there are potential renters willing and able to pay P_B. This often leads to various "backdoor" payments to apartment owners, such as large security deposits, payments for things renters may not want (such as furniture), so-called "key" payments ("The monthly rent is $500 and the key price is $3,000"), or simple bribes.

In the end, rent controls and other price ceilings often end up hurting some of the people they are intended to help. Many people will have trouble finding apartments to rent. Ironically, some of those who do find apartments may actually end up paying more than they would have paid in the absence of rent control. And many of the people that the rent controls do help (primarily current occupants, regardless of their income, and those lucky enough to find apartments) are not those they are intended to help (the poor). There are also costs in government administration and enforcement.

Because New York City has the longest history of rent controls of any city in the United States, its program has been widely studied. There is general agreement that the rent control program has reduced tenant mobility, led to a substantial gap between rents on controlled and uncontrolled units, and favored long-term residents at the expense of newcomers to the city (Arnott, R., 1995). These distortions have grown over time, another frequent consequence of price controls.

A more direct means of helping poor tenants, one that would avoid interfering with the functioning of the market, would be to subsidize their incomes. As with price floors, interfering with the market mechanism may solve one problem, but it creates many others at the same time.

Key Takeaways

- Price floors create surpluses by fixing the price above the equilibrium price. At the price set by the floor, the quantity supplied exceeds the quantity demanded.
- In agriculture, price floors have created persistent surpluses of a wide range of agricultural commodities. Governments typically purchase the amount of the surplus or impose production restrictions in an attempt to reduce the surplus.
- Price ceilings create shortages by setting the price below the equilibrium. At the ceiling price, the quantity demanded exceeds the quantity supplied.
- Rent controls are an example of a price ceiling, and thus they create shortages of rental housing.
- It is sometimes the case that rent controls create "backdoor" arrangements, ranging from requirements that tenants rent items that they do not want to outright bribes, that result in rents higher than would exist in the absence of the ceiling.

Try It!

A minimum wage law is another example of a price floor. Draw demand and supply curves for unskilled labor. The horizontal axis will show the quantity of unskilled labor per period and the vertical axis will show the hourly wage rate for unskilled workers, which is the price of unskilled labor. Show and explain the effect of a minimum wage that is above the equilibrium wage.

Case in Point: Corn: It Is Not Just Food Any More

Figure 4.12

Herry Lawford – Stocks – CC BY 2.0.

Government support for corn dates back to the Agricultural Act of 1938 and, in one form or another, has been part of agricultural legislation ever since. Types of supports have ranged from government purchases of surpluses to target pricing, land set asides, and loan guarantees. According to one estimate, the U.S. government spent nearly $42 billion to support corn between 1995 and 2004.

Then, during the period of rising oil prices of the late 1970s and mounting concerns about dependence on foreign oil from volatile regions in the world, support for corn, not as a food, but rather as an input into the production of ethanol—an alternative to oil-based fuel—began. Ethanol tax credits were part of the Energy Act of 1978. Since 1980, a tariff of 50¢ per gallon against imported ethanol, even higher today, has served to protect domestic corn-based ethanol from imported ethanol, in particular from sugar-cane-based ethanol from Brazil.

The Energy Policy Act of 2005 was another milestone in ethanol legislation. Through loan guarantees, support for research and development, and tax credits, it mandated that 4 billion gallons of ethanol be used by 2006 and 7.5 billion gallons by 2012. Ethanol production had already reached 6.5 billion gallons by 2007, so new legislation in 2007 upped the ante to 15 billion gallons by 2015.

Beyond the increased amount the government is spending to support corn and corn-based ethanol, criticism of the policy has three major prongs:

1. Corn-based ethanol does little to reduce U.S. dependence on foreign oil because the energy required to produce a gallon of corn-based ethanol is quite high. A 2006 National Academy of Sciences paper estimated that one gallon of ethanol is needed to bring 1.25 gallons of it to market. Other studies show an even less favorable ratio.
2. Biofuels, such as corn-based ethanol, are having detrimental effects on the environment, with increased deforestation, stemming from more land being used to grow fuel inputs, contributing to global warming.
3. The diversion of corn and other crops from food to fuel is contributing to rising food prices and an increase in world hunger. C. Ford Runge and Benjamin Senauer wrote in *Foreign Affairs* that even small

> increases in prices of food staples have severe consequences on the very poor of the world, and “Filling the 25-gallon tank of an SUV with pure ethanol requires over 450 pounds of corn—which contains enough calories to feed one person for a year.”

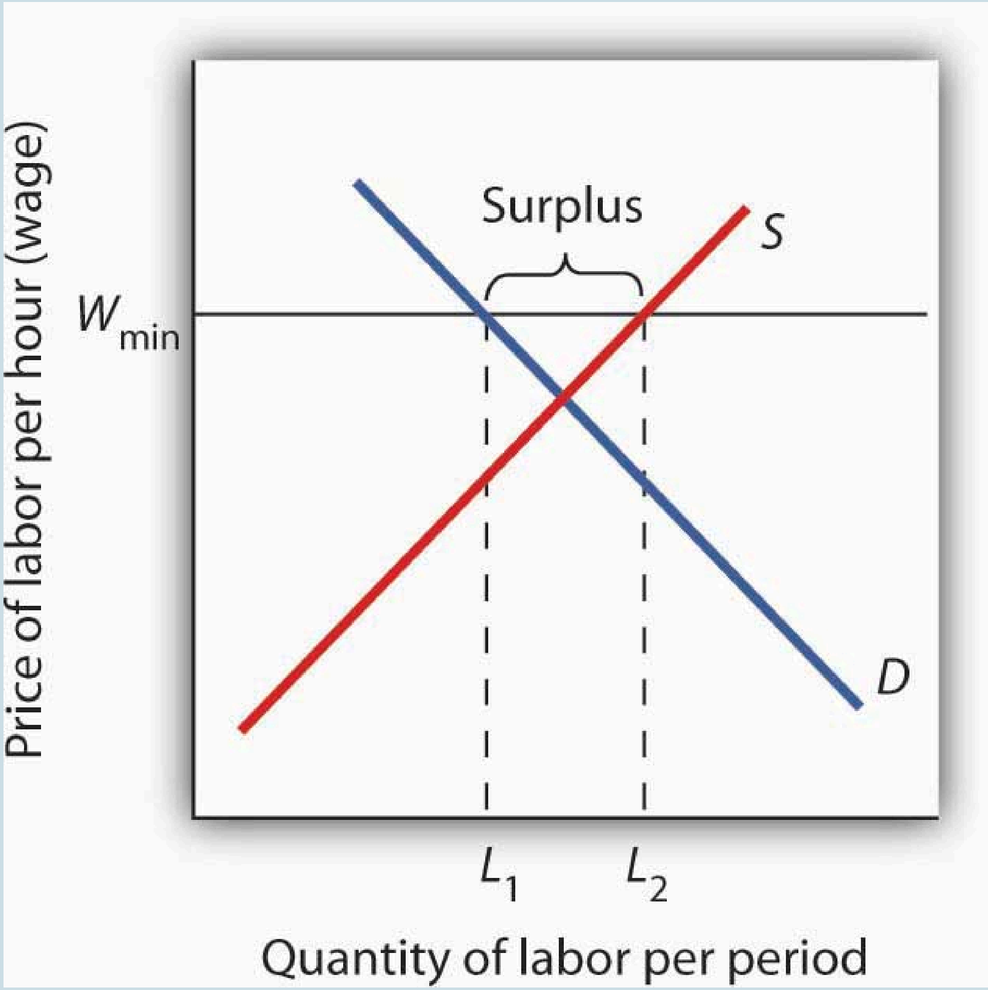
Some of these criticisms may be contested as exaggerated: Will the ratio of energy-in to energy-out improve as new technologies emerge for producing ethanol? Did not other factors, such as weather and rising food demand worldwide, contribute to higher grain prices? Nonetheless, it is clear that corn-based ethanol is no free lunch. It is also clear that the end of government support for corn is nowhere to be seen.

Sources: Alexei Barrionuevo, “Mountains of Corn and a Sea of Farm Subsidies,” *New York Times*, November 9, 2005, online version; David Freddoso, “Children of the Corn,” National Review Online, May 6, 2008; C. Ford Runge and Benjamin Senauer, “How Biofuels Could Starve the Poor,” *Foreign Affairs*, May/June 2007, online version; Michael Grunwald, “The Clean Energy Scam,” *Time* 171:14 (April 7, 2008): 40–45.

Answer to Try It! Problem

A minimum wage (W_{min}) that is set above the equilibrium wage would create a surplus of unskilled labor equal to ($L_2 - L_1$). That is, L_2 units of unskilled labor are offered at the minimum wage, but companies only want to use L_1 units at that wage. Because unskilled workers are a substitute for a skilled workers, forcing the price of unskilled workers higher would increase the demand for skilled labor and thus increase their wages.

Figure 4.13



References

Arnott, R., "Time for Revisionism on Rent Control," *Journal of Economic Perspectives* 9(1) (Winter, 1995): 99–120.

4.3 Review and Practice

Summary

In this chapter we used the tools of demand and supply to understand a wide variety of market outcomes. We learned that technological change and the entry of new sellers has caused the supply curve of personal computers to shift markedly to the right, thereby reducing equilibrium price and increasing equilibrium quantity. Market forces have made personal computers a common item in offices and homes.

Crude oil and gasoline prices soared in 2008 and then fell back. We looked at the causes of these increases as well as their impacts. Crude oil prices rose in large part As a result of increased demand, particularly from China. Higher prices for crude oil led to higher prices for gasoline. Those higher prices not only hurt consumers of gasoline, they also put upward pressure on the prices of a wide range of goods and services. Crude oil and gasoline prices then decreased dramatically in the last part of 2008, as world growth declined.

The model of demand and supply also explains the determination of stock prices. The price per share of corporate stock reflects the market's estimate of the expected profitability of the firm. Any information about the firm that causes potential buyers or current owners of corporate stock to reevaluate how profitable they think the firm is, or will be, will cause the equilibrium price of the stock to change.

We then examined markets in which some form of government price control keeps price permanently above or below equilibrium. A price floor leads to persistent surpluses because it is set above the equilibrium price, whereas a price ceiling, because it is set below the equilibrium price, leads to persistent shortages. We saw that interfering with the market mechanism may solve one problem but often creates other problems at the same time. We discussed what some of these unintended consequences might be. For example, agricultural price floors aimed at boosting farm income have also raised prices for consumers and cost taxpayers dearly, and the bulk of government payments have gone to large farms. Rent controls have lowered rents, but they have also reduced the quantity of rental housing supplied, created shortages, and sometimes led to various forms of "backdoor" payments, which sometimes force the price of rental housing above what would exist in the absence of controls.

Concept Problems

1. Like personal computers, digital cameras have become a common household item. Digital camera prices have plunged in the last 10 years. Use the model of demand and supply to explain the fall in price and increase in quantity.
2. Enron Corp. was one of several corporations convicted of fraud in its accounting practices during the early part of this decade. It had created dummy corporations to hide massive borrowing and to give it the appearance of extraordinary profitability. Use the model of demand and supply to explain the likely

impact of such convictions on the stocks of other corporations.

3. During World War II there was a freeze on wages, and corporations found they could evade the freeze by providing other fringe benefits such as retirement funds for their employees. The Office of Price Administration, which administered the wage freeze, ruled that the offer of retirement funds was not a violation of the freeze. The Internal Revenue Service went along with this and ruled that employer-financed retirement plans were not taxable income. Was the wage freeze an example of a price floor or a price ceiling? Use the model of demand and supply to explain why employers began to offer such benefits to their employees.
4. The text argues that political instability in potential suppliers of oil such as Iraq and Venezuela accounts for a relatively steep supply curve for crude oil such as the one shown in Figure 4.2 "The Increasing Demand for Crude Oil" Suppose that this instability eases considerably and that the world supply curve for crude oil becomes much flatter. Draw such a curve, and explain its implications for the world economy and for typical consumers.
5. Suppose that technological change affects the dairy industry in the same way it has affected the computer industry. However, suppose that dairy price supports remain in place. How would this affect government spending on the dairy program? Use the model of demand and supply to support your answer.
6. People often argue that there is a "shortage" of child care. Using the model of demand and supply, evaluate whether this argument is likely to be correct.
7. "During most of the past 50 years the United States has had a surplus of farmers, and this has been the root of the farm problem." Comment.
8. Suppose the Department of Agriculture ordered all farmers to reduce the acreage they plant by 10%. Would you expect a 10% reduction in food production? Why or why not?
9. The text argues that the increase in gasoline prices had a particularly strong impact on low-income people. Name some other goods and services for which a sharp increase in price would have a similar impact on people with low incomes.
10. Suppose that the United States and the European Union impose a price ceiling on crude oil of $25 per barrel. Explain, and illustrate graphically, how this would affect the markets for crude oil and for gasoline in the United States and in the European Union.
11. Given that rent controls can actually hurt low-income people, devise a housing strategy that would provide affordable housing for those whose incomes fall below the poverty line (in 2004, this was about $19,000 for a family of four).

Numerical Problems

Problems 1–4 are based on the following demand and supply schedules for corn (all quantities are in millions of bushels per year).

Price per bushel	Quantity demanded	Quantity supplied
$0	6	0
1	5	1
2	4	2
3	3	3
4	2	4
5	1	5
6	0	6

1. Draw the demand and supply curves for corn. What is the equilibrium price? The equilibrium quantity?
2. Suppose the government now imposes a price floor at $4 per bushel. Show the effect of this program graphically. How large is the surplus of corn?
3. With the price floor, how much do farmers receive for their corn? How much would they have received if there were no price floor?
4. If the government buys all the surplus wheat, how much will it spend?

Problems 5–9 are based on the following hypothetical demand and supply curves for apartments

Rent/Month	Number of Apts. Demanded/Month	Number of Apts. Supplied/Month
$0	120,000	0
200	100,000	20,000
400	80,000	40,000
600	60,000	60,000
800	40,000	80,000
1000	20,000	100,000
1200	0	120,000

1. Draw the demand and supply curves for apartments.
2. What is the equilibrium rent per month? At this rent, what is the number of apartments demanded and supplied per month?
3. Suppose a ceiling on rents is set at $400 per month. Characterize the situation that results from this policy.
4. At the rent ceiling, how many apartments are demanded? How many are supplied?
5. How much are people willing to pay for the number of apartments supplied at the ceiling? Describe the arrangements to which this situation might lead.

CHAPTER 5: MACROECONOMICS: THE BIG PICTURE

Start Up: Financial Crisis Batters Economy

The world economy recovered after the recession of 2001. Growth had been fairly rapid, unemployment had stayed low, and inflation seemed to be under control. However, at the end of 2007 the US economy started to unravel, taking with it the largest economies around the world. The economy of the United States and those of much of the world were rocked by the worst financial crisis in nearly 80 years. That crisis plunged the economy into a downturn in total output and employment that seemed likely to last a long time.

In the US, a good deal of the economy's momentum when things were going well had been fueled by rising house prices. Between 1995 and 2007, housing prices in the United States more than doubled. As house prices rose, consumers who owned houses grew richer and increased their consumption purchases. That helped fuel economic growth. The boom in housing prices had been encouraged by policies of the nation's monetary authority, the Federal Reserve, which had shifted to an expansionary monetary policy that held short-term interest rates below the inflation rate. Another development, subprime mortgages—mortgage loans to buyers whose credit or income would not ordinarily qualify for mortgage loans—helped bring on the ultimate collapse. When they were first developed, subprime mortgage loans seemed a hugely profitable investment for banks and a good deal for home buyers. Financial institutions developed a wide range of instruments based on "mortgage-backed securities." As long as house prices kept rising, the system worked and was profitable for virtually all players in the mortgage market. Many firms undertook investments in mortgage-backed securities that assumed house prices would keep rising. Large investment banks bet heavily that house prices would continue rising. Powerful members of Congress pressured two government-sponsored enterprises, Fannie Mae (the Federal National Mortgage Association) and Freddie Mac (the Federal Home Loan Mortgage Corporation), to be even more aggressive in encouraging banks to make mortgage loans to low-income families. The pressure came from the executive branch of government as well—under both Democratic and Republican administrations. In 1996, the Department of Housing and Urban Development (under Bill Clinton, a Democrat) required that 12% of mortgages purchased by Fannie Mae and Freddie Mac be for households with incomes less than 60% of the median income in their region. That target was increased to 20% in 2000, 22% in 2005 (now under George W. Bush, a Republican), and was to have increased to 28% in 2008.Russell Roberts, "How Government Stoked the Mania," Wall Street Journal, October 3, 2008, p. A21. But that final target would not be reached, as both Fannie Mae and Freddie Mac were seized by the government in 2008. To top things off, a loosening in bank and investment bank regulations gave financial institutions greater leeway in going overboard with purchases of mortgage-backed securities. As house prices began falling in 2007, a system based on the assumption they would continue rising began to unravel very fast. The investment bank Bear Stearns and insurance company American International Group (AIG) required massive infusions of federal money to keep them afloat. In September of 2008, firm after firm with assets tied to mortgage-backed securities began to fail. In some cases, the government rescued them; in other cases, such as Lehman Brothers, they were allowed to fail.

The financial crisis had dramatic and immediate effects on the economy. The economy's total output, which had been growing through the first half of 2008, fell at an annual rate of 0.5% in the third quarter, according to advance estimates by the Bureau of Economic Analysis. Consumers, having weathered higher gasoline prices and higher food prices for most of the year, reduced their consumption expenditures as the value of their houses and the stocks they held plunged—consumption fell at an annual rate of 3.7% in the third quarter. As cold fear gripped financial markets and expectations of further slowdown ensued, firms cut down on investment spending, which includes spending on plant and equipment used in production. While this nonresidential investment component of output fell at an annual rate of 1.5%, the residential component fell even faster as housing investment sank at an annual rate of 17.6%. Government purchases and net exports rose, but not enough to offset reductions in consumption and private investment.Bureau of Economic Analysis, press release, November 25, 2008. As output shrank, unemployment rose. Through the first nine months of 2008 there was concern that price levels in the United States and in most of the world economies were rising rapidly, but toward the end of the year the concern shifted to whether or not the price level might fall.

This recession, which officially began in December 2007 and ended in June 2009, was brutal: At 18 months in length, it was the longest U.S. recession since World War II. The nation's output fell over 4%. The unemployment rate rose dramatically, hitting 10% at the end of 2009, and remaining above 9% throughout 2010. In 2010, to many people it certainly did not feel that the so-called Great Recession had really ended.

And in Canada………

The following remarks are from "The "Great" Recession in Canada: Perception vs. Reality" by the Former Deputy Governor Jean Boivin (2010-2012). Published March 28, 2011

Phase One: Sudden Slowdown

The financial crisis was expected to have a significant impact in Canada, and for the first phase of the cycle, this was certainly the case.

During the this recession, GDP declined by 3.3 per cent over three quarters. In contrast, over the same period of time in the 1980s and the 1990s, it fell by 2.2 per cent and 1.9 per cent, respectively.

A prominent feature of the recent recession was the **spectacular drop in exports.** Exports were harder hit than in any previous recession, decreasing by 16 per cent over three quarters, while the most significant drop during the recessions of the 1980s and 1990s was only 8 per cent (Chart 4).

Investments were equally hard hit by the recession. There was a 22 per cent downturn in investments over just three quarters (Chart 5). Nothing like this has ever been seen. It took two years during the 1980s recession, and three years during the 1990s recession, before a downturn of comparable magnitude was recorded. This recent decline in investment is partly due to the exceptionally high levels of uncertainty haunting the global economy.

In sum, the recent recession was different from previous ones, owing to a more pronounced slowdown triggered by unusually steep drops in exports and investment. During its initial phase, the effects of the crisis in Canada—albeit to a somewhat lesser degree—were comparable to those in the United States and showed real signs of becoming a "Great Recession" (Chart 6).

Phase Two: Rapid Recovery

Despite the rapid slowdown, the recovery was faster than those that followed previous recessions. Why?

Neither exports nor investments can provide the answer. While GDP has recovered to pre-recession levels, business investment and exports have only recovered 45 per cent and 67 per cent, respectively, of the losses incurred during the recession.

If the recovery was speedier, despite weaker contributions from investment and exports, support for the recovery must have come from household and government spending. This was indeed the case. Household spending declined by only 2 per cent between 2009 and 2010, compared with 6 per cent during the previous two recessions. The contribution of government spending to growth was more than one percentage point in each year.

The greater strength of household and government spending reflects Canada's favourable position at the outset of the recession. Major adjustments had been made to the structure of the Canadian economy. Business and household balance sheets were relatively sound, and the banking system was robust, managed prudently, and sufficiently capitalized. Canada's monetary policy framework had been effective and was credible. The fiscal situation was favourable, and the social safety net and regulatory framework were effective. As well, household spending was boosted by the prosperity arising from strong demand for our natural resources and by improved terms of trade.

This favourable position gave Canada the flexibility it needed to respond strongly to the crisis without compromising the credibility of our public policy frameworks. Thanks to the expansionary monetary and fiscal measures adopted in concert with other G-20 countries, Canada was able to support domestic demand which contributed significantly to the economic recovery.

Important Lingering Issues

In Canada, then, we had room to manoeuvre to help us effectively absorb the aftershocks of the global economic crisis. It is essential to maintain this buffer in light of the elevated risks that still exist worldwide and the structural issues that persist in the Canadian economy, even after the recession. The standard of living that we will be able to sustain in the medium term will depend, in fact, on our ability to address these issues.

Allow me to address three of these issues: household indebtedness, international competitiveness and, more importantly, our productivity.

Household Indebtedness

Since the beginning of the recovery, household credit has increased at twice the rate of personal disposable income. In the autumn of 2010, Canadian household debt climbed to an unprecedented level of 147 per cent of disposable income (Chart 7).

The relatively healthy financial condition of Canadian households at the beginning of the "Great" Recession helped the Canadian economy to better withstand the initial shocks of the crisis. However, going forward, it is essential to maintain the necessary room to manoeuvre to keep household spending on a viable path. This leads us to believe that the rate of household spending will more closely correspond to future earnings, and certain signs to that effect have already been observed.

Canada's International Competitiveness

The second issue is our ability to compete internationally. The slow recovery of exports is due in part to the sluggishness of global economic activity. It is also due to the continued erosion of Canadian business competitiveness over the past ten years. This erosion can be attributed to the appreciation of the Canadian dollar and Canada's poor productivity performance. Thus, Canadian exporters are seeing their market shares for a wide range of goods drop in the U.S.

market—by far the most important market for Canada—while exporters in other countries, such as China and Mexico, are gaining ground (Chart 8).

As global economic growth continues to take root, we are seeing early evidence of a recovery in net exports. But, at this point, exports are still weak when compared with previous recessions. And in a world of growing international competition, we should not assume that the forces causing the erosion of competitiveness through the previous decade will simply fade away because of a global recovery.

This situation highlights the need to diversify our export markets and increase our ability to compete, not only with American producers, but also with other foreign exporters.

Productivity and Investment

This brings us to the third issue. As I just discussed, international competitiveness is based on our ingenuity, the efficiency with which we produce, or, for short, productivity. But beyond its influence on international competitiveness, productivity is a fundamental determinant of our economic well-being. To improve productivity, we need investment.

The slow recovery of investment in this cycle is particularly surprising in light of relatively favourable financial conditions: interest rates remain low, and the exchange rate facilitates imports of machinery and equipment.

The elevated level of uncertainty experienced during the recession, especially from a global perspective, was a major hindrance to business investment. This uncertainty was not confined to our borders: the link between uncertainty and business investment was clearly evident in the economies of the United States, Germany and the United Kingdom. [7]

Yet heightened uncertainty is only part of the explanation. Although the recession in the United States was more serious and Americans faced at least the same degree of global uncertainty as we experienced in Canada, Canadian business investment in machinery and equipment lags behind that of the United States (Chart 9). In 2009, Canadian workers had access, on average, to approximately half the capital expenditures in machinery and equipment and information and communication technologies (ICT) of those available to their American counterparts. This is not a new phenomenon. In fact, between 1987 and 2009, Canadian investment in machinery and equipment and ICT per worker represented, on average, 77 per cent and 59 per cent, respectively, of similar American investments.

It is true that business investment started to recover at the end of 2009. Yet much progress remains to be made: less than half of the extraordinary drop in investments of the last recession have been recovered. With the increasing globalization of markets and the demographic challenges we face, maintaining our standard of living will require improved productivity. We must continue to innovate and to invest in promising projects.

Conclusion: Perception vs. Reality

We are fond of repeating the old adage: “An ounce of prevention is worth a pound of cure.” Recent experience expands the notion and shows that good prevention measures can also make the cure more effective. Before the Great Recession, Canada was able to protect itself by ensuring that it had room to manoeuvre to absorb the shocks of the crisis. The lessons we learned from the past were reflected in the adoption of sound public policy frameworks. A solid position, combined with the relatively healthy state of Canadian households, gave us the flexibility to withstand the worst effects of the global shock.

Future economists studying the 2007–09 recession in Canada may find it difficult to go beyond their first impressions and assess its true impact. Some will undoubtedly surmise that the economic activity of this time did indeed reflect, not only the extent of the shock, but also our ability to absorb it. The storm we weathered was a major one. We should not forget that it could have struck at a time when we were more vulnerable and less flexible. Things could have unfolded very differently, with disastrous results.

It is some comfort to know that, collectively, we were able to limit the damage. We must proceed with the strategy that has served us so well: continue to learn from our experiences to ensure better prevention and, when necessary, a better cure. For this, we must strive to deal with the issues that confront us with strength and determination.

Source: https://www.bankofcanada.ca/2011/03/great-recession-canada-perception-reality/

The Study of Macroeconomics

Output, employment, and the price level are the key variables in the study of macroeconomics, which is the analysis of aggregate values of economic variables. What determines a country's output, and why does output in some economies expand while in others it contracts? Why do some economies grow faster than others? What causes prices throughout an economy to fluctuate, and how do such fluctuations affect people? What causes employment and unemployment? Why does a country's unemployment rate fluctuate? Why do different countries have different unemployment rates?

We would pronounce an economy "healthy" if its annual output of goods and services were growing at a rate it can sustain, its price level stable, and its unemployment rate low. What would constitute "good" numbers for each of these variables depends on time and place, but those are the outcomes that most people would agree are desirable for the aggregate economy. When the economy deviates from what is considered good performance, there are often calls for the government to "do something" to improve performance. How government policies affect economic performance is a major topic of macroeconomics. When the financial and economic crises struck throughout the world in 2008, there was massive intervention from world central banks and from governments throughout the world in an effort to stimulate their economies.

This chapter provides a preliminary sketch of the most important macroeconomic issues: growth of total output and the business cycle, changes in the price level, and unemployment. Grappling with these issues will be important to you not only in your exploration of macroeconomics but throughout your life.

5.1 Growth of Real GDP and Business Cycles

Learning Objectives

1. Define real gross domestic product and explain how its calculation avoids both double-counting and the effects of changes in the price level.
2. Identify the phases of a business cycle.
3. Relate business cycles to the overall long-run trend in real GDP.

To determine whether the economy of a nation is growing or shrinking in size, economists use a measure of total output called real GDP. Real GDP, short for real gross domestic product, is the total value of all final goods and services produced during a particular year or period, adjusted to eliminate the effects of changes in prices. Let us break that definition up into parts.

Notice that only "final" goods and services are included in GDP. Many goods and services are purchased for use as inputs in producing something else. For example, a pizza parlor buys flour to make pizzas. If we counted the value of the flour and the value of the pizza, we would end up counting the flour twice and thus overstating the value of total production. Including only final goods avoids double-counting. If the flour is produced during a particular period but has not been sold, then it is a "final good" for that period and is counted.

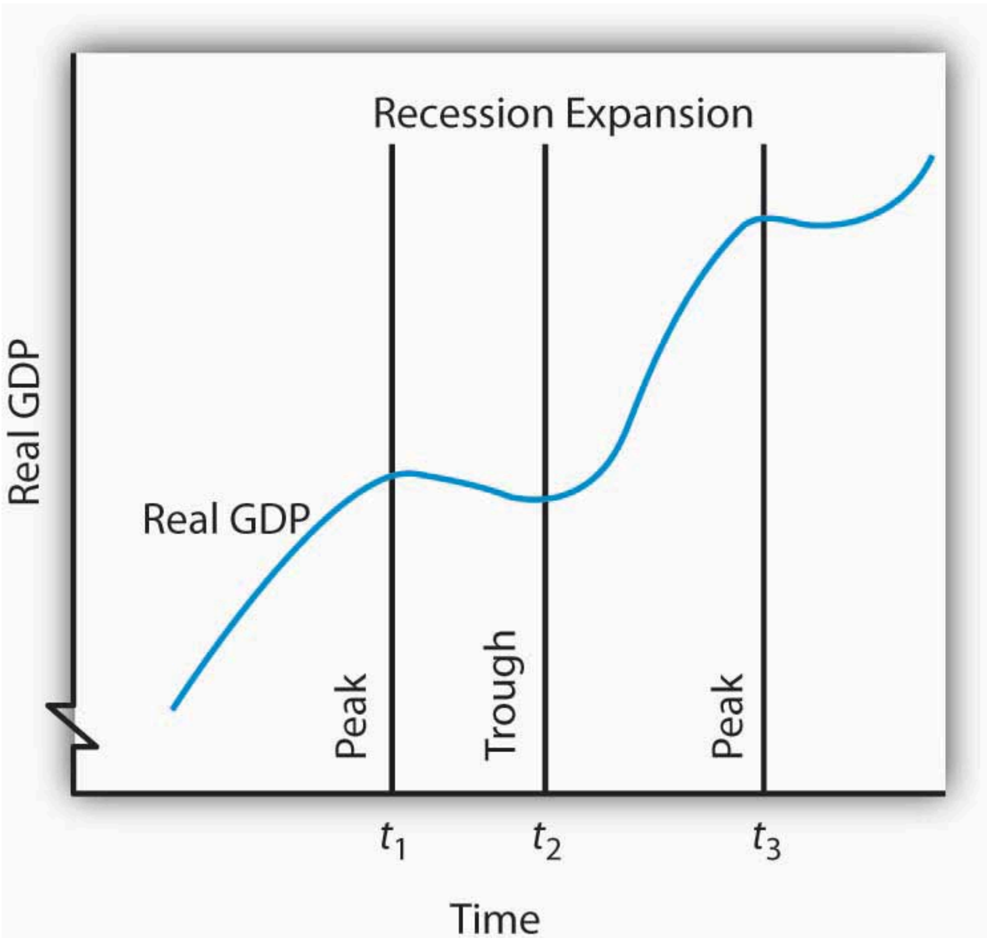
We want to determine whether the economy's output is growing or shrinking. If each final good or service produced, from hammers to haircuts, were valued at its current market price, and then we were to add the values of all such items produced, we would not know if the total had changed because output changed or because prices changed or both. The market value of all final goods and services produced can rise even if total output falls. To isolate the behaviour of total output only, we must hold prices constant at some level. For example, if we measure the value of basketball output over time using a fixed price for valuing the basketballs, then only an increase in the number of basketballs produced could increase the value of the contribution made by basketballs to total output. By making such an adjustment for basketballs and all other goods and services, we obtain a value for real GDP. In contrast, nominal GDP, usually just referred to as gross domestic product (GDP), is the total value of final goods and services for a particular period valued in terms of prices for that period. For example, real GDP fell in the third quarter of 2008. However, if the price level in a particular country is rising, nominal GDP will also rise.

We will save a detailed discussion of the computation of GDP for another chapter. In this section, our goal is to use the concept of real GDP to look at the business cycle—the economy's pattern of expansion, then contraction, then expansion again—and at growth of real GDP.

Phases of the Business Cycle

Figure 5.1 "Phases of the Business Cycle" shows a stylized picture of a typical business cycle. It shows that economies go through periods of increasing and decreasing real GDP, but that over time they generally move in the direction of increasing levels of real GDP. A sustained period in which real GDP is rising is an expansion; a sustained period in which real GDP is falling is a recession. Typically, an economy is said to be in a recession when real GDP drops for two consecutive quarters.

Figure 5.1 Phases of the Business Cycle



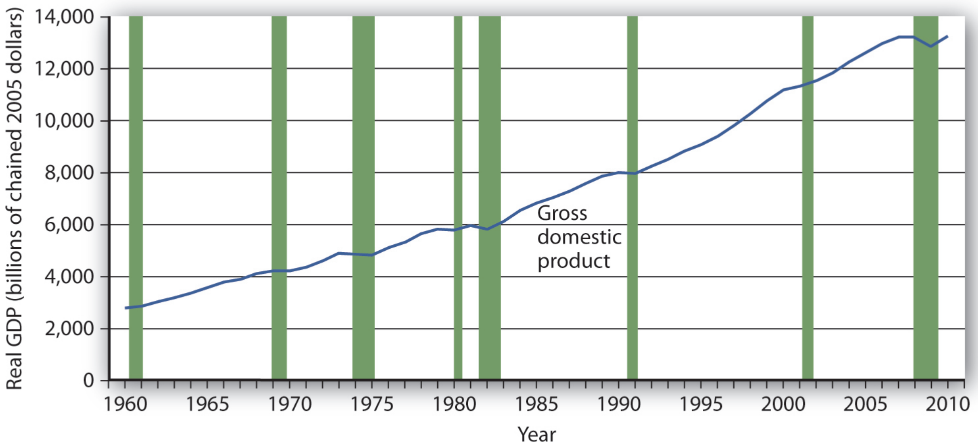
The business cycle is a series of expansions and contractions in real GDP. The cycle begins at a peak and continues through a recession, a trough, and an expansion. A new cycle begins at the next peak. Here, the first peak occurs at time t_1, the trough at time t_2, and the next peak at time t_3. Notice that there is a tendency for real GDP to rise over time.

At time t_1 in Figure 5.1 "Phases of the Business Cycle", an expansion ends and real GDP turns downward. The point at which an expansion ends and a recession begins is called the peak of the business cycle. Real GDP then falls during a period of recession. Eventually it starts upward again (at time t_2). The point at which a recession ends and an expansion begins is called the trough of the business cycle. The expansion continues until another peak is reached at time t_3[1]. A complete business cycle is defined by the passage from one peak to the next.

Business Cycles and the Growth of Real GDP in the United States

Figure 5.2 "Expansions and Recessions, 1960–2010" shows movements in real GDP in the United States from 1960 to 2010. Canada's pattern looks much the same, except the Canadian economy is about one-tenth the size of the U.S. Over this 50 year period, the economy experienced eight recessions shown by the shaded areas in the chart. Although periods of expansion have been more prolonged than periods of recession, we see the cycle of economic activity that characterizes economic life.

Figure 5.2 Expansions and Recessions, 1960–2010



The chart shows movements in real GDP since 1960. Recessions—periods of falling real GDP—are shown as shaded areas. On average, the annual rate of growth of real GDP over the period was 3.2% per year.

Source: Bureau of Economic Analysis, NIPA
Table 1.1.6. Real Gross Domestic Product, Chained Dollars [Billions of chained (2005) dollars]. Seasonally adjusted at annual rates.
Data for 2010 is through 3rd quarter.

Some interesting data on Canada's historical GDP levels can be found here:

https://tradingeconomics.com/canada/gdp

Real GDP clearly grew between 1960 and 2010. While the economy experienced expansions and recessions, its general trend during the period was one of rising real GDP. The average annual rate of growth of real GDP was about 3.2%.

During the post–World War II period, the average expansion has lasted 58 months, and the average recession has lasted about 11 months. The 2001 recession, which lasted eight months, was thus slightly shorter than the average. The 2007-2009 recession lasted 18 months; it was the longest of the post-World War II period.

Economists have sought for centuries to explain the forces at work in a business cycle. Not only are the currents that move the economy up or down intellectually fascinating but also an understanding of them is of tremendous practical importance. A business cycle is not just a movement along a curve in a textbook. It is new jobs for people, or the loss of them. It is new income, or the loss of it. It is the funds to build new schools or to provide better health care—or the lack of funds to do all those things. The story of the business cycle is the story of progress and plenty, of failure and sacrifice.

The effects of recessions extend beyond the purely economic realm and influence the social fabric of society as well. Suicide rates and property crimes—burglary, larceny, and motor vehicle theft tend to rise during recessions. Even popular music appears to be affected.

In our study of macroeconomics, we will gain an understanding of the forces at work in the business cycle. We will also explore policies through which the public sector might act to make recessions less severe and, perhaps, to prolong expansions. We turn next to an examination of price-level changes and unemployment.

Key Takeaways

- Real gross domestic product (real GDP) is a measure of the value of all final goods and services produced during a particular year or period, adjusted to eliminate the effects of price changes.
- The economy follows a path of expansion, then contraction, then expansion again. These fluctuations make up the business cycle.
- The point at which an expansion becomes a recession is called the peak of a business cycle; the point at which a recession becomes an expansion is called the trough.
- Over time, the general trend for most economies is one of rising real GDP. On average, real GDP in the United States has grown at a rate of over 3% per year since 1960.

Try It!

The data below show the behavior of real GDP in Turkey from the first quarter of 2001 through the third quarter of 2002. Use the data to plot real GDP in Turkey and indicate the phases of the business cycle.

Period	Real GDP (billions of New Turkish lira, 1987 prices)
First quarter, 2001	24.1
Second quarter, 2001	26.0
Third quarter, 2001	33.1
Fourth quarter, 2001	27.1
First quarter, 2002	24.6
Second quarter, 2002	28.3
Third quarter, 2002	35.7

Case in Point: The Art of Predicting Recessions

Figure 20.3

Reynermedia – Numbers And Finance – CC BY 2.0.

People who make a living tracking the economy and trying to predict its future do not do a very good job at predicting turning points in economic activity. The 52 economists surveyed by the *Wall Street Journal* each month did predict that the economy would slip into a recession in the third quarter of 2008. They made that prediction, however, in October—after the third quarter had ended. In September, the last month of the third quarter, the average forecast among the 52 economists had the economy continuing to grow through the third and fourth quarters of 2008. That September survey was taken before the financial crisis hit, a crisis that took virtually everyone by surprise. Of course, as we have already noted, the third-quarter downturn had not been identified as a recession by the NBER's Business Cycle Dating Committee as of November of 2008.

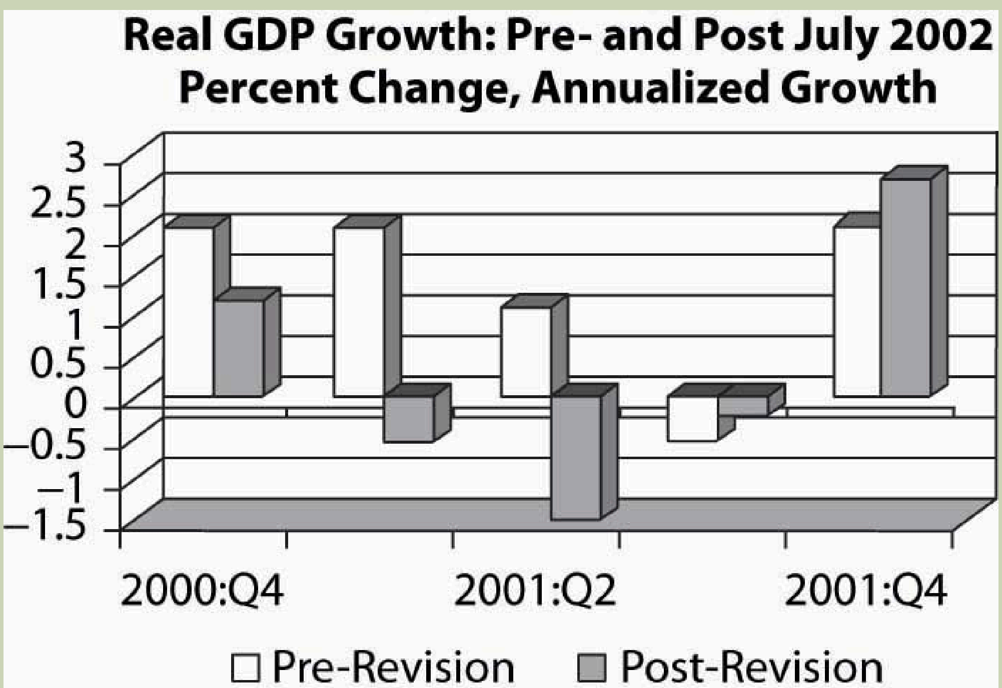
Predicting business cycle turning points has always been a tricky business. The experience of the recession of 2001 illustrates this. As the accompanying table shows, even as late as September 10, 2001, only 13 out of the 100 Blue Chip forecasters had answered in the affirmative to the question, "Has the United States slipped into a recession?" even though we now know the recession had begun the previous March. Comparing the data that were originally released by the U.S. Bureau of Economic Analysis shortly after the end of each quarter with the revised data that were released after July 2002 provides an important insight into explaining why the forecasters seem to have done so badly. As the graph on pre-revision and post-revision estimates of real GDP growth shows, the data released shortly after the end of each quarter showed an economy expanding through the second quarter of 2001, whereas the revised data show the economy contracting modestly in the first quarter of 2001 and then more forcefully in the second quarter. Only after the attacks on the World Trade Center in New York City and the Pentagon in Washington, D.C., on September 11, 2001, did most of the Blue Chip forecasters realize the economy was in recession.

The National Bureau of Economic Research (NBER) Business Cycle Dating Committee in November 2001 released a press announcement dating the onset of the recession as March 2001. The committee argued that "before the attacks of September 11, it is possible that the decline in the economy would have been too mild to qualify as a recession. The attacks clearly deepened the contraction and may have been an important factor in turning the episode into a recession." While surprising at the time, the revised data suggest that the committee made a good call.

This episode in economic history also points out the difference between the common definition of a recession as two consecutive quarters of declining real GDP and the NBER Dating Committee's continued insistence that it does not define a recession in this way. Rather the committee looks not only at real GDP but also at employment, income, and other factors. The behavior of employment during 2001 seems to have been an important factor in the November 2001 decision to proclaim March 2001 as the peak despite the misleading information on real GDP coming out of the Bureau of Economic Analysis at the time. The slow pickup in employment may also, though, have made it hesitate to call November 2001 the trough until July 2003.

Question posed: "Has the United States slipped into a recession?"		
Date	**Percent of Blue Chip responders answering "Yes"**	**Percent of Blue Chip responders answering "No"**
February 2001	5	95
June 2001	7	93
July 2001	13	87
August 2001	5	85
September 10, 2001	13	87
September 19, 2001	82	18

Figure 20.4 Real GDP Growth: Pre- and Post-July, 2002

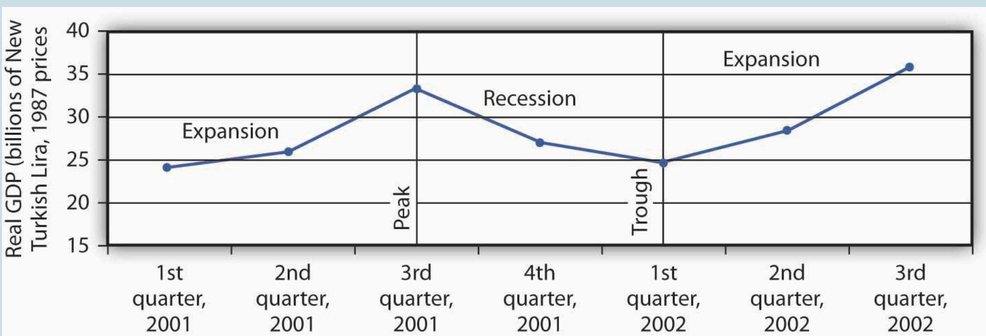


Sources: Phil Izzo, "Economists Expect Crisis to Deepen," *Wall Street Journal Online*, October 10, 2008; Kevin L.

Kliesen, "The 2001 Recession: How Was It Different and What Developments May Have Caused It?" *Federal Reserve Bank of St. Louis Review*, September/October 2003: 23–37; http://www.nber.org/cycles/; "Press Release," Business Cycle Dating Committee, National Bureau of Economic Research, press release, Cambridge, Massachusetts, July 17, 2002.

Answer to Try It! Problem

Figure 5.5



[1]Some economists prefer to break the expansion phase into two parts. The recovery phase is said to be the period between the previous trough and the time when the economy achieves its previous peak level of real GDP. The "expansion" phase is from that point until the following peak.

References

Greenhouse, S., "Job Market Gets Better for U.S. Graduates, But Only Slightly: Offers Will Increase 5% Over Last Year; Average Starting Salaries Are Down," *The International Herald Tribune*, May 26, 2010, Finance 16.

Lewin, T., "A Hemline Index, Updated," *New York Times*, October 19, 2008, Section WK, 1.

5.2 Price-Level Changes

Learning Objectives

1. Define inflation and deflation, explain how their rates are determined, and articulate why price-level changes matter.
2. Explain what a price index is and outline the general steps in computing a price index.
3. Describe and compare different price indexes.
4. Explain how to convert nominal values to real values and explain why it is useful to make this calculation.

Concern about changes in the price level has always dominated economic discussion. With inflation in Canada generally averaging only between 1% and 3% each year since 1996, it may seem surprising how much attention the behaviour of the price level still commands. Yet inflation was a concern in 2004 when there was fear that the rising price of oil could trigger higher prices in other areas. Just the year before, when inflation fell below 2%, there was talk about the risk of deflation. That did not happen; prices continued rising. Inflation rose substantially in the first half of 2008, renewing fears about subsequent further increases. 2010 brought on renewed concern of possible deflation. Just what are inflation and deflation? How are they measured? And most important, why do we care? These are some of the questions we will explore in this section.

Inflation is an increase in the average level of prices, and deflation is a decrease in the average level of prices. In an economy experiencing inflation, most prices are likely to be rising, whereas in an economy experiencing deflation, most prices are likely to be falling.

There are two key points in these definitions:

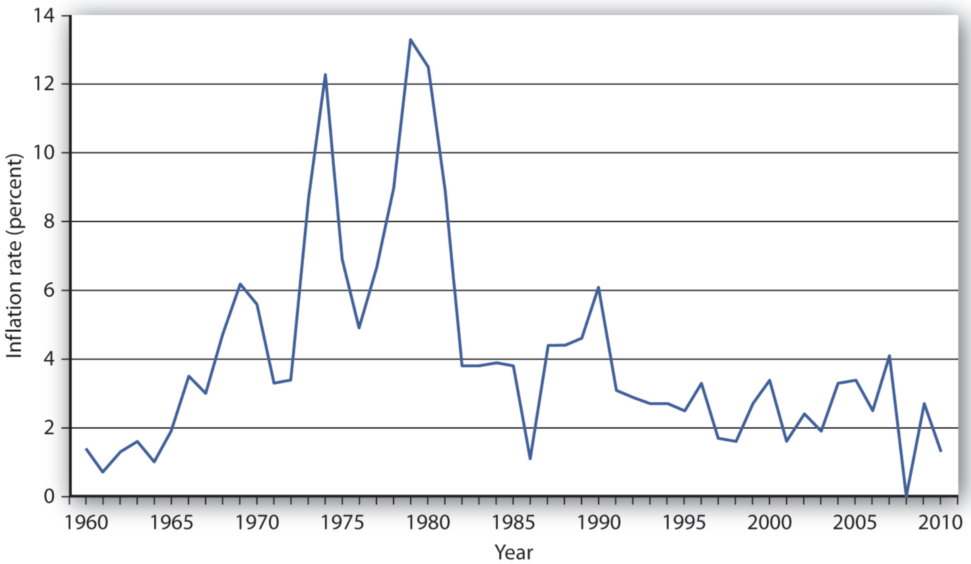
1. Inflation and deflation refer to changes in the average level of prices, not to changes in particular prices. An increase in medical costs is not inflation. A decrease in gasoline prices is not deflation. Inflation means the average level of prices is rising, and deflation means the average level of prices is falling.
2. Inflation and deflation refer to *rising* prices and *falling* prices, respectively; therefore, they do not have anything to do with the *level* of prices at any one time. "High" prices do not imply the presence of inflation, nor do "low" prices imply deflation. Inflation means a positive *rate of change* in average prices, and deflation means a negative *rate of change* in average prices.

Why Do We Care?

What difference does it make if the average level of prices changes? First, consider the impact of inflation.

Inflation is measured as the annual rate of increase in the average level of prices. Figure 5.6 "Inflation, 1960–2010" shows how volatile inflation has been in the United States over the past four decades. In the 1960s the inflation rate rose, and it became dramatically worse in the 1970s. The inflation rate plunged in the 1980s and continued to ease downward in the 1990s. It remained low in the early 2000s and began to accelerate in 2007 and has remained low since.

Figure 5.6 Inflation, 1960–2010

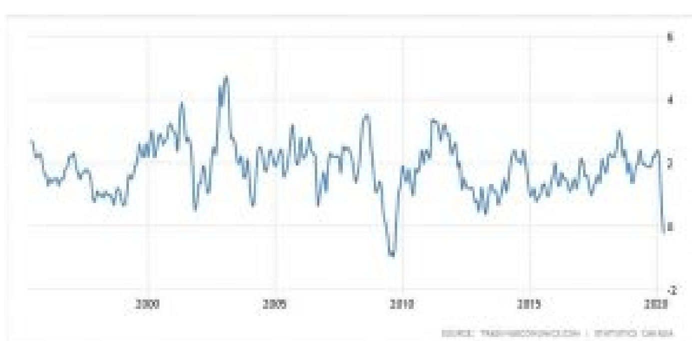


The U.S. inflation rate, measured as the annual rate of change in the average level of prices paid by consumers, varied considerably over the 1960–2010 period.

Source: Bureau of Labor Statistics, All Urban Consumers CPI-U, 1982–84 = 100, Dec.–Dec. inflation rate. Data for 2010 is through October. and https://takloo.wordpress.com/2010/07/20/canadas-historical-inflation-rate-1916-to-2010/

Here is a look at inflation in Canada from 1960 through 2018:

[Historical Inflation Rates in Canada]



Whether one regards inflation as a "good" thing or a "bad" thing depends very much on one's economic situation. If you are a borrower, unexpected inflation is a good thing—it reduces the value of money that you must repay. If you are a lender, it is a bad thing because it reduces the value of future payments you will receive. Whatever any particular person's situation may be, inflation always produces the following effects on the economy: it reduces the value of money and it reduces the value of future monetary obligations. It can also create uncertainty about the future.

Suppose that you have just found a $10 bill you stashed away in 1990. Prices have increased by about 50% since then; your money will buy less than what it would have purchased when you put it away. Your money has thus lost value.

Money loses value when its purchasing power falls. Since inflation is a rise in the level of prices, the amount of goods and services a given amount of money can buy falls with inflation.

Just as inflation reduces the value of money, it reduces the value of future claims on money. Suppose you have borrowed $100 from a friend and have agreed to pay it back in one year. During the year, however, prices double. That means that when you pay the money back, it will buy only half as much as it could have bought when you borrowed it. That is good for you but tough on the person who lent you the money. Of course, if you and your friend had anticipated such rapid inflation, you might have agreed to pay back a larger sum to adjust for it. When people anticipate inflation, they can adjust for its consequences in determining future obligations. But *unanticipated* inflation helps borrowers and hurts lenders.

Inflation's impact on future claims can be particularly hard on people who must live on a fixed income, that is, on an income that is predetermined through some contractual arrangement and does not change with economic conditions. An annuity, for example, typically provides a fixed stream of money payments. Retirement pensions sometimes generate fixed income. Inflation erodes the value of such payments.

Given the danger posed by inflation for people on fixed incomes, many retirement plans provide for indexed payments. An indexed payment is one whose dollar amount changes with the rate of change in the price level. If a payment changes at the same rate as the rate of change in the price level, the purchasing power of the payment remains constant. Social Security payments, for example, are indexed to maintain their purchasing power.

Because inflation reduces the purchasing power of money, the threat of future inflation can make people reluctant to lend for long periods. From a lender's point of view, the danger of a long-term commitment of funds is that future inflation will wipe out the value of the amount that will eventually be paid back. Lenders are reluctant to make such commitments.

Uncertainty can be particularly pronounced in countries where extremely high inflation is a threat. Hyperinflation is generally defined as an inflation rate in excess of 200% per year. Inflation of that magnitude erodes the value of money very quickly. Hyperinflations occurred in Germany in the 1920s and in Yugoslavia in the early 1990s. There are stories about how people in Germany during the hyperinflation brought wheelbarrows full of money to stores to pay for ordinary items. In Yugoslavia in 1993 there was a report of a shop owner barring the entrance to his store with a mop while he changed his prices.

The inflation rate rose to an astronomical rate in 2008 in Zimbabwe. As the government printed more money and put it in circulation, prices rose. When inflation began to accelerate, the government found it "necessary" to print more and more money, causing prices to rise very fast. The inflation rate in Zimbabwe reached an astonishing 11.2 million percent in July of 2008, according to Zimbabwe's Central Statistics Office. A loaf of bread cost 200,000 Zimbabwe dollars in February 2008. That same loaf cost 1.6 trillion Zimbabwe dollars by August (CNN, 2008).

Do the problems associated with inflation imply that deflation would be a good thing? The answer is simple: no. Like inflation, deflation changes the value of money and the value of future obligations. It also creates uncertainty about the future.

If there is deflation, the real value of a given amount of money rises. In other words, if there had been deflation since 2000, a $10 bill you had stashed away in 2000 would buy more goods and services today. That sounds good, but should you buy $10 worth of goods and services now when you would be able to buy even more for your $10 in the future if the deflation continues? When Japan experienced deflation in the late 1990s and early 2000s, Japanese consumers seemed to be doing just that—waiting to see if prices would fall further. They were spending less per person and, as we will see throughout our study of macroeconomics, less consumption often meant less output, fewer jobs, and the prospect of a recurring recessions.

And, if you had to use the $10 to pay back a debt you owed, the purchasing power of your money would be higher than when you borrowed the money. The lender would feel good about being able to buy more with the $10 than you were able to, but you would feel like you had gotten a raw deal.

Unanticipated deflation hurts borrowers and helps lenders. If the parties anticipate the deflation, a loan agreement can be written to reflect expected changes in the price level.

The threat of deflation can make people reluctant to borrow for long periods. Borrowers become reluctant to enter into long-term contracts because they fear that deflation will raise the value of the money they must pay back in the future. In such an environment, firms may be reluctant to borrow to build new factories, for example. This is because they fear that the prices at which they can sell their output will drop, making it difficult for them to repay their loans.

Deflation was common in the United States in the latter third of the 19th century. In the 20th century, there was a period of deflation after World War I and again during the Great Depression in the 1930s.

Price Indexes

How do we actually measure inflation and deflation (that is, changes in the price level)? Price-level change is measured as the percentage rate of change in the level of prices. But how do we find a price level?

Economists measure the price level with a price index. A price index is a number whose movement reflects movement in the average level of prices. If a price index rises 10%, it means the average level of prices has risen 10%.

There are four steps one must take in computing a price index:

1. Select the kinds and quantities of goods and services to be included in the index. A list of these goods and services, and the quantities of each, is the “market basket” for the index.
2. Determine what it would cost to buy the goods and services in the market basket in some period that is the base period for the index. A base period is a time period against which costs of the market basket in other periods will be compared in computing a price index. Most often, the base period for an index is a single year. If, for example, a price index had a base period of 1990, costs of the basket in other periods would be compared to the cost of the basket in 1990. We will encounter one index, however, whose base period stretches over three years.
3. Compute the cost of the market basket in the current period.
4. Compute the price index. It equals the current cost divided by the base-period cost of the market basket.

Equation 20.1

$$Priceindex = currentcostofbasket/base - periodcostofbasket$$

(While published price indexes are typically reported with this number multiplied by 100, our work with indexes will be simplified by omitting this step.)

Suppose that we want to compute a price index for movie fans, and a survey of movie watchers tells us that a typical fan rents 4 movies on DVD and sees 3 movies in theaters each month. At the theater, this viewer consumes a medium-sized soft drink and a medium-sized box of popcorn. Our market basket thus might include 4 DVD rentals, 3 movie admissions, 3 medium soft drinks, and 3 medium servings of popcorn.

Our next step in computing the movie price index is to determine the cost of the market basket. Suppose we surveyed movie theaters and DVD-rental stores in 2007 to determine the average prices of these items, finding the values given in Table 5.1 “Pricing a Market Basket”. At those prices, the total monthly cost of our movie market basket in 2007 was $48. Now suppose that in 2008 the prices of movie admissions and DVD rentals rise, soft-drink prices at movies fall, and popcorn prices remain unchanged. The combined effect of these changes pushes the 2008 cost of the basket to $50.88.

Table 5.1 Pricing a Market Basket

Item	**Quantity in Basket**	**2007 Price**	**Cost in 2007 Basket**	**2008 Price**	**Cost in 2008 Basket**
DVD rental	4	$2.25	$9.00	$2.97	$11.88
Movie admission	3	7.75	23.25	8.00	24.00
Popcorn	3	2.25	6.75	2.25	6.75
Soft drink	3	3.00	9.00	2.75	8.25
Total cost of basket		**2007**	$48.00	**2008**	$50.88

To compute a price index, we need to define a market basket and determine its price. The table gives the composition of the movie market basket and prices for 2007 and 2008. The cost of the entire basket rises from $48 in 2007 to $50.88 in 2008.

Using the data in Table 5.1 “Pricing a Market Basket”, we could compute price indexes for each year. Recall that a price index is the ratio of the current cost of the basket to the base-period cost. We can select any year we wish as the base year; take 2007. The 2008 movie price index (MPI) is thus

$$MPI_{2008} = \$50.88/\$48 = 1.06$$

The value of any price index in the base period is always 1. In the case of our movie price index, the 2007 index would be the current (2007) cost of the basket, $48, divided by the base-period cost, which is the same thing: $48/$48 = 1.

The Consumer Price Index (CPI)

One widely used price index in Canada is the consumer price index (CPI), a price index whose movement reflects changes in the prices of goods and services typically purchased by consumers. When the media report the U.S. inflation rate, the number cited is usually a rate computed using the CPI. The CPI is also used to determine whether people’s incomes are keeping up with the costs of the things they buy. The CPI is often used to measure changes in the cost of living, though as we shall see, there are problems in using it for this purpose.

The market basket for the CPI contains thousands of goods and services. The composition of the basket is determined by Statistics Canada. Surveyors tally the prices of the goods and services in the basket each month in cities all over Canada to determine the current cost of the basket. The major categories of items in the CPI are food and beverages, housing, apparel, transportation, medical care, recreation, education and communication, and other goods and services.

The current cost of the basket of consumer goods and services is then compared to the base-period cost of that same basket. The **base year** is currently 2002, and the basket for that **year** is given the value of 100. In 2012 the **CPI** averaged 121.7, which means that what you could buy for $100 in 2002 cost $121.70 in 2012. Each month's CPI thus reflects the ratio of the current cost of the basket divided by its base-period cost.

Equation 5.2

$$CPI = current\ cost\ of\ basket / 1982 - 1984\ cost\ of\ basket$$

Like many other price indexes, the CPI is computed with a fixed market basket. The composition of the basket generally remains unchanged from one period to the next. Because buying patterns change, however, the basket is revised accordingly.

Computing the Rate of Inflation or Deflation

The rate of inflation or deflation is the percentage rate of change in a price index between two periods. Given price-index values for two periods, we can calculate the rate of inflation or deflation as the change in the index divided by the initial value of the index, stated as a percentage:

Equation 5.4

$$Rate\ of\ inflation\ or\ deflation = percentage\ change\ in\ index / initial\ value\ of\ index$$

To calculate inflation in movie prices over the 2007–2008 period, for example, we could apply Equation 5.4 to the price indexes we computed for those two years as follows:

$$Movie\ inflation\ rate\ in\ 2008 = (1.06 - 1.00)/1.00 = -.06 = 6$$

The CPI is often used for calculating price-level change for the economy. For example, the rate of inflation in 2007 can be computed from the December 2006 price level (2.016) and the December 2007 level (2.073):

$$Inflation\ rate = (2.073 - 2.016)/2.016 = 0.028 = 2.8$$

Price indexes are useful. They allow us to see how the general level of prices has changed. They allow us to estimate the rate of change in prices, which we report as the rate of inflation or deflation. And they give us a tool for converting nominal values to real values so we can make better comparisons of economic performance across time.

Key Takeaways

- Inflation is an increase in the average level of prices, and deflation is a decrease in the average level of prices. The rate of inflation or deflation is the percentage rate of change in a price index.
- The consumer price index (CPI) is the most widely used price index in Canada.

- Nominal values can be converted to real values by dividing by a price index.
- Inflation and deflation affect the real value of money, of future obligations measured in money, and of fixed incomes. Unanticipated inflation and deflation create uncertainty about the future.

Case in Point: Take Me Out to the Ball Game …

Figure 5.7

Keith Allison – Boston Red Sox Celebrating a Win – CC BY-SA 2.0.

The cost of a trip to the old ball game jumped 7.9% in 2008, according to *Team Marketing Report*, a Chicago-based newsletter. The report bases its estimate on its fan price index, whose market basket includes two adult average-priced tickets, two child average-priced tickets, two small draft beers, four small soft drinks, four regular-sized hot dogs, parking for one car, two game programs, and two least expensive, adult-sized adjustable baseball caps. The average price of the market basket was $191.92 in 2008.

Team Marketing compiles the cost of the basket for each of major league baseball's 30 teams. According to this compilation, the Boston Red Sox was the most expensive team to watch in 2008; the Tampa Bay Rays was the cheapest. The Rays made it to the World Series in 2008; the Red Sox did not. By that measure, the Rays were something of a bargain. The table shows the cost of the fan price index market basket for 2008.

Team	Basket Cost	Team	Basket Cost
Boston Red Sox	$320.71	San Francisco Giants	$183.74
New York Yankees	$275.10	Cincinnati Reds	$167.14
Chicago Cubs	$251.96	Minnesota Twins	$165.71
New York Mets	$251.19	Baltimore Orioles	$165.40
Toronto Blue Jays	$230.46	Florida Marlins	$164.26
Los Angeles Dodgers	$229.14	AZ Diamondbacks	$162.84
St. Louis Cardinals	$217.28	Colorado Rockies	$160.00
Houston Astros	$215.45	Atlanta Braves	$157.15
Chicago White Sox	$214.51	Kansas City Royals	$151.16
Oakland Athletics	$206.80	Texas Rangers	$148.04
San Diego Padres	$201.72	Pittsburgh Pirates	$146.32
Philadelphia Phillies	$199.56	Milwaukee Brewers	$141.52
Washington Nationals	$195.50	Los Angeles Angels	$140.42
Cleveland Indians	$192.38	Tampa Bay Rays	$136.31
Seattle Mariners	$191.16	**MLB Average**	**$191.92**
Detroit Tigers	$190.13		

Sources: Team Marketing Report, TMR's Fan Cost Index Major League Baseball 2008 at http://www.teammarketing.com and personal interview.

References

Cage, R., John Greenlees, and Patrick Jackman, "Introducing the Chained Consumer Price Index" (paper, Seventh Meeting of the International Working Group on Price Indices, Paris, France, May 2003), available at http://stats.bls.gov/cpi/superlink.htm.

CNN, "Zimbabwe Inflation Hits 11,200,000%," CNN.com, August 19, 2008.

5.3 Unemployment

Learning Objectives

1. Explain how unemployment is measured in Canada.
2. Define three different types of unemployment.
3. Define and illustrate graphically what is meant by the natural level of employment. Relate the natural level of employment to the natural rate of unemployment.

For an economy to produce all it can and achieve a solution on its production possibilities curve, the factors of production in the economy must be fully employed. Failure to fully employ these factors leads to a solution inside the production possibilities curve in which society is not achieving the output it is capable of producing.

In thinking about the employment of society's factors of production, we place special emphasis on labour. The loss of a job can wipe out a household's entire income; it is a more compelling human problem than, say, unemployed capital, such as a vacant apartment. In measuring unemployment, we thus focus on labour rather than on capital and natural resources.

Measuring Unemployment

Statistics Canada defines a person as unemployed if he or she is not working but is looking for and available for work. The labour force is the total number of people working or unemployed. The unemployment rate is the percentage of the labour force that is unemployed.

To estimate the unemployment rate, government uses the information they collect in surveys of various Canadian households. At each of these randomly selected households, the surveyor asks about the employment status of each adult (everyone age 15 or over) who lives there. Many households include more than one adult; the survey gathers information on about roughly 100,000 adults. The surveyor asks if each adult is working. If the answer is yes, the person is counted as employed. If the answer is no, the surveyor asks if that person has looked for work at some time during the previous four weeks and is available for work at the time of the survey. If the answer to that question is yes, the person is counted as unemployed. If the answer is no, that person is not counted as a member of the labour force. Figure 5.8 "Computing the Unemployment Rate" shows the survey's results for the civilian (nonmilitary) population for November 2010. The unemployment rate is then computed as the number of people unemployed divided by the labour force—the sum of the number of people not working but available and looking for work plus the number of people working.

A monthly survey of households divides the civilian adult population into three groups. Those who have jobs are counted as employed; those who do not have jobs but are looking for them and are available for work are counted as unemployed; and those who are not working and are not looking for work are not counted as members of the labor force. The unemployment rate equals the number of people looking for work divided by the sum of the number of people looking for work and the number of people employed.

The problem of understating unemployment among women has been fixed, but others remain. A worker who has been cut back to part-time work still counts as employed, even if that worker would prefer to work full time. A person who is out of work, would like to work, has looked for work in the past year, and is available for work, but who has given up looking, is considered a discouraged worker. Discouraged workers are not counted as unemployed, but a tally is kept each month of the number of discouraged workers.

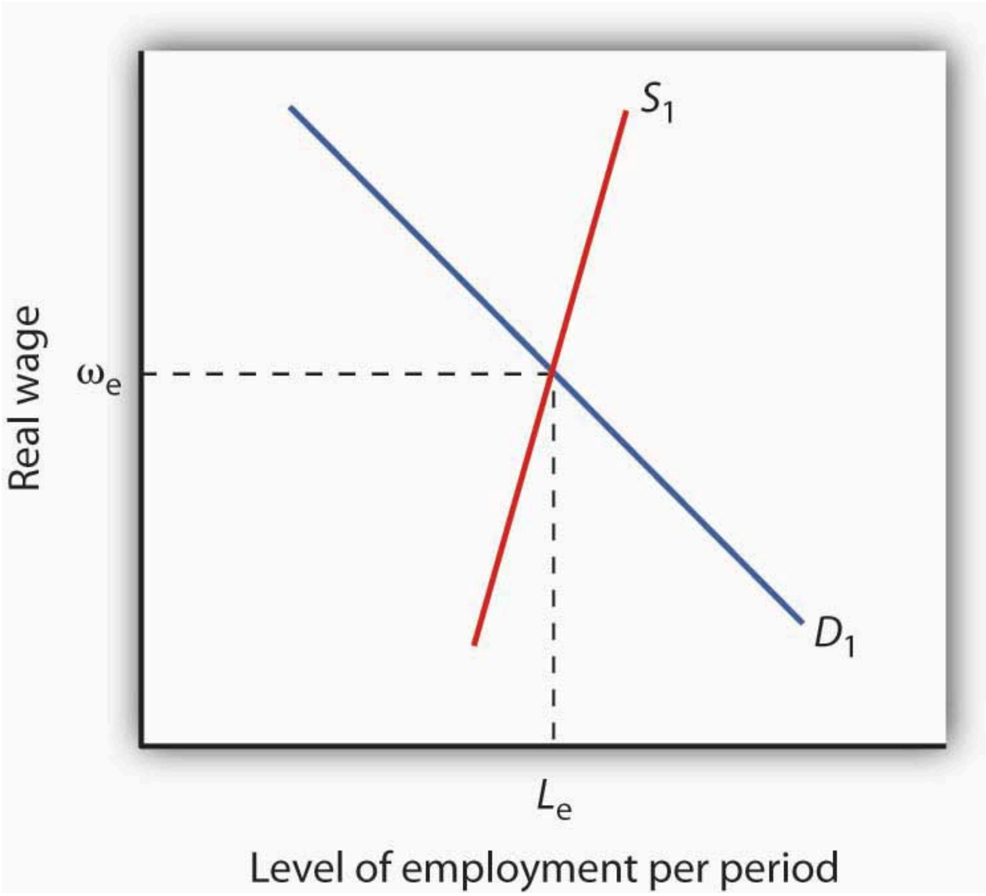
The official measures of employment and unemployment can yield unexpected results. For example, when firms expand output, they may be reluctant to hire additional workers until they can be sure the demand for increased output will be sustained. They may respond first by extending the hours of employees previously reduced to part-time work or by asking full-time personnel to work overtime. None of that will increase employment, because people are simply counted as "employed" if they are working, regardless of how much or how little they are working. In addition, an economic expansion may make discouraged workers more optimistic about job prospects, and they may resume their job searches. Engaging in a search makes them unemployed again—and increases unemployment. Thus, an economic expansion may have little effect initially on employment and may even increase unemployment.

Types of Unemployment

Workers may find themselves unemployed for different reasons. Each source of unemployment has quite different implications, not only for the workers it affects but also for public policy.

Figure 5.9 "The Natural Level of Employment" applies the demand and supply model to the labour market. The price of labour is taken as the real wage, which is the nominal wage divided by the price level; the symbol used to represent the real wage is the Greek letter omega, ω. The supply curve is drawn as upward sloping, though steep, to reflect studies showing that the quantity of labour supplied at any one time is nearly fixed. Thus, an increase in the real wage induces a relatively small increase in the quantity of labour supplied. The demand curve shows the quantity of labor demanded at each real wage. The lower the real wage, the greater the quantity of labour firms will demand. In the case shown here, the real wage, ω_e, equals the equilibrium solution defined by the intersection of the demand curve D_1 and the supply curve S_1. The quantity of labour demanded, L_e, equals the quantity supplied. The employment level at which the quantity of labour demanded equals the quantity supplied is called the natural level of employment. It is sometimes referred to as full employment.

Figure 5.9 The Natural Level of Employment



The employment level at which the quantity of labour demanded equals the quantity supplied is called the natural level of employment. Here, the natural level of employment is L_e, which is achieved at a real wage ω_e.

Even if the economy is operating at its natural level of employment, there will still be some unemployment. The rate of unemployment consistent with the natural level of employment is called the natural rate of unemployment. Business cycles may generate additional unemployment. We discuss these various sources of unemployment below.

Frictional Unemployment

Even when the quantity of labour demanded equals the quantity of labour supplied, not all employers and potential workers have found each other. Some workers are looking for jobs, and some employers are looking for workers. During the time it takes to match them up, the workers are unemployed. Unemployment that occurs because it takes time for employers and workers to find each other is called frictional unemployment.

The case of college graduates engaged in job searches is a good example of frictional unemployment. Those who did not land a job while still in school will seek work. Most of them will find jobs, but it will take time. During that time, these new graduates will be unemployed. If information about the labor market were costless, firms and potential workers would instantly know everything they needed to know about each other and there would be no need for searches on the part of workers and firms. There would be no frictional unemployment. But information is costly. Job searches are needed to produce this information, and frictional unemployment exists while the searches continue.

Structural Unemployment

Another reason there can be unemployment even if employment equals its natural level stems from potential mismatches between the skills employers seek and the skills potential workers offer. Every worker is different; every job has its special characteristics and requirements. The qualifications of job seekers may not match those that firms require. Even if the number of employees firms demand equals the number of workers available, people whose qualifications do not satisfy what firms are seeking will find themselves without work. Unemployment that results from a mismatch between worker qualifications and the characteristics employers require is called structural unemployment.

Structural unemployment emerges for several reasons. Technological change may make some skills obsolete or require new ones. The widespread introduction of personal computers since the 1980s, for example, has lowered demand for typists who lacked computer skills.

Structural unemployment can occur if too many or too few workers seek training or education that matches job requirements. Students cannot predict precisely how many jobs there will be in a particular category when they graduate, and they are not likely to know how many of their fellow students are training for these jobs. Structural unemployment can easily occur if students guess wrong about how many workers will be needed or how many will be supplied.

Structural unemployment can also result from geographical mismatches. Economic activity may be booming in one region and slumping in another. It will take time for unemployed workers to relocate and find new jobs. And poor or costly transportation may block some urban residents from obtaining jobs only a few miles away.

Public policy responses to structural unemployment generally focus on job training and education to equip workers with the skills firms demand. The government publishes regional labour-market information, helping to inform unemployed workers of where jobs can be found. The North American Free Trade Agreement (NAFTA) which is now called USMCA, created a free trade region encompassing Mexico, the United States, and Canada, has created some structural unemployment in the three countries.

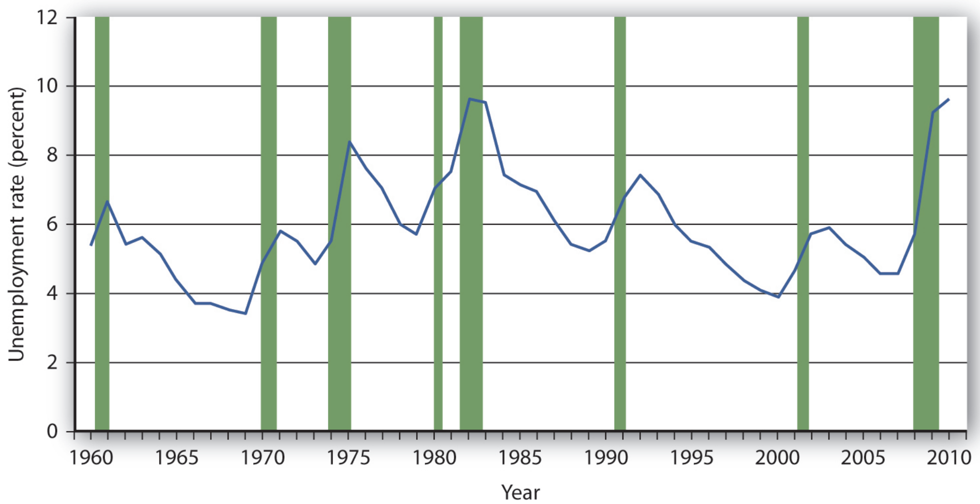
Although government programs may reduce frictional and structural unemployment, they cannot eliminate it. Information in the labour market will always have a cost, and that cost creates frictional unemployment. An economy with changing demands for goods and services, changing technology, and changing production costs will always have some sectors expanding and others contracting—structural unemployment is inevitable. An economy at its natural level of employment will therefore have frictional and structural unemployment.

Cyclical Unemployment

Of course, the economy may not be operating at its natural level of employment, so unemployment may be above or below its natural level. In a later chapter we will explore what happens when the economy generates employment greater or less than the natural level. Cyclical unemployment is unemployment in excess of the unemployment that exists at the natural level of employment.

Figure 5.10 "Unemployment Rate, 1960–2010" shows the unemployment rate in the United States for the period from 1960 through November 2010. We see that it has fluctuated considerably. How much of it corresponds to the natural rate of unemployment varies over time with changing circumstances. For example, in a country with a demographic "bulge" of new entrants into the labour force, frictional unemployment is likely to be high, because it takes the new entrants some time to find their first jobs. This factor alone would raise the natural rate of unemployment. A demographic shift toward more mature workers would lower the natural rate. During recessions, highlighted in [Figure 5.10 "Unemployment Rate, 1960–2010"], the part of unemployment that is cyclical unemployment grows. The analysis of fluctuations in the unemployment rate, and the government's responses to them, will occupy center stage in much of the remainder of this book.

Figure 5.10 Unemployment Rate in the US, 1960–2010

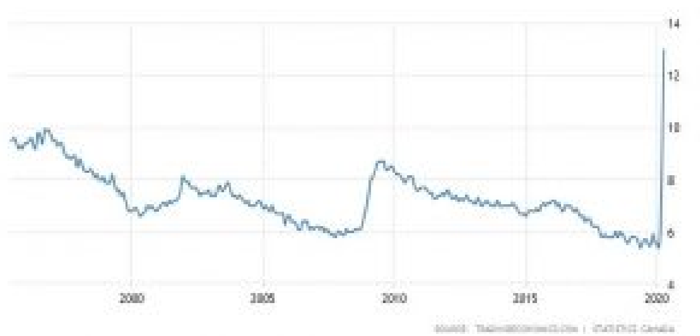


The chart shows the unemployment rate for each year from 1960 to 2010. Recessions are shown as shaded areas.

Source: *Economic Report of the President*, 2010, Table B-42. Data for 2010 is average of first eleven months from the Bureau of Labor Statistics home page.

UNEMPLOYMENT RATES IN CANADA

[Historical Unemployment Rates in Canada]



Key Takeaways

- People who are not working but are looking and available for work at any one time are considered unemployed. The unemployment rate is the percentage of the labor force that is unemployed.
- When the labor market is in equilibrium, employment is at the natural level and the unemployment rate equals the natural rate of unemployment.
- Even if employment is at the natural level, the economy will experience frictional and structural unemployment. Cyclical unemployment is unemployment in excess of that associated with the natural level of employment.

Try It!

Given the data in the table, compute the unemployment rate in Year 1 and in Year 2. Explain why, in this example, both the number of people employed and the unemployment rate increased.

Year	Number employed (in millions)	Number unemployed (in millions)
1	20	2
2	21	2.4

Read it: [Joblessness in Canada Due to Covid-19]

Canada lost nearly 2 million jobs in April amid COVID-19 crisis: Statistics Canada

CBC News · Posted: May 08, 2020 8:41 AM ET | Last Updated: May 8

Canada lost almost two million jobs during the month of April, a record high, as the impact of COVID-19 on the economy made itself known.

Statistics Canada's Labour Force Survey data released Friday brings the total number of jobs lost during the crisis to more than three million.

The closure of non-essential services to slow the spread of COVID-19 has devastated the economy and forced businesses to shutter temporarily.

Statistics Canada says the unemployment rate soared to 13 per cent as the full force of the pandemic hit, compared with 7.8 per cent in March.

Economists on average had expected the loss of four million jobs and an unemployment rate of 18 per cent, according to financial markets data firm Refinitiv.

Douglas Porter, chief economist with BMO Economics, said those high projections could be tied to reports that more than seven million Canadians had applied for the Canada Emergency Response Benefit (CERB).

"The 5.2 percentage point rise in the jobless rate was considerably less than expected, and massively smaller than the 10.3 percentage point spike in the U.S. jobless rate," he said in an emailed statement.

Since comparable data became available in 1976, the April unemployment rate was second-highest only to December 1982, when it reached 13.1 per cent.

"By most accounts, unemployment soared to over 25 per cent during the worst of the Great Depression, as high as 30 per cent by some measures," said Porter.

The rapid decline in employment is unprecedented, Statistics Canada says. The decline since February (-15.7 per cent) outpaces previous financial crises, including the 1981-1982 recession, which resulted in a drop of -5.4 per cent over 17 months.

"Today's job numbers start to complete the picture of just how devastating the COVID-19 crisis has been to the Canadian labour market," said economist Brendon Bernard of job site Indeed Canada in an email.

"While April's decline was a bit more modest than expected, that 6.4 per cent of all Canadian adults could lose employment in a single month is staggering."

Fewer hours for many who are still working

Job losses are not the only way the COVID-19 crisis has impacted employment. In addition to those who are now out of work, the number of people who were employed but worked less than half their usual hours because of the pandemic increased by 2.5 million from February to April.

That means the cumulative effect of the economic shutdown – people both no longer employed or working far less – was a staggering 5.5 million as of the week of April 12.

All provinces have been hard hit by the crisis. Employment dropped in all provinces for the second month running, with losses of more than 10 per cent everywhere.

Quebec had the worst losses in April at -18.7 per cent, or 821,000 jobs.

Here are the jobless rates last month by province (numbers from the previous month in brackets):

- Newfoundland and Labrador 16.0 per cent (11.7).
- Prince Edward Island 10.8 (8.6).
- Nova Scotia 12.0 (9.0).
- New Brunswick 13.2 (8.8).
- Quebec 17.0 (8.1).
- Ontario 11.3 (7.6).
- Manitoba 11.4 (6.4).
- Saskatchewan 11.3 (7.3).

- Alberta 13.4 (8.7).
- British Columbia 11.5 (7.2).

As for Canada's three territories, due to logistical challenges, Statistics Canada's Labour Force Survey methodology is somewhat different there than in the rest of Canada, said Bernard.

"Employment rates in the territories haven't been hit as hard as the rest of Canada. That said, Nunavut saw a noticeable drop in the share of the population with a job in April, a sign of how public health concerns and economic uncertainty have extended beyond just the pandemic's hot spots."

There is evidence to suggest, however, that many pandemic-related job losses are temporary.

In the month of April, almost all of the newly unemployed – 97 per cent – were on temporary layoff, indicating they expect to return to their previous places of work as the shutdown is relaxed.

At the same time, workplace participation declined to 3.7 percentage points, bringing Canada's participation rate to 59.8 per cent, as more than 1 million dropped out of the labour force, said Porter.

Among those still employed, the Labour Force Survey found that an additional 3.3 million workers were working from home in April.

Vulnerable workers most affected

Losses continued to be more rapid in jobs with less security and poorer pay.

Over the two-month period since February, overall employment (not adjusted for seasonality) declined 17.8 per cent.

But it was above average among employees with temporary jobs (-30.2 per cent) and those who had been in their jobs one year or less (-29.5 per cent).

Declines were sharper for employees earning less than two-thirds of the median hourly wage of $25.04 (-38.1 per cent) and those paid by the hour (-25.1 per cent).

Statistics Canada said this is consistent with job declines observed in accommodation and food services, and wholesale and retail trade, which tend to be more precarious and lower paying.

The number of people age 15 and older living households where no one is employed increased 23.5 per cent – or more than 1.6 million – in the two months between February and April.

Among couples, the number where neither partner is employed increased by 22.5 per cent, or 845,000, while single parents who are not employed increased by 53.9 per cent, or 126,000.

Small companies hardest hit

Smaller companies – defined as those with fewer than 20 employees – have shed 30.8 per cent of their workers, medium-sized firms have let 25.1 per cent of workers go, and large companies have seen employment decline by 12.6 per cent.

- U.S. economy lost 20.5 million jobs in April

Hard-hit sectors at the outset include retail, hotels, restaurants and bars, which continued to see losses in April. The losses in the service sector also continued in April, down 1.4 million, or 9.6 per cent, Statistics Canada says.

Proportionally, the losses were greater in goods-producing sectors like construction and manufacturing, which combined lost 621,000 jobs for a drop of 15.8 per cent after being virtually unchanged in March.

The federal government announced Friday it will extend its emergency wage-subsidy to the end of June. The program — which covers 75 per cent of employees' pay, up to $847, to help employers keep their workers on the payroll for the duration of the COVID-19 crisis — was previously set to end June 6.

Average wages in April were up 10.8 per cent compared with a year ago. Economists say the increase is because a disproportionate share of job losses in low-paying positions has allowed the average wage to rise.

"At the same time, relatively more people remained employed in industries where work can be done from home, such as public administration and professional, scientific and technical services, two of the highest-paying industries," the report says.

Answer to Try It! Problem

In Year 1 the total labour force includes 22 million workers, and so the unemployment rate is 2/22 = 9.1%. In Year 2 the total labour force numbers 23.4 million workers; therefore the unemployment rate is 2.4/23.4 = 10.3%. In this example, both the number of people employed and the unemployment rate rose, because more people (23.4 – 22 = 1.4 million) entered the labour force, of whom 1 million found jobs and 0.4 million were still looking for jobs.

5.4 Review and Practice

Summary

In this chapter we examined growth in real GDP and business cycles, price-level changes, and unemployment. We saw how these phenomena are defined and looked at their consequences.

Examining real GDP, rather than nominal GDP, over time tells us whether the economy is expanding or contracting. Real GDP in Canada and the US shows a long upward trend, but with the economy going through phases of expansion and recession around that trend. These phases make up the business cycle. An expansion reaches a peak, and the economy falls into a recession. The recession reaches a trough and begins an expansion again.

Inflation is an increase in the price level and deflation is a decrease in the price level. The rate of inflation or deflation is the percentage rate of change in a price index. We looked at the calculation of the consumer price index (CPI). The CPI is widely used in the calculation of price-level changes.

Inflation and deflation affect economic activity in several ways. They change the value of money and of claims on money. Unexpected inflation benefits borrowers and hurts lenders. Unexpected deflation benefits lenders and hurts borrowers. Both inflation and deflation create uncertainty and make it difficult for individuals and firms to enter into long-term financial commitments.

The unemployment rate is measured as the percentage of the labour force not working but seeking work. Frictional unemployment occurs because information about the labour market is costly; it takes time for firms seeking workers and workers seeking firms to find each other. Structural unemployment occurs when there is a mismatch between the skills offered by potential workers and the skills sought by firms. Both frictional and structural unemployment occur even if employment and the unemployment rate are at their natural levels. Cyclical unemployment is unemployment that is in excess of that associated with the natural level of employment.

Concept Problems

1. Describe the phases of a business cycle.
2. On the basis of recent news reports, what phase of the business cycle do you think the economy is in now? What is the inflation or deflation rate? The unemployment rate?
3. Suppose you compare your income this year and last year and find that your nominal income fell but your real income rose. How could this have happened?
4. Suppose you calculate a grocery price inflation rate. Using the arguments presented in the chapter, explain possible sources of upward bias in the rate you calculate, relative to the actual trend of food prices.

5. Name three items you have purchased during the past year that have increased in quality during the year. What kind of adjustment would you make in assessing their prices for the CPI?
6. Why do some people gain and other people lose from inflation and deflation?
7. Suppose unemployed people leave a province to obtain jobs in other provinces. What do you predict will happen to the unemployment rate in the province experiencing the out-migration? What might happen to the unemployment rates in the provinces experiencing in-migration?
8. Minority teenagers have the highest unemployment rates of any group. One reason for this phenomenon is high transportation costs for many minority teens. What form of unemployment (cyclical, frictional, or structural) do high transportation costs suggest?
9. Welfare reforms put more pressure on welfare recipients to look for work. The new law mandated cutting off benefits after a certain length of time. How do you think this provision might affect the unemployment rate?
10. Canadian workers work more hours than their European counterparts. Should Parliament legislate a shorter workweek?

Numerical Problems

1. Plot the quarterly data for real GDP in Canada for the last two years. (You can find the data at Statistics Canada. Relate recent changes in real GDP to the concept of the phases of the business cycle.)
2. Suppose that in 2009, the items in the market basket for our movie price index cost $53.40. Use the information in the chapter to compute the price index for that year. How does the rate of movie price inflation from 2008 to 2009 compare with the rate from 2007 to 2008?
3. Recompute the movie price indexes for 2007 and 2008 using 2008 as the base year. Now compute the rate of inflation for the 2007–2008 period. Compare your result to the inflation rate calculated for that same period using 2007 as the base year.
4. Here are some hypothetical statistics for a country in August 2018. Compute the unemployment rate for that month (all figures are in thousands).

Population (Civilian, noninstitutional)	229,167
Civilian Labour Force	151,698
Participation Rate	66.2%
Not in Labour Force	77,469
Employed	144,579
Unemployed	7,119

5. Suppose an economy has 10,000 people who are not working but looking and available for work and 90,000 people who are working. What is its unemployment rate? Now suppose 4,000 of the people looking for work get discouraged and give up their searches. What happens to the unemployment rate? Would you interpret this as good news for the economy or bad news? Explain.

6. Suppose you are given the following data for a small economy:

   Number of unemployed workers: 1,000,000.

   Labour force: 10,000,000.

   Based on this data, answer following:

   1. What is the unemployment rate?
   2. Can you determine whether the economy is operating at its full employment level?
   3. Now suppose people who had been seeking jobs become discouraged, and give up their job searches. The labour force shrinks to 900,500 workers, and unemployment falls to 500,000 workers. What is the unemployment rate now? Has the economy improved?

7. Suppose you are given the following data for an economy:

Month	**Real GDP**	**Employment**
1	$10.0 trillion	100 million
2	$10.4 trillion	104 million
3	$10.5 trillion	105 million
4	$10.3 trillion	103 million
5	$10.2 trillion	102 million
6	$10.3 trillion	103 million
7	$10.6 trillion	106 million
8	$10.7 trillion	107 million
9	$10.6 trillion	106 million

   1. Plot the data for real GDP, with the time period on the horizontal axis and real GDP on the vertical axis.
   2. There are two peaks. When do they occur?
   3. When does the trough occur?

8. The Consumer Price Index in Period 1 is 107.5. It is 103.8 in Period 2. Is this a situation of inflation for deflation? What is the rate?

CHAPTER 6: MEASURING TOTAL OUTPUT AND INCOME

Start Up:

It is early morning when a half-dozen senior officials enter a room at the Finance Department in Ottawa. Once inside, they get to work to calculate one of the most important indicators of economic activity we have: the official estimate of the value of the economy's total output, known as its gross domestic product (GDP).

When the team has finished its computations, the results will be placed in an envelope and shared with other finance executives and the Prime Minister. The senior officials who meet to compute GDP are economists. The adviser who delivers the estimate to the PM is on his Council of Economic Advisers.

Once this information is shared it will appear in major stories on the Internet and in the next editions of the nation's newspapers; the estimate of the previous quarter's GDP will be one of the lead items on television and radio news broadcasts that day. It is a very important statistic. The estimate of GDP provides the best available reading of macroeconomic performance. It will affect government policy, and it will influence millions of decisions in the private sector. Prior knowledge of the GDP estimate could be used to anticipate the response in the stock and bond markets, so great care is taken that only a handful of trusted officials have access to the information until it is officially released.

The GDP estimate took on huge significance in the fall of 2008 as the Canada, the United States, and much of the rest of the world went through the wrenching experience of the worst financial crisis since the Great Depression of the 1930s. The expectation that the financial crisis would lead to an economic collapse was widespread, and the quarterly announcements of GDP figures were more anxiously awaited than ever.

The primary measure of the ups and downs of economic activity—the business cycle—is real GDP. When an economy's output is rising, the economy creates more jobs, more income, and more opportunities for people. In the long run, an economy's output and income, relative to its population, determine the material standard of living of its people.

Clearly GDP is an important indicator of macroeconomic performance. It is the topic we will consider in this chapter. We will learn that GDP can be measured either in terms of the total value of output produced or as the total value of income generated in producing that output. We will begin with an examination of measures of GDP in terms of output. Our initial focus will be on nominal GDP: the value of total output measured in current prices. We will turn to real GDP—a measure of output that has been adjusted for price level changes—later in the chapter. We will refer to nominal GDP simply as GDP. When we discuss the real value of the measure, we will call it real GDP.

6.1 Measuring Total Output

Learning Objectives

1. Define gross domestic product and its four major spending components and illustrate the various flows using the circular flow model.
2. Distinguish between measuring GDP as the sum of the values of final goods and services and as the sum of values added at each stage of production.
3. Distinguish between gross domestic product and gross national product.

An economy produces a mind-boggling array of goods and services. In 2007, for example, Domino's Pizza produced 19 million slices of pizzas in it's Canadian stores. Western Forest Products, based in Vancouver, BC, produced 864 million board feet of lumber in 2016. The Queens University football team drew in tens of thousands of fans to Richardson Stadium to watch its home games—and won the national championship. One pediatric nurse in Toronto's Sunnybrook Hospital, helped to deliver 250 babies and took care of 233 additional patients. A list of all the goods and services produced in any year would be virtually endless.

So—what kind of year was 2007? We would not get very far trying to wade through a list of all the goods and services produced that year. It is helpful to have instead a single number that measures total output in the economy; that number is GDP.

The Components of GDP

We can divide the goods and services produced during any period into four broad components, based on who buys them. These components of GDP are personal consumption (C), gross private domestic investment (I), government purchases (G), and net exports (X_n). Thus

Equation 21.1

$$GDP = consumption\ (C) + private\ investment\ (I) + government\ purchases\ (G) + net\ exports\ (X_n)$$

or

$$GDP = C + I + G + X_n$$

We will examine each of these components, and we will see how each fits into the pattern of macroeconomic activity. Before we begin, it will be helpful to distinguish between two types of variables: stocks and flows. A flow variable is a variable that is measured over a specific period of time. A stock variable is a variable that is independent of time. Income is an example of a flow variable. To say one's income is, for example, $1,000 is meaningless without a time dimension. Is it $1,000 per hour? Per day? Per week? Per month? Until we know the time period, we have no idea what the income figure means. The balance in a checking account is an example of a stock variable. When we learn that the balance in a checking account is $1,000, we know precisely what that means; we do not need a time dimension. We will see that stock and flow variables play very different roles in macroeconomic analysis.

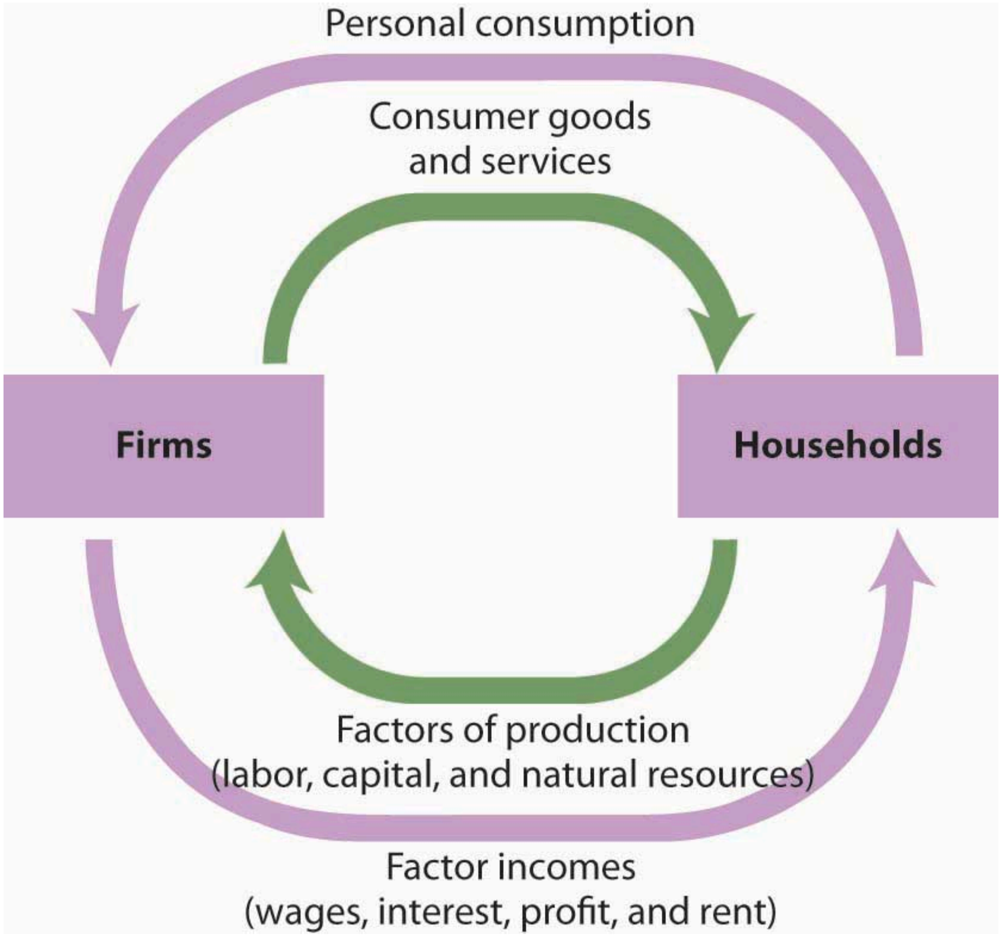
Personal Consumption

Personal consumption is a flow variable that measures the value of goods and services purchased by households during a time period. Purchases by households of groceries, health-care services, clothing, and automobiles—all are counted as consumption.

The production of consumer goods and services accounts for about 70% of total output. Because consumption is such a large part of GDP, economists seeking to understand the determinants of GDP must pay special attention to the determinants of consumption. In a later chapter we will explore these determinants and the impact of consumption on economic activity.

Personal consumption represents a demand for goods and services placed on firms by households. In the chapter on demand and supply, we saw how this demand could be presented in a circular flow model of the economy. Figure 6.1 "Personal Consumption in the Circular Flow" presents a circular flow model for an economy that produces only personal consumption goods and services. (We will add the other components of GDP to the circular flow as we discuss them.) Spending for these goods flows from households to firms; it is the arrow labeled "Personal consumption." Firms produce these goods and services using factors of production: labour, capital, and natural resources. These factors are ultimately owned by households. The production of goods and services thus generates income to households; we see this income as the flow from firms to households labeled "Factor incomes" in the exhibit.

Figure 6.1 Personal Consumption in the Circular Flow



Personal consumption spending flows from households to firms. In return, consumer goods and services flow from firms to households. To produce the goods and services households demand, firms employ factors of production owned by households. There is thus a flow of factor services from households to firms, and a flow of payments of factor incomes from firms to households.

In exchange for payments that flow from households to firms, there is a flow of consumer goods and services from firms to households. This flow is shown in Figure 6.1 "Personal Consumption in the Circular Flow" as an arrow going from firms to households. When you buy a soda, for example, your payment to the store is part of the flow of personal consumption; the soda is part of the flow of consumer goods and services that goes from the store to a household—yours.

Similarly, the lower arrow in Figure 6.1 "Personal Consumption in the Circular Flow" shows the flow of factors of production—labour, capital, and natural resources—from households to firms. If you work for a firm, your labor is part of this flow. The wages you receive are part of the factor incomes that flow from firms to households.

There is a key difference in our interpretation of the circular flow picture in Figure 6.1 "Personal Consumption in the Circular Flow" from our analysis of the same model in the demand and supply chapter. There, our focus was microeconomics, which examines individual units of the economy. In thinking about the flow of consumption spending from households to firms, we emphasized demand and supply in particular markets—markets for such things as blue jeans, haircuts, and apartments. In thinking about the flow of income payments from firms to households, we focused on the demand and supply for particular factors of production, such as textile workers, barbers, and apartment buildings. Because our focus now is macroeconomics, the study of aggregates of economic activity, we will think in terms of the *total* of personal consumption and the *total* of payments to households.

Private Investment

Gross private domestic investment is the value of all goods produced during a period for use in the production of other goods and services. Like personal consumption, gross private domestic investment is a flow variable. It is often simply referred to as "private investment." A hammer produced for a carpenter is private investment. A printing press produced for a magazine publisher is private investment, as is a conveyor-belt system produced for a manufacturing firm. Capital includes all the goods that have been produced for use in producing other goods; it is a stock variable. Private investment is a flow variable that adds to the stock of capital during a period.

Heads Up!

The term "*investment*" can generate confusion. In everyday conversation, we use the term "*investment*" to refer to uses of money to earn income. We say we have invested in a stock or invested in a bond. Economists, however, restrict "*investment*" to activities that increase the economy's stock of capital. The purchase of a share of stock does not add to the capital stock; it is not investment in the economic meaning of the word. We refer to the exchange of financial assets, such as stocks or bonds, as financial investment to distinguish it from the creation of capital that occurs as the result of investment. Only when new capital is produced does investment occur. Confusing the economic concept of private investment with the concept of financial investment can cause misunderstanding of the way in which key components of the economy relate to one another.

Gross private domestic investment includes three flows that add to or maintain the nation's capital stock: expenditures by business firms on new buildings, plants, tools, equipment, and software that will be used in the production of goods and services; expenditures on new residential housing; and changes in business inventories. Any addition to a firm's inventories represents an addition to investment; a reduction subtracts from investment. For example, if a clothing store stocks 1,000 pairs of jeans, the jeans represent an addition to inventory and are part of gross private domestic investment. As the jeans are sold, they are subtracted from inventory and thus subtracted from investment.

By recording additions to inventories as investment and reductions from inventories as subtractions from investment, the accounting for GDP records production in the period in which it occurs. Suppose, for example, that Levi Strauss manufactures 1 million pairs of jeans late in 2007 and distributes them to stores at the end of December. The jeans will be added to inventory; they thus count as investment in 2007 and enter GDP for that year. Suppose they are sold in January 2008. They will be counted as consumption in GDP for 2008 but subtracted from inventory, and from investment. Thus, the production of the jeans will add to GDP in 2007, when they were produced. They will not count in 2008, save for any increase in the value of the jeans resulting from the services provided by the retail stores that sold them.

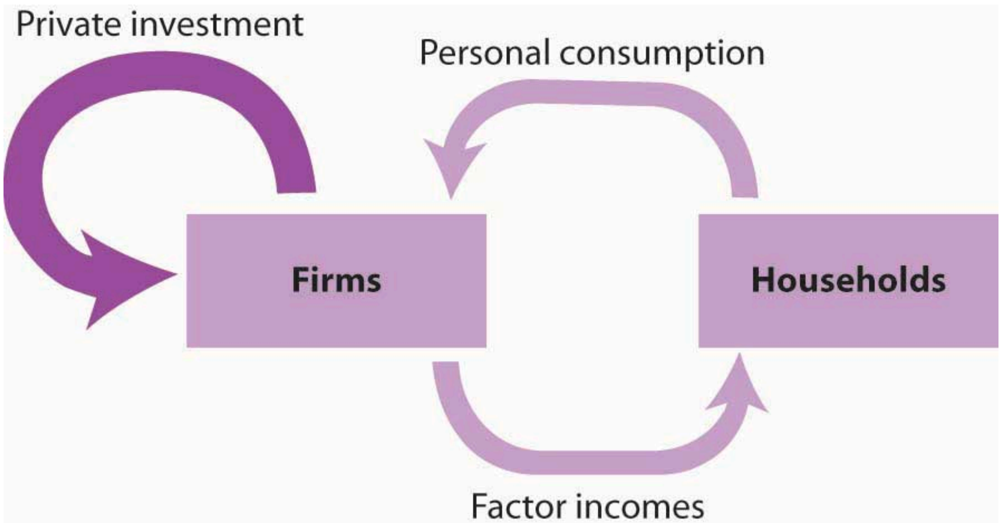
Private investment accounts for about 16% of GDP and at times even less. Despite its relatively small share of total economic activity, private investment plays a crucial role in the macroeconomy for two reasons:

1. Private investment represents a choice to forgo current consumption in order to add to the capital stock of the economy. Private investment therefore adds to the economy's capacity to produce and shifts its production possibilities curve outward. Investment is thus one determinant of economic growth, which is explored in another chapter.
2. Private investment is a relatively volatile component of GDP; it can change dramatically from one year to the next. Fluctuations in GDP are often driven by fluctuations in private investment. We will examine the determinants of

private investment in a chapter devoted to the study of investment.

Private investment represents a demand placed on firms for the production of capital goods. While it is a demand placed on firms, it flows from firms. In the circular flow model in Figure 6.2 "Private Investment in the Circular Flow", we see a flow of investment going from firms to firms. The production of goods and services for consumption generates factor incomes to households; the production of capital goods for investment generates income to households as well.

Figure 6.2 Private Investment in the Circular Flow



Private investment constitutes a demand placed on firms by other firms. It also generates factor incomes for households. To simplify the diagram, only the spending flows are shown—the corresponding flows of goods and services have been omitted.

Figure 6.2 "Private Investment in the Circular Flow" shows only spending flows and omits the physical flows represented by the arrows in Figure 6.1 "Personal Consumption in the Circular Flow". This simplification will make our analysis of the circular flow model easier. It will also focus our attention on spending flows, which are the flows we will be studying.

Government Purchases

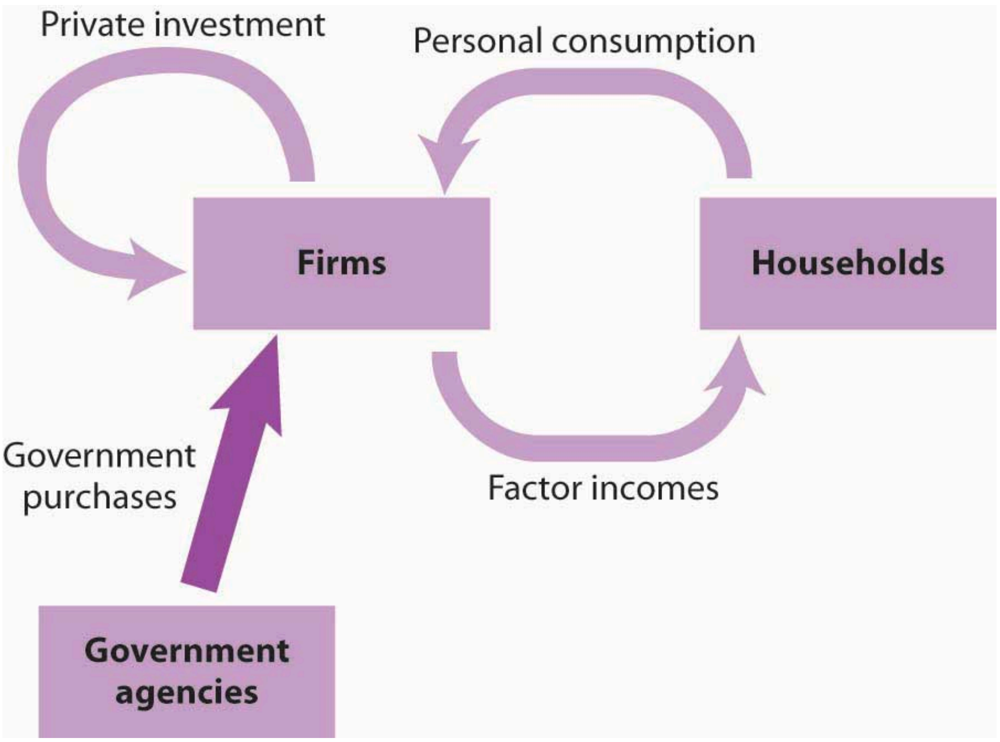
Government agencies at all levels purchase goods and services from firms. They purchase office equipment, vehicles, buildings, janitorial services, and so on. Many government agencies also produce goods and services. Police departments produce police protection. Public schools produce education. The Lakeridge Health System provides medical services.

Government purchases are the sum of purchases of goods and services from firms by government agencies plus the total value of output produced by government agencies themselves during a time period. Government purchases make up about 20-25% of GDP.

Government purchases are not the same thing as government spending. Much government spending takes the form of transfer payments, which are payments that do not require the recipient to produce a good or service in order to receive them. Transfer payments include Social Security and other types of assistance to retired people, welfare payments to poor people, and unemployment compensation to people who have lost their jobs. Transfer payments are certainly significant—they account for roughly half of all federal government spending in the United States. They do not count in a nation's GDP, because they do not reflect the production of a good or service.

Government purchases represent a demand placed on firms, represented by the flow shown in Figure 6.3 "Government Purchases in the Circular Flow". Like all the components of GDP, the production of goods and services for government agencies creates factor incomes for households.

Figure 6.3 Government Purchases in the Circular Flow



Purchases of goods and services by government agencies create demands on firms. As firms produce these goods and services, they create factor incomes for households.

Net Exports

Sales of a country's goods and services to buyers in the rest of the world during a particular time period represent its exports. A purchase by a Japanese buyer of a Ford Taurus produced in Canada is a Canadian export. Exports also include such transactions as the purchase of accounting services from a Toronto accounting firm by a shipping line based in Hong Kong or the purchase of a ticket to Canada's Wonderland by a tourist from the United States. Imports are purchases of foreign-produced goods and services by a country's residents during a period. Canadian imports include such transactions as the purchase by Canadians of cars produced in Japan or tomatoes grown in Mexico or a stay in a French hotel by a tourist from Canada. Subtracting imports from exports yields net exports.

Equation 6.2

$$Exports\ (X) - imports\ (M) = net\ exports\ (X_n)$$

Canada's services trade with the world in 2017: $251.1 billion

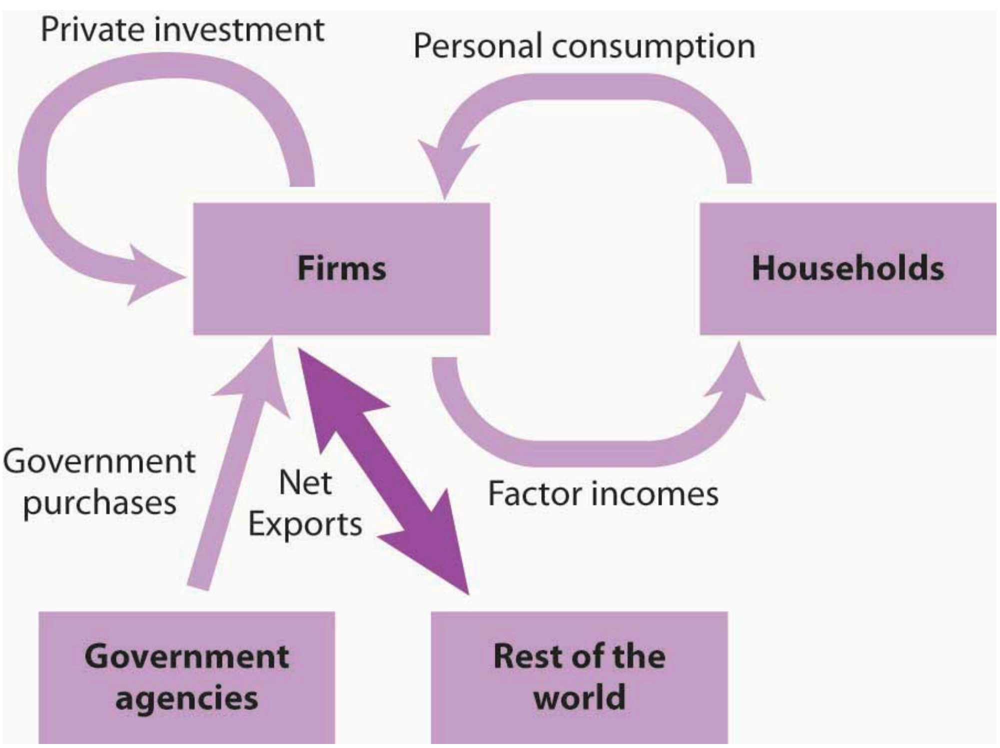
- Exports: $113.1 billion, a 3.9% increase from 2016
- Imports: $138.0 billion, a 4.4% increase from 2016

Services trade deficit in 2017: $24.8 billion, an increase from $23.3 billion in 2016

Notice that imports were greater than exports in this area. The difference between these two figures reveals that net exports were negative because imports exceeded exports. Negative net exports constitute a trade deficit. The amount of the deficit is the amount by which imports exceed exports. When exports exceed imports there is a trade surplus. The magnitude of the surplus is the amount by which exports exceed imports.

In the circular flow diagram in Figure 6.4 "Net Exports in the Circular Flow", net exports are shown with an arrow connecting firms to the rest of the world. The balance between the flows of exports and imports is net exports. When there is a trade surplus, net exports are positive and add spending to the circular flow. A trade deficit implies negative net exports; spending flows from firms to the rest of the world.

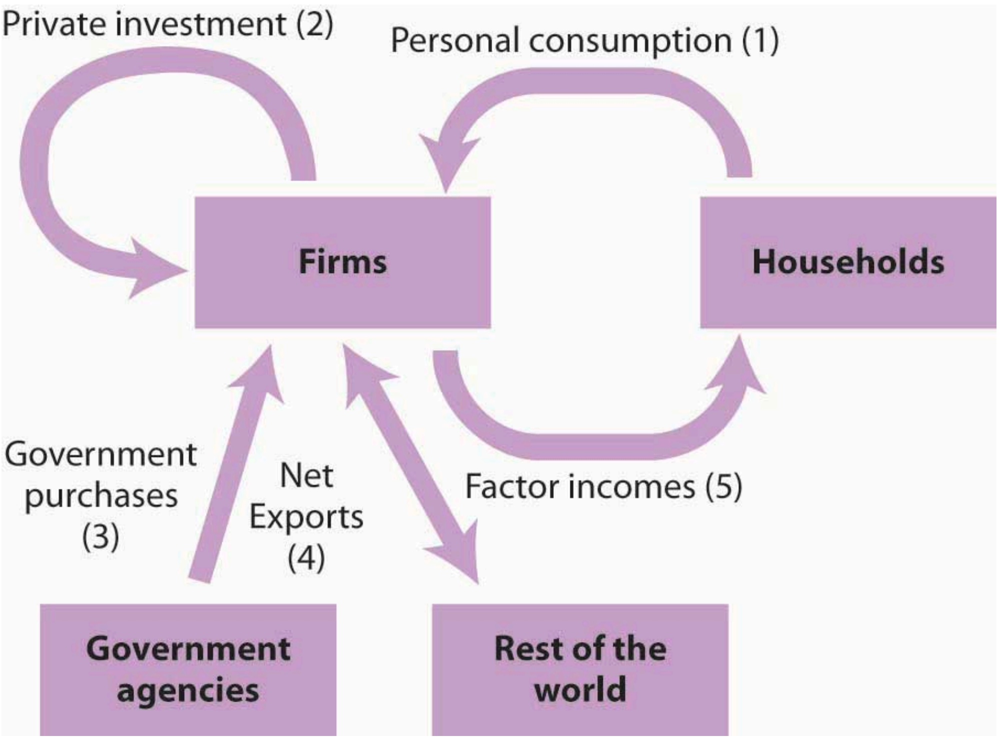
Figure 6.4 Net Exports in the Circular Flow



Net exports represent the balance between exports and imports. Net exports can be positive or negative. If they are positive, net export spending flows from the rest of the world to firms. If they are negative, spending flows from firms to the rest of the world.

The production of goods and services for personal consumption, private investment, government purchases, and net exports makes up a nation's GDP. Firms produce these goods and services in response to demands from households (personal consumption), from other firms (private investment), from government agencies (government purchases), and from the rest of the world (net exports). All of this production creates factor income for households. Figure 6.5 "Spending in the Circular Flow Model" shows the circular flow model for all the spending flows we have discussed. Each flow is numbered for use in the exercise at the end of this section.

Figure 6.5 Spending in the Circular Flow Model

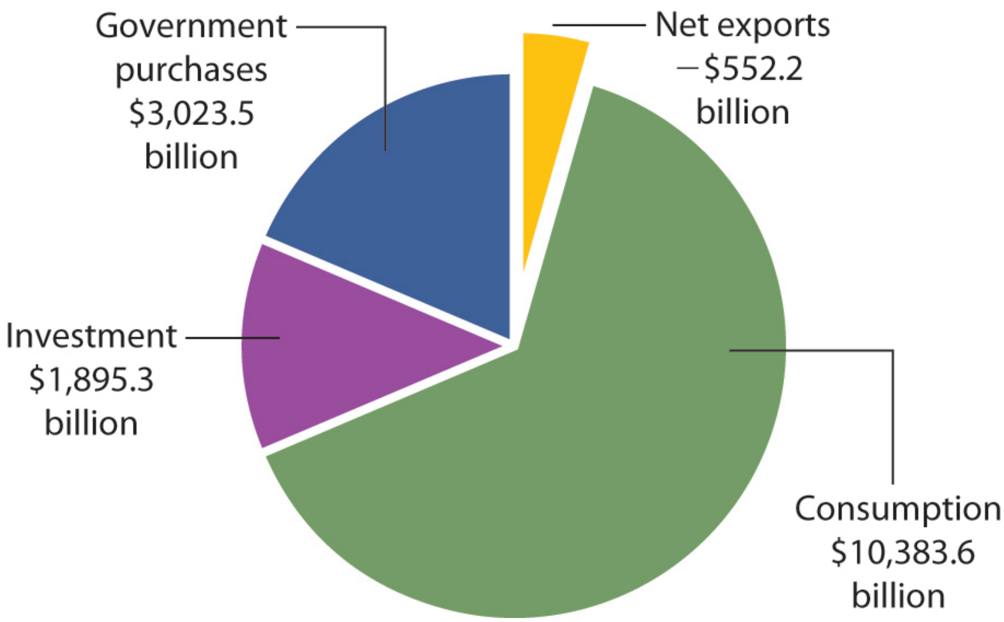


GDP equals the sum of production by firms of goods and services for personal consumption (1), private investment (2), government purchases (3), and net exports (4). The circular flow model shows these flows and shows that the production of goods and services generates factor incomes (5) to households.

The circular flow model identifies some of the forces at work in the economy, forces that we will be studying in later chapters. For example, an increase in any of the flows that place demands on firms (personal consumption, private investment, government purchases, and exports) will induce firms to expand their production. This effect is characteristic of the expansion phase of the business cycle. An increase in production will require firms to employ more factors of production, which will create more income for households. Households are likely to respond with more consumption, which will induce still more production, more income, and still more consumption. Similarly, a reduction in any of the demands placed on firms will lead to a reduction in output, a reduction in firms' use of factors of production, a reduction in household incomes, a reduction in income, and so on. This sequence of events is characteristic of the contraction phase of the business cycle. Much of our work in macroeconomics will involve an analysis of the forces that prompt such changes in demand and an examination of the economy's response to them.

Figure 6.6 "Components of GDP, 2010 (Q3) in Billions of Dollars" shows the size of the components of GDP in the United States in 2010. The statistics for GDP in the US are roughly ten times that which are recorded in Canada. Remember however, that the US has about ten times the population as Canada as well. We see that the production of goods and services for personal consumption accounted for about 70% of GDP. Imports exceeded exports, so net exports were negative.

Figure 6.6 Components of GDP, 2010 (Q3) in Billions of Dollars



Consumption makes up the largest share of GDP. Net exports were negative in 2010. Total GDP—the sum of personal consumption, private investment, government purchases, and net exports—equaled $14,750.2 billion in 2010.

Source:

Bureau of Economic Analysis, National Income and Product Accounts, Table 1.1.5 (December 2010)

Final Goods and Value Added

GDP is the total value of all *final* goods and services produced during a particular period valued at prices in that period. That is not the same as the total value of all goods and services produced during a period. This distinction gives us another method of estimating GDP in terms of output.

Suppose, for example, that a logger cuts some trees and sells the logs to a sawmill. The mill makes lumber and sells it to a construction firm, which builds a house. The market price for the lumber includes the value of the logs; the price of the house includes the value of the lumber. If we try to estimate GDP by adding the value of the logs, the lumber, and the house, we would be counting the lumber twice and the logs three times. This problem is called "double counting," and the economists who compute GDP seek to avoid it.

In the case of logs used for lumber and lumber produced for a house, GDP would include the value of the house. The lumber and the logs would not be counted as additional production because they are intermediate goods that were produced for use in building the house.

Another approach to estimating the value of final production is to estimate for each stage of production the value added, the amount by which the value of a firm's output exceeds the value of the goods and services the firm purchases from other firms. Table 6.1 "Final Value and Value Added" illustrates the use of value added in the production of a house.

Table 6.1 Final Value and Value Added

Good	**Produced by**	**Purchased by**	**Price**	**Value Added**
Logs	Logger	Sawmill	$12,000	$12,000
Lumber	Sawmill	Construction firm	$25,000	$13,000
House	Construction firm	Household	$125,000	$100,000
		Final Value	**$125,000**	
		Sum of Values Added		**$125,000**

If we sum the value added at each stage of the production of a good or service, we get the final value of the item. The example shown here involves the construction of a house, which is produced from lumber that is, in turn, produced from logs.

Suppose the logs produced by the logger are sold for $12,000 to a mill, and that the mill sells the lumber it produces from these logs for $25,000 to a construction firm. The construction firm uses the lumber to build a house, which it sells to a household for $125,000. (To simplify the example, we will ignore inputs other than lumber that are used to build the house.) The value of the final product, the house, is $125,000. The value added at each stage of production is estimated as follows:

1. The logger adds $12,000 by cutting the logs.
2. The mill adds $13,000 ($25,000 – $12,000) by cutting the logs into lumber.
3. The construction firm adds $100,000 ($125,000 – $25,000) by using the lumber to build a house.

The sum of values added at each stage ($12,000 + $13,000 + $100,000) equals the final value of the house, $125,000.

The value of an economy's output in any period can thus be estimated in either of two ways. The values of final goods and services produced can be added directly, or the values added at each stage in the production process can be added. The Commerce Department uses both approaches in its estimate of the nation's GDP.

GNP: An Alternative Measure of Output

While GDP represents the most commonly used measure of an economy's output, economists sometimes use an alternative measure. Gross national product (GNP) is the total value of final goods and services produced during a particular period with factors of production owned by the residents of a particular country.

The difference between GDP and GNP is a subtle one. The GDP of a country equals the value of final output produced within the borders of that country; the GNP of a country equals the value of final output produced using factors owned by residents of the country. Most production in a country employs factors of production owned by residents of that country, so the two measures overlap. Differences between the two measures emerge when production in one country employs factors of production *owned* by residents of other countries.

Suppose, for example, that a resident of Montreal, Quebec owns and operates a watch repair shop across the Canadian–U.S. border in Vermont. The value of watch repair services produced at the shop would be counted as part of US's GDP because they are produced in the USA. That value would not, however, be part of Canadian GDP. But, because the watch repair services were produced using capital and labour provided by a resident of Canada, they would be counted as part of GNP in Canada and not as part of GNP in the US.

Because most production fits in both a country's GDP as well as its GNP, there is seldom much difference between the two measures. The relationship between GDP and GNP is given by

Equation 6.3

$$\text{GDP} + \text{net income received from abroad by residents of a nation} = \text{GNP}$$

Key Takeaways

- GDP is the sum of final goods and services produced for consumption (C), private investment (I), government purchases (G), and net exports (X_n). Thus GDP = C + I + G + X_n.
- GDP can be viewed in the context of the circular flow model. Consumption goods and services are produced in response to demands from households; investment goods are produced in response to demands for new capital by firms; government purchases include goods and services purchased by government agencies; and net exports equal exports less imports.
- Total output can be measured two ways: as the sum of the values of final goods and services produced and as the sum of values added at each stage of production.
- GDP plus net income received from other countries equals GNP. GNP is the measure of output typically used to compare incomes generated by different economies.

Try It!

Here is a two-part exercise.

1. Suppose you are given the following data for an economy:

Personal consumption	$1,000
Home construction	100
Increase in inventories	40
Equipment purchases by firms	60
Government purchases	100
Social Security payments to households	40
Government welfare payments	100
Exports	50
Imports	150

Identify the number of the flow in Figure 6.5 "Spending in the Circular Flow Model" to which each of these items corresponds. What is the economy's GDP?

2. Suppose a dairy farm produces raw milk, which it sells for $1,000 to a dairy. The dairy produces cream, which it sells for $3,000 to an ice cream manufacturer. The ice cream manufacturer uses the cream to make ice cream, which it sells for $7,000 to a grocery store. The grocery store sells the ice cream to consumers for $10,000. Compute the value added at each stage of production, and compare this figure to the final value of the product produced. Report your results in a table similar to that given in Table 6.1 "Final Value and Value Added".

Case in Point: The Spread of the Value Added Tax

Figure 21.7

401(K) 2012 – Tax – CC BY-SA 2.0.

Outside the United States, the value added tax (VAT) has become commonplace. Governments of more than 120 countries use it as their primary means of raising revenue. While the concept of the VAT originated in France in the 1920s, no country adopted it until after World War II. In 1948, France became the first country in the world to use the VAT. In 1967, Brazil became the first country in the Western Hemisphere to do so. The VAT spread to other western European and Latin American countries in the 1970s and 1980s and then to countries in the Asia/Pacific region, central European and former Soviet Union area, and Africa in the 1990s and early 2000s.

What is the VAT? It is equivalent to a sales tax on final goods and services but is collected at each stage of production.

Take the example given in Table 6.1 "Final Value and Value Added", which is a simplified illustration of a house built in three stages. If there were a sales tax of 10% on the house, the household buying it would pay $137,500, of which the construction firm would keep $125,000 of the total and turn $12,500 over to the government.

With a 10% VAT, the sawmill would pay the logger $13,200, of which the logger would keep $12,000 and turn $1,200 over to the government. The sawmill would sell the lumber to the construction firm for $27,500—keeping $26,200, which is the $25,000 for the lumber itself and $1,200 it already paid in tax. The government at this stage would get $1,300, the difference between the $2,500 the construction firm collected as tax and the $1,200 the sawmill already paid in tax to the logger at the previous stage. The household would pay the construction firm $137,500. Of that total, the construction firm would turn over to the government $10,000, which is the difference between the $12,500 it collected for the government in tax from the household and the $2,500 in tax that it already paid when it bought the lumber from the sawmill. The table below shows that in the end, the tax revenue generated by a 10% VAT is the same as that generated by a 10% tax on final sales.

Why bother to tax in stages instead of just on final sales? One reason is simply record keeping, since it may be difficult to determine in practice if any particular sale is the final one. In the example, the construction firm does not need to know if it is selling the house to a household or to some intermediary business.

Also, the VAT may lead to higher revenue collected. For example, even if somehow the household buying the house avoided paying the tax, the government would still have collected some tax revenue at earlier stages of production. With a tax on retail sales, it would have collected nothing. The VAT has another advantage from the point of view of government agencies. It has the appearance at each stage of taking a smaller share. The individual amounts collected are not as obvious to taxpayers as a sales tax might be.

Good	**Price**	**Value Added**	**Tax Collected**	**– Tax Already Paid**	**= Value Added Tax**
Logs	$12,000	$12,000	$1,200	– $0	= $1,200
Lumber	$25,000	$13,000	$2,500	– $1,200	= $1,300
House	$125,000	$100,000	$12,500	– $2,500	= $10,000
Total			**$16,200**	**–$3,700**	**= $12,500**

Answer to Try It! Problem

1. GDP equals $1,200 and is computed as follows (the numbers in parentheses correspond to the flows in Figure 6.5 “Spending in the Circular Flow Model”):

Personal consumption (1)	$1,000
Private investment (2)	200
Housing	100
Equipment and software	60
Inventory change	40
Government purchases (3)	100
Net exports (4)	–100
GDP	$1,200

Notice that neither welfare payments nor Social Security payments to households are included. These are transfer payments, which are not part of the government purchases component of GDP.

2. Here is the table of value added.

Good	Produced by	Purchased by	Price	Value Added
Raw milk	Dairy farm	Dairy	$1,000	$1,000
Cream	Dairy	Ice cream maker	3,000	2,000
Ice cream	Ice cream manufacturer	Grocery store	7,000	4,000
Retail ice cream	Grocery store	Consumer	10,000	3,000
		Final Value	**$10,000**	
		Sum of Values Added		**$10,000**

6.2 Measuring Total Income

Learning Objectives

1. Define gross domestic income and explain its relationship to gross domestic product.
2. Discuss the components of gross domestic income.
3. Define disposable personal income and explain how to calculate it from GDP.

We saw in the last section that the production of goods and services generates factor incomes to households. The production of a given value of goods and services generates an equal value of total income. Gross domestic income (GDI) equals the total income (Y) generated in an economy by the production of final goods and services during a particular period. It is a flow variable. Because an economy's total output equals the total income generated in producing that output, GDP = GDI (or Y). We can estimate GDP either by measuring total output or by measuring total income.

Consider a $4 box of Cheerios. It is part of total output and thus is part of GDP. Who gets the $4? Part of the answer to that question can be found by looking at the cereal box. Cheerios are made from oat flour, wheat starch, sugar, salt, and a variety of vitamins and minerals. Therefore, part of the $4 goes to the farmers who grew the oats, the wheat, and the beets or cane from which the sugar was extracted. Workers and machines at General Mills combined the ingredients, crafted all those little O's, toasted them, and put them in a box. The workers were paid part of the $4 as wages. The owners of General Mills and the capital it used received part of the $4 as profit. The box containing the Cheerios was made from a tree, so a lumber company somewhere received part of the $4. The truck driver who brought the box of cereal to the grocery store got part of the $4, as did the owner of the truck itself and the owner of the oil that fueled the truck. The clerk who rang up the sale at the grocery store was paid part of the $4. And so on.

How much of the $4 was income generated in the production of the Cheerios? The answer is simple: all of it. Some of the money went to workers as wages. Some went to owners of the capital and natural resources used to produce it. Profits generated along the way went to the owners of the firms involved. All these items represent costs of producing the Cheerios and also represent income to households.

Part of the $4 cost of the Cheerios, while it makes up a portion of GDI (or Y), does not represent ordinary income actually earned by households. That part results from two other production costs: depreciation and taxes related to the production of the Cheerios. Nevertheless, they are treated as a kind of income; we will examine their role in GDI (or Y) below.

As it is with Cheerios, so it is with everything else. The value of output equals the income generated as the output is produced.

The Components of GDI (or Y)

Employee compensation is the largest among the components of factor income (called 'wages and salaries'). Factor income also includes profit (the payment to entrepreneurs), rent (the payment to natural resources), and interest (the payment to capital). In addition, GDI (or Y) includes charges for depreciation and taxes associated with production.

Depreciation and production-related taxes, such as sales taxes, make up part of the cost of producing goods and services and must be accounted for in estimating GDI (or Y). We will discuss each of these components of GDI (or Y) next.

Employee Compensation

Compensation of employees in the form of wages, salaries, and benefits makes up the largest single component of income generated in the production of GDP. In the second quarter of 2008, employee compensation represented approximately 57% of GDI (or Y).

The structure of employee compensation has changed dramatically in the last several decades. In 1950, virtually all employee compensation—95% of it—came in the form of wages and salaries. The remainder, about 5%, came in the form of additional benefits such as employer contributions to retirement programs and health insurance. In 2008, the share of benefits was roughly 20% of total employee compensation.

Profits

The profit component of income earned by households equals total revenues of firms less costs as measured by conventional accounting. Profits amounted to about 16% of GDI, or $2,332.2 billion in 2010, down sharply from five decades earlier, when profits represented about 25% of the income generated in GDI[1].

Profits are the reward the owners of firms receive for being in business. The opportunity to earn profits is the driving force behind production in a market economy.

Rental Income

Rental income, such as the income earned by owners of rental housing or payments for the rent of natural resources, is the smallest component of GDI (about 2%); it is the smallest of the income flows to households. The meaning of rent in the computation of GDI is the same as its meaning in conventional usage; it is a charge for the temporary use of some capital asset or natural resource[2].

Net Interest

Businesses both receive and pay interest. GDI includes net interest, which equals interest paid less interest received by domestic businesses, plus interest received from foreigners less interest paid to foreigners. Interest payments on mortgage and home improvement loans are counted as interest paid by business, because homeowners are treated as businesses in the income accounts. In 2010 net interest accounted for 6.3% of GDI.

Depreciation

Over time the machinery and buildings that are used to produce goods and services wear out or become obsolete. A farmer's tractor, for example, wears out as it is used. A technological change may make some equipment obsolete. The introduction of personal computers, for example, made the electric typewriters used by many firms obsolete. Depreciation is a measure of the amount of capital that wears out or becomes obsolete during a period. Depreciation is referred to in official reports as the consumption of fixed capital.

Depreciation is a cost of production, so it represents part of the price charged for goods and services. It is therefore counted as part of the income generated in the production of those goods and services. Depreciation represented about 13% of GDI in 2008.

Indirect Taxes

The final component of the income measure of GDI is indirect business taxes[3]. Indirect taxes are taxes imposed on the production or sale of goods and services or on other business activity. (By contrast, a direct tax is a tax imposed directly on income; the personal income and corporate income taxes are direct taxes.) Indirect taxes, which include sales and excise taxes and property taxes, make up part of the cost to firms of producing goods and services. Like depreciation, they are part of the price of those goods and services and are therefore treated as part of the income generated in their production. Indirect business taxes amounted to 7.7% of GDI in 2010.

Table 6.2 "GDP and GDI (or Y), 2010" shows the components of GDI (or Y) in 2010. Employee compensation represented the largest share of GDI. The exhibit also shows the components of GDP for the same year.

In principle, GDP and GDI (or Y) should be equal, but their estimated values never are, because the data come from different sources. Output data from a sample of firms are used to estimate GDP, while income data from a sample of households are used to estimate GDI. The difference is the statistical discrepancy shown in the right-hand column of Table 6.2 "GDP and GDI, 2010". Some of the difficulties with these data are examined in the Case in Point feature on discrepancies between GDP and GDI.

Table 6.2 GDP and GDI (or Y), 2010, USA

Gross domestic product	$14,750.2	Gross Domestic Income	$14,584.8
Personal Consumption Expenditures	10,383.6	Compensation of Employees	8,040.8
Gross Private Domestic Investment	1,895.3	Profits[4]	2,332.3
Government consumption expenditures and gross investment	3,023.5	Rental income of persons	304.7
Net exports of goods and services	− 552.2	Net interest	913.9
		Taxes on production and imports[5]	1,121.0
		Consumption of fixed capital (depreciation)	1872.1
		Statistical discrepancy	165.4

The table shows the composition of GDP and GDI in the third quarter of 2010 (in billions of dollars at an annual rate) for the USA. Notice the rough equality of the two measures. (They are not quite equal because of measurement errors; the difference is due to a statistical discrepancy and is reduced significantly over time as the data are revised.)

Sources: Bureau of Economic Analysis National Income and Product Accounts, Tables 1.10 and 1.1.5 (December, 2010). See Brent R. Moulton and Eugene P. Seskin, "Preview of the 2003 Comprehensive Revision of the National Income and Product Accounts," Bureau of Economic Analysis, *Survey of Current Business*, June 2003, pp. 17–34.

Key Takeaways

- Gross domestic product, GDP, equals gross domestic income, GDI (or Y), which includes compensation, profits, rental income, indirect taxes, and depreciation.
- We can use GDP, a measure of total output, to compute disposable personal income, a measure of income received by households and available for them to spend.

Try It!

The following income data refer to the same economy for which you had output data in the first part of the previous Try It! Compute GDI (or Y) from the data below and confirm that your result equals the GDP figure you computed in the previous Try It! Assume that GDP = GNP for this problem (that is, assume all factor incomes are earned and paid in the domestic economy).

Employee compensation	$700
Social Security payments to households	40
Welfare payments	100
Profits	200
Rental income	50
Net interest	25
Depreciation	50
Indirect taxes	175

A simplifying assumption can be made for the purposes of this course. That is, to simplify, one can think of indirect taxes and depreciation as technical adjustments to reconcile the expenditure approach to measuring GDP with the Income Approach to measuring GDP. Simplifying allows us to see the Income Approach as the sum of the factor payments:

$$\mathbf{Y = W + R + i + \pi}$$

Answer to Try It! Problem

GDI (or Y) equals $1,200. Note that this value equals the value for GDP obtained from the estimate of output in the first part of the previous Try It! Here is the computation:

Employee compensation	$700
Profits	200
Rental income	50
Net interest	25
Depreciation	50
Indirect taxes	175
GDI	$1,200

Once again, note that Social Security and welfare payments to households are transfer payments. They do not represent payments to household factors of production for current output of goods and services, and therefore are not included in GDI.

[1]Although reported separately by the Department of Finance, we have combined proprietors' income (typically independent business owners and farmers) with corporate profits to simplify the discussion.

[2]If you have studied microeconomics, you know that the term "*rent*" in economics has a quite different meaning. The national income and product accounts use the accounting, not the economic, meaning of "*rent*."

[3]The adjustment for indirect business taxes includes two other minor elements: transfer payments made by business firms and surpluses or deficits of government enterprises.

[4]Profit is corporate profit ($1,274.7) plus proprietors' income ($1,057.6), both with inventory valuation and capital consumption adjustment.

[5]Indirect taxes include taxes on production and imports of $1,060.6 plus business transfer payments ($133.1) less subsidies ($58.5) and current surplus of government enterprise ($14.1). Prior to the 2003 National Income and Product Accounts (NIPA) revisions, the category "taxes on production and imports" was, with some technical and other minor adjustments, referred to as "indirect business taxes." See Brent R. Moulton and Eugene P. Seskin, "Preview of the 2003 Comprehensive Revision of the National Income and Product Accounts," Bureau of Economic Analysis, *Survey of Current Business*, June 2003, pp. 17-34.

6.3 GDP and Economic Well-Being

Learning Objectives

1. Discuss and give examples of measurement and conceptual problems in using real GDP as a measure of economic performance and of economic well-being.
2. Explain the use of per capita real GNP or GDP to compare economic performance across countries and discuss its limitations.

GDP is the measure most often used to assess the economic well-being of a country. Besides measuring the pulse of a country, it is the figure used to compare living standards in different countries.

Of course, to use GDP as an indicator of overall economic performance, we must convert nominal GDP to real GDP, since nominal values can rise or fall simply as a result of changes in the price level. For example, the movie *Titanic*, released in 1997, brought in $601 million—the highest amount ever in gross box office receipts, while *Gone with the Wind*, released in 1939, earned only $199 million and ranks 49th in terms of nominal receipts. But does that mean that *Titanic* actually did better than *Gone with the Wind*? After all, the average price of a movie ticket in 1939 was 25 cents. At the time of *Titanic*, the average ticket price was about $5. A better way to compare these two movies in terms of popularity is to control for the price of movie tickets—the same strategy that economists use with real GDP in order to determine whether output is rising or falling. Adjusting the nominal box-office receipts using 1998 movie prices to obtain real revenue reveals that in real terms *Gone with the Wind* continues to be the top real grosser of all time with real box-office receipts of about $1.3 billion. As illustrated by this example on revenues from popular movies, we might draw erroneous conclusions about performance if we base them on nominal values instead of on real values. In contrast, real GDP, despite the problems with price indexes that were explained in another chapter, provides a reasonable measure of the total output of an economy, and changes in real GDP provide an indication of the direction of movement in total output.

We begin this section by noting some of the drawbacks of using real GDP as a measure of the economic welfare of a country. Despite these shortcomings, we will see that it probably remains our best single indicator of macroeconomic performance.

Measurement Problems in Real GDP

There are two measurement problems, other than those associated with adjusting for price level changes, in using real GDP to assess domestic economic performance.

Revisions

The first estimate of real GDP for a calendar quarter is called the advance estimate. It is issued about a month after the quarter ends. To produce a measure so quickly, officials at the Department of Commerce must rely on information from relatively few firms and households. One month later, it issues a revised estimate, and a month after that it issues its final estimate. Often the advance estimate of GDP and the final estimate do not correspond. The recession of 2001, for example, began in March of that year. But the first estimates of real GDP for the second and third quarters of 2001 showed output continuing to rise. It was not until later revisions that it became clear that a recession had been under way.

But the revision story does not end there. Every summer, the Commerce Department issues revised figures for the previous two or three years. Once every five years, the department conducts an extensive analysis that traces flows of inputs and outputs throughout the economy. It focuses on the outputs of some firms that are inputs to other firms. In the process of conducting this analysis, the department revises real GDP estimates for the previous five years. Sometimes the revisions can paint a picture of economic activity that is quite different from the one given even by the revised estimates of GDP. For example, revisions of the data for the 1990–1991 recession issued several years later showed that the recession had been much more serious than had previously been apparent, and the recovery was more pronounced.

The Service Sector

Another problem lies in estimating production in the service sector. The output of goods in an economy is relatively easy to compute. There are so many bushels of corn, so many pounds of beef. But what is the output of a bank? Of a hospital? It is easy to record the dollar value of output to enter in nominal GDP, but estimating the quantity of output to use in real GDP is quite another matter. In some cases, the Department of Commerce estimates service sector output based on the quantity of labor used. For example, if this technique were used in the banking industry and banking used 10% more labor, the department would report that production has risen 10%. If the number of employees remains unchanged, reported output remains unchanged. In effect, this approach assumes that output per worker—productivity—in those sectors remains constant when studies have indicated that productivity has increased greatly in the service sector. Since 1990 progress has been made in measurement in this area, which allows in particular for better estimation of productivity changes and price indexes for different service sector industries, but more remains to be done in this large sector of the U.S. economy (Triplett & Bosworth, 2008).

Conceptual Problems with Real GDP

A second set of limitations of real GDP stems from problems inherent in the indicator itself. Real GDP measures market activity. Goods and services that are produced and exchanged in a market are counted; goods and services that are produced but that are not exchanged in markets are not[1].

Household Production

Suppose you are considering whether to eat at home for dinner tonight or to eat out. You could cook dinner for yourself at a cost of $5 for the ingredients plus an hour or so of your time. Alternatively, you could buy an equivalent meal at a restaurant for perhaps $15. Your decision to eat out rather than cook would add $10 to the GDP.

But that $10 addition would be misleading. After all, if you had stayed home you might have produced an equivalent meal. The only difference is that the value of your time would not have been counted. But surely your time is not worthless; it is just not counted. Similarly, GDP does not count the value of your efforts to clean your own house, to wash your own car, or to grow your own vegetables. In general, GDP omits the entire value added by members of a household who do household work themselves.

There is reason to believe this omission is serious. Economists J. Steven Landefeld and Stephanie H. McCulla of the U.S. Bureau of Economic Analysis estimated in a 2000 paper the value of household output from 1946 to 1997. Their estimate of household output in 1946 was 50% of reported GDP. Since then, that percentage has fallen, because more women have entered the workforce, so that more production that once took place in households now occurs in the market. Households now eat out more, purchase more prepared foods at the grocery store, hire out child-care services they once performed themselves, and so on. Their estimate for 1997, for example, suggests that household production amounted to 36% of reported GDP (Landefeld & McCulla, 2000).

This problem is especially significant when GDP is used to make comparisons across countries. In low-income countries, a much greater share of goods and services is not exchanged in a market. Estimates of GDP in such countries are adjusted to reflect nonmarket production, but these adjustments are inevitably imprecise.

Underground and Illegal Production

Some production goes unreported in order to evade taxes or the law. It is not likely to be counted in GDP. Legal production for which income is unreported in order to evade taxes generally takes place in what is known as the "underground economy." For example, a carpenter might build a small addition to a dentist's house in exchange for orthodontic work for the carpenter's children. Although income has been earned and output generated in this example of bartering, the transaction is unlikely to be reported for income tax or other purposes and thus is not counted in GDP. Illegal activities are not reported for income taxes for obvious reasons and are thus difficult to include in GDP.

Leisure

Leisure is an economic good. All other things being equal, more leisure is better than less leisure.

But all other things are not likely to be equal when it comes to consuming leisure. Consuming more leisure means supplying less work effort. And that means producing less GDP. If everyone decided to work 10% fewer hours, GDP would fall. But that would not mean that people were worse off. In fact, their choice of more leisure would suggest they prefer the extra leisure to the goods and services they give up by consuming it. Consequently, a reduction in GDP would be accompanied by an increase in satisfaction, not a reduction.

The GDP Accounts Ignore "Bads"

Suppose a wave of burglaries were to break out across the United States. One would expect people to respond by buying more and louder burglar alarms, better locks, fiercer German shepherds, and more guard services than they had before. To the extent that they pay for these by dipping into savings rather than replacing other consumption, GDP increases. An epidemic might have much the same effect on GDP by driving up health-care spending. But that does not mean that crime and disease are good things; it means only that crime and disease may force an increase in the production of goods and services counted in the GDP.

Similarly, the GDP accounts ignore the impact of pollution on the environment. We might produce an additional $200 billion in goods and services but create pollution that makes us feel worse off by, say, $300 billion. The GDP accounts simply report the $200 billion in increased production. Indeed, some of the environmental degradation might itself boost GDP. Dirtier air may force us to wash clothes more often, to paint buildings more often, and to see the doctor more often, all of which will tend to increase GDP!

Conclusion: GDP and Human Happiness

More GDP cannot necessarily be equated with more human happiness. But more GDP does mean more of the goods and services we measure. It means more jobs. It means more income. And most people seem to place a high value on these things. For all its faults, GDP does measure the production of most goods and services. And goods and services get produced, for the most part, because we want them. We might thus be safe in giving two cheers for GDP—and holding back the third in recognition of the conceptual difficulties that are inherent in using a single number to summarize the output of an entire economy.

International Comparisons of Real GDP and GNP

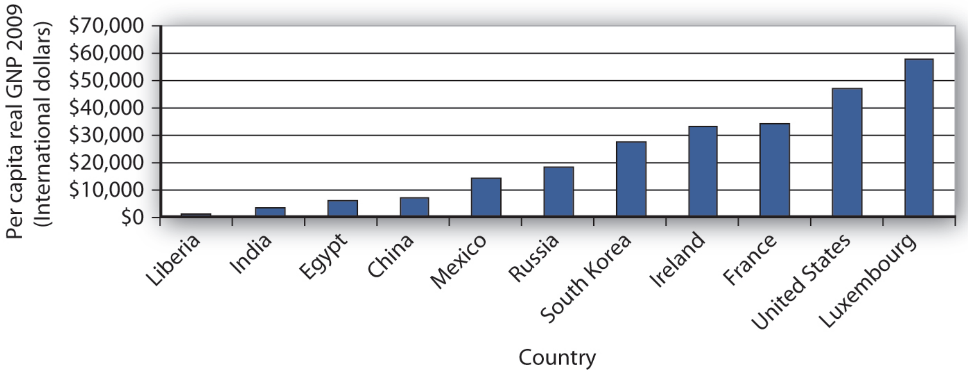
Real GDP or GNP estimates are often used in comparing economic performance among countries. In making such comparisons, it is important to keep in mind the general limitations to these measures of economic performance that we noted earlier. Further, countries use different methodologies for collecting and compiling data.

Three other issues are important in comparing real GDP or GNP for different countries: the adjustment of these figures for population, adjusting to a common currency, and the incorporation of nonmarket production.

In international comparisons of real GNP or real GDP, economists generally make comparisons not of real GNP or GDP but of per capita real GNP or GDP, which equals a country's real GNP or GDP divided by its population. In 2009, for example, Japan had a real GDP of about $4,000 billion and Luxembourg had a real GDP of about $29 billion. We can conclude that Japan's economy produced far more goods and services than did Luxembourg's. But Japan had almost 300 times as many people as did Luxembourg. Japan's per capita real GDP in 2009 was $33,280; Luxembourg's was $57,640, the highest in the world that year.

Figure 21.9 "Comparing Per Capita Real GNP, 2009" compares per capita real GNP for 11 countries in 2009. It is based on data that uses a measure called "international dollars" in order to correct for differences in the purchasing power of $1 across countries. The data also attempt to adjust for nonmarket production (such as that of rural families that grow their own food, make their own clothing, and produce other household goods and services themselves).

Figure 21.9 Comparing Per Capita Real GNP, 2009



There is a huge gap between per capita income in one of the poorest countries in the world, Liberia, and wealthier nations such as the United States and Luxembourg.

Source: World Bank, World Development Indicators Online

The disparities in income are striking; Luxembourg, the country with the highest per capita real GNP, had an income level more than 200 times greater than Liberia, the country with the lowest per capita real GNP.

What can we conclude about international comparisons in levels of GDP and GNP? Certainly we must be cautious. There are enormous difficulties in estimating any country's total output. Comparing one country's output to another presents additional challenges. But the fact that a task is difficult does not mean it is impossible. When the data suggest huge disparities in levels of GNP per capita, for example, we observe real differences in living standards.

Key Takeaways

- Real GDP or real GNP is often used as an indicator of the economic well-being of a country.
- Problems in the measurement of real GDP, in addition to problems encountered in converting from nominal to real GDP, stem from revisions in the data and the difficulty of measuring output in some sectors, particularly the service sector.
- Conceptual problems in the use of real GDP as a measure of economic well-being include the facts that it does not include nonmarket production and that it does not properly adjust for "bads" produced in the economy.
- Per capita real GDP or GNP can be used to compare economic performance in different countries.

Try It!

What impact would each of the following have on real GDP? Would economic well-being increase or decrease as a result?

1. On average, people in a country decide to increase the number of hours they work by 5%.
2. Spending on homeland security increases in response to a terrorist attack.
3. The price level and nominal GDP increase by 10%.

Case in Point: Per Capita Real GDP and Olympic Medal Counts

Figure 21.10

Aurelien Guichard – [Olympics]() – CC BY-SA 2.0.

In the popular lore, the Olympics provide an opportunity for the finest athletes in the world to compete with each other head-to-head on the basis of raw talent and hard work. And yet, contenders from Laos tend to finish last or close to it in almost any event in which they compete. One Laotian athlete garnered the unenviable record of having been the slowest entrant in the nearly half-century long history of the 20-kilometer walk. In contrast, U.S. athletes won 103 medals at the 2004 Athens Olympics and 110 medals at the 2008 Beijing Olympics. Why do Laotians fare so poorly and Americans so well, with athletes from other countries falling in between?

Economists Daniel K. N. Johnson and Ayfer Ali have been able to predict with astonishing accuracy the number of medals different countries will win on the basis of a handful of factors, including population, climate, political structure, and real per capita GDP. For example, they predicted that the United States would win 103 medals in Athens and that is precisely how many the United States won. They predicted 103 medals for the United States in Beijing; 110 were won. They did not expect the Laotians to win any medals in either Athens or Beijing, and that was indeed the outcome.

Johnson and Ali estimated that summer game participant nations average one more medal per additional $1,000 of per capita real GDP. With per capita real GDP in Laos less than the equivalent of $500 compared to per capita real GDP in the United States of about $38,000, the results for these two nations could be considered a foregone conclusion. According to Johnson and Ali, "High productive capacity or income per person displays an ability to pay the costs necessary to send athletes to the Games, and may also be associated with a higher quality of training and better equipment." For example, a Laotian swimmer at Athens, Vilayphone Vongphachanh, had never practiced in an Olympic-size pool, and a runner, Sirivanh Ketavong, had worn the same running shoes for four years.

The good news is that as the per capita real GDP in some relatively poor countries has risen, the improved living standards have led to increased Olympic medal counts. China, for instance, won 28 medals in 1988 and 63 in 2004. As the host for the 2008 games, it won an impressive total of 100 medals.

While not a perfect measure of the well-being of people in a country, per capita real GDP does tell us about the opportunities available to the average citizen in a country. Americans would surely find it hard to imagine living at the level of consumption of the average Laotian. In *The Progress Paradox: How Life Gets Better While People Feel Worse*, essayist Gregg Easterbrook notes that a higher material standard of living is not associated with higher reported happiness. But, he concludes, the problems of prosperity seem less serious than those of poverty, and prosperity gives people and nations the means to address problems. The Olympic medal count for each nation strongly reflects its average standard of living and hence the opportunities available to its citizens.

Sources: Gregg Easterbrook, *The Progress Paradox: How Life Gets Better While People Feel Worse* (New York: Random House, 2003); Daniel K. N. Johnson and Ayfer Ali, "A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympic Games," *Social Science Quarterly* 84, no. 4 (December 2004): 974–93; David Wallechinsky, "Why I'll Cheer for Laos," *Parade Magazine*, August 8, 2004, p. 8.

Answer to Try It! Problem

1. Real GDP would increase. Assuming the people chose to increase their work effort and forgo the extra leisure, economic well-being would increase as well.
2. Real GDP would increase, but the extra expenditure in the economy was due to an increase in something "bad," so economic well-being would likely be lower.
3. No change in real GDP. For some people, economic well-being might increase and for others it might decrease, since inflation does not affect each person in the same way.

[1]There are two exceptions to this rule. The value of food produced and consumed by farm households is counted in GDP. More important, an estimate of the rental values of owner-occupied homes is included. If a family rents a house, the rental payments are included in GDP. If a family lives in a house it owns, the Department of Commerce estimates what

the house would rent for and then includes that rent in the GDP estimate, even though the house's services were not exchanged in the marketplace.

References

Landefeld, J. S., and Stephanie H. McCulla, "Accounting for Nonmarket Household Production within a National Accounts Framework," *Review of Income & Wealth* 46, no. 3 (September 2000): 289–307.

Triplett, J. E., and Barry P. Bosworth, "The State of Data for Services Productivity Measurement in the United States," *International Productivity Monitor* 16 (Spring 2008): 53–70.

6.4 Review and Practice

Summary

This chapter focused on the measurement of GDP. The total value of output (GDP) equals the total value of income generated in producing that output (GDI or Y). We can illustrate the flows of spending and income through the economy with the circular flow model. Firms produce goods and services in response to demands from households (personal consumption), other firms (private investment), government agencies (government purchases), and the rest of the world (net exports). This production, in turn, creates a flow of factor incomes to households. Thus, GDP can be estimated using two types of data: (1) data that show the total value of output and (2) data that show the total value of income generated by that output.

In looking at GDP as a measure of output, we divide it into four components: consumption, investment, government purchases, and net exports. GDP equals the sum of final values produced in each of these areas. It can also be measured as the sum of values added at each stage of production. The components of GDP measured in terms of income (GDI or Y) are employee compensation, profits, rental income, net interest, depreciation, and indirect taxes.

We also explained other measures of income such as GNP and disposable personal income. Disposable personal income is an important economic indicator, because it is closely related to personal consumption, the largest component of GDP.

GDP is often used as an indicator of how well off a country is. Of course, to use it for this purpose, we must be careful to use real GDP rather than nominal GDP. Still, there are problems with our estimate of real GDP. Problems encountered in converting nominal GDP to real GDP were discussed in the previous chapter. In this chapter we looked at additional measurement problems and conceptual problems.

Frequent revisions in the data sometimes change our picture of the economy considerably. Accounting for the service sector is quite difficult. Conceptual problems include the omission of nonmarket production and of underground and illegal production. GDP ignores the value of leisure and includes certain “bads.”

We cannot assert with confidence that more GDP is a good thing and that less is bad. However, real GDP remains our best single indicator of economic performance. It is used not only to indicate how any one economy is performing over time but also to compare the economic performance of different countries.

Concept Problems

1. GDP is used as a measure of macroeconomic performance. What, precisely, does it measure?
2. Many economists have attempted to create a set of social accounts that would come closer to measuring the economic well-being of the society than does GDP. What modifications of the current approach would you recommend to them?

3. Every good produced creates income for the owners of the factors of production that created the product or service. For a recent purchase you made, try to list all the types of factors of production involved in making the product available, and try to determine who received income as a result of your purchase.
4. Explain how the sale of used textbooks in your campus bookstore affects the GDP calculation.
5. Look again at the circular flow diagram in Figure 6.5 "Spending in the Circular Flow Model" and assume it is drawn for Canada. State the flows in which each of the following transactions would be entered.
    1. A consumer purchases fresh fish at a local fish market.
    2. A grocery store acquires 1,000 rolls of paper towels for later resale.
    3. People in France flock to Canada to see the latest play.
    4. A construction firm builds a new house.
    5. A couple from British Columbia visits Guadalajara and stays in a hotel there.
    6. The city of Toronto purchases computer paper from a local firm.
6. Suggest an argument for and an argument against counting in GDP all household-produced goods and services that are not sold, such as the value of child care or home-cooked meals.
7. Suppose a nation's firms make heavy use of factors of production owned by residents of foreign countries, while foreign firms make relatively little use of factors owned by residents of that nation. How does the nation's GDP compare to its GNP?
8. Suppose Country A has the same GDP as Country B, and that neither nation's residents own factors of production used by foreign firms, nor do either nation's firms use factors of production owned by foreign residents. Suppose that, relative to Country B, depreciation, indirect business taxes, and personal income taxes in Country A are high, while welfare and Social Security payments to households in Country A are relatively low. Which country has the higher disposable personal income? Why?
9. Suppose that virtually everyone in Canada decides to take life a little easier, and the length of the average workweek falls by 25%. How will that affect GDP? Per capita GDP? How will it affect economic welfare?
10. Comment on the following statement: "It does not matter that the value of the labour people devote to producing things for themselves is not counted in GDP; because we make the same 'mistake' every year, relative values are unaffected."
11. Name some of the services, if any, you produced at home that do get counted in GDP. Are there any goods you produce that are not counted?

Numerical Problems

1. Given the following nominal data, compute GDP. Assume net factor incomes from abroad = 0 (that is, GDP = GNP).

Nominal Data for GDP and NNP	$ Billions
Consumption	2,799.8
Depreciation	481.6
Exports	376.2
Gross private domestic investment	671.0
Indirect taxes	331.4
Government purchases	869.7
Government transfer payments	947.8
Imports	481.7

2. Find data for each of the following countries on real GDP and population. Use the data to calculate the GDP per capita for each of the following countries:
    1. Mozambique
    2. India
    3. Pakistan
    4. United States
    5. Canada
    6. Russia
    7. Brazil
    8. Iran
    9. Colombia
3. Now construct a bar graph showing your results in the previous problem, organizing the countries from the highest to the lowest GNP per capita, with countries on the horizontal axis and GNP per capita on the vertical axis.
4. Suppose Country A has a GDP of $4 trillion. Residents of this country earn $500 million from assets they own in foreign countries. Residents of foreign countries earn $300 million from assets they own in Country A. Compute:
    1. Country A's net foreign income.
    2. Country A's GNP.
5. Suppose a country's GDP equals $500 billion for a particular year. Economists in the country estimate that household production equals 40% of GDP.
    1. What is the value of the country's household production for that year?
    2. Counting both GDP and household production, what is the country's total output for the year?
6. A miner extracts iron from the earth. A steel mill converts the iron to steel beams for use in construction. A construction company uses the steel beams to make a building. Assume that the total product of these firms represents the only components of the building and that they will have no other uses. Complete the following table:

Company	Product	Total Sales	Value Added
Acme Mining	iron ore	$100,000	?
Fuller Mill	steel beams	$175,000	?
Crane Construction	building	$1,100,000	?
Total Value Added			?

7. You are given the data below for 2008 for the imaginary country of Amagre, whose currency is the G.

Consumption	350 billion G
Transfer payments	100 billion G
Investment	100 billion G
Government purchases	200 billion G
Exports	50 billion G
Imports	150 billion G
Bond purchases	200 billion G
Earnings on foreign investments	75 billion G
Foreign earnings on Amagre investment	25 billion G

   1. Compute net foreign investment.
   2. Compute net exports.
   3. Compute GDP.
   4. Compute GNP.

CHAPTER 7: AGGREGATE DEMAND AND AGGREGATE SUPPLY

Start Up: The Great Warning

The first warning came from the Harvard Economic Society, an association of Harvard economics professors, early in 1929. The society predicted in its weekly newsletter that the seven-year-old expansion was coming to an end. Recession was ahead. Almost no one took the warning seriously. The economy, fueled by soaring investment, had experienced stunning growth. The 1920s had seen the emergence of many entirely new industries—automobiles, public power, home appliances, synthetic fabrics, radio, and motion pictures. The decade seemed to have acquired a momentum all its own. Prosperity was not about to end, no matter what a few economists might say.

Summer came, and no recession was apparent. The Harvard economists withdrew their forecast. As it turned out, they lost their nerve too soon. Indeed, industrial production had already begun to fall. The worst downturn in our history, the Great Depression, had begun.

The collapse was swift. The stock market crashed in October 1929. Real GDP plunged nearly 10% by 1930. By the time the economy hit bottom in 1933, real GDP had fallen 30%, unemployment had increased from 3.2% in 1929 to 25% in 1933, and prices, measured by the implicit price deflator, had plunged 23% from their 1929 level. The depression held the economy in its cruel grip for more than a decade; it was not until World War II that full employment was restored.

In this chapter we go beyond explanations of the main macroeconomic variables to introduce a model of macroeconomic activity that we can use to analyze problems such as fluctuations in gross domestic product (real GDP), the price level, and employment: the model of aggregate demand and aggregate supply. We will use this model throughout our exploration of macroeconomics. In this chapter we will present the broad outlines of the model; greater detail, more examples, and more thorough explanations will follow in subsequent chapters.

We will examine the concepts of the aggregate demand curve and the short- and long-run aggregate supply curves. We will identify conditions under which an economy achieves an equilibrium level of real GDP that is consistent with full employment of labor. Potential output is the level of output an economy can achieve when labor is employed at its natural level. **Potential output** is also called the natural level of real GDP. When an economy fails to produce at its potential, there may be actions that the government or the central bank can take to push the economy toward it, and in this chapter we will begin to consider the pros and cons of doing so.

7.1 Aggregate Demand

Learning Objectives

1. Define potential output, also called full employment GDP.
2. Define aggregate demand, represent it using a hypothetical aggregate demand curve, and identify and explain the three effects that cause this curve to slope downward.
3. Distinguish between a change in the aggregate quantity of goods and services demanded and a change in aggregate demand.
4. Use examples to explain how each component of aggregate demand can be a possible aggregate demand shifter.
5. Explain what a multiplier is and tell how to calculate it.

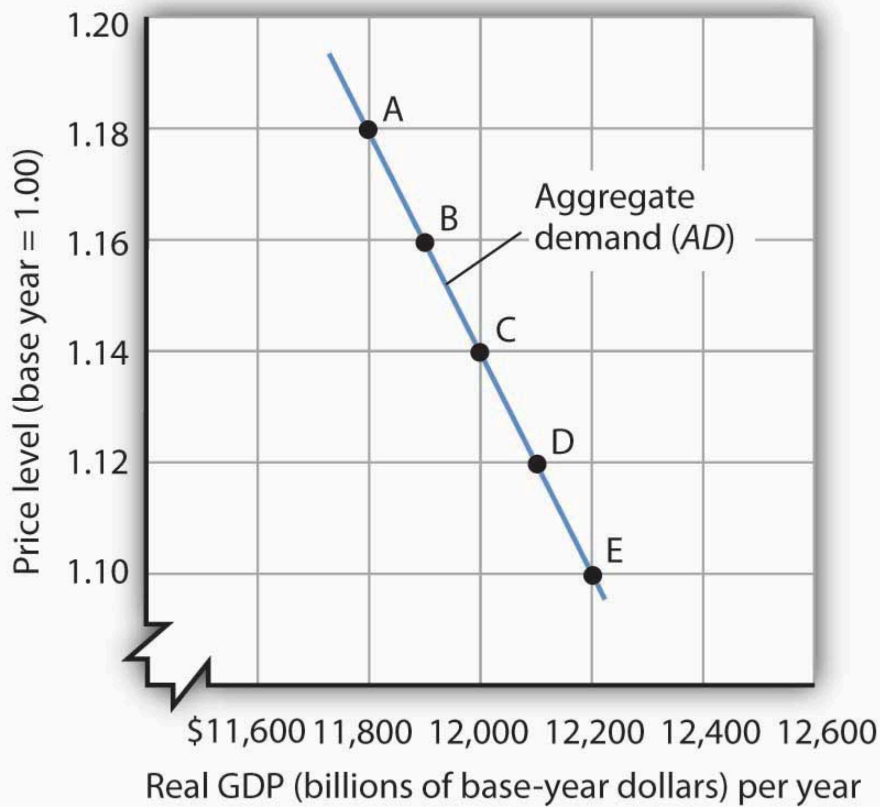
Firms face four sources of demand: households (personal consumption), other firms (investment), government agencies (government purchases), and foreign markets (net exports). Aggregate demand is the relationship between the total quantity of goods and services demanded (from all the four sources of demand) and the price level, all other determinants of spending unchanged. The aggregate demand curve is a graphical representation of aggregate demand.

The Slope of the Aggregate Demand Curve

We will use the implicit price deflator as our measure of the price level; the aggregate quantity of goods and services demanded is measured as real GDP. We can also use the CPI as the price level on the vertical axis. The table in Figure 7.1 "Aggregate Demand" gives values for each component of aggregate demand at each price level for a hypothetical economy. Various points on the aggregate demand curve are found by adding the values of these components at different price levels. The aggregate demand curve for the data given in the table is plotted on the graph in Figure 7.1 "Aggregate Demand". At point A, at a price level of 1.18, $11,800 billion worth of goods and services will be demanded; at point C, a reduction in the price level to 1.14 increases the quantity of goods and services demanded to $12,000 billion; and at point E, at a price level of 1.10, $12,200 billion will be demanded.

Figure 7.1 Aggregate Demand

Point on aggregate demand curve	Price level	$C+$	$I+$	$G+$	$X_n=$	Aggregate demand
A	1.18	8,400	1,820	2,150	−570	11,800
B	1.16	8,450	1,860	2,150	−560	11,900
C	1.14	8,500	1,900	2,150	−550	12,000
D	1.12	8,550	1,940	2,150	−540	12,100
E	1.10	8,600	1,980	2,150	−530	12,200



An aggregate demand curve (AD) shows the relationship between the total quantity of output demanded (measured as real GDP) and the price level (measured as the implicit price deflator). At each price level, the total quantity of goods and services demanded is the sum of the components of real GDP, as shown in the table. There is a negative relationship between the price level and the total quantity of goods and services demanded, all other things unchanged.

The negative slope of the aggregate demand curve suggests that it behaves in the same manner as an ordinary demand curve. But we *cannot* apply the reasoning we use to explain downward-sloping demand curves in individual markets to explain the downward-sloping aggregate demand curve. There are two reasons for a negative relationship between price and quantity demanded in individual markets. First, a lower price induces people to substitute more of the good whose price has fallen for other goods, increasing the quantity demanded. Second, the lower price creates a higher real income. This normally increases quantity demanded further.

Neither of these effects is relevant to a change in prices in the aggregate. When we are dealing with the average of all prices—the price level—we can no longer say that a fall in prices will induce a change in relative prices that will lead consumers to buy more of the goods and services whose prices have fallen and less of the goods and services whose prices have not fallen. The price of corn may have fallen, but the prices of wheat, sugar, tractors, steel, and most other goods or services produced in the economy are likely to have fallen as well.

Furthermore, a reduction in the price level means that it is not just the prices consumers pay that are falling. It means the prices people receive—their wages, the rents they may charge as landlords, the interest rates they earn—are likely to be falling as well. A falling price level means that goods and services are cheaper, but incomes are lower, too. There is no reason to expect that a change in real income will boost the quantity of goods and services demanded—indeed, no change in real income would occur. If nominal incomes and prices all fall by 10%, for example, real incomes do not change.

Why, then, does the aggregate demand curve slope downward? One reason for the downward slope of the aggregate demand curve lies in the relationship between real wealth (the stocks, bonds, and other assets that people have accumulated) and consumption (one of the four components of aggregate demand). When the price level falls, the real value of wealth increases—it packs more purchasing power. For example, if the price level falls by 25%, then $10,000 of wealth could purchase more goods and services than it would have if the price level had not fallen. An increase in wealth will induce people to increase their consumption. The consumption component of aggregate demand will thus be greater at lower price levels than at higher price levels. The tendency for a change in the price level to affect real wealth and thus alter consumption is called the wealth effect; it suggests a negative relationship between the price level and the real value of consumption spending.

A second reason the aggregate demand curve slopes downward lies in the relationship between interest rates and investment. A lower price level lowers the demand for money, because less money is required to buy a given quantity of goods. What economists mean by money demand will be explained in more detail in a later chapter. But, as we learned in studying demand and supply, a reduction in the demand for something, all other things unchanged, lowers its price. In this case, the "something" is money and its price is the interest rate. A lower price level thus reduces interest rates. Lower interest rates make borrowing by firms to build factories or buy equipment and other capital more attractive. A lower interest rate means lower mortgage payments, which tends to increase investment in residential houses. Investment thus rises when the price level falls. The tendency for a change in the price level to affect the interest rate and thus to affect the quantity of investment demanded is called the interest rate effect. John Maynard Keynes, a British economist whose analysis of the Great Depression and what to do about it led to the birth of modern macroeconomics, emphasized this effect. For this reason, the interest rate effect is sometimes called the Keynes effect.

A third reason for the rise in the total quantity of goods and services demanded as the price level falls can be found in changes in the net export component of aggregate demand. All other things unchanged, a lower price level in an economy reduces the prices of its goods and services relative to foreign-produced goods and services. A lower price level makes that economy's goods more attractive to foreign buyers, increasing exports. It will also make foreign-produced goods and services less attractive to the economy's buyers, reducing imports. The result is an increase in net exports. The international trade effect is the tendency for a change in the price level to affect net exports.

Taken together, then, a fall in the price level means that the quantities of consumption, investment, and net export components of aggregate demand may all rise. Since government purchases are determined through a political process, we assume there is no causal link between the price level and the real volume of government purchases. Therefore, this component of GDP does not contribute to the downward slope of the curve.

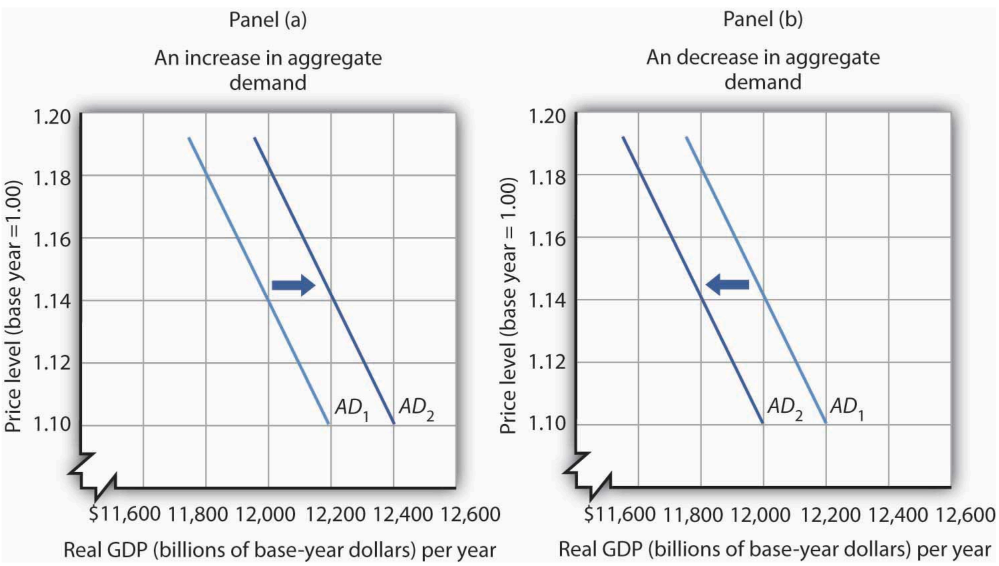
In general, a change in the price level, with all other determinants of aggregate demand unchanged, causes a movement along the aggregate demand curve. A movement along an aggregate demand curve is a change in the aggregate quantity of goods and services demanded. A movement from point A to point B on the aggregate demand curve in Figure 7.1 "Aggregate Demand" is an example. Such a change is a response to a change in the price level.

Notice that the axes of the aggregate demand curve graph are drawn with a break near the origin to remind us that the plotted values reflect a relatively narrow range of changes in real GDP and the price level. We do not know what might happen if the price level or output for an entire economy approached zero. Such a phenomenon has never been observed.

Changes in Aggregate Demand

Aggregate demand changes in response to a change in any of its components. An increase in the total quantity of consumer goods and services demanded at every price level, for example, would shift the aggregate demand curve to the right. A change in the aggregate quantity of goods and services demanded at every price level is a change in aggregate demand, which shifts the aggregate demand curve. Increases and decreases in aggregate demand are shown in Figure 7.2 "Changes in Aggregate Demand".

Figure 7.2 Changes in Aggregate Demand



An increase in consumption, investment, government purchases, or net exports shifts the aggregate demand curve AD_1 to the right as shown in Panel (a). A reduction in one of the components of aggregate demand shifts the curve to the left, as shown in Panel (b).

What factors might cause the aggregate demand curve to shift? Each of the components of aggregate demand is a possible aggregate demand shifter. We shall look at some of the events that can trigger changes in the components of aggregate demand and thus shift the aggregate demand curve.

Changes in Consumption

Several events could change the quantity of consumption at each price level and thus shift aggregate demand. One determinant of consumption is consumer confidence. If consumers expect good economic conditions and are optimistic about their economic prospects, they are more likely to buy major items such as cars or furniture. The result would be an increase in the real value of consumption at each price level and an increase in aggregate demand. In the second half of the 1990s, sustained economic growth and low unemployment fueled high expectations and consumer optimism. Surveys revealed consumer confidence to be very high. That consumer confidence translated into increased consumption and increased aggregate demand. In contrast, a decrease in consumption would accompany diminished consumer expectations and a decrease in consumer confidence, as happened after the stock market crash of 1929. The same problem has plagued the economies of most Western nations in 2008 as declining consumer confidence has tended to reduce consumption. A survey by the Conference Board in September of 2008 showed that just 13.5% of consumers surveyed expected economic conditions in North America to improve in the next six months. Similarly pessimistic views prevailed in the previous two months. That contributed to the decline in consumption that occurred in the third quarter of the year.

Another factor that can change consumption and shift aggregate demand is tax policy. A cut in personal income taxes leaves people with more after-tax income, which may induce them to increase their consumption.

Transfer payments such as welfare and social assistance also affect the income people have available to spend. At any given price level, an increase in transfer payments raises consumption and aggregate demand, and a reduction lowers consumption and aggregate demand.

Changes in Investment

Investment is the production of new capital that will be used for future production of goods and services. Firms make investment choices based on what they think they will be producing in the future. The expectations of firms thus play a critical role in determining investment. If firms expect their sales to go up, they are likely to increase their investment so that they can increase production and meet consumer demand. Such an increase in investment raises the aggregate quantity of goods and services demanded at each price level; it increases aggregate demand.

Changes in interest rates also affect investment and thus affect aggregate demand. We must be careful to distinguish such changes from the interest rate effect, which causes a movement along the aggregate demand curve. A change in interest rates that results from a change in the price level affects investment in a way that is already captured in the downward slope of the aggregate demand curve; it causes a movement along the curve. A change in interest rates for some other reason shifts the curve. We examine reasons interest rates might change in another chapter.

Investment can also be affected by tax policy. For example, there could be a reduction in the tax rate on certain capital gains. Capital gains result when the owner of an asset, such as a house or a factory, sells the asset for more than its purchase price (less any depreciation claimed in earlier years). The lower capital gains tax could stimulate investment, because the owners of such assets know that they will lose less to taxes when they sell those assets, thus making assets subject to the tax more attractive.

Changes in Government Purchases

Any change in government purchases, all other things unchanged, will affect aggregate demand. An increase in government purchases increases aggregate demand; a decrease in government purchases decreases aggregate demand.

Some good examples of government spending exist in the US. Many economists argued that reductions in defense spending in the wake of the collapse of the Soviet Union in 1991 tended to reduce aggregate demand. Similarly, increased defense spending for the wars in Afghanistan and Iraq increased aggregate demand. Dramatic increases in defense spending to fight World War II accounted in large part for the rapid recovery from the Great Depression. Canada experienced similar changes, although to a lesser extent.

Changes in Net Exports

A change in the value of net exports at each price level shifts the aggregate demand curve. A major determinant of net exports is foreign demand for a country's goods and services; that demand will vary with foreign incomes. An increase in foreign incomes increases a country's net exports and aggregate demand; a slump in foreign incomes reduces net exports and aggregate demand. For example, several major U.S. trading partners in Asia suffered recessions in 1997 and 1998. Lower real incomes in those countries reduced U.S. exports and tended to reduce aggregate demand.

Exchange rates also influence net exports, all other things unchanged. A country's exchange rate is the price of its currency in terms of another currency or currencies. A rise in the U.S. exchange rate means that it takes more Japanese yen, for example, to purchase one dollar. That also means that U.S. traders get more yen per dollar. Since prices of goods produced in Japan are given in yen and prices of goods produced in the United States are given in dollars, a rise in the U.S. exchange rate increases the price to foreigners for goods and services produced in the United States, thus reducing U.S. exports; it reduces the price of foreign-produced goods and services for U.S. consumers, thus increasing imports to the United States. A higher exchange rate tends to reduce net exports, reducing aggregate demand. A lower exchange rate tends to increase net exports, increasing aggregate demand.

Foreign price levels can affect aggregate demand in the same way as exchange rates. For example, when foreign price levels fall relative to the price level in the United States, U.S. goods and services become relatively more expensive, reducing exports and boosting imports in the United States. Such a reduction in net exports reduces aggregate demand. An increase in foreign prices relative to U.S. prices has the opposite effect.

The trade policies of various countries can also affect net exports. A policy by Japan to increase its imports of goods and services from India, for example, would increase net exports in India.

The Multiplier

A change in any component of aggregate demand shifts the aggregate demand curve. Generally, the aggregate demand curve shifts by more than the amount by which the component initially causing it to shift changes.

Suppose that net exports increase due to an increase in foreign incomes. As foreign demand for domestically made products rises, a country's firms will hire additional workers or perhaps increase the average number of hours that their employees work. In either case, incomes will rise, and higher incomes will lead to an increase in consumption. Taking into account these other increases in the components of aggregate demand, the aggregate demand curve will shift by more than the initial shift caused by the initial increase in net exports.

The multiplierThe ratio of the change in the quantity of real GDP demanded at each price level to the initial change in one or more components of aggregate demand that produced it. is the ratio of the change in the quantity of real GDP demanded at each price level to the initial change in one or more components of aggregate demand that produced it:

Equation 22.1

$$Multiplier = \frac{\Delta(real\ GDP\ demanded\ at\ each\ price\ level)}{initial\Delta(component\ of\ AD)}$$

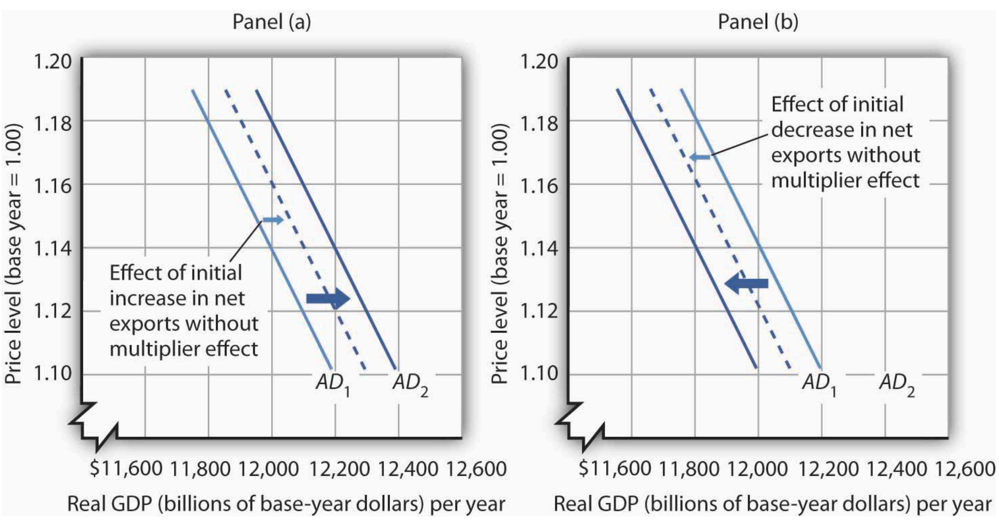
We use the capital Greek letter delta (Δ) to mean "change in." In the aggregate demand–aggregate supply model presented in this chapter, it is the number by which we multiply an initial change in aggregate demand to obtain the amount by which the aggregate demand curve shifts as a result of the initial change. In other words, we can use Equation 22.1 to solve for the change in real GDP demanded at each price level:

Equation 22.2

$$\Delta(real\ GDP\ demanded\ at\ each\ price\ level) = multiplier \times initial\ \Delta(component\ of\ AD)$$

Suppose that the initial increase in net exports is $100 billion and that the initial $100-billion increase generates additional consumption of $100 billion at each price level. In Panel (a) of Figure 22.3 "The Multiplier", the aggregate demand curve shifts to the right by $200 billion—the amount of the initial increase in net exports times the multiplier of 2. We obtained the value for the multiplier in this example by plugging $200 billion (the initial $100-billion increase in net exports plus the $100-billion increase that it generated in consumption) into the numerator of Equation 22.1 and $100 billion into the denominator. Similarly, a decrease in net exports of $100 billion leads to a decrease in aggregate demand of $200 billion at each price level, as shown in Panel (b).

Figure 22.3 The Multiplier



A change in one component of aggregate demand shifts the aggregate demand curve by more than the initial change. In Panel (a), an initial increase of $100 billion of net exports shifts the aggregate demand curve to the right by $200 billion at each price level. In Panel (b), a decrease of net exports of $100 billion shifts the aggregate demand curve to the left by $200 billion. In this example, the multiplier is 2.

Key Takeaways

- Potential output is the level of output an economy can achieve when labor is employed at its natural level. When an economy fails to produce at its potential, the government or the central bank may try to push the economy toward its potential.
- The aggregate demand curve represents the total of consumption, investment, government purchases, and net exports at each price level in any period. It slopes downward because of the wealth effect on consumption, the interest rate effect on investment, and the international trade effect on net exports.
- The aggregate demand curve shifts when the quantity of real GDP demanded at each price level changes.
- The multiplier is the number by which we multiply an initial change in aggregate demand to obtain the amount by which the aggregate demand curve shifts at each price level as a result of the initial change.

Try It!

Explain the effect of each of the following on the aggregate demand curve for the United States:

1. A decrease in consumer optimism
2. An increase in real GDP in the countries that buy U.S. exports
3. An increase in the price level
4. An increase in government spending on highways

Case in Point: The Multiplied Economic Impact of SARS on China's Economy

Figure 22.4

Zervas – wrong face mask – CC BY-NC-ND 2.0.

Severe Acute Respiratory Syndrome (SARS), an atypical pneumonia-like disease, broke onto the world scene in late 2002. In March 2003, the World Health Organization (WHO) issued its first worldwide alert and a month later its first travel advisory, which recommended that travelers avoid Hong Kong and the southern province of China, Guangdong. Over the next few months, additional travel advisories were issued for other parts of China, Taiwan, and briefly for Toronto, Canada. By the end of June, all WHO travel advisories had been removed.

To estimate the overall impact of SARS on the Chinese economy in 2003, economists Wen Hai, Zhong Zhao, and Jian Want of Peking University's China Center for Economic Research conducted a survey of Beijing's tourism industry in April 2003. Based on findings from the Beijing area, they projected the tourism sector of China as a whole would lose $16.8 billion—of which $10.8 billion came from an approximate 50% reduction in foreign tourist revenue and $6 billion from curtailed domestic tourism, as holiday celebrations were cancelled and domestic travel restrictions imposed.

To figure out the total impact of SARS on China's economy, they argued that the multiplier for tourism revenue in China is between 2 and 3. Since the SARS outbreak only began to have a major economic impact after March, they assumed a smaller multiplier of 1.5 for all of 2003. They thus predicted that the Chinese economy would be $25.3 billion smaller in 2003 as a result of SARS:

Source: Wen Hai, Zhong Zhao, and Jian Wan, "The Short-Term Impact of SARS on the Chinese Economy," *Asian Economic Papers* 3, no. 1 (Winter 2004): 57–61.

Answer to Try It! Problem

1. A decline in consumer optimism would cause the aggregate demand curve to shift to the left. If consumers are more pessimistic about the future, they are likely to cut purchases, especially of major items.
2. An increase in the real GDP of other countries would increase the demand for U.S. exports and cause the aggregate demand curve to shift to the right. Higher incomes in other countries will make consumers in those countries more willing and able to buy U.S. goods.
3. An increase in the price level corresponds to a movement up along the unchanged aggregate demand curve. At the higher price level, the consumption, investment, and net export components of aggregate demand will all fall; that is, there will be a reduction in the total quantity of goods and services demanded, but not a shift of the aggregate demand curve itself.
4. An increase in government spending on highways means an increase in government purchases. The aggregate demand curve would shift to the right.

7.2 Aggregate Demand and Aggregate Supply: The Long Run and the Short Run

Learning Objectives

1. Distinguish between the short run and the long run, as these terms are used in macroeconomics.
2. Draw a hypothetical long-run aggregate supply curve and explain what it shows about the natural levels of employment and output at various price levels, given changes in aggregate demand.
3. Draw a hypothetical short-run aggregate supply curve, explain why it slopes upward, and explain why it may shift; that is, distinguish between a change in the aggregate quantity of goods and services supplied and a change in short-run aggregate supply.
4. Discuss various explanations for wage and price stickiness.
5. Explain and illustrate what is meant by equilibrium in the short run and relate the equilibrium to potential output.

In macroeconomics, we seek to understand two types of equilibria, one corresponding to the short run and the other corresponding to the long run. The short run in macroeconomic analysis is a period in which wages and some other prices do not respond to changes in economic conditions. In certain markets, as economic conditions change, prices (including wages) may not adjust quickly enough to maintain equilibrium in these markets. A sticky price is a price that is slow to adjust to its equilibrium level, creating sustained periods of shortage or surplus. Wage and price stickiness prevent the economy from achieving its natural level of employment and its potential output. In contrast, the long run in macroeconomic analysis is a period in which wages and prices are flexible. In the long run, employment will move to its natural level and real GDP to potential.

We begin with a discussion of long-run macroeconomic equilibrium, because this type of equilibrium allows us to see the macroeconomy after full market adjustment has been achieved. In contrast, in the short run, price or wage stickiness is an obstacle to full adjustment. Why these deviations from the potential level of output occur and what the implications are for the macroeconomy will be discussed in the section on short-run macroeconomic equilibrium.

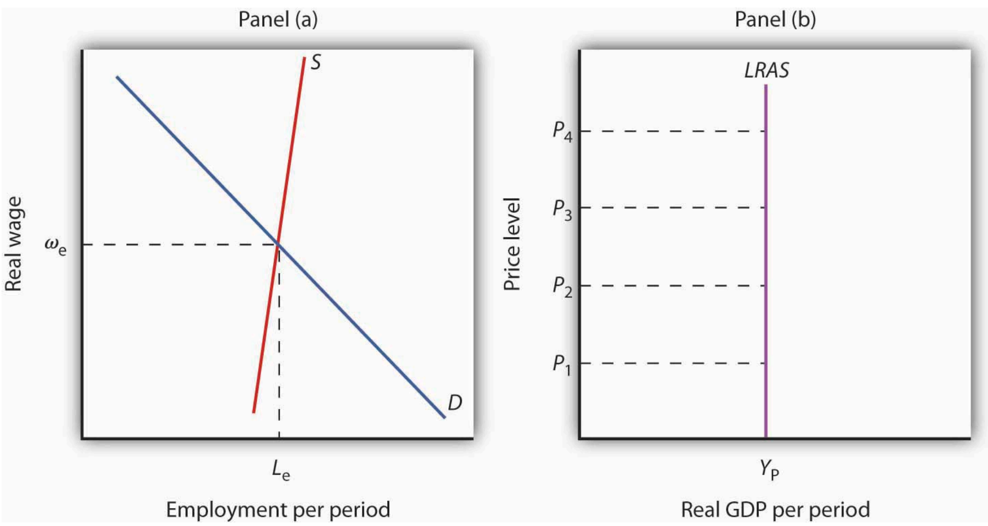
The Long Run

As explained in a previous chapter, the natural level of employment occurs where the real wage adjusts so that the quantity of labor demanded equals the quantity of labor supplied. When the economy achieves its natural level of employment, it achieves its potential level of output. We will see that real GDP eventually moves to potential, because all wages and prices are assumed to be flexible in the long run.

Long-Run Aggregate Supply

The long-run aggregate supply (LRAS) curve relates the level of output produced by firms to the price level in the long run. In Panel (b) of Figure 22.5 "Natural Employment and Long-Run Aggregate Supply", the long-run aggregate supply curve is a vertical line at the economy's potential level of output. There is a single real wage at which employment reaches its natural level. In Panel (a) of Figure 22.5 "Natural Employment and Long-Run Aggregate Supply", only a real wage of ω_e generates natural employment L_e. The economy could, however, achieve this real wage with any of an infinitely large set of nominal wage and price-level combinations. Suppose, for example, that the equilibrium real wage (the ratio of wages to the price level) is 1.5. We could have that with a nominal wage level of 1.5 and a price level of 1.0, a nominal wage level of 1.65 and a price level of 1.1, a nominal wage level of 3.0 and a price level of 2.0, and so on.

Figure 22.5 Natural Employment and Long-Run Aggregate Supply



When the economy achieves its natural level of employment, as shown in Panel (a) at the intersection of the demand and supply curves for labor, it achieves its potential output, as shown in Panel (b) by the vertical long-run aggregate supply curve *LRAS* at Y_P.

In Panel (b) we see price levels ranging from P_1 to P_4. Higher price levels would require higher nominal wages to create a real wage of ω_e, and flexible nominal wages would achieve that in the long run.

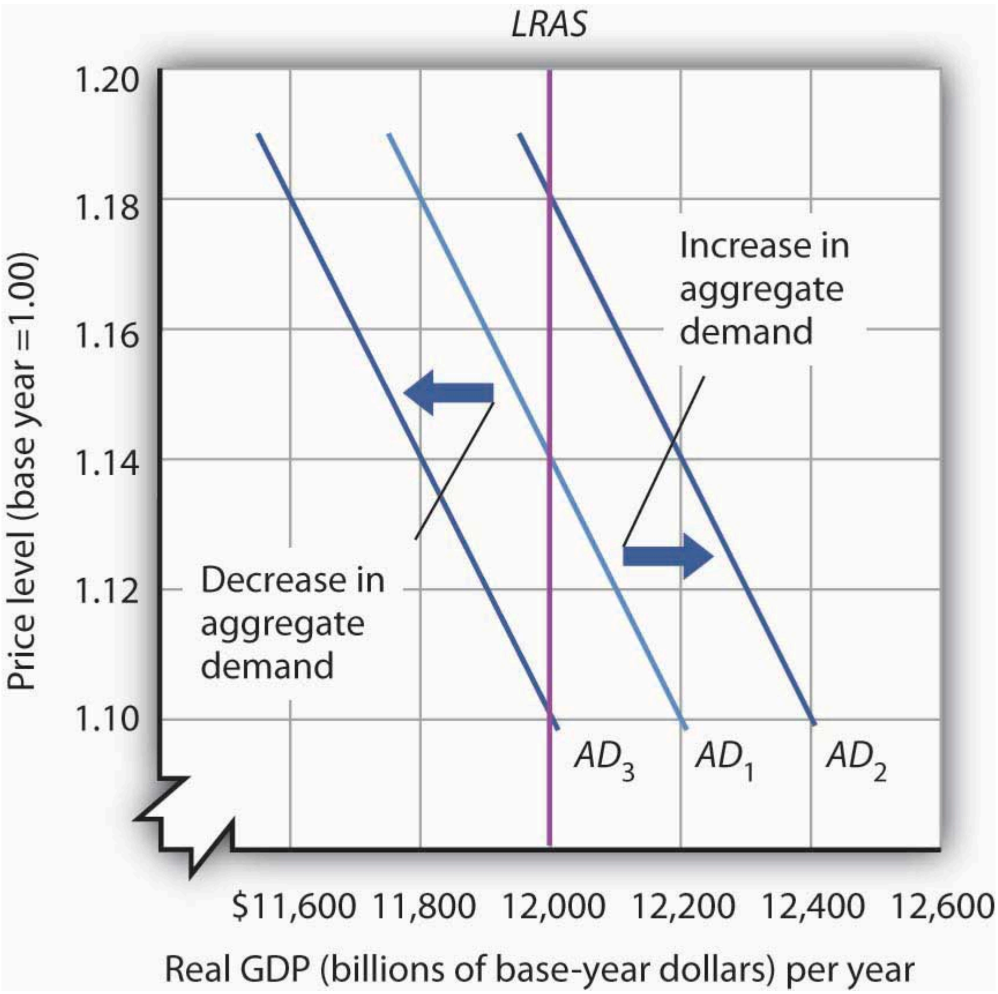
In the long run, then, the economy can achieve its natural level of employment and potential output at any price level. This conclusion gives us our long-run aggregate supply curve. With only one level of output at any price level, the long-run aggregate supply curve is a vertical line at the economy's potential level of output of Y_P.

Equilibrium Levels of Price and Output in the Long Run

The intersection of the economy's aggregate demand curve and the long-run aggregate supply curve determines its equilibrium real GDP and price level in the long run. Figure 22.6 "Long-Run Equilibrium" depicts an economy in long-run equilibrium. With aggregate demand at AD_1 and the long-run aggregate supply curve as shown, real GDP is

$12,000 billion per year and the price level is 1.14. If aggregate demand increases to AD_2, long-run equilibrium will be reestablished at real GDP of $12,000 billion per year, but at a higher price level of 1.18. If aggregate demand decreases to AD_3, long-run equilibrium will still be at real GDP of $12,000 billion per year, but with the now lower price level of 1.10.

Figure 22.6 Long-Run Equilibrium



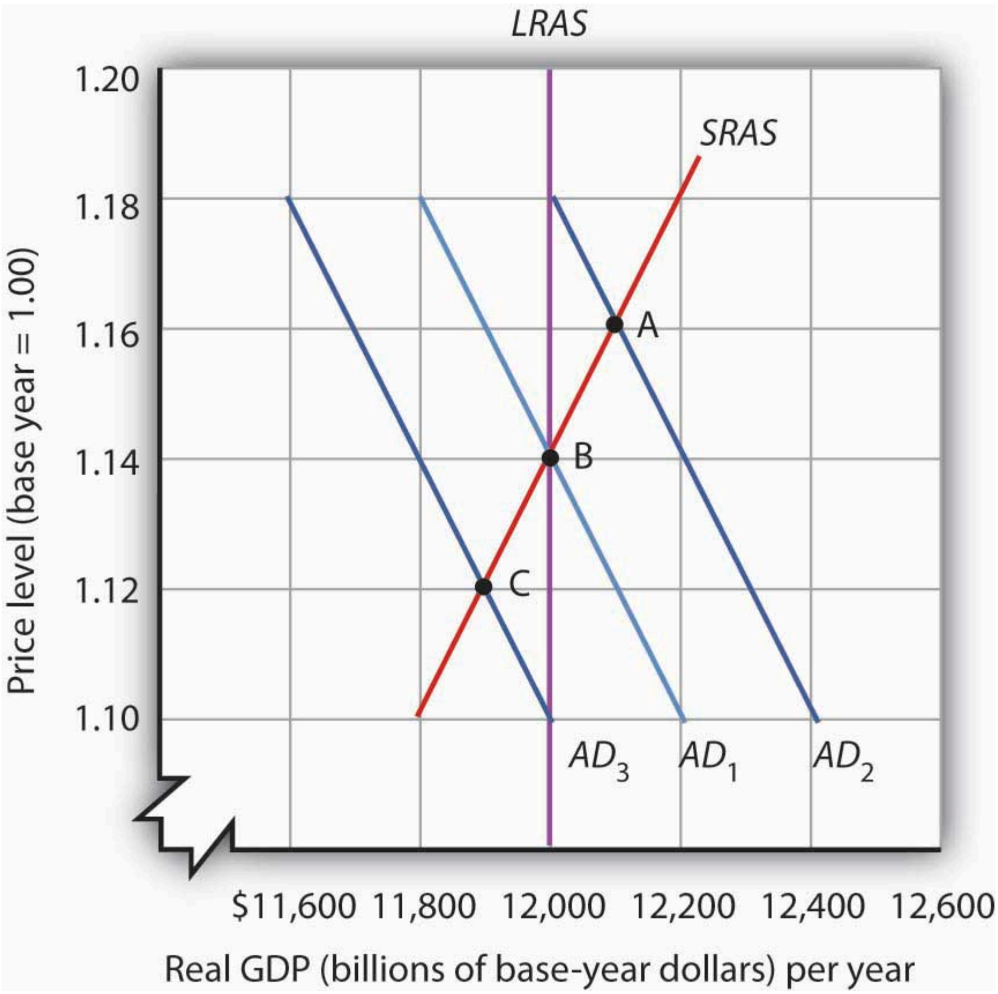
Long-run equilibrium occurs at the intersection of the aggregate demand curve and the long-run aggregate supply curve. For the three aggregate demand curves shown, long-run equilibrium occurs at three different price levels, but always at an output level of $12,000 billion per year, which corresponds to potential output.

The Short Run

Analysis of the macroeconomy in the short run—a period in which stickiness of wages and prices may prevent the economy from operating at potential output—helps explain how deviations of real GDP from potential output can and do occur. We will explore the effects of changes in aggregate demand and in short-run aggregate supply in this section.

Short-Run Aggregate Supply

Figure 22.7 Deriving the Short-Run Aggregate Supply Curve



The economy shown here is in long-run equilibrium at the intersection of AD_1 with the long-run aggregate supply curve. If aggregate demand increases to AD_2, in the short run, both real GDP and the price level rise. If aggregate demand decreases to AD_3, in the short run, both real GDP and the price level fall. A line drawn through points A, B, and C traces out the short-run aggregate supply curve SRAS.

The model of aggregate demand and long-run aggregate supply predicts that the economy will eventually move toward its potential output. To see how nominal wage and price stickiness can cause real GDP to be either above or below potential in the short run, consider the response of the economy to a change in aggregate demand. Figure 22.7 "Deriving the Short-Run Aggregate Supply Curve" shows an economy that has been operating at potential output of $12,000 billion and a price level of 1.14. This occurs at the intersection of AD_1 with the long-run aggregate supply curve at point B. Now suppose that the aggregate demand curve shifts to the right (to AD_2). This could occur as a result of an increase in exports. (The shift from AD_1 to AD_2 includes the multiplied effect of the increase in exports.) At the price level of 1.14, there is now excess demand and pressure on prices to rise. If all prices in the economy adjusted quickly, the economy would quickly settle at potential output of $12,000 billion, but at a higher price level (1.18 in this case).

Is it possible to expand output above potential? Yes. It may be the case, for example, that some people who were in the labor force but were frictionally or structurally unemployed find work because of the ease of getting jobs at the going nominal wage in such an environment. The result is an economy operating at point A in Figure 22.7 "Deriving the Short-Run Aggregate Supply Curve" at a higher price level and with output temporarily above potential.

Consider next the effect of a reduction in aggregate demand (to AD_3), possibly due to a reduction in investment. As the price level starts to fall, output also falls. The economy finds itself at a price level–output combination at which real GDP is below potential, at point C. Again, price stickiness is to blame. The prices firms receive are falling with the reduction in demand. Without corresponding reductions in nominal wages, there will be an increase in the real wage. Firms will employ less labor and produce less output.

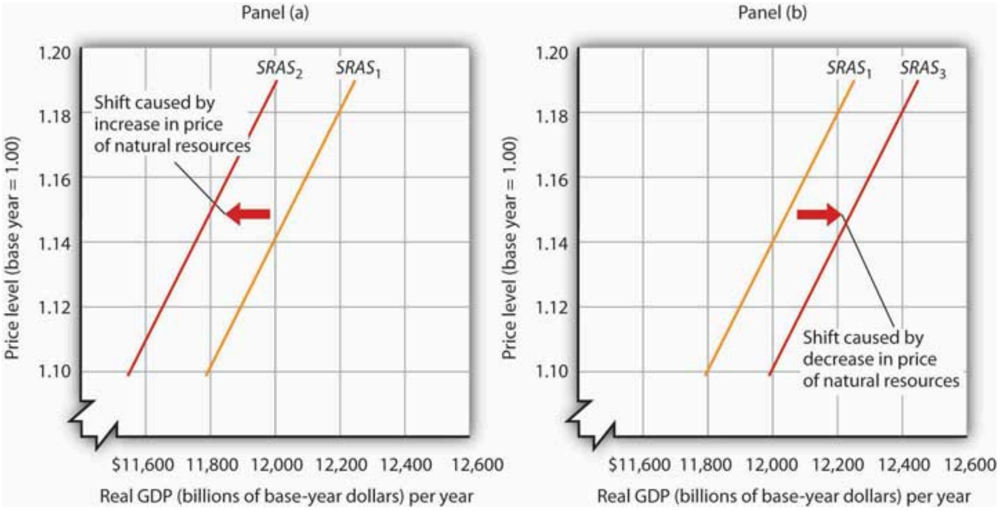
By examining what happens as aggregate demand shifts over a period when price adjustment is incomplete, we can trace out the short-run aggregate supply curve by drawing a line through points A, B, and C. The short-run aggregate supply (SRAS) curve is a graphical representation of the relationship between production and the price level in the short run. Among the factors held constant in drawing a short-run aggregate supply curve are the capital stock, the stock of natural resources, the level of technology, and the prices of factors of production.

A change in the price level produces a change in the aggregate quantity of goods and services supplied and is illustrated by the movement along the short-run aggregate supply curve. This occurs between points A, B, and C in Figure 22.7 "Deriving the Short-Run Aggregate Supply Curve".

A change in the quantity of goods and services supplied at every price level in the short run is a change in short-run aggregate supply. Changes in the factors held constant in drawing the short-run aggregate supply curve shift the curve. (These factors may also shift the long-run aggregate supply curve; we will discuss them along with other determinants of long-run aggregate supply in the next chapter.)

One type of event that would shift the short-run aggregate supply curve is an increase in the price of a natural resource such as oil. An increase in the price of natural resources or any other factor of production, all other things unchanged, raises the cost of production and leads to a reduction in short-run aggregate supply. In Panel (a) of Figure 22.8 "Changes in Short-Run Aggregate Supply", $SRAS_1$ shifts leftward to $SRAS_2$. A decrease in the price of a natural resource would lower the cost of production and, other things unchanged, would allow greater production from the economy's stock of resources and would shift the short-run aggregate supply curve to the right; such a shift is shown in Panel (b) by a shift from $SRAS_1$ to $SRAS_3$.

Figure 22.8 Changes in Short-Run Aggregate Supply



A reduction in short-run aggregate supply shifts the curve from $SRAS_1$ to $SRAS_2$ in Panel (a). An increase shifts it to the right to $SRAS_3$, as shown in Panel (b).

Reasons for Wage and Price Stickiness

Wage or price stickiness means that the economy may not always be operating at potential. Rather, the economy may operate either above or below potential output in the short run. Correspondingly, the overall unemployment rate will be below or above the natural level.

Many prices observed throughout the economy do adjust quickly to changes in market conditions so that equilibrium, once lost, is quickly regained. Prices for fresh food and shares of common stock are two such examples.

Other prices, though, adjust more slowly. Nominal wages, the price of labor, adjust very slowly. We will first look at why nominal wages are sticky, due to their association with the unemployment rate, a variable of great interest in macroeconomics, and then at other prices that may be sticky.

Wage Stickiness

Wage contracts fix nominal wages for the life of the contract. The length of wage contracts varies from one week or one month for temporary employees, to one year (teachers and professors often have such contracts), to three years (for most union workers employed under major collective bargaining agreements). The existence of such explicit contracts means that both workers and firms accept some wage at the time of negotiating, even though economic conditions could change while the agreement is still in force.

Think about your own job or a job you once had. Chances are you go to work each day knowing what your wage will be. Your wage does not fluctuate from one day to the next with changes in demand or supply. You may have a formal contract with your employer that specifies what your wage will be over some period. Or you may have an informal understanding that sets your wage. Whatever the nature of your agreement, your wage is “stuck” over the period of the agreement. Your wage is an example of a sticky price.

One reason workers and firms may be willing to accept long-term nominal wage contracts is that negotiating a contract is a costly process. Both parties must keep themselves adequately informed about market conditions. Where unions are involved, wage negotiations raise the possibility of a labor strike, an eventuality that firms may prepare for by accumulating additional inventories, also a costly process. Even when unions are not involved, time and energy spent discussing wages takes away from time and energy spent producing goods and services. In addition, workers may simply prefer knowing that their nominal wage will be fixed for some period of time.

Some contracts do attempt to take into account changing economic conditions, such as inflation, through cost-of-living adjustments, but even these relatively simple contingencies are not as widespread as one might think. One reason might be that a firm is concerned that while the aggregate price level is rising, the prices for the goods and services it sells might not be moving at the same rate. Also, cost-of-living or other contingencies add complexity to contracts that both sides may want to avoid.

Even markets where workers are not employed under explicit contracts seem to behave as if such contracts existed. In these cases, wage stickiness may stem from a desire to avoid the same uncertainty and adjustment costs that explicit contracts avert.

Finally, minimum wage laws prevent wages from falling below a legal minimum, even if unemployment is rising. Unskilled workers are particularly vulnerable to shifts in aggregate demand.

Price Stickiness

Rigidity of other prices becomes easier to explain in light of the arguments about nominal wage stickiness. Since wages are a major component of the overall cost of doing business, wage stickiness may lead to output price stickiness. With nominal wages stable, at least some firms can adopt a “wait and see” attitude before adjusting their prices. During this time, they can evaluate information about why sales are rising or falling (Is the change in demand temporary or permanent?) and try to assess likely reactions by consumers or competing firms in the industry to any price changes they might make (Will consumers be angered by a price increase, for example? Will competing firms match price changes?).

In the meantime, firms may prefer to adjust output and employment in response to changing market conditions, leaving product price alone. Quantity adjustments have costs, but firms may assume that the associated risks are smaller than those associated with price adjustments.

Another possible explanation for price stickiness is the notion that there are adjustment costs associated with changing prices. In some cases, firms must print new price lists and catalogs, and notify customers of price changes. Doing this too often could jeopardize customer relations.

Yet another explanation of price stickiness is that firms may have explicit long-term contracts to sell their products to other firms at specified prices. For example, electric utilities often buy their inputs of coal or oil under long-term contracts.

Taken together, these reasons for wage and price stickiness explain why aggregate price adjustment may be incomplete in the sense that the change in the price level is insufficient to maintain real GDP at its potential level. These reasons do not lead to the conclusion that no price adjustments occur. But the adjustments require some time. During this time, the economy may remain above or below its potential level of output.

Equilibrium Levels of Price and Output in the Short Run

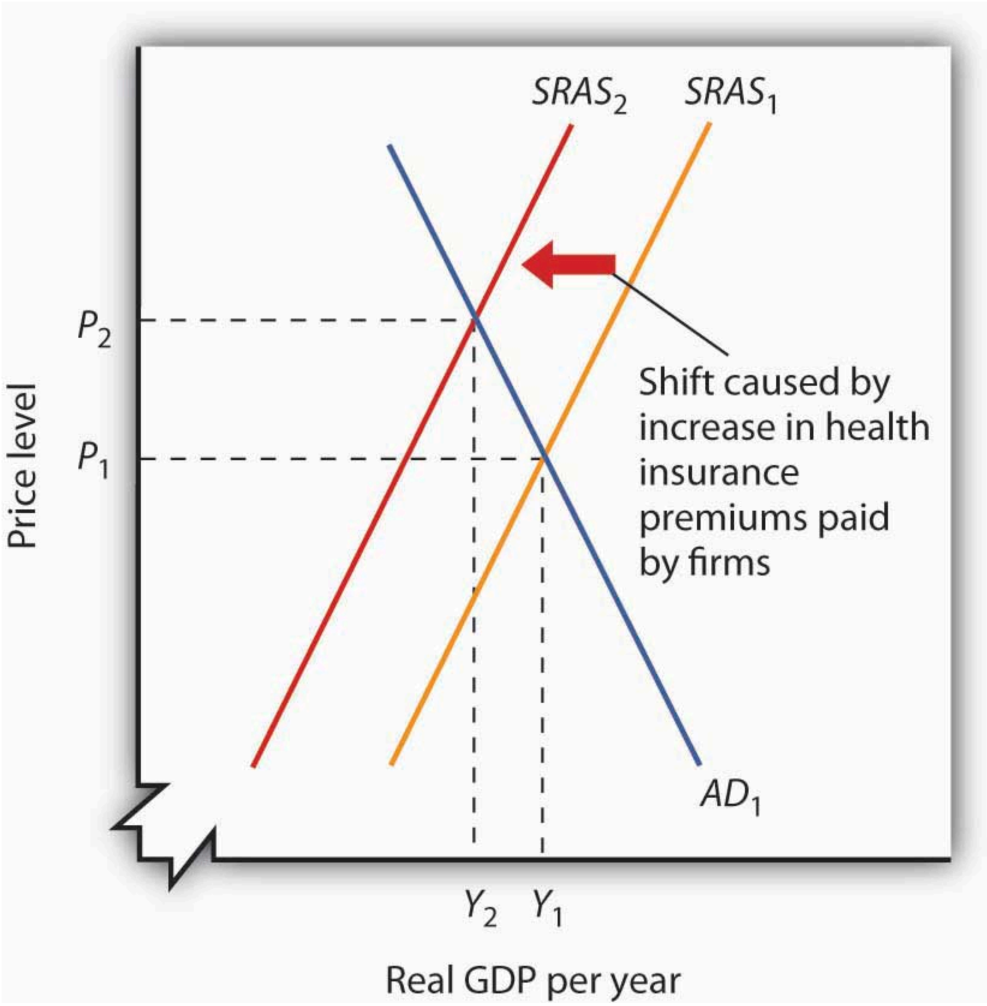
To illustrate how we will use the model of aggregate demand and aggregate supply, let us examine the impact of two events: an increase in the cost of health care and an increase in government purchases. The first reduces short-run aggregate supply; the second increases aggregate demand. Both events change equilibrium real GDP and the price level in the short run.

A Change in the Cost of Health Care

In the United States, most people receive health insurance for themselves and their families through their employers. In fact, it is quite common for employers to pay a large percentage of employees' health insurance premiums, and this benefit is often written into labor contracts. As the cost of health care has gone up over time, firms have had to pay higher and higher health insurance premiums. With nominal wages fixed in the short run, an increase in health insurance premiums paid by firms raises the cost of employing each worker. It affects the cost of production in the same way that higher wages would. The result of higher health insurance premiums is that firms will choose to employ fewer workers.

Suppose the economy is operating initially at the short-run equilibrium at the intersection of AD_1 and $SRAS_1$, with a real GDP of Y_1 and a price level of P_1, as shown in Figure 22.9 "An Increase in Health Insurance Premiums Paid by Firms". This is the initial equilibrium price and output in the short run. The increase in labor cost shifts the short-run aggregate supply curve to $SRAS_2$. The price level rises to P_2 and real GDP falls to Y_2.

Figure 22.9 An Increase in Health Insurance Premiums Paid by Firms



An increase in health insurance premiums paid by firms increases labor costs, reducing short-run aggregate supply from $SRAS_1$ to $SRAS_2$. The price level rises from P_1 to P_2 and output falls from Y_1 to Y_2.

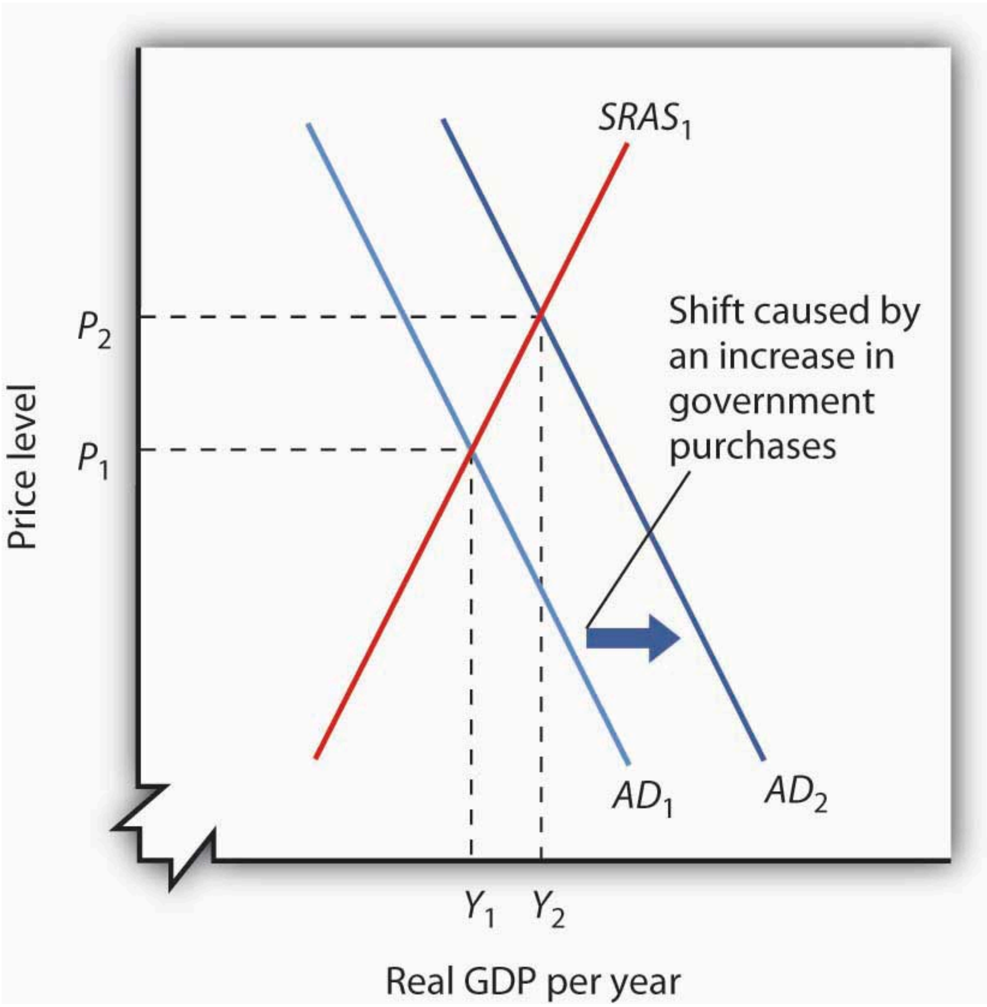
A reduction in health insurance premiums would have the opposite effect. There would be a shift to the right in the short-run aggregate supply curve with pressure on the price level to fall and real GDP to rise.

A Change in Government Purchases

Suppose the federal government increases its spending for highway construction. This circumstance leads to an increase in U.S. government purchases and an increase in aggregate demand.

Assuming no other changes affect aggregate demand, the increase in government purchases shifts the aggregate demand curve by a multiplied amount of the initial increase in government purchases to AD_2 in Figure 22.10 "An Increase in Government Purchases". Real GDP rises from Y_1 to Y_2, while the price level rises from P_1 to P_2. Notice that the increase in real GDP is less than it would have been if the price level had not risen.

Figure 22.10 An Increase in Government Purchases



An increase in government purchases boosts aggregate demand from AD_1 to AD_2. Short-run equilibrium is at the intersection of AD_2 and the short-run aggregate supply curve $SRAS_1$. The price level rises to P_2 and real GDP rises to Y_2.

In contrast, a reduction in government purchases would reduce aggregate demand. The aggregate demand curve shifts to the left, putting pressure on both the price level and real GDP to fall.

In the short run, real GDP and the price level are determined by the intersection of the aggregate demand and short-run aggregate supply curves. Recall, however, that the short run is a period in which sticky prices may prevent the economy from reaching its natural level of employment and potential output. In the next section, we will see how the model adjusts to move the economy to long-run equilibrium and what, if anything, can be done to steer the economy toward the natural level of employment and potential output.

Key Takeaways

- The short run in macroeconomics is a period in which wages and some other prices are sticky. The long run is a period in which full wage and price flexibility, and market adjustment, has been achieved, so that the economy is at the natural level of employment and potential output.
- The long-run aggregate supply curve is a vertical line at the potential level of output. The intersection of the economy's aggregate demand and long-run aggregate supply curves determines its equilibrium real GDP and price level in the long run.
- The short-run aggregate supply curve is an upward-sloping curve that shows the quantity of total output that will be produced at each price level in the short run. Wage and price stickiness account for the short-run aggregate supply curve's upward slope.
- Changes in prices of factors of production shift the short-run aggregate supply curve. In addition, changes in the capital stock, the stock of natural resources, and the level of technology can also cause the short-run aggregate supply curve to shift.
- In the short run, the equilibrium price level and the equilibrium level of total output are determined by the intersection of the aggregate demand and the short-run aggregate supply curves. In the short run, output can be either below or above potential output.

Try It!

The tools we have covered in this section can be used to understand the Great Depression of the 1930s. We know that investment and consumption began falling in late 1929. The reductions were reinforced by plunges in net exports and government purchases over the next four years. In addition, nominal wages plunged 26% between 1929 and 1933. We also know that real GDP in 1933 was 30% below real GDP in 1929. Use the tools of aggregate demand and short-run aggregate supply to graph and explain what happened to the economy between 1929 and 1933.

Case in Point: The U.S. Recession of 2001

Figure 22.11

What were the causes of the U.S. recession of 2001? Economist Kevin Kliesen of the Federal Reserve Bank of St. Louis points to four factors that, taken together, shifted the aggregate demand curve to the left and kept it there for a long enough period to keep real GDP falling for about nine months. They were the fall in stock market prices, the decrease in business investment both for computers and software and in structures, the decline in the real value of exports, and the aftermath of 9/11. Notable exceptions to this list of culprits were the behavior of consumer spending during the period and new residential housing, which falls into the investment category.

During the expansion in the late 1990s, a surging stock market probably made it easier for firms to raise funding for investment in both structures and information technology. Even though the stock market bubble burst well before the actual recession, the continuation of projects already underway delayed the decline in the investment component of GDP. Also, spending for information technology was probably prolonged as firms dealt with Y2K computing issues, that is, computer problems associated with the change in the date from 1999 to 2000. Most computers used only two digits to indicate the year, and when the year changed from '99 to '00, computers did not know how to interpret the change, and extensive reprogramming of computers was required.

Real exports fell during the recession because (1) the dollar was strong during the period and (2) real GDP growth in the rest of the world fell almost 5% from 2000 to 2001.

Then, the terrorist attacks of 9/11, which literally shut down transportation and financial markets for several days, may have prolonged these negative tendencies just long enough to turn what might otherwise have been a mild decline into enough of a downtown to qualify the period as a recession.

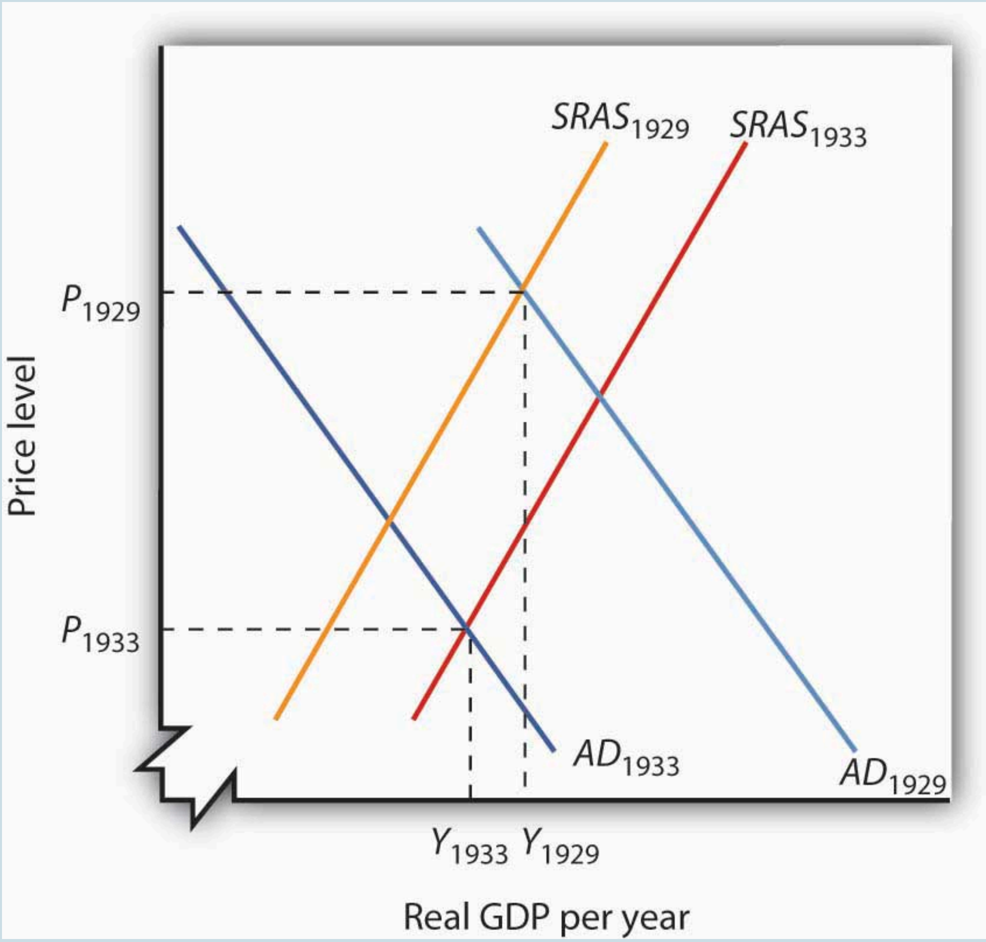
During this period the measured price level was essentially stable—with the implicit price deflator rising by less than 1%. Thus, while the aggregate demand curve shifted left as a result of all the reasons given above, there was also a leftward shift in the short-run aggregate supply curve.

Source: Kevin L. Kliesen, "The 2001 Recession: How Was It Different and What Developments May Have Caused It?" *The Federal Reserve Bank of St. Louis Review*, September/October 2003: 23–37.

Answer to Try It! Problem

All components of aggregate demand (consumption, investment, government purchases, and net exports) declined between 1929 and 1933. Thus the aggregate demand curve shifted markedly to the left, moving from AD_{1929} to AD_{1933}. The reduction in nominal wages corresponds to an increase in short-run aggregate supply from $SRAS_{1929}$ to $SRAS_{1933}$. Since real GDP in 1933 was less than real GDP in 1929, we know that the movement in the aggregate demand curve was greater than that of the short-run aggregate supply curve.

Figure 22.12


$SRAS_{1929}$
$SRAS_{1933}$
P_{1929}
Price level
P_{1933}
AD_{1933}
AD_{1929}
Y_{1933} Y_{1929}
Real GDP per year

7.3 Recessionary and Inflationary Gaps and Long-Run Macroeconomic Equilibrium

Learning Objectives

1. Explain and illustrate graphically recessionary and inflationary gaps and relate these gaps to what is happening in the labor market.
2. Identify the various policy choices available when an economy experiences an inflationary or recessionary gap and discuss some of the pros and cons that make these choices controversial.

The intersection of the economy's aggregate demand and short-run aggregate supply curves determines equilibrium real GDP and price level in the short run. The intersection of aggregate demand and long-run aggregate supply determines its long-run equilibrium. In this section we will examine the process through which an economy moves from equilibrium in the short run to equilibrium in the long run.

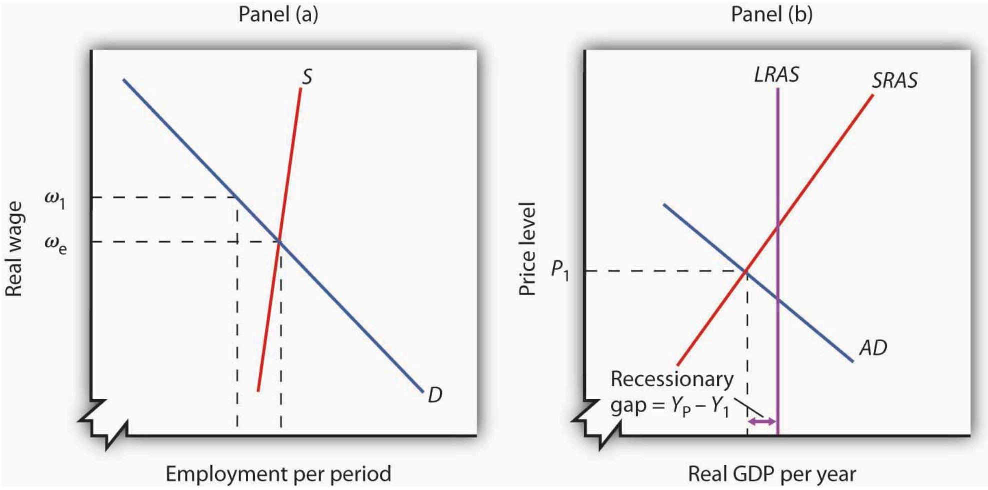
The long run puts a nation's macroeconomic house in order: only frictional and structural unemployment remain, and the price level is stabilized. In the short run, stickiness of nominal wages and other prices can prevent the economy from achieving its potential output. Actual output may exceed or fall short of potential output. In such a situation the economy operates with a gap. When output is above potential, employment is above the natural level of employment. When output is below potential, employment is below the natural level.

Recessionary and Inflationary Gaps

At any time, real GDP and the price level are determined by the intersection of the aggregate demand and short-run aggregate supply curves. If employment is below the natural level of employment, real GDP will be below potential. The aggregate demand and short-run aggregate supply curves will intersect to the left of the long-run aggregate supply curve.

Suppose an economy's natural level of employment is L_e, shown in Panel (a) of Figure 7.13 "A Recessionary Gap". This level of employment is achieved at a real wage of ω_e. Suppose, however, that the initial real wage ω_1 exceeds this equilibrium value. Employment at L_1 falls short of the natural level. A lower level of employment produces a lower level of output; the aggregate demand and short-run aggregate supply curves, *AD* and *SRAS*, intersect to the left of the long-run aggregate supply curve *LRAS* in Panel (b). The gap between the level of real GDP and potential output, when real GDP is less than potential, is called a recessionary gapThe gap between the level of real GDP and potential output, when real GDP is less than potential..

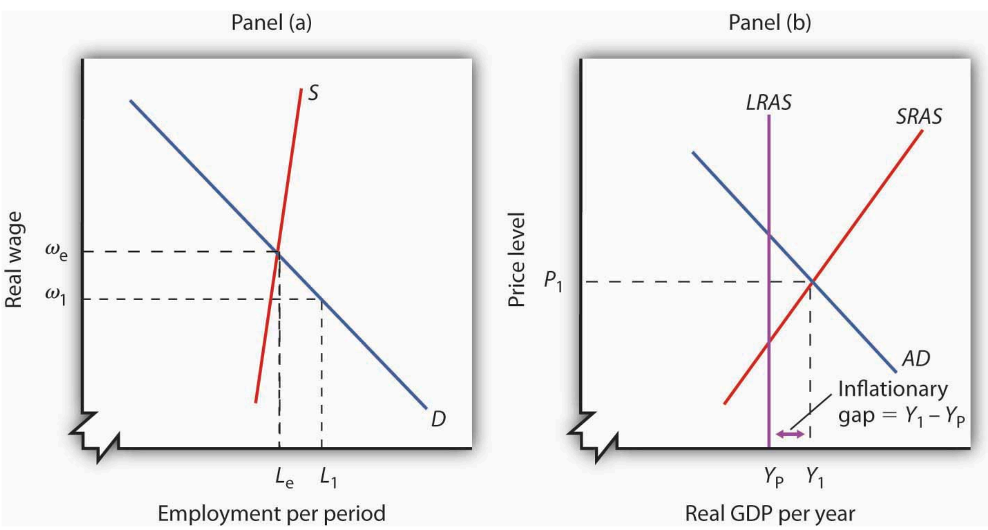
Figure 7.13 A Recessionary Gap



If employment is below the natural level, as shown in Panel (a), then output must be below potential. Panel (b) shows the recessionary gap $Y_P - Y_1$, which occurs when the aggregate demand curve *AD* and the short-run aggregate supply curve *SRAS* intersect to the left of the long-run aggregate supply curve *LRAS*.

Just as employment can fall short of its natural level, it can also exceed it. If employment is greater than its natural level, real GDP will also be greater than its potential level. [Figure 7.14 "An Inflationary Gap"](#) shows an economy with a natural level of employment of L_e in Panel (a) and potential output of Y_P in Panel (b). If the real wage ω_1 is less than the equilibrium real wage ω_e, then employment L_1 will exceed the natural level. As a result, real GDP, Y_1, exceeds potential. The gap between the level of real GDP and potential output, when real GDP is greater than potential, is called an inflationary gap. The gap between the level of real GDP and potential output, when real GDP is greater than potential.. In Panel (b), the inflationary gap equals $Y_1 - Y_P$.

Figure 7.14 An Inflationary Gap



Panel (a) shows that if employment is above the natural level, then output must be above potential. The inflationary gap, shown in Panel (b), equals $Y_1 - Y_P$. The aggregate demand curve *AD* and the short-run aggregate supply curve *SRAS* intersect to the right of the long-run aggregate supply curve *LRAS*.

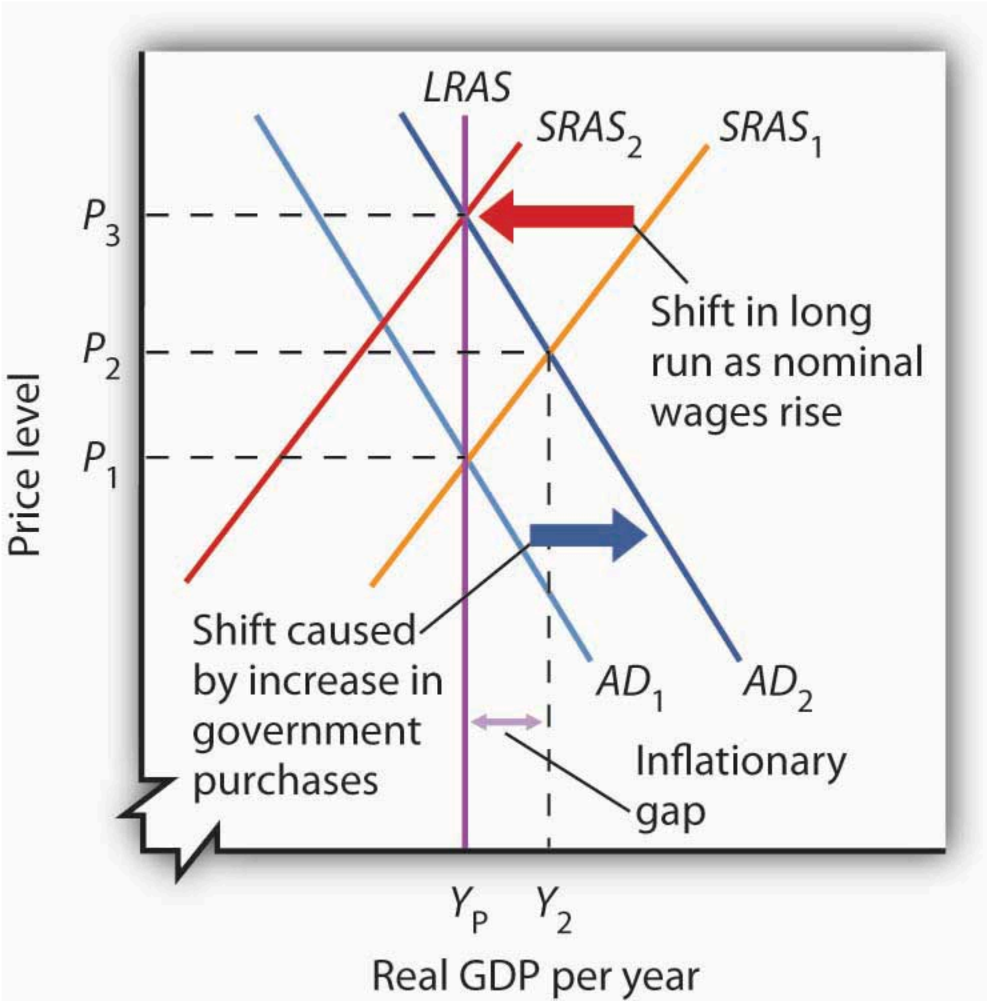
Restoring Long-Run Macroeconomic Equilibrium

We have already seen that the aggregate demand curve shifts in response to a change in consumption, investment, government purchases, or net exports. The short-run aggregate supply curve shifts in response to changes in the prices of factors of production, the quantities of factors of production available, or technology. Now we will see how the economy responds to a shift in aggregate demand or short-run aggregate supply using two examples presented earlier: a change in government purchases and a change in health-care costs. By returning to these examples, we will be able to distinguish the long-run response from the short-run response.

A Shift in Aggregate Demand: An Increase in Government Purchases

Suppose an economy is initially in equilibrium at potential output YP as in Figure 7.15 "Long-Run Adjustment to an Inflationary Gap". Because the economy is operating at its potential, the labor market must be in equilibrium; the quantities of labour demanded and supplied are equal.

Figure 7.15 Long-Run Adjustment to an Inflationary Gap



An increase in aggregate demand to AD_2 boosts real GDP to Y_2 and the price level to P_2, creating an inflationary gap of $Y_2 - Y_P$. In the long run, as price and nominal wages increase, the short-run aggregate supply curve moves to $SRAS_2$. Real GDP returns to potential.

Now suppose aggregate demand increases because one or more of its components (consumption, investment, government purchases, and net exports) has increased at each price level. For example, suppose government purchases increase. The aggregate demand curve shifts from AD_1 to AD_2 in Figure 7.15 "Long-Run Adjustment to an Inflationary Gap". That will increase real GDP to Y_2 and force the price level up to P_2 in the short run. The higher price level, combined with a fixed nominal wage, results in a lower real wage. Firms employ more workers to supply the increased output.

The economy's new production level Y_2 exceeds potential output. Employment exceeds its natural level. The economy with output of Y_2 and price level of P_2 is only in short-run equilibrium; there is an inflationary gap equal to the difference between Y_2 and Y_P. Because real GDP is above potential, there will be pressure on prices to rise further.

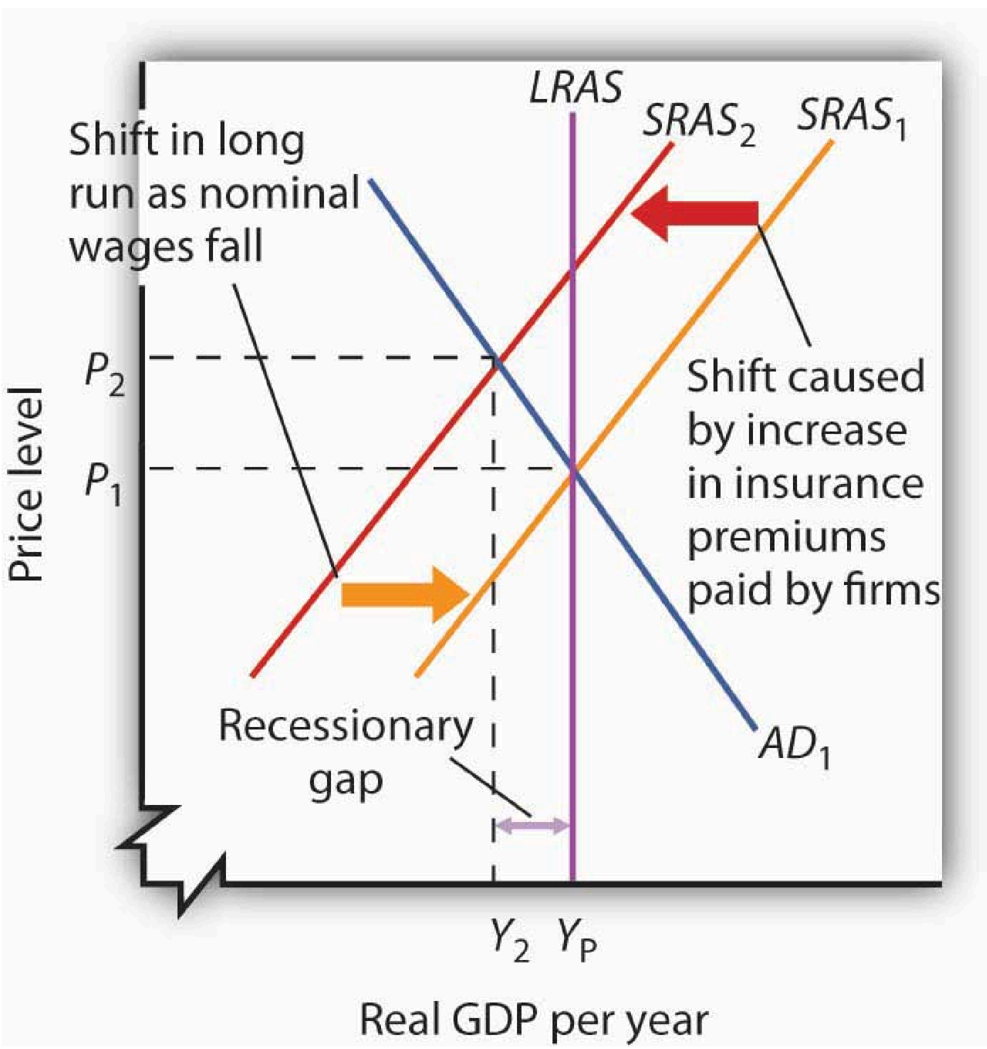
Ultimately, the nominal wage will rise as workers seek to restore their lost purchasing power. As the nominal wage rises, the short-run aggregate supply curve will begin shifting to the left. It will continue to shift as long as the nominal wage rises, and the nominal wage will rise as long as there is an inflationary gap. These shifts in short-run aggregate supply, however, will reduce real GDP and thus begin to close this gap. When the short-run aggregate supply curve reaches

$SRAS_2$, the economy will have returned to its potential output, and employment will have returned to its natural level. These adjustments will close the inflationary gap.

A Shift in Short-Run Aggregate Supply: An Increase in the Cost of Health Care

Again suppose, with an aggregate demand curve at AD_1 and a short-run aggregate supply at $SRAS_1$, an economy is initially in equilibrium at its potential output Y_P, at a price level of P_1, as shown in Figure 22.16 "Long-Run Adjustment to a Recessionary Gap". Now suppose that the short-run aggregate supply curve shifts owing to a rise in the cost of health care. As we explained earlier, because health insurance premiums are paid primarily by firms for their workers, an increase in premiums raises the cost of production and causes a reduction in the short-run aggregate supply curve from $SRAS_1$ to $SRAS_2$.

Figure 7.16 Long-Run Adjustment to a Recessionary Gap



A decrease in aggregate supply from $SRAS_1$ to $SRAS_2$ reduces real GDP to Y_2 and raises the price level to P_2, creating a recessionary gap of $Y_P - Y_2$. In the long run, as prices and nominal wages decrease, the short-run aggregate supply curve moves back to $SRAS_1$ and real GDP returns to potential.

As a result, the price level rises to P_2 and real GDP falls to Y_2. The economy now has a recessionary gap equal to the difference between Y_P and Y_2. Notice that this situation is particularly disagreeable, because both unemployment and the price level rose.

With real GDP below potential, though, there will eventually be pressure on the price level to fall. Increased unemployment also puts pressure on nominal wages to fall. In the long run, the short-run aggregate supply curve shifts back to $SRAS_1$. In this case, real GDP returns to potential at Y_P, the price level falls back to P_1, and employment returns to its natural level. These adjustments will close the recessionary gap.

How sticky prices and nominal wages are will determine the time it takes for the economy to return to potential. People often expect the government or the central bank to respond in some way to try to close gaps. This issue is addressed next.

Gaps and Public Policy

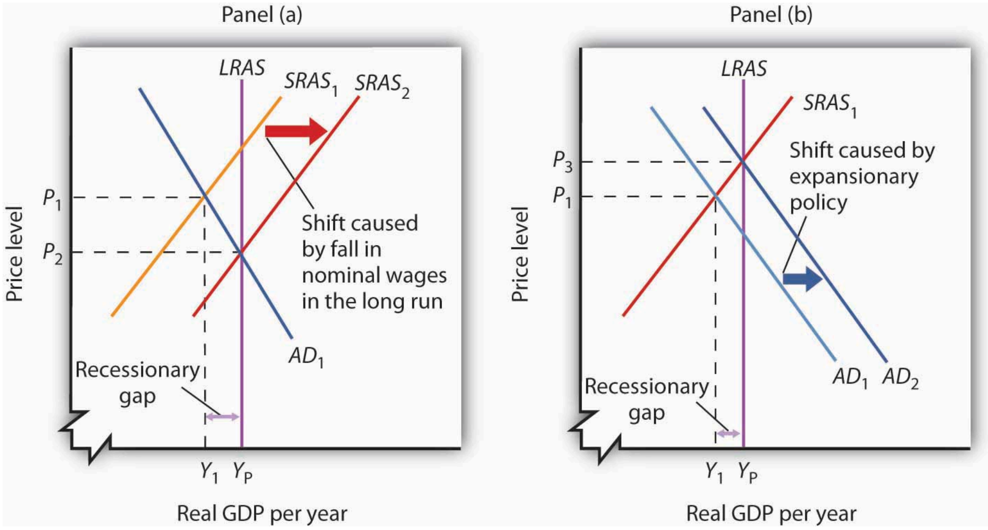
If the economy faces a gap, how do we get from that situation to potential output?

Gaps present us with two alternatives. First, we can do nothing. In the long run, real wages will adjust to the equilibrium level, employment will move to its natural level, and real GDP will move to its potential. Second, we can do something. Faced with a recessionary or an inflationary gap, policy makers can undertake policies aimed at shifting the aggregate demand or short-run aggregate supply curves in a way that moves the economy to its potential. A policy choice to take no action to try to close a recessionary or an inflationary gap, but to allow the economy to adjust on its own to its potential output, is a nonintervention policy. A policy choice to take no action to try to close a recessionary or an inflationary gap, but to allow the economy to adjust on its own to its potential output.. A policy in which the government or central bank acts to move the economy to its potential output is called a stabilization policy. A policy in which the government or central bank acts to move the economy to its potential output..

Nonintervention or Expansionary Policy?

Figure 7.17 "Alternatives in Closing a Recessionary Gap" illustrates the alternatives for closing a recessionary gap. In both panels, the economy starts with a real GDP of Y_1 and a price level of P_1. There is a recessionary gap equal to $Y_P - Y_1$. In Panel (a), the economy closes the gap through a process of self-correction. Real and nominal wages will fall as long as employment remains below the natural level. Lower nominal wages shift the short-run aggregate supply curve. The process is a gradual one, however, given the stickiness of nominal wages, but after a series of shifts in the short-run aggregate supply curve, the economy moves toward equilibrium at a price level of P_2 and its potential output of Y_P.

Figure 7.17 Alternatives in Closing a Recessionary Gap



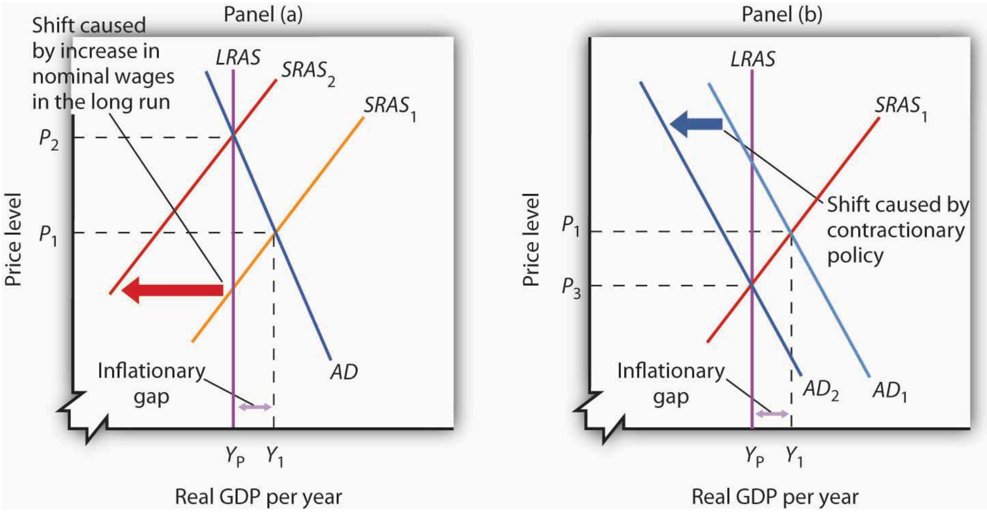
Panel (a) illustrates a gradual closing of a recessionary gap. Under a nonintervention policy, short-run aggregate supply shifts from $SRAS_1$ to $SRAS_2$. Panel (b) shows the effects of expansionary policy acting on aggregate demand to close the gap.

Panel (b) illustrates the stabilization alternative. Faced with an economy operating below its potential, public officials act to stimulate aggregate demand. For example, the government can increase government purchases of goods and services or cut taxes. Tax cuts leave people with more after-tax income to spend, boost their consumption, and increase aggregate demand. As AD_1 shifts to AD_2 in Panel (b) of Figure 7.17 "Alternatives in Closing a Recessionary Gap", the economy achieves output of Y_P, but at a higher price level, P_3. A stabilization policy designed to increase real GDP is known as an expansionary policyA stabilization policy designed to increase real GDP..

Nonintervention or Contractionary Policy?

Figure 7.18 "Alternatives in Closing an Inflationary Gap" illustrates the alternatives for closing an inflationary gap. Employment in an economy with an inflationary gap exceeds its natural level—the quantity of labor demanded exceeds the long-run supply of labor. A nonintervention policy would rely on nominal wages to rise in response to the shortage of labor. As nominal wages rise, the short-run aggregate supply curve begins to shift, as shown in Panel (a), bringing the economy to its potential output when it reaches $SRAS_2$ and P_2.

Figure 22.18 Alternatives in Closing an Inflationary Gap



Panel (a) illustrates a gradual closing of an inflationary gap. Under a nonintervention policy, short-run aggregate supply shifts from $SRAS_1$ to $SRAS_2$. Panel (b) shows the effects of contractionary policy to reduce aggregate demand from AD_1 to AD_2 in order to close the gap.

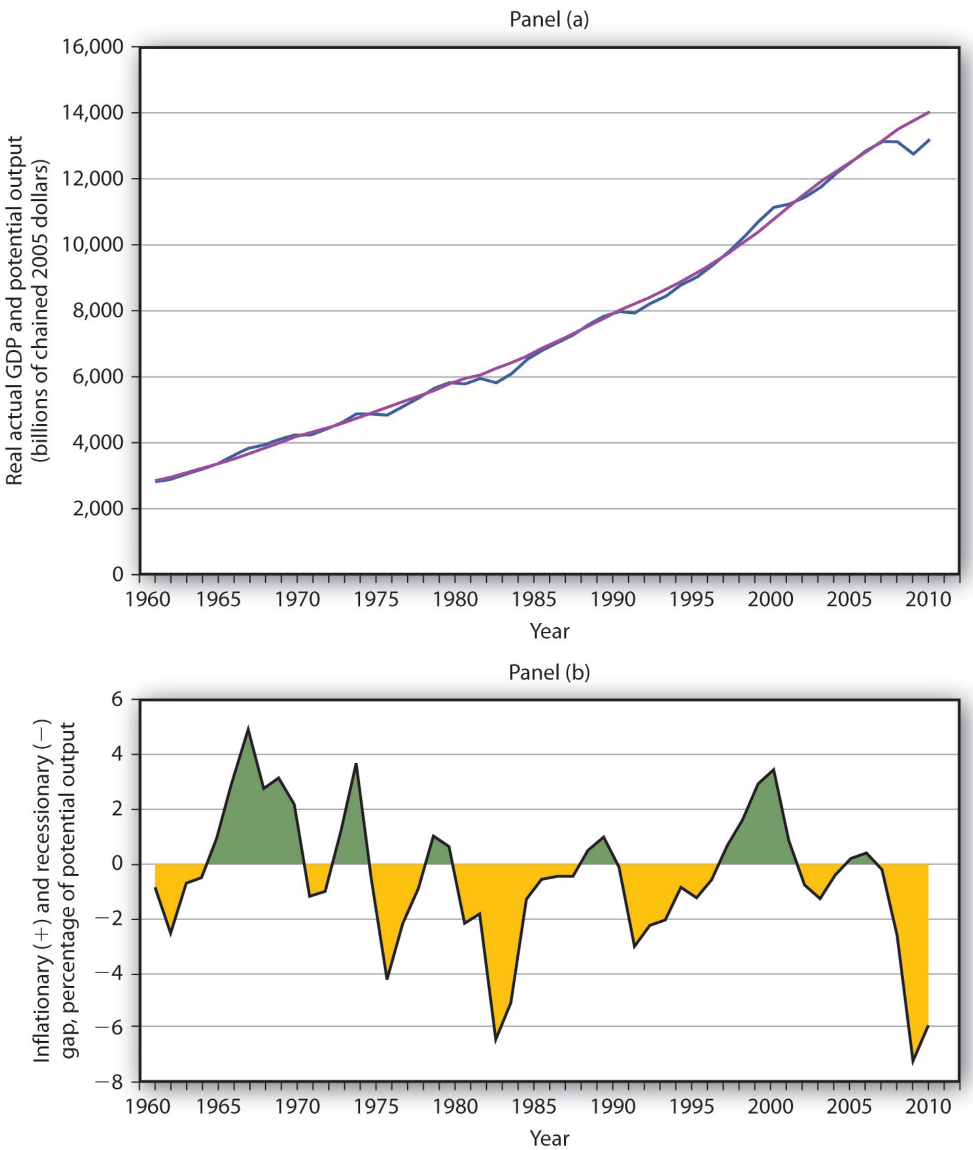
A stabilization policy that reduces the level of GDP is a contractionary policy A stabilization policy designed to reduce real GDP.. Such a policy would aim at shifting the aggregate demand curve from AD_1 to AD_2 to close the gap, as shown in Panel (b). A policy to shift the aggregate demand curve to the left would return real GDP to its potential at a price level of P_3.

For both kinds of gaps, a combination of letting market forces in the economy close part of the gap and of using stabilization policy to close the rest of the gap is also an option. Later chapters will explain stabilization policies in more detail, but there are essentially two types of stabilization policy: fiscal policy and monetary policy. Fiscal policy The use of government purchases, transfer payments, and taxes to influence the level of economic activity. is the use of government purchases, transfer payments, and taxes to influence the level of economic activity. Monetary policy The use of central bank policies to influence the level of economic activity. is the use of central bank policies to influence the level of economic activity.

To Intervene or Not to Intervene: An Introduction to the Controversy

How large are inflationary and recessionary gaps? Panel (a) of Figure 7.19 "Real GDP and Potential Output" shows potential output versus the actual level of real GDP in the United States since 1960. Real GDP appears to follow potential output quite closely, although you see some periods where there have been inflationary or recessionary gaps. Panel (b) shows the sizes of these gaps expressed as percentages of potential output. The percentage gap is positive during periods of inflationary gaps and negative during periods of recessionary gaps. The economy seldom departs by more than 5% from its potential output.

Figure 7.19 Real GDP and Potential Output



Panel (a) shows potential output (the blue line) and actual real GDP (the purple line) since 1960. Panel (b) shows the gap between potential and actual real GDP expressed as a percentage of potential output. Inflationary gaps are shown in green and recessionary gaps are shown in yellow.

Source: Bureau of Economic Analysis, NIPA Table 1.1.6. Real Gross Domestic Product, Chained Dollars [Billions of chained (2005) dollars]. Seasonally adjusted at annual rates 2010 is through 3rd quarter; Congressional Budget Office, The Budget and Economic Outlook, August 2010.

Panel (a) gives a long-run perspective on the economy. It suggests that the economy generally operates at about potential output. In Panel (a), the gaps seem minor. Panel (b) gives a short-run perspective; the view it gives emphasizes the gaps. Both of these perspectives are important. While it is reassuring to see that the economy is often close to potential, the years in which there are substantial gaps have real effects: Inflation or unemployment can harm people.

Some economists argue that stabilization policy can and should be used when recessionary or inflationary gaps exist. Others urge reliance on the economy's own ability to correct itself. They sometimes argue that the tools available to the public sector to influence aggregate demand are not likely to shift the curve, or they argue that the tools would shift the curve in a way that could do more harm than good.

Economists who advocate stabilization policies argue that prices are sufficiently sticky that the economy's own adjustment to its potential will be a slow process—and a painful one. For an economy with a recessionary gap, unacceptably high levels of unemployment will persist for too long a time. For an economy with an inflationary gap, the increased prices that occur as the short-run aggregate supply curve shifts upward impose too high an inflation rate in the short run. These economists believe it is far preferable to use stabilization policy to shift the aggregate demand curve in an effort to shorten the time the economy is subject to a gap.

Economists who favour a nonintervention approach accept the notion that stabilization policy can shift the aggregate demand curve. They argue, however, that such efforts are not nearly as simple in the real world as they may appear on paper. For example, policies to change real GDP may not affect the economy for months or even years. By the time the impact of the stabilization policy occurs, the state of the economy might have changed. Policy makers might choose an expansionary policy when a contractionary one is needed or vice versa. Other economists who favor nonintervention also question how sticky prices really are and if gaps even exist.

The debate over how policy makers should respond to recessionary and inflationary gaps is an ongoing one. These issues of nonintervention versus stabilization policies lie at the heart of the macroeconomic policy debate. We will return to them as we continue our analysis of the determination of output and the price level.

Key Takeaways

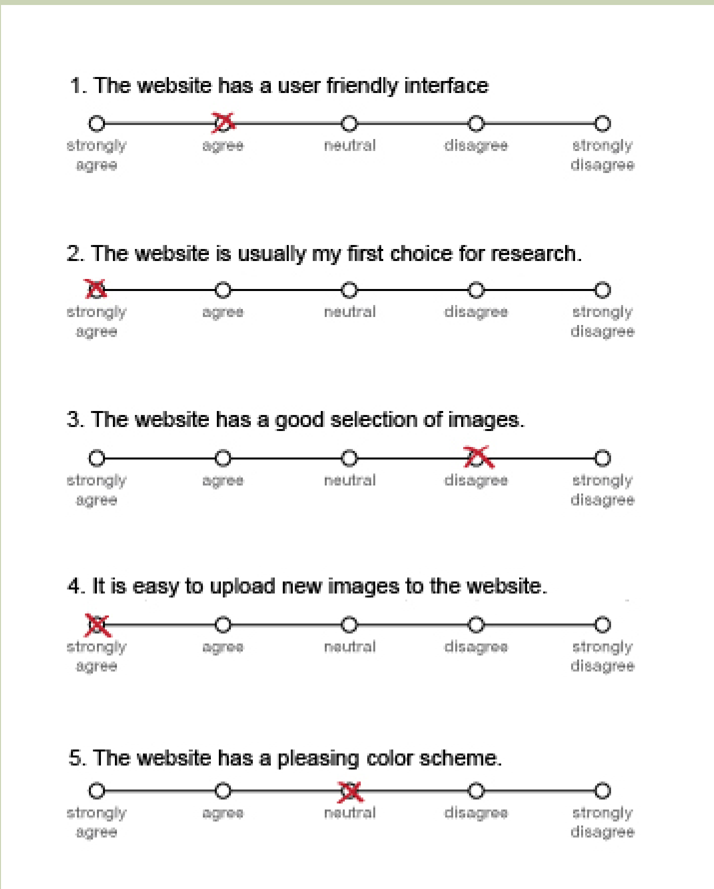
- When the aggregate demand and short-run aggregate supply curves intersect below potential output, the economy has a recessionary gap. When they intersect above potential output, the economy has an inflationary gap.
- Inflationary and recessionary gaps are closed as the real wage returns to equilibrium, where the quantity of labour demanded equals the quantity supplied. Because of nominal wage and price stickiness, however, such an adjustment takes time.
- When the economy has a gap, policy makers can choose to do nothing and let the economy return to potential output and the natural level of employment on its own. A policy to take no action to try to close a gap is a nonintervention policy.
- Alternatively, policy makers can choose to try to close a gap by using stabilization policy. Stabilization policy designed to increase real GDP is called expansionary policy. Stabilization policy designed to decrease real GDP is called contractionary policy.

Try It!

Using the scenario of the Great Depression of the 1930s, as analyzed in the previous Try It!, tell what kind of gap the U.S. economy faced in 1933, assuming the economy had been at potential output in 1929. Do you think the unemployment rate was above or below the natural rate of unemployment? How could the economy have been brought back to its potential output?

Case in Point: Survey of Economists Reveals Little Consensus on Macroeconomic Policy Issues

Figure 7.20



"An economy in short-run equilibrium at a real GDP below potential GDP has a self-correcting mechanism that will eventually return it to potential real GDP."

Of economists surveyed, 36% disagreed, 33% agreed with provisos, 25% agreed, and 5% did not respond. So, only about 60% of economists responding to the survey agreed that the economy would adjust on its own.

"Changes in aggregate demand affect real GDP in the short run but not in the long run."

On this statement, 36% disagreed, 31% agreed with provisos, 29% agreed, and 4% did not respond. Once again, about 60% of economists accepted the conclusion of the aggregate demand–aggregate supply model.

This level of disagreement on macroeconomic policy issues among economists, based on a fall 2000 survey of members of the American Economic Association, stands in sharp contrast to their more harmonious responses to questions on international economics and microeconomics. For example,

"Tariffs and import quotas usually reduce the general welfare of society."

Seventy-two percent of those surveyed agreed with this statement outright and another 21% agreed with provisos. So, 93% of economists generally agreed with the statement.

"Minimum wages increase unemployment among young and unskilled workers."

On this, 45% agreed and 29% agreed with provisos.

"Pollution taxes or marketable pollution permits are a more economically efficient approach to pollution control than emission standards."

On this environmental question, only 6% disagreed and 63% wholeheartedly agreed.

The relatively low degree of consensus on macroeconomic policy issues and the higher degrees of consensus on other economic issues found in this survey concur with results of other periodic surveys since 1976.

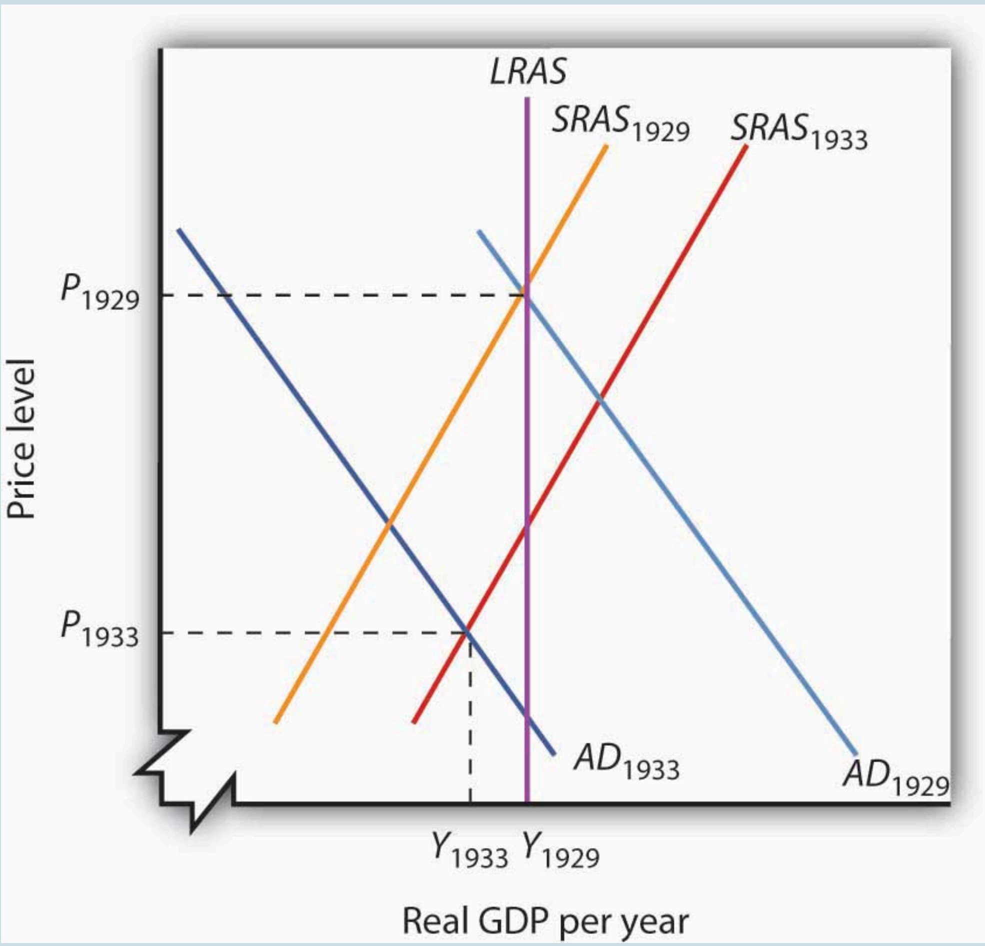
So, as textbook authors, we will not hide the dirty laundry from you. Fortunately, though, the model of aggregate demand–aggregate supply we present throughout the macroeconomic chapters can handle most of these disagreements. For example, economists who agree with the first proposition quoted above, that an economy operating below potential has self-correcting mechanisms to bring it back to potential, are probably assuming that wages and prices are not very sticky and hence that the short-run aggregate supply curve will shift rather easily to the right, as shown in Panel (a) of Figure 7.17 "Alternatives in Closing a Recessionary Gap". In contrast, economists who disagree with the statement are saying that the movement of the short-run aggregate supply curve is likely to be slow. This latter group of economists probably advocates expansionary policy as shown in Panel (b) of Figure 7.17 "Alternatives in Closing a Recessionary Gap". Both groups of economists can use the same model and its constructs to analyze the macroeconomy, but they may disagree on such things as the slopes of the various curves, on how fast these various curves shift, and on the size of the underlying multiplier. The model allows economists to speak the same language of analysis even though they disagree on some specifics.

Source: Dan Fuller and Doris Geide-Stevenson, "Consensus on Economic Issues: A Survey of Republicans, Democrats and Economists," *Eastern Economic Journal* 33, no. 1 (Winter 2007): 81–94.

Answer to Try It! Problem

To the graph in the previous Try It! problem we add the long-run aggregate supply curve to show that, with output below potential, the U.S. economy in 1933 was in a recessionary gap. The unemployment rate was above the natural rate of unemployment. Indeed, real GDP in 1933 was about 30% below what it had been in 1929, and the unemployment rate had increased from 3% to 25%. Note that during the period of the Great Depression, wages did fall. The notion of nominal wage and other price stickiness discussed in this section should not be construed to mean complete wage and price inflexibility. Rather, during this period, nominal wages and other prices were not flexible enough to restore the economy to the potential level of output. There are two basic choices on how to close recessionary gaps. Nonintervention would mean waiting for wages to fall further. As wages fall, the short-run aggregate supply curve would continue to shift to the right. The alternative would be to use some type of expansionary policy. This would shift the aggregate demand curve to the right. These two options were illustrated in Figure 7.18 "Alternatives in Closing an Inflationary Gap".

Figure 7.21



7.4 Review and Practice

Summary

In this chapter, we outlined the model of aggregate demand and aggregate supply. We saw that the aggregate demand curve slopes downward, reflecting the tendency for the aggregate quantity of goods and services demanded to rise as the price level falls and to fall as the price level rises. The negative relationship between the price level and the quantity of goods and services demanded results from the wealth effect for consumption, the interest rate effect for investment, and the international trade effect for net exports. We examined the factors that can shift the aggregate demand curve as well. Generally, the aggregate demand curve shifts by a multiple of the initial amount by which the component causing it to shift changes.

We distinguished between two types of equilibria in macroeconomics—one corresponding to the short run, a period of analysis in which nominal wages and some prices are sticky, and the other corresponding to the long run, a period in which full wage and price flexibility, and hence market adjustment, have been achieved. Long-run equilibrium occurs at the intersection of the aggregate demand curve with the long-run aggregate supply curve. The long-run aggregate supply curve is a vertical line at the economy's potential level of output. Short-run equilibrium occurs at the intersection of the aggregate demand curve with the short-run aggregate supply curve. The short-run aggregate supply curve relates the quantity of total output produced to the price level in the short run. It is upward sloping because of wage and price stickiness. In short-run equilibrium, output can be below or above potential.

If an economy is initially operating at its potential output, then a change in aggregate demand or short-run aggregate supply will induce a recessionary or inflationary gap. Such a gap will be closed in the long run by changes in the nominal wage, which will shift the short-run aggregate supply curve to the left (to close an inflationary gap) or to the right (to close a recessionary gap). Policy makers might respond to a recessionary or inflationary gap with a nonintervention policy, or they could use stabilization policy.

Concept Problems

1. Explain how the following changes in aggregate demand or short-run aggregate supply, other things held unchanged, are likely to affect the level of total output and the price level in the short run.
   1. An increase in aggregate demand
   2. A decrease in aggregate demand
   3. An increase in short-run aggregate supply
   4. A reduction in short-run aggregate supply
2. Explain why a change in one component of aggregate demand will cause the aggregate demand curve to

shift by a multiple of the initial change.

3. Use the model of aggregate demand and short-run aggregate supply to explain how each of the following would affect real GDP and the price level in the short run.
    1. An increase in government purchases
    2. A reduction in nominal wages
    3. A major improvement in technology
    4. A reduction in net exports
4. How would an increase in the supply of labour affect the natural level of employment and potential output? How would it affect the real wage, the level of real GDP, and the price level in the short run? How would it affect long-run aggregate supply? What kind of gaps would be created?
5. Give three reasons for the downward slope of the aggregate demand curve.
6. "When the price level falls, people's wealth increases. When wealth increases, the real volume of consumption increases. Therefore, a decrease in the price level will cause the aggregate demand curve to shift to the right." Do you agree? Explain.
7. Suppose the economy has a recessionary gap. We know that if we do nothing, the economy will close the gap on its own. Alternatively, we could arrange for an increase in aggregate demand (say, by increasing government spending) to close the gap. How would your views about the degree of price stickiness in the economy influence your views on whether such a policy would be desirable?
8. The cost of hiring workers includes not only payments made directly to workers, that is, wages, but payments made on behalf of workers as well, such as contributions by employers to pension plans and to health-care insurance for employees. How would a decrease in the cost of employer-provided health insurance affect the economy? Using Figure 7.9 "An Increase in Health Insurance Premiums Paid by Firms" as a guide, draw a graph to illustrate your answer.
9. Suppose nominal wages never changed. What would be the significance of such a characteristic?
10. Suppose the minimum wage were increased sharply. How would this affect the equilibrium price level and output level in the model of aggregate demand and aggregate supply in the short run? In the long run?
11. Explain the short-run impact of each of the following.
    1. A discovery that makes cold fusion a reality, greatly reducing the cost of producing energy
    2. An increase in the payroll tax

Numerical Problems

1. Suppose the aggregate demand and short-run aggregate supply schedules for an economy whose potential output equals $2,700 are given by the table.

	Aggregate Quantity of Goods and Services	
Price Level	**Demanded**	**Supplied**
0.50	$3,500	$1,000
0.75	3,000	2,000
1.00	2,500	2,500
1.25	2,000	2,700
1.50	1,500	2,800

1. Draw the aggregate demand, short-run aggregate supply, and long-run aggregate supply curves.
2. State the short-run equilibrium level of real GDP and the price level.
3. Characterize the current economic situation. Is there an inflationary or a recessionary gap? If so, how large is it?
4. Now suppose aggregate demand increases by $700 at each price level; for example, the aggregate quantity of goods and services demanded at a price level of 0.50 now equals $4,200. Show the new aggregate demand curve, state the new short-run equilibrium price level and real GDP, and state whether there is an inflationary or a recessionary gap and give its size.

2. An economy is characterized by the values in the table for aggregate demand and short-run aggregate supply. Its potential output is $1,500.

	Aggregate Quantity of Goods and Services	
Price Level	**Demanded**	**Supplied**
0.50	$2,500	$1,500
0.75	2,000	2,000
1.00	1,500	2,300
1.25	1,000	2,500
1.50	500	2,600

1. Draw the aggregate demand, short-run aggregate supply, and long-run aggregate supply curves.
2. State the equilibrium level of real GDP and the price level.
3. Characterize the current economic situation. Is there an inflationary or a recessionary gap? If so, how large is it?
4. Now suppose that nominal wages rise and that the price level required to induce a particular level of total output rises by 0.50. For example, a price level of 1.00 is now required to induce producers to produce a real GDP of $1,500. Show the new short-run aggregate supply curve, state the new equilibrium price level and real GDP, and state whether there is an inflationary or a recessionary gap and give its size. Why might such a change occur?

3. Suppose the price level in a particular economy equals 1.3 and that the quantity of real GDP demanded at that price level is $1,200. An increase of 0.1 point in the price level reduces the quantity of real GDP demanded by $220, and a reduction of 0.1 point would produce an increase in the quantity of real GDP demanded of $220. Draw the aggregate demand curve and show the price level and quantity of real GDP demanded at three points.

4. Suppose an economy is described by the following aggregate demand and short-run aggregate supply curves. The potential level of output is $10 trillion.

	Aggregate Quantity of Goods and Services	
Price Level	**Demanded**	**Supplied**
3.0	$11.0 trillion	$9.0 trillion
3.4	$10.8 trillion	$9.2 trillion
3.8	$10.6 trillion	$9.4 trillion
4.2	$10.4 trillion	$9.6 trillion
4.6	$10.2 trillion	$9.8 trillion
5.0	$10.0 trillion	$10.0 trillion
5.4	$9.8 trillion	$10.2 trillion
5.8	$9.6 trillion	$10.4 trillion
6.2	$9.4 trillion	$10.6 trillion
6.6	$9.2 trillion	$10.8 trillion
7.0	$9.0 trillion	$11.0 trillion

   1. Draw the aggregate demand and short-run aggregate supply curves.
   2. What is the initial real GDP?
   3. What is the initial price level?
   4. What kind of gap, if any, exists?
   5. After the increase in health-care costs, each level of real GDP requires an increase in the price level of 0.8. For example, producing $9.0 trillion worth of goods and services now requires a price level of 3.8. What is the short-run equilibrium level of real GDP?
   6. After the health-care cost increase, what is the new equilibrium price level in the short run?
   7. What sort of gap, if any, now exists?

5. According to Alaskan state economist Mark Edwards, the multiplier effect of Alaska's trade with Japan is such that for every $1 billion exported from Alaska to Japan another $600 million is added to the state's economy (Volz, M., 2004). Calculate the size of the export multiplier.
6. The Nottinghamshire Research Observatory in England calculated that students who attend Nottingham Technical University spend about £2,760 each in the local economy for a total of £50.45 million. In total, the impact of their spending on the local economy is £63 million (Nottingham Evening Post, 2004). Calculate the size of the student spending multiplier.
7. In Goa, India, the multiplier effect of iron ore exports is calculated to be 1.62 (Ta, V. K., 2003). Calculate the impact of an additional 1,000 rupees of iron ore exports on the economy of Goa.

References

Nottingham Evening Post, "University Brings in £250m to Economy," November 4, 2004, p. 37.

Ta, V. K., "Iron Ore Mining Gives Impetus to Goa's Economy," *Times of India*, April 30, 2003.

Volz, M., “Trade Officials Hopeful for Japanese Recovery,” *Associated Press and Local Wire*, June 22, 2004, BC cycle.

CHAPTER 8: ECONOMIC GROWTH

Start Up: How Important Is Economic Growth?

How important is economic growth? The best way to answer that question is to imagine life without growth—to imagine that we did not have the gains growth brings.

For starters, divide your family's current income by six and imagine what your life would be like. Think about the kind of housing your family could afford, the size of your entertainment budget, whether you could still attend school. That will give you an idea of life a century ago in the United States, when average household incomes, adjusted for inflation, were about one-sixth what they are today. People had far smaller homes, they rarely had electricity in their homes, and only a tiny percentage of the population could even consider a college education.

To get a more recent perspective, consider how growth has changed living standards over the past half-century or so. In 1950, the United States was the world's richest nation. But if households were rich then, subsequent economic growth has made them far richer. Average per capita real disposable personal income has tripled since then. Indeed, the average household income in 1950, which must have seemed lofty then, was below what we now define as the poverty line for a household of four, even after adjusting for inflation. Economic growth during the last half-century has dramatically boosted our standard of living—and our standard of what it takes to get by.

One gauge of rising living standards is housing. A half-century ago, most families did not own homes. Today, about two-thirds do. Those homes have gotten a lot bigger: new homes built today are more than twice the size of new homes built 50 years ago. Some household appliances, such as telephones or washing machines, that we now consider basic, were luxuries a half-century ago. In 1950, less than two-thirds of housing units had complete plumbing facilities. Today, over 99% do.

Economic growth has brought gains in other areas as well. For one thing, we are able to afford more schooling. In 1950, the median number of years of school completed by adults age 25 or over was 6.8. Today, about 85% have completed 12 years of schooling and about 28% have completed four years of college. We also live longer. A baby born in 1950 had a life expectancy of 68 years. A baby born in 2004 had an expected life of nearly 10 years longer.

Of course, while economic growth can improve our material well-being, it is no panacea for all the ills of society. Americans today worry about the level of violence in society, environmental degradation, and what seems to be a loss of basic values. But while it is easy to be dismayed about many challenges of modern life, we can surely be grateful for our material wealth. Our affluence gives us the opportunity to grapple with some of our most difficult problems and to enjoy a range of choices that people only a few decades ago could not have imagined.

We learned a great deal about economic growth in the context of the production possibilities curve. Our purpose in this chapter is to relate the concept of economic growth to the model of aggregate demand and aggregate supply that we developed in the previous chapter and will use throughout our exploration of macroeconomics. We will review the forces that determine a nation's economic growth rate and examine the prospects for growth in the future. We begin by looking at the significance of growth to the overall well-being of society.

8.1 The Significance of Economic Growth

Learning Objectives

1. Define economic growth and explain it using the production possibilities model and the concept of potential output.
2. State the rule of 72 and use it to show how even small differences in growth rates can have major effects on a country's potential output over time.
3. Calculate the percentage rate of growth of output per capita.

To demonstrate the impact of economic growth on living standards of a nation, we must start with a clear definition of economic growth and then study its impact over time. We will also see how population growth affects the relationship between economic growth and the standard of living an economy is able to achieve.

Defining Economic Growth

Economic growth is a long-run process that occurs as an economy's potential output increases. Changes in real GDP from quarter to quarter or even from year to year are short-run fluctuations that occur as aggregate demand and short-run aggregate supply change. Regardless of media reports stating that the economy grew at a certain rate in the last quarter or that it is expected to grow at a particular rate during the next year, short-run changes in real GDP say little about economic growth. In the long run, economic activity moves toward its level of potential output. Increases in potential constitute economic growth.

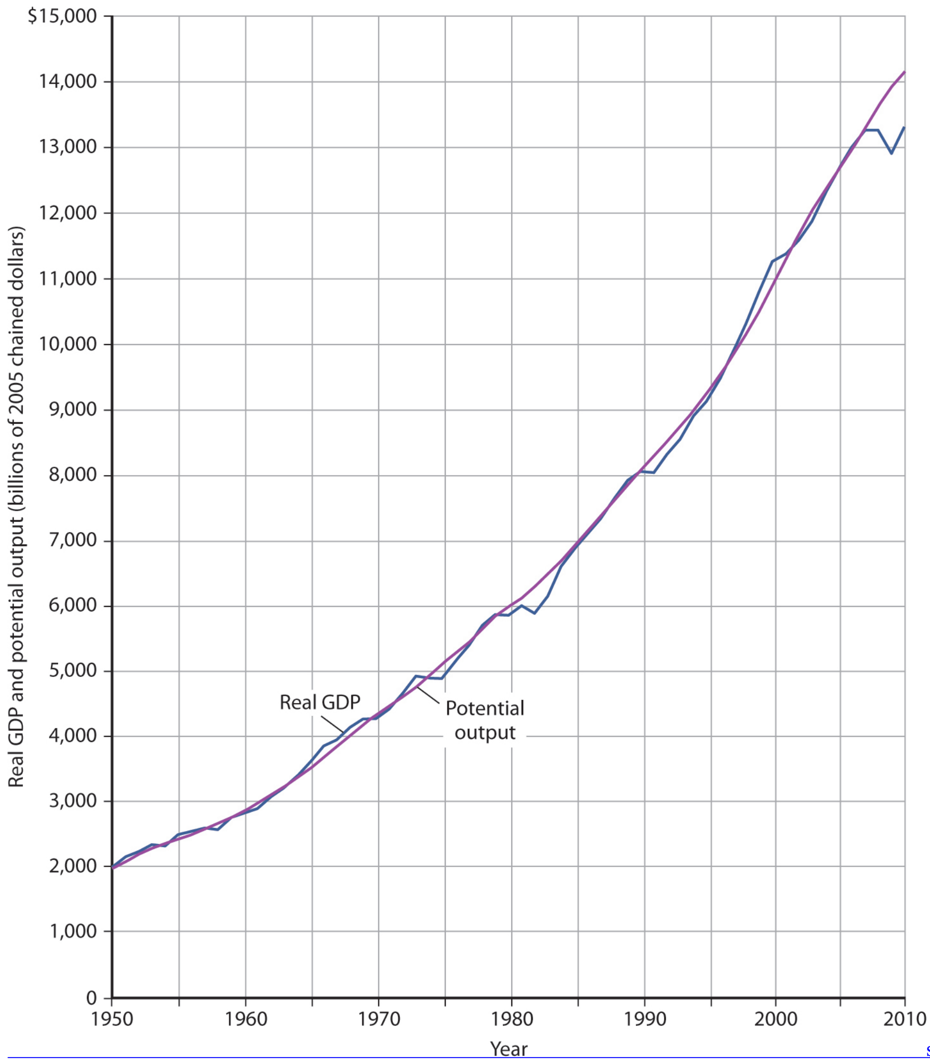
Earlier we defined economic growth as the process through which an economy achieves an outward shift in its production possibilities curve. How does a shift in the production possibilities curve relate to a change in potential output? To produce its potential level of output, an economy must operate on its production possibilities curve. An increase in potential output thus implies an outward shift in the production possibilities curve. In the framework of the macroeconomic model of aggregate demand and aggregate supply, we show economic growth as a shift to the right in the long-run aggregate supply curve.

There are three key points about economic growth to keep in mind:

1. Growth is a process. It is not a single event; rather, it is an unfolding series of events.
2. We define growth in terms of the economy's ability to produce goods and services, as indicated by its level of potential output.
3. Growth suggests that the economy's ability to produce goods and services is rising. A discussion of economic growth is thus a discussion of the series of events that increase the economy's ability to produce goods and services.

Figure 23.1 "A Century of Economic Growth" shows the record of economic growth for the U.S. economy over the past century. The graph shows annual levels of actual real GDP and of potential output. We see that the economy has experienced dramatic growth over the past century; potential output has soared more than 30-fold. The figure also reminds us of a central theme of our analysis of macroeconomics: real GDP fluctuates about potential output. Real GDP sagged well below its potential during the Great Depression of the 1930s and rose well above its potential as the nation mobilized its resources to fight World War II. With the exception of these two periods, real GDP has remained close to the economy's potential output. Since 1950, the actual level of real GDP has deviated from potential output by an average of less than 2%.

Figure 23.1 A Century of Economic Growth



st century, the level of potential output reached a level nearly 30 times its level a century earlier. Over the years, actual real GDP fluctuated about a rising level of potential output." width="497"/>

At the start of the 21st century, the level of potential output reached a level nearly 30 times its level a century earlier. Over the years, actual real GDP fluctuated about a rising level of potential output.

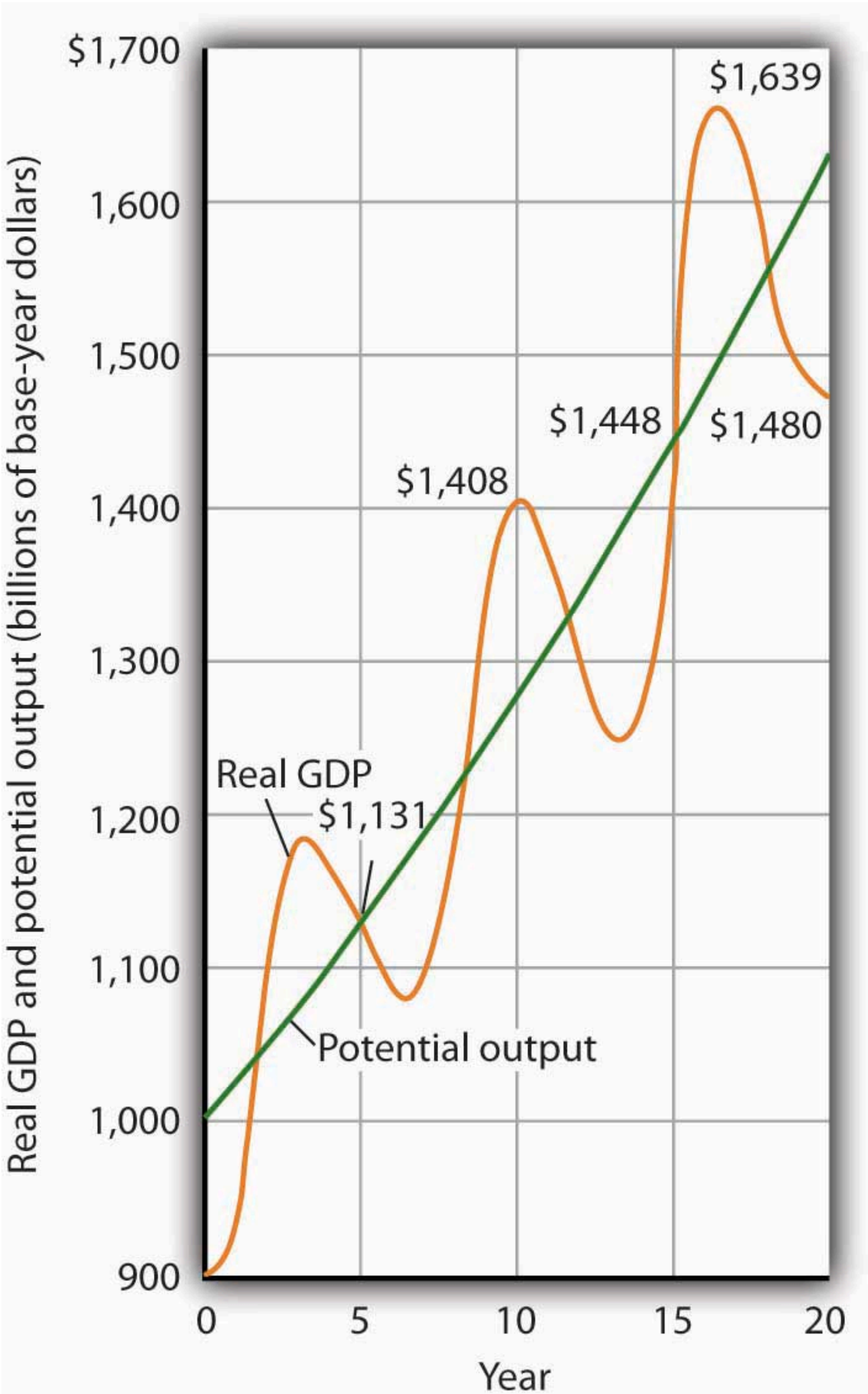
Source: 1900–1949 data from Robert Gordon, *Macroeconomics*, 6th ed. (New York: HarperCollins, 1993), Table A-1, pp. A1–A3; data for 1950–2010 from Congressional Budget Office, The Budget and Economic Outlook, August 2010.

We urge you to take some time with Figure 23.1 "A Century of Economic Growth". Over the course of the last century, it is economic growth that has taken center stage. Certainly, the fluctuations about potential output have been important. The recessionary gaps—periods when real GDP slipped below its potential—were often wrenching experiences in which millions of people endured great hardship. The inflationary gaps—periods when real GDP rose above its potential level—often produced dramatic increases in price levels. Those fluctuations mattered. It was the unemployment and/or the inflation that came with them that made headlines. But it was the quiet process of economic growth that pushed living standards ever higher. We must understand growth if we are to understand how we got where we are, and where we are likely to be going during the 21st century.

Figure 23.2 "Cyclical Change Versus Growth" tells us why we use changes in potential output, rather than actual real GDP, as our measure of economic growth. Actual values of real GDP are affected not just by changes in the potential level of output, but also by the cyclical fluctuations about that level of output.

Given our definition of economic growth, we would say that the hypothetical economy depicted in Figure 23.2 "Cyclical Change Versus Growth" grew at a 2.5% annual rate throughout the period. If we used actual values of real GDP, however, we would obtain quite different interpretations. Consider, for example, the first decade of this period: it began with a real GDP of $900 billion and a recessionary gap, and it ended in year 10 with a real GDP of $1,408 billion and an inflationary gap. If we record growth as the annual rate of change between these levels, we find an annual rate of growth of 4.6%—a rather impressive performance.

Figure 23.2 Cyclical Change Versus Growth



The use of actual values of real GDP to measure growth can give misleading results. Here, an economy's potential output (shown in green) grows at a steady rate of 2.5% per year, with actual values of real GDP fluctuating about that trend. If we measure growth in the first 10 years as the annual rate of change between beginning and ending values of real GDP, we get a growth rate of 4.6%. The rate for the second decade is 0.5%. Growth estimates based on changes in real GDP are affected by cyclical changes that do not represent economic growth.

Now consider the second decade shown in Figure 23.2 "Cyclical Change Versus Growth". It began in year 10, and it ended in year 20 with a recessionary gap. If we measure the growth rate over that period by looking at beginning and ending values of actual real GDP, we compute an annual growth rate of 0.5%. Viewed in this way, performance in the first decade is spectacular while performance in the second is rather lackluster. But these figures depend on the starting and ending points we select; the growth rate of potential output was 2.5% throughout the period.

By measuring economic growth as the rate of increase in potential output, we avoid such problems. One way to do this is to select years in which the economy was operating at the natural level of employment and then to compute the annual rate of change between those years. The result is an estimate of the rate at which potential output increased over the period in question. For the economy shown in Figure 23.2 "Cyclical Change Versus Growth", for example, we see that real GDP equaled its potential in years 5 and 15. Real GDP in year 5 was $1,131, and real GDP in year 15 was $1,448. The annual rate of change between these two years was 2.5%. If we have estimates of potential output, of course, we can simply compute annual rates of change between any two years.

The Rule of 72 and Differences in Growth Rates

The Case in Point on presidents and growth at the end of this section suggests a startling fact: the U.S. growth rate began slowing in the 1970s, did not recover until the mid-1990s, only to slow down again in the 2000s. The question we address here is: does it matter? Does a percentage point drop in the growth rate make much difference? It does. To see why, let us investigate what happens when a variable grows at a particular percentage rate.

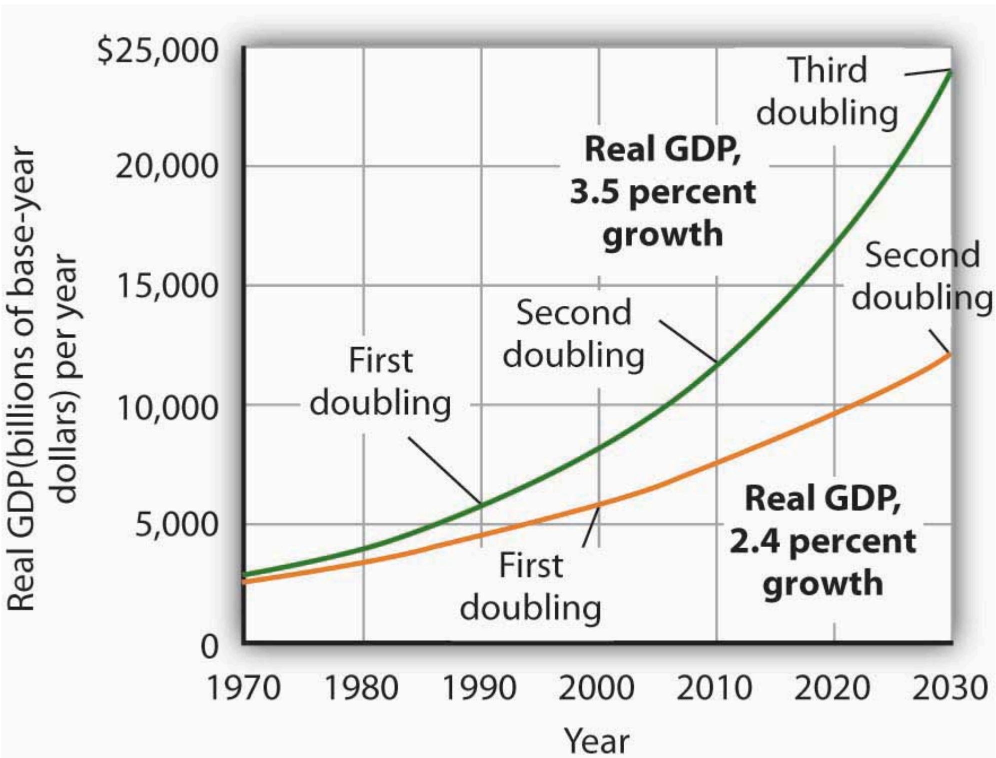
Suppose two economies with equal populations start out at the same level of real GDP but grow at different rates. Economy A grows at a rate of 3.5%, and Economy B grows at a rate of 2.4%. After a year, the difference in real GDP will hardly be noticeable. After a decade, however, real GDP in Economy A will be 11% greater than in Economy B. Over longer periods, the difference will be more dramatic. After 100 years, for example, income in Economy A will be nearly three times as great as in Economy B. If population growth in the two countries has been the same, the people of Economy A will have a far higher standard of living than those in Economy B. The difference in real GDP per person will be roughly equivalent to the difference that exists today between Great Britain and Mexico.

Over time, small differences in growth rates create large differences in incomes. An economy growing at a 3.5% rate increases by 3.5% of its initial value in the first year. In the second year, the economy increases by 3.5% of that new, higher value. In the third year, it increases by 3.5% of a still higher value. When a quantity grows at a given percentage rate, it experiences exponential growthWhen a quantity grows at a given percentage rate.. A variable that grows exponentially follows a path such as those shown for potential output in Figure 23.1 "A Century of Economic Growth" and Figure 23.2 "Cyclical Change Versus Growth". These curves become steeper over time because the growth rate is applied to an ever-larger base.

A variable growing at some exponential rate doubles over fixed intervals of time. The doubling time is given by the rule of 72A variable's approximate doubling time equals 72 divided by the growth rate, stated as a whole number., which states that a variable's approximate doubling time equals 72 divided by the growth rate, stated as a whole number. If the level of income were increasing at a 9% rate, for example, its doubling time would be roughly 72/9, or 8 years[1].

Let us apply this concept of a doubling time to the reduction in the U.S. growth rate. Had the U.S. economy continued to grow at a 3.5% rate after 1970, then its potential output would have doubled roughly every 20 years (72/3.5 = 20). That means potential output would have doubled by 1990, would double again by 2010, and would double again by 2030. Real GDP in 2030 would thus be eight times as great as its 1970 level. Growing at a 2.4% rate, however, potential output doubles only every 30 years (72/2.4 = 30). It would take until 2000 to double once from its 1970 level, and it would double once more by 2030. Potential output in 2030 would thus be four times its 1970 level if the economy grew at a 2.4% rate (versus eight times its 1970 level if it grew at a 3.5% rate). The 1.1% difference in growth rates produces a 100% difference in potential output by 2030. The different growth paths implied by these growth rates are illustrated in Figure 23.3 "Differences in Growth Rates".

Figure 23.3 Differences in Growth Rates



The chart suggests the significance in the long run of a small difference in the growth rate of real GDP. We begin in 1970, when real GDP equaled $2,873.9 billion. If real GDP grew at an annual rate of 3.5% from that year, it would double roughly every 20 years: in 1990, 2010, and 2030. Growth at a 2.4% rate, however, implies doubling every 30 years: in 2000 and 2030. By 2030, the 3.5% growth rate leaves real GDP at twice the level that would be achieved by 2.4% growth.

Growth in Output per Capita

Of course, it is not just how fast potential output grows that determines how fast the average person's material standard of living rises. For that purpose, we examine economic growth on a per capita basis. An economy's output per capitaReal GDP per person. equals real GDP per person. If we let N equal population, then

Equation 23.1

$$Output\ per\ capita = \frac{real\ GDP}{N}$$

In the United States in the third quarter of 2010, for example, real GDP was $13,277.4 billion (annual rate). The U.S. population was 311.0 million. Real U.S. output per capita thus equaled $42,693.

We use output per capita as a gauge of an economy's material standard of living. If the economy's population is growing, then output must rise as rapidly as the population if output per capita is to remain unchanged. If, for example, population increases by 2%, then real GDP would have to rise by 2% to maintain the current level of output per capita. If real GDP rises by less than 2%, output per capita will fall. If real GDP rises by more than 2%, output per capita will rise. More generally, we can write:

Equation 23.2

$$\%\ rate\ of\ growth\ of\ output\ per\ capita\ \cong\ \%\ rate\ of\ growth\ of\ output\ -\ \%\ rate\ of\ population$$

For economic growth to translate into a higher standard of living on average, economic growth must exceed population growth. From 1970 to 2004, for example, Sierra Leone's population grew at an annual rate of 2.1% per year, while its real GDP grew at an annual rate of 1.4%; its output per capita thus fell at a rate of 0.7% per year. Over the same period, Singapore's population grew at an annual rate of 2.1% per year, while its real GDP grew 7.4% per year. The resultant 5.3% annual growth in output per capita transformed Singapore from a relatively poor country to a country with the one of the highest per capita incomes in the world.

Key Takeaways

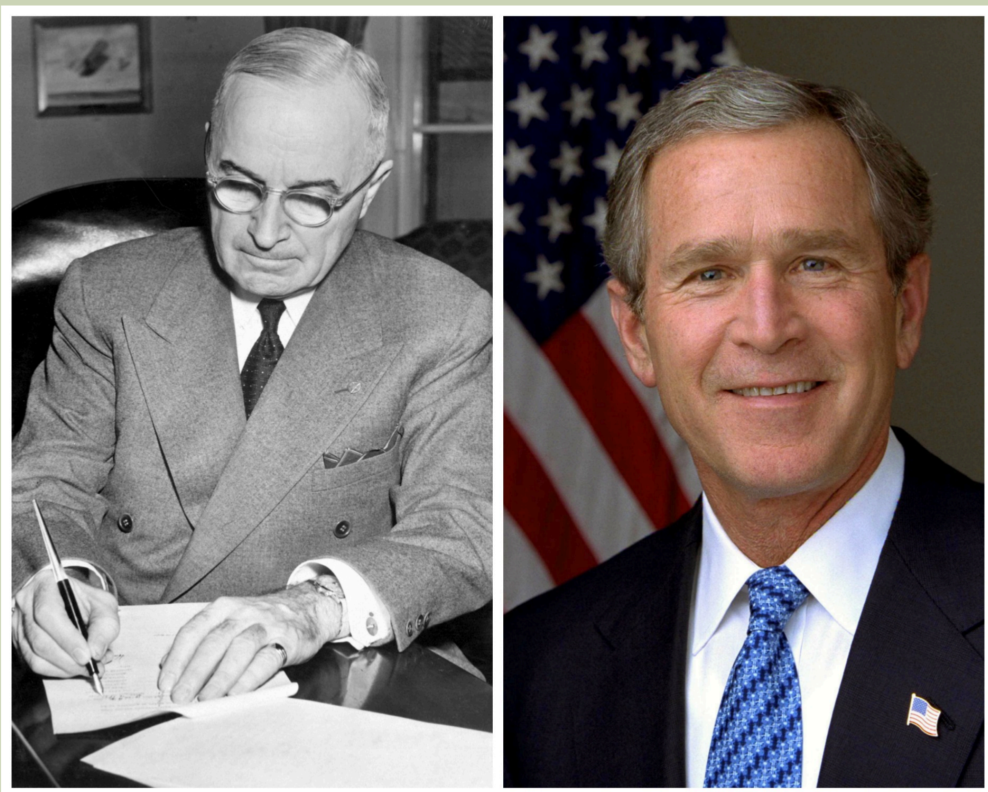
- Economic growth is the process through which an economy's production possibilities curve shifts outward. We measure it as the rate at which the economy's potential level of output increases.
- Measuring economic growth as the rate of increase of the actual level of real GDP can lead to misleading results due to the business cycle.
- Growth of a quantity at a particular percentage rate implies exponential growth. When something grows exponentially, it doubles over fixed intervals of time; these intervals may be computed using the rule of 72.
- Small differences in rates of economic growth can lead to large differences in levels of potential output over long periods of time.
- To assess changes in average standards of living, we subtract the percentage rate of growth of population from the percentage rate of growth of output to get the percentage rate of growth of output per capita.

Try It!

Suppose an economy's potential output and real GDP is $5 million in 2000 and its rate of economic growth is 3% per year. Also suppose that its population is 5,000 in 2000, and that its population grows at a rate of 1% per year. Compute GDP per capita in 2000. Now estimate GDP and GDP per capita in 2072, using the rule of 72. At what rate does GDP per capita grow? What is its doubling time? Is this result consistent with your findings for GDP per capita in 2000 and in 2072?

Case in Point: Presidents and Economic Growth

Figure 23.4



Wikimedia Commons – public domain; Wikimedia Commons – public domain.

President	Annual Increase in Real GDP (%)	Growth Rate (%)
Truman 1949–1952	5.4	4.4
Eisenhower 1953–1960	2.4	3.4
Kennedy-Johnson 1961–1968	5.1	4.3
Nixon-Ford 1969–1976	2.7	3.4
Carter 1977–1980	3.2	3.1
Reagan 1981–1988	3.5	3.1
G. H. W. Bush 1989–1992	2.4	2.7
Clinton 1992–2000	3.6	3.2
G. W. Bush 2001–2008 (Q3)	2.1	2.7

Presidents are often judged by the rate at which the economy grew while they were in office. This test is unfair on two counts. First, a president has little to do with the forces that determine growth. And second, such tests simply compute the annual rate of growth in real GDP over the course of a presidential term, which we know can be affected by cyclical factors. A president who takes office when the economy is down and goes out with the economy up will look like an economic star; a president with the bad luck to have reverse circumstances will seem like a dud. Here are annual rates of change in real GDP for each of the postwar presidents, together with rates of economic growth, measured as the annual rate of change in potential output.

The presidents' economic records are clearly affected by luck. Presidents Truman, Kennedy, Reagan, and Clinton, for example, began their terms when the economy had a recessionary gap and ended them with an inflationary gap or at about potential output. Real GDP thus rose faster than potential output during their presidencies. The Eisenhower, Nixon-Ford, H. W. Bush, and G. W. Bush administrations each started with an inflationary gap or at about potential and ended with a recessionary gap, thus recording rates of real GDP increase below the rate of gain in potential. Only Jimmy Carter, who came to office and left it with recessionary gaps, presided over a relatively equivalent rate of increase in actual GDP versus potential output.

Answer to Try It! Problem

GDP per capita in 2000 equals $1,000 ($5,000,000/5,000). If GDP rises 3% per year, it doubles every 24 years (= 72/3). Thus, GDP will be $10,000,000 in 2024, $20,000,000 in 2048, and $40,000,000 in 2072. Growing at a rate of 1% per year, population will have doubled once by 2072 to 10,000. GDP per capita will thus be $4,000 (= $40,000,000/10,000). Notice that GDP rises by eight times its original level, while the increase in GDP per capita is fourfold. The latter value represents a growth rate in output per capita of 2% per year, which implies a doubling time of 36 years. That gives two doublings in GDP per capita between 2000 and 2072 and confirms a fourfold increase.

[1]Notice the use of the words *roughly* and *approximately*. The actual value of an income of $1,000 growing at rate r for a period of n years is $1,000 × $(1 + r)^n$. After 8 years of growth at a 9% rate, income would thus be $1,000 $(1 + 0.09)^8$

= \$1,992.56. The rule of 72 predicts that its value will be \$2,000. The rule of 72 gives an approximation, not an exact measure, of the impact of exponential growth.

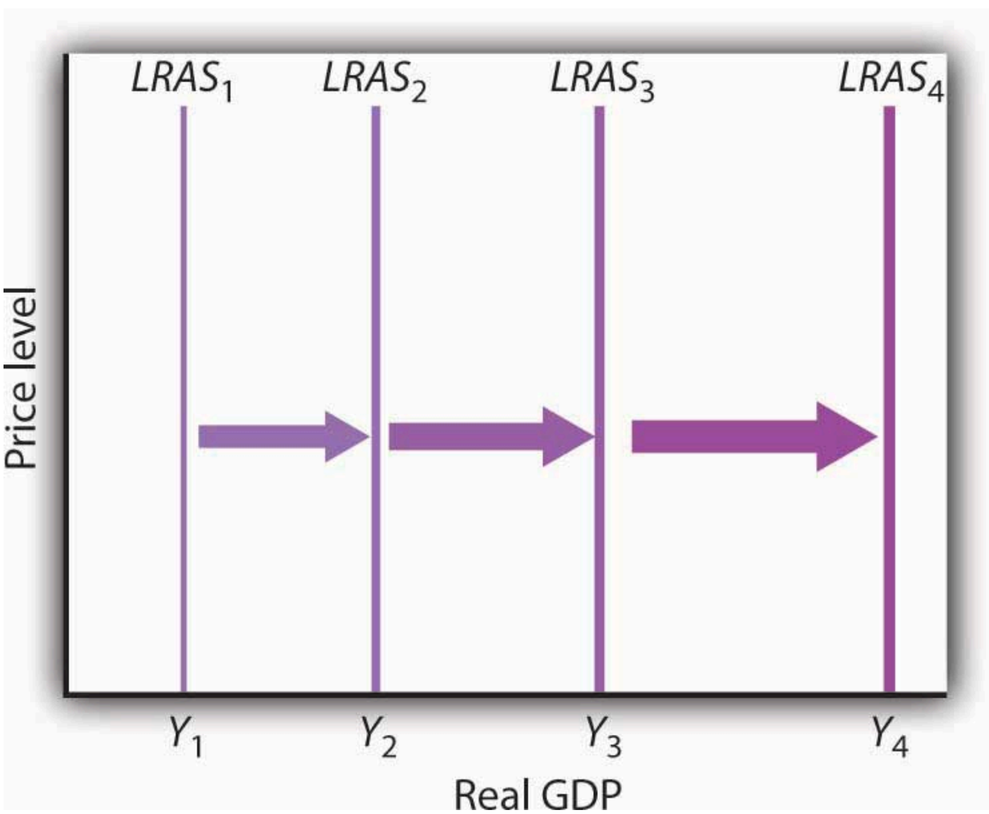
8.2 Growth and the Long-Run Aggregate Supply Curve

Learning Objectives

1. Explain and illustrate graphically the concept of the aggregate production function. Explain how its shape relates to the concept of diminishing marginal returns.
2. Derive the long-run aggregate supply curve from the model of the labor market and the aggregate production function.
3. Explain how the long-run aggregate supply curve shifts in responses to shifts in the aggregate production function or to shifts in the demand for or supply of labor.

Economic growth means the economy's potential output is rising. Because the long-run aggregate supply curve is a vertical line at the economy's potential, we can depict the process of economic growth as one in which the long-run aggregate supply curve shifts to the right.

Figure 23.5 Economic Growth and the Long-Run Aggregate Supply Curve



Because economic growth is the process through which the economy's potential output is increased, we can depict it as a series of rightward shifts in the long-run aggregate supply curve. Notice that with exponential growth, each successive shift in *LRAS* is larger and larger.

Figure 23.5 "Economic Growth and the Long-Run Aggregate Supply Curve" illustrates the process of economic growth. If the economy begins at potential output of Y_1, growth increases this potential. The figure shows a succession of increases in potential to Y_2, then Y_3, and Y_4. If the economy is growing at a particular percentage rate, and if the levels shown represent successive years, then the size of the increases will become larger and larger, as indicated in the figure.

Because economic growth can be considered as a process in which the long-run aggregate supply curve shifts to the right, and because output tends to remain close to this curve, it is important to gain a deeper understanding of what determines long-run aggregate supply (*LRAS*). We shall examine the derivation of *LRAS* and then see what factors shift the curve. We shall begin our work by defining an aggregate production function.

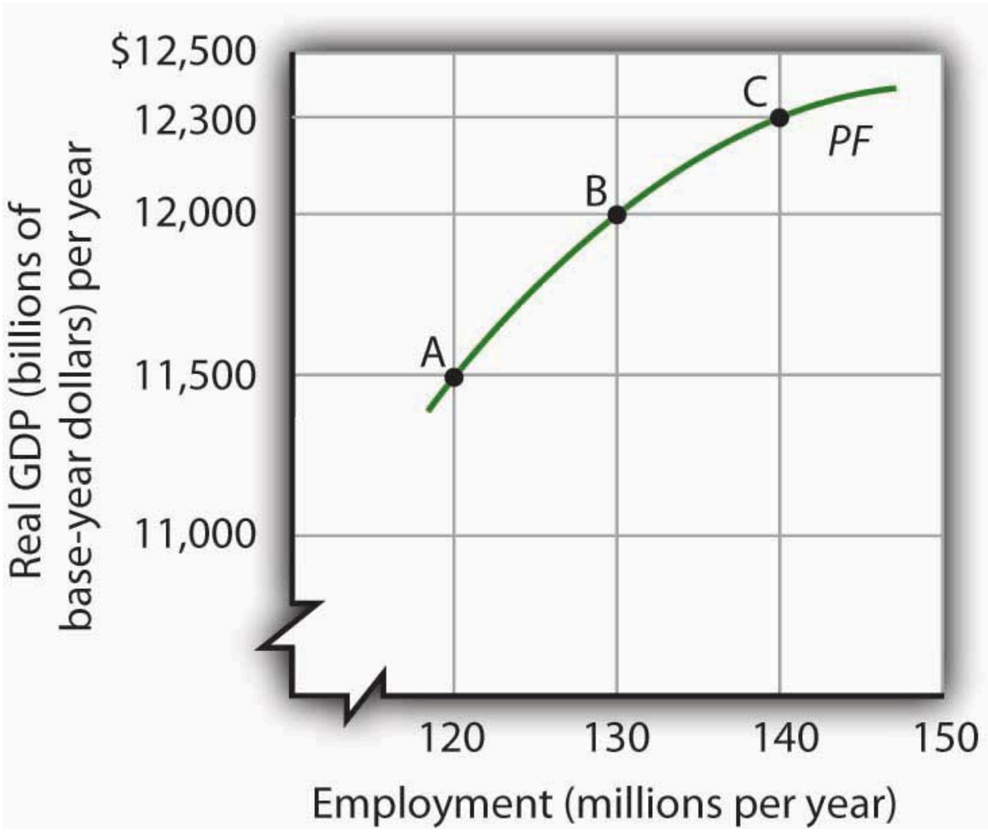
The Aggregate Production Function

An aggregate production functionFunction that relates the total output of an economy to the total amount of labor employed in the economy, all other determinants of production (capital, natural resources, and technology) being unchanged. relates the total output of an economy to the total amount of labor employed in the economy, all other determinants of production (that is, capital, natural resources, and technology) being unchanged. An economy operating on its aggregate production function is producing its potential level of output.

Figure 23.6 "The Aggregate Production Function" shows an aggregate production function (*PF*). It shows output levels for a range of employment between 120 million and 140 million workers. When the level of employment is 120 million, the economy produces a real GDP of $11,500 billion (point A). A level of employment of 130 million produces a real GDP of $12,000 billion (point B), and when 140 million workers are employed, a real GDP of $12,300 billion is produced (point C). In drawing the aggregate production function, the amount of labor varies, but everything else that could affect output, specifically the quantities of other factors of production and technology, is fixed.

The shape of the aggregate production function shows that as employment increases, output increases, but at a decreasing rate. Increasing employment from 120 million to 130 million, for example, increases output by $500 billion to $12,000 billion at point B. The next 10 million workers increase production by $300 billion to $12,300 billion at point C. This example illustrates diminishing marginal returns. Diminishing marginal returnsSituation that occurs when additional units of a variable factor add less and less to total output, given constant quantities of other factors. occur when additional units of a variable factor add less and less to total output, given constant quantities of other factors.

Figure 23.6 The Aggregate Production Function



An aggregate production function (*PF*) relates total output to total employment, assuming all other factors of production and technology are fixed. It shows that increases in employment lead to increases in output but at a decreasing rate.

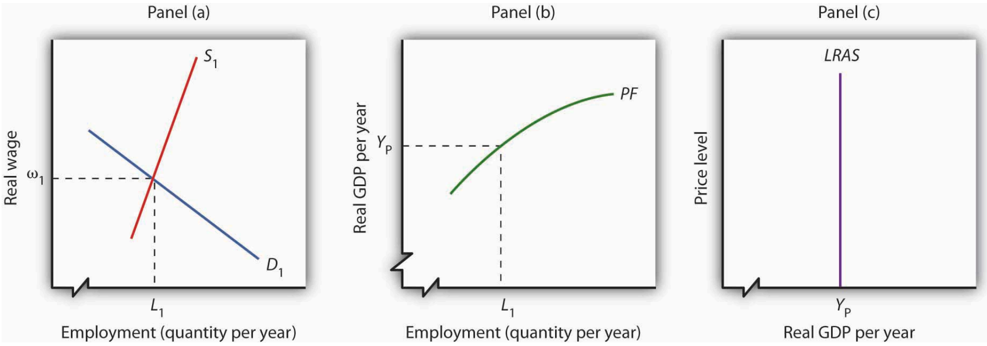
It is easy to picture the problem of diminishing marginal returns in the context of a single firm. The firm is able to increase output by adding workers. But because the firm's plant size and stock of equipment are fixed, the firm's capital per worker falls as it takes on more workers. Each additional worker adds less to output than the worker before. The firm, like the economy, experiences diminishing marginal returns.

The Aggregate Production Function, the Market for Labor, and Long-Run Aggregate Supply

To derive the long-run aggregate supply curve, we bring together the model of the labor market, introduced in the first macro chapter and the aggregate production function.

As we learned, the labor market is in equilibrium at the natural level of employment. The demand and supply curves for labor intersect at the real wage at which the economy achieves its natural level of employment. We see in Panel (a) of Figure 23.7 "Deriving the Long-Run Aggregate Supply Curve" that the equilibrium real wage is ω_1 and the natural level of employment is L_1. Panel (b) shows that with employment of L_1, the economy can produce a real GDP of Y_P. That output equals the economy's potential output. It is that level of potential output that determines the position of the long-run aggregate supply curve in Panel (c).

Figure 23.7 Deriving the Long-Run Aggregate Supply Curve



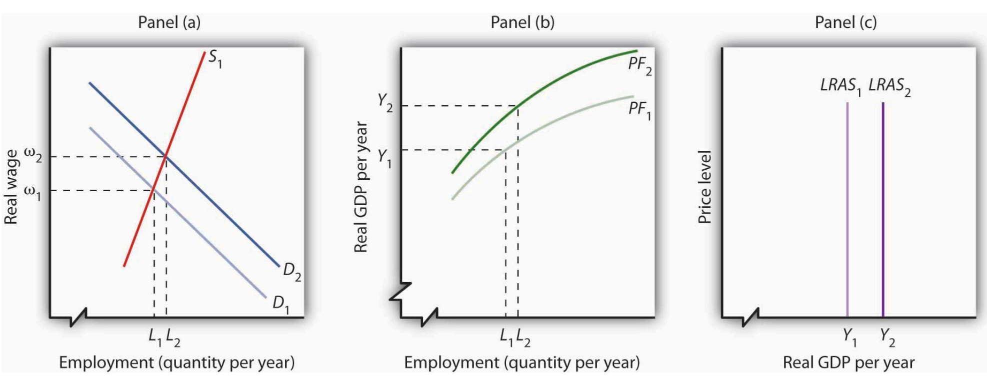
Panel (a) shows that the equilibrium real wage is ω_1, and the natural level of employment is L_1. Panel (b) shows that with employment of L_1, the economy can produce a real GDP of Y_P. That output equals the economy's potential output. It is at that level of potential output that we draw the long-run aggregate supply curve in Panel (c).

Changes in Long-Run Aggregate Supply

The position of the long-run aggregate supply curve is determined by the aggregate production function and the demand and supply curves for labor. A change in any of these will shift the long-run aggregate supply curve.

Figure 23.8 "Shift in the Aggregate Production Function and the Long-Run Aggregate Supply Curve" shows one possible shifter of long-run aggregate supply: a change in the production function. Suppose, for example, that an improvement in technology shifts the aggregate production function in Panel (b) from PF_1 to PF_2. Other developments that could produce an upward shift in the curve include an increase in the capital stock or in the availability of natural resources.

Figure 23.8 Shift in the Aggregate Production Function and the Long-Run Aggregate Supply Curve



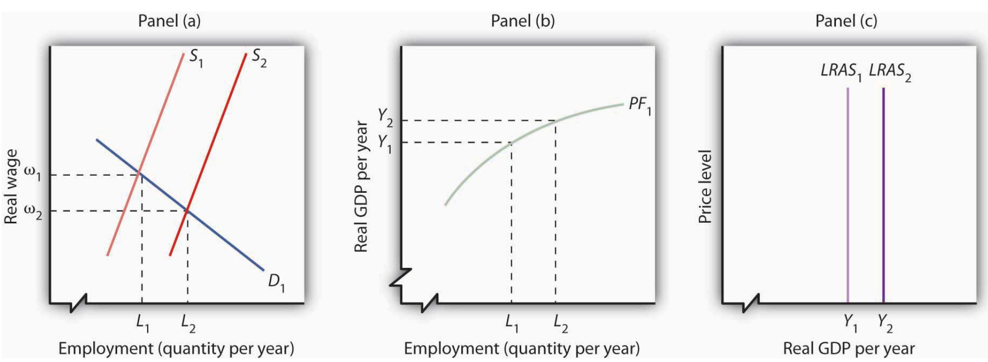
An improvement in technology shifts the aggregate production function upward in Panel (b). Because labor is more productive, the demand for labor shifts to the right in Panel (a), and the natural level of employment increases to L_2. In Panel (c) the long-run aggregate supply curve shifts to the right to Y_2.

The shift in the production function to PF_2 means that labor is now more productive than before. This will affect the demand for labor in Panel (a). Before the technological change, firms employed L_1 workers at a real wage ω_1. If workers are more productive, firms will find it profitable to hire more of them at ω_1. The demand curve for labor thus shifts to D_2 in Panel (a). The real wage rises to ω_2, and the natural level of employment rises to L_2. The increase in the real wage reflects labor's enhanced productivityThe amount of output per worker., the amount of output per worker. To see how potential output changes, we see in Panel (b) how much output can be produced given the new natural level of employment and the new aggregate production function. The real GDP that the economy is capable of producing rises from Y_1 to Y_2. The higher output is a reflection of a higher natural level of employment, along with the fact that labor has become more productive as a result of the technological advance. In Panel (c) the long-run aggregate supply curve shifts to the right to the vertical line at Y_2.

This analysis dispels a common misconception about the impact of improvements in technology or increases in the capital stock on employment. Some people believe that technological gains or increases in the stock of capital reduce the demand for labor, reduce employment, and reduce real wages. Certainly the experience of the United States and most other countries belies that notion. Between 1990 and 2007, for example, the U.S. capital stock and the level of technology increased dramatically. During the same period, employment and real wages rose, suggesting that the demand for labor increased by more than the supply of labor. As some firms add capital or incorporate new technologies, some workers at those firms may lose their jobs. But for the economy as a whole, new jobs become available *and* they generally offer higher wages. The demand for labor rises.

Another event that can shift the long-run aggregate supply curve is an increase in the supply of labor, as shown in Figure 23.9 "Increase in the Supply of Labor and the Long-Run Aggregate Supply Curve". An increased supply of labor could result from immigration, an increase in the population, or increased participation in the labor force by the adult population. Increased participation by women in the labor force, for example, has tended to increase the supply curve for labor during the past several decades.

Figure 23.9 Increase in the Supply of Labor and the Long-Run Aggregate Supply Curve



An increase in the supply of labor shifts the supply curve in Panel (a) to S_2, and the natural level of employment rises to L_2. The real wage falls to ω_2. With increased labor, the aggregate production function in Panel (b) shows that the economy is now capable of producing real GDP at Y_2. The long-run aggregate supply curve in Panel (c) shifts to $LRAS_2$.

In Panel (a), an increase in the labor supply shifts the supply curve to S_2. The increase in the supply of labor does not change the stock of capital or natural resources, nor does it change technology—it therefore does not shift the aggregate production function. Because there is no change in the production function, there is no shift in the demand for labor.

The real wage falls from ω_1 to ω_2 in Panel (a), and the natural level of employment rises from L_1 to L_2. To see the impact on potential output, Panel (b) shows that employment of L_2 can produce real GDP of Y_2. The long-run aggregate supply curve in Panel (c) thus shifts to $LRAS_2$. Notice, however, that this shift in the long-run aggregate supply curve to the right is associated with a reduction in the real wage to ω_2.

Of course, the aggregate production function and the supply curve of labor can shift together, producing higher real wages at the same time population rises. That has been the experience of most industrialized nations. The increase in real wages in the United States between 1990 and 2007, for example, came during a period in which an increasing population increased the supply of labor. The demand for labor increased by more than the supply, pushing the real wage up. The accompanying Case in Point looks at gains in real wages in the face of technological change, an increase in the stock of capital, and rapid population growth in the United States during the 19th century.

Our model of long-run aggregate supply tells us that in the long run, real GDP, the natural level of employment, and the real wage are determined by the economy's production function and by the demand and supply curves for labor. Unless an event shifts the aggregate production function, the demand curve for labor, or the supply curve for labor, it affects neither the natural level of employment nor potential output. Economic growth occurs only if an event shifts the economy's production function or if there is an increase in the demand for or the supply of labor.

Key Takeaways

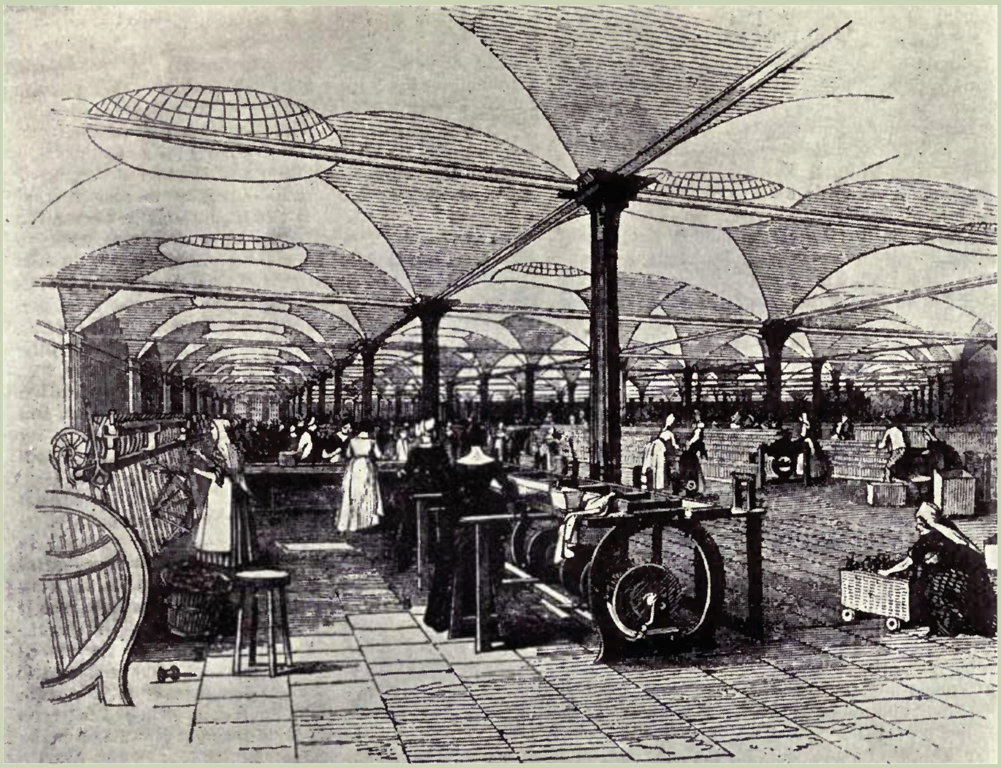
- The aggregate production function relates the level of employment to the level of real GDP produced per period.
- The real wage and the natural level of employment are determined by the intersection of the demand and supply curves for labor. Potential output is given by the point on the aggregate production function corresponding to the natural level of employment. This output level is the same as that shown by the long-run aggregate supply curve.
- Economic growth can be shown as a series of shifts to the right in *LRAS*. Such shifts require either upward shifts in the production function or increases in demand for or supply of labor.

Try It!

Suppose that the quantity of labor supplied is 50 million workers when the real wage is $20,000 per year and that potential output is $2,000 billion per year. Draw a three-panel graph similar to the one presented in Figure 23.9 "Increase in the Supply of Labor and the Long-Run Aggregate Supply Curve" to show the economy's long-run equilibrium. Panel (a) of your graph should show the demand and supply curves for labor, Panel (b) should show the aggregate production function, and Panel (c) should show the long-run aggregate supply curve. Now suppose a technological change increases the economy's output with the same quantity of labor as before to $2,200 billion, and the real wage rises to $21,500. In response, the quantity of labor supplied increases to 51 million workers. In the same three panels you have already drawn, sketch the new curves that result from this change. Explain what happens to the level of employment, the level of potential output, and the long-run aggregate supply curve. (Hint: you have information for only one point on each of the curves you draw—two for the supply of labor; simply draw curves of the appropriate shape. Do not worry about getting the scale correct.)

Case in Point: Technological Change, Employment, and Real Wages During the Industrial Revolution

Figure 23.10



Wikimedia Commons – public domain.

Technological change and the capital investment that typically comes with it are often criticized because they replace labor with machines, reducing employment. Such changes, critics argue, hurt workers. Using the model of aggregate demand and aggregate supply, however, we arrive at a quite different conclusion. The model predicts that improved technology will increase the demand for labor and boost real wages.

The period of industrialization, generally taken to be the time between the Civil War and World War I, was a good test of these competing ideas. Technological changes were dramatic as firms shifted toward mass production and automation. Capital investment soared. Immigration increased the supply of labor. What happened to workers?

Employment more than doubled during this period, consistent with the prediction of our model. It is harder to predict, from a theoretical point of view, the consequences for real wages. The latter third of the 19th century was a period of massive immigration to the United States. Between 1865 and 1880, more than 5 million people came to the United States from abroad; most were of working age. The pace accelerated between 1880 and 1923, when more than 23 million people moved to the United States from other countries. Immigration increased the supply of labor, which should reduce the real wage. There were thus two competing forces at work: Technological change and capital investment tended to increase real wages, while immigration tended to reduce them by increasing the supply of labor.

The evidence suggests that the forces of technological change and capital investment proved far more powerful than increases in labor supply. Real wages soared 60% between 1860 and 1890. They continued to increase after that. Real wages in manufacturing, for example, rose 37% from 1890 to 1914.

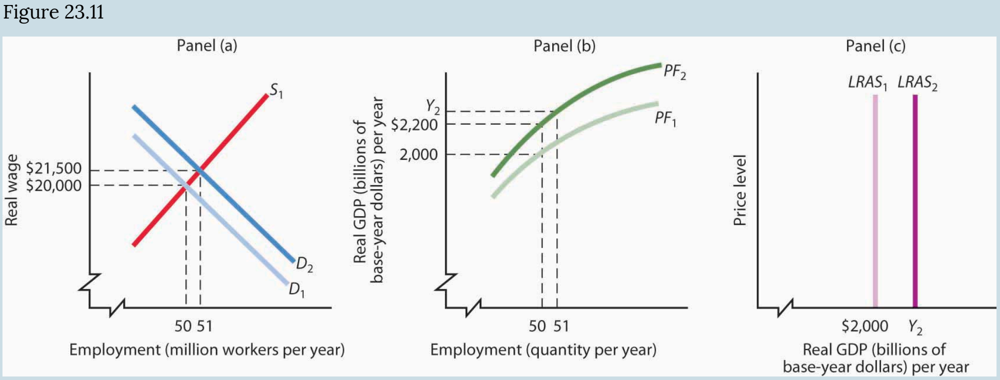
Technological change and capital investment displace workers in some industries. But for the economy as a whole, they increase worker productivity, increase the demand for labor, and increase real wages.

Sources: Wage data taken from Clarence D. Long, *Wages and Earnings in the United States*, 1860–1990 (Princeton, NJ: Princeton University Press, 1960), p. 109, and from Albert Rees, *Wages in Manufacturing*, 1890–1914 (Princeton, NJ: Princeton University Press, 1961), pp. 3–5. Immigration figures taken from Gary M. Walton and Hugh Rockoff, *History of the American Economy*, 6th ed. (New York: Harcourt Brace Jovanovich, 1990), p. 371.

Answer to Try It! Problem

The production function in Panel (b) shifts up to PF_2. Because it reflects greater productivity of labor, firms will increase their demand for labor, and the demand curve for labor shifts to D_2 in Panel (a). $LRAS_1$ shifts to $LRAS_2$ in Panel (c). Employment and potential output rise. Potential output will be greater than $2,200 billion.

Figure 23.11



8.3 Determinants of Economic Growth

Learning Objectives

1. Discuss the sources of economic growth.
2. Discuss possible reasons why countries grow at different rates.

In this section, we review the main determinants of economic growth. We also examine the reasons for the widening disparities in economic growth rates among countries in recent years.

The Sources of Economic Growth

As we have learned, there are two ways to model economic growth: (1) as an outward shift in an economy's production possibilities curve, and (2) as a shift to the right in its long-run aggregate supply curve. In drawing either one at a point in time, we assume that the economy's factors of production and its technology are unchanged. Changing these will shift both curves. Therefore, anything that increases the quantity or quality of factors of production or that improves the technology available to the economy contributes to economic growth.

The sources of growth for the U.S. economy in the 20th century were presented in the chapter on sources of production. There we learned that the main sources of growth for the United States from 1948 to 2002 were divided between increases in the quantities of labor and of physical capital (about 60%) and in improvements in the qualities of the factors of production and technology (about 40%). Since 1995, however, improvements in factor quality and technology have been the main drivers of economic growth in the United States.

In order to devote resources to increasing physical and human capital and to improving technology—activities that will enhance future production—society must forgo using them now to produce consumer goods. Even though the people in the economy would enjoy a higher standard of living today without this sacrifice, they are willing to reduce present consumption in order to have more goods and services available for the future.

As a college student, you personally made such a choice. You decided to devote time to study that you could have spent earning income. With the higher income, you could enjoy greater consumption today. You made this choice because you expect to earn higher income in the future and thus to enjoy greater consumption in the future. Because many other people in the society also choose to acquire more education, society allocates resources to produce education. The education produced today will enhance the society's human capital and thus its economic growth.

All other things equal, higher saving allows more resources to be devoted to increases in physical and human capital and technological improvement. In other words, saving, which is income not spent on consumption, promotes economic growth by making available resources that can be channeled into growth-enhancing uses.

Explaining Recent Disparities in Growth Rates

Toward the end of the 20th century, it appeared that some of the world's more affluent countries were growing robustly while others were growing more slowly or even stagnating. This observation was confirmed in a major study by the Organization for Economic Co-operation and Development (OECD)[1], whose members are listed in Table 23.1 "Growing Disparities in Rates of Economic Growth". The table shows that for the OECD countries as a whole, economic growth per capita fell from an average of 2.2% per year in the 1980s to an average of 1.9% per year in the 1990s. The higher standard deviation in the latter period confirms an increased disparity of growth rates in the more recent period. Moreover, the data on individual countries show that per capita growth in some countries (specifically, the United States, Canada, Ireland, Netherlands, Norway, and Spain) picked up, especially in the latter half of the 1990s, while it decelerated in most of the countries of continental Europe and Japan.

Table 23.1 Growing Disparities in Rates of Economic Growth

Trend Growth of GDP per Capita			
Country	**1980–1990**	**1990–2000**	**1996–2000**
United States	2.1	2.3	2.8
Japan	3.3	1.4	0.9
Germany	1.9	1.2	1.7
France	1.6	1.5	1.9
Italy	2.3	1.5	1.7
United Kingdom	2.2	2.1	2.3
Canada	1.4	1.7	2.6
Austria	2.1	1.9	2.3
Belgium	2.0	1.9	2.3
Denmark	1.9	1.9	2.3
Finland	2.2	2.1	3.9
Greece	0.5	1.8	2.7
Iceland	1.7	1.5	2.6
Ireland	3.0	6.4	7.9
Luxembourg	4.0	4.5	4.6
Netherlands	1.6	2.4	2.7
Portugal	3.1	2.8	2.7
Spain	2.3	2.7	3.2
Sweden	1.7	1.5	2.6
Switzerland	1.4	0.4	1.1
Turkey	2.1	2.1	1.9
Australia	1.6	2.4	2.8
New Zealand	1.4	1.2	1.8
Mexico	0.0	1.6	2.7
Korea	7.2	5.1	4.2
Hungary	–	2.3	3.5
Poland	–	4.2	4.8
Czech Republic	–	1.7	1.4
OECD24[2]	2.2	1.9	2.2
Standard Deviation of OECD24	0.74	1.17	1.37

Variation in the growth in real GDP per capita has widened among the world's leading industrialized economies.

Source: Excerpted from Table 1.1 Organization for Economic Co-operation and Development, *Sources of Economic Growth in OECD Countries*, 2003: p. 32–33.

The study goes on to try to explain the reasons for the divergent growth trends. The main findings were:

- In general, countries with accelerating per capita growth rates also experienced significant increases in employment, while those with stagnant or declining employment generally experienced reductions in per capita growth rates.
- Enhancements in human capital contributed to labor productivity and economic growth, but in slower growing

countries such improvements were not enough to offset the impact of reduced or stagnant labor utilization.
- Information and communication technology has contributed to economic growth both through rapid technological progress within the information and communication technology industry itself as well as, more recently, through the use of information and communication technology equipment in other industries. This has made an important contribution to growth in several of the faster growing countries.
- Other factors associated with more growth include: investments in physical and human capital, sound macroeconomic policies (especially low inflation), private sector research and development, trade exposure, and better developed financial markets. Results concerning the impact of the size of the government and of public sector research and development on growth were more difficult to interpret.
- With qualifications, the study found that strict regulation of product markets (for example, regulations that reduce competition) and strict employment protection legislation (for example, laws that make hiring and firing of workers more difficult) had negative effects on growth.
- All countries show a large number of firms entering and exiting markets. But, a key difference between the United States and Europe is that new firms in the United States start out smaller and less productive than those of Europe but grow faster when they are successful. The report hypothesizes that lower start-up costs and less strict labor market regulations may encourage U.S. entrepreneurs to enter a market and then to expand, if warranted. European entrepreneurs may be less willing to experiment in a market in the first place.

The general concern in the second half of the 1970s and the 1980s was that economic growth was slowing down and that it might not be possible to reverse this pattern. The 1990s and early 2000s, in which growth picked up in some countries but not in others, suggested that the problem was not universal and led to a search for the reasons for the disparities in growth rates that emerged. The OECD study described above gives some possible explanations. The findings of that study practically beg countries to examine closely their economic policies at a variety of levels and to consider changes that may add flexibility to their economies.

In closing, it is worth reiterating that economic freedom and higher incomes tend to go together. Countries could not have attained high levels of income if they had not maintained the economic freedom that contributed to high incomes in the first place. Thus, it is also likely that rates of economic growth in the future will be related to the amount of economic freedom countries choose. We shall see in later chapters that monetary and fiscal policies that are used to stabilize the economy in the short run can also have an impact on long-run economic growth.

Key Takeaways

- The main sources of growth for the United States from 1948 to 2002 were divided between increases in the quantities of labor and of physical capital (about 60%) and in improvements in the qualities of the factors of production and technology (about 40%). Since 1995, however, improvements in factor quality and technology have been the main drivers of economic growth in the United States.
- There has been a growing disparity in the rates of economic growth in industrialized countries in the last decade, which may reflect various differences in economic structures and policies.

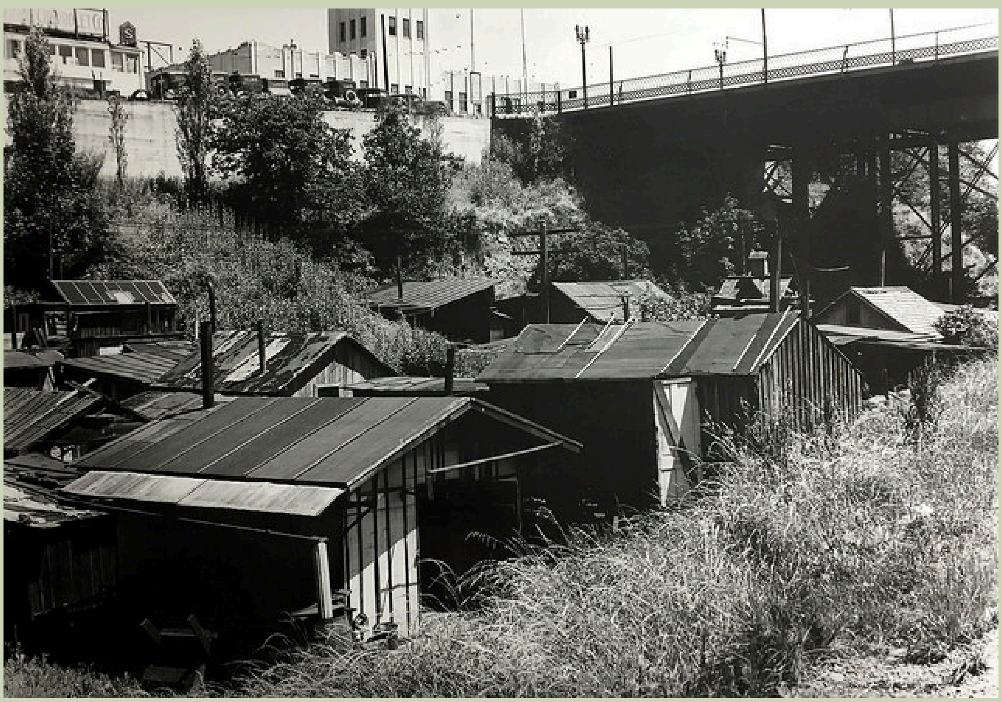
Try It!

All other things unchanged, compare the position of a country's expected production possibility curve and the expected position of its long-run aggregate supply curve if:

1. Its labor force increases in size by 3% per year compared to 2% per year.
2. Its saving rate falls from 15% to 10%.
3. It passes a law making it more difficult to fire workers.
4. Its level of education rises more quickly than it has in the past.

Case in Point: Economic Growth in Poor Countries … or Lack Thereof

Figure 23.12



Alan – [A Home in Sullivan's Gulch (A Portland Hooverville)] –

Economist William Easterly in his aptly named book *The Elusive Quest for Growth: Economists' Adventures and*

Misadventures in the Tropics admits that after 50 years of searching for the magic formula for turning poor countries into rich ones, the quest remains elusive.

Poor countries just need more physical capital, you say? Easterly points out that between 1960 and 1985, the capital stock per work in both Gambia and Japan rose by over 500%. The result? In Gambia, output per worker over the 25-year period rose 2%; in Japan, output per worker rose 260%.

So, it must be that poor countries need more human capital? Again, he finds startling comparisons. For example, human capital expanded faster in Zambia than in Korea, but Zambia's annual growth rate is 7 percentage points below Korea's.

Too much population growth? Too little? More foreign aid? Too much? As Easterly proceeds, writing a prescription for growth seems ever more difficult: "None has delivered as promised," he concludes (p. xi).

While Easterly does not offer his own new panacea for how to move countries to a higher level of per capita GDP, where the model presented in this chapter does seem to provide some explanations of why a country's growth rate may vary over time or differ from another country's, he does argue that creating incentives for growth in poor countries is crucial. Acknowledging a role for plain luck, Easterly argues that good government policies—ones that keep low such negatives as inflation, corruption, and red tape—and quality institutions and laws—ones that, for example, honor contracts and reward merit—will help.

How to actually improve such incentives might constitute the next great quest:

"We have learned once and for all that there are no magical elixirs to bring a happy ending to our quest for growth. Prosperity happens when all the players in the development game have the right incentives. It happens when government incentives induce technological adaptation, high-quality investment in machines, and high-quality schooling. It happens when donors face incentives that induce them to give aid to countries with good policies where aid will have high payoffs, not to countries with poor policies where aid is wasted. It happens when the poor get good opportunities and incentives, which requires government welfare programs that reward rather than penalize earning income. It happens when politics is not polarized between antagonistic interest groups. . . . The solutions are a lot more difficult to describe than the problems. The way forward must be to create incentives for growth for the trinity of governments, donors, and individuals." (p. 289–90)

Source: William Easterly, *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics* (Cambridge: MIT Press, 2002).

Answer to Try It! Problem

Situations 1 and 4 should lead to a shift further outward in the country's production possibility curve and further to the right in its long-run aggregate supply curve. Situations 2 and 3 should lead to smaller outward shifts in the country's production possibility curve and smaller rightward shifts in its long-run aggregate supply curve.

[1]The material in this section is based on Organization for Economic Co-operation and Development, *The Sources of Economic Growth in OECD Countries*, 2003.

[2]Excludes Czech Republic, Hungary, Korean, Mexico, Poland, and Slovak Republic

8.4 Review and Practice

Summary

We saw that economic growth can be measured by the rate of increase in potential output. Measuring the rate of increase in actual real GDP can confuse growth statistics by introducing elements of cyclical variation.

Growth is an exponential process. A variable increasing at a fixed percentage rate doubles over fixed intervals. The doubling time is approximated by the rule of 72. The exponential nature of growth means that small differences in growth rates have large effects over long periods of time. Per capita rates of increase in real GDP are found by subtracting the growth rate of the population from the growth rate of GDP.

Growth can be shown in the model of aggregate demand and aggregate supply as a series of rightward shifts in the long-run aggregate supply curve. The position of the *LRAS* is determined by the aggregate production function and by the demand and supply curves for labor. A rightward shift in *LRAS* results either from an upward shift in the production function, due to increases in factors of production other than labor or to improvements in technology, or from an increase in the demand for or the supply of labor.

Saving plays an important role in economic growth, because it allows for more capital to be available for future production, so the rate of economic growth can rise. Saving thus promotes growth.

In recent years, rates of growth among the world's industrialized countries have grown more disparate. Recent research suggests this may be related to differing labor and product market conditions, differences in the diffusion of information and communications technologies, as well as differences in macroeconomic and trade policies. Evidence on the role that government plays in economic growth was less conclusive.

Concept Problems

1. Suppose the people in a certain economy decide to stop saving and instead use all their income for consumption. They do nothing to add to their stock of human or physical capital. Discuss the prospects for growth of such an economy.
2. Singapore has a saving rate that is roughly three times greater than that of the United States. Its greater saving rate has been one reason why the Singapore economy has grown faster than the U.S. economy. Suppose that if the United States increased its saving rate to, say, twice the Singapore level, U.S. growth would surpass the Singapore rate. Would that be a good idea?
3. Suppose an increase in air pollution causes capital to wear out more rapidly, doubling the rate of depreciation. How would this affect economic growth?
4. Some people worry that increases in the capital stock will bring about an economy in which everything is done by machines, with no jobs left for people. What does the model of economic growth presented in this chapter predict?

5. China's annual rate of population growth was 1.2% from 1975 to 2003 and is expected to be 0.6% from 2003 through 2015. How do you think this will affect the rate of increase in real GDP? How will this affect the rate of increase in per capita real GDP?
6. Suppose technology stops changing. Explain the impact on economic growth.
7. Suppose a series of terrorist attacks destroys half the capital in the United States but does not affect the population. What will happen to potential output and to the real wage?
8. "Given the rate at which scientists are making new discoveries, we will soon reach the point that no further discoveries can be made. Economic growth will come to a stop." Discuss.
9. Suppose real GDP increases during President Obama's term in office at a 5% rate. Would that imply that his policies were successful in "growing the economy"?
10. Suppose that for some country it was found that its economic growth was based almost entirely on increases in quantities of factors of production. Why might such growth be difficult to sustain?

Numerical Problems

1. The population of the world in 2003 was 6.314 billion. It grew between 1975 and 2003 at an annual rate of 1.6%. Assume that it continues to grow at this rate.
   1. Compute the doubling time.
   2. Estimate the world population in 2048 and 2093 (assuming all other things remain unchanged).
2. With a world population in 2003 of 6.314 billion and a projected population growth rate of 1.1% instead (which is the United Nations' projection for the period 2003 to 2015).
   1. Compute the doubling time.
   2. State the year in which the world's population would be 12.628 billion.
3. Suppose a country's population grows at the rate of 2% per year and its output grows at the rate of 3% per year.
   1. Calculate its rate of growth of per capita output.
   2. If instead its population grows at 3% per year and its output grows at 2% per year, calculate its rate of growth of per capita output.
4. The rate of economic growth per capita in France from 1996 to 2000 was 1.9% per year, while in Korea over the same period it was 4.2%. Per capita real GDP was $28,900 in France in 2003, and $12,700 in Korea. Assume the growth rates for each country remain the same.
   1. Compute the doubling time for France's per capita real GDP.
   2. Compute the doubling time for Korea's per capita real GDP.
   3. What will France's per capita real GDP be in 2045?
   4. What will Korea's per capita real GDP be in 2045?

5. Suppose real GDPs in country A and country B are identical at $10 trillion dollars in 2005. Suppose country A's economic growth rate is 2% and country B's is 4% and both growth rates remain constant over time.

    1. On a graph, show country A's potential output until 2025.
    2. On the same graph, show country B's potential output.
    3. Calculate the percentage difference in their levels of potential output in 2025.

    Suppose country A's population grows 1% per year and country B's population grows 3% per year.

    1. On a graph, show country A's potential output per capita in 2025.
    2. On the same graph, show country B's potential output per capita in 2025.
    3. Calculate the percentage difference in their levels of potential output per capita in 2025.

6. Two countries, A and B, have identical levels of real GDP per capita. In Country A, an increase in the capital stock increases the potential output by 10%. Country B also experiences a 10% increase in its potential output, but this increase is the result of an increase in its labor force. Using aggregate production functions and labor-market analyses for the two countries, illustrate and explain how these events are likely to affect living standards in the two economies.

7. Suppose the information below characterizes an economy:

Employment (in millions)	**Real GDP (in billions)**
1	200
2	700
3	1,100
4	1,400
5	1,650
6	1,850
7	2,000
8	2,100
9	2,170
10	2,200

    1. Construct the aggregate production function for this economy.
    2. What kind of returns does this economy experience? How do you know?
    3. Assuming that total available employment is 7 million, draw the economy's long-run aggregate supply curve.

    Suppose that improvement in technology means that real GDP at each level of employment rises by $200 billion.

    1. Construct the new aggregate production function for this economy.
    2. Construct the new long-run aggregate supply curve for the economy.

8. In Table 23.1 "Growing Disparities in Rates of Economic Growth", we can see that Japan's growth rate of per capita real GDP fell from 3.3% per year in the 1980s to 1.4% per year in the 1990s.

    1. Compare the percent increase in its per capita real GDP over the 20-year period to what it would have been if it had maintained the 3.3% per capita growth rate of the 1980s.
    2. Japan's per capita GDP in 1980 was about $24,000 (in U.S. 2000 dollars) in 1980. Calculate what it would have been if the growth rate of the 1980s had been maintained Calculate about how much it is, given the actual growth rates over the two decades.
    3. In Table 23.1 "Growing Disparities in Rates of Economic Growth", we can see that Ireland's growth rate of per capita real GDP grew from 3.0% per year in the 1980s to 6.4% per year in the 1990s.
    4. Compare the percent increase in its per capita real GDP over the 20-year period to what it would have been if it had maintained the 3.0% per capita growth rate of the 1980s.
    5. Ireland's per capita GDP in 1980 was about $10,000 (in U.S. 2000 dollars). Calculate what it would have been if the growth rate of the 1980s had been maintained. Calculate about how much it is, given the actual growth rates over the two decades.

CHAPTER 9: THE NATURE AND CREATION OF MONEY

You can see the variety of things that have been used as money in Canada in James Powell's *A History of the Canadian Dollar*, available at the Bank of Canada's website: http://goo.gl/BnIgpK. Our money in the seventeenth and eighteenth centuries was silver and gold coins from many countries, and playing cards. The British pound sterling, the Spanish silver dollar, and the US dollar were the main moneys in Canada in the nineteenth century, followed by paper currencies issued by banks and by the government since the late nineteenth century.

It is not the commodity or token used as money that matters, but the social convention that it is accepted without question as a means of payment. Money makes it easier for everyone to buy and sell goods and services and economize on the use of scarce resources.

Money is defined by four important functions. It provides:

1. A means of payment as a medium of exchange
2. A unit of account
3. A store of value
4. A standard of deferred payments

As a means of payment money is involved in most exchanges. We use money to pay for the goods and services – from food and clothing to transportation, to rent, to fees and taxes. People receive wages and salaries and other types of income in money. Money is not consumed in these transactions. It is used as a medium of exchange.

Means of payment: a commodity or token generally accepted in payment for goods and services or the repayment of debt.

Exchange transactions without money are barter exchanges, a direct exchange of one good for another. These exchanges depend on a *double coincidence of wants*. Each party to the exchange has a good the other party wants and is willing to trade one good for another. This means exchange transactions are expensive as people must find others who have what they want and want what they have. Using a money as a medium of exchange dramatically lowers the cost and increases the ease and efficiency of trade.

Barter exchanges: direct exchanges of goods or services for goods or services without the use of money.

Money also serves as a unit of account. Prices in Canada are quoted in Canadian dollars. Similarly in other countries prices are quoted in domestic currency. In much of Europe prices are in euros, in the United States in US dollars and in Japan in yen. This reflects the convenience of using the same units for the means of payment and the unit of account. However, there are exceptions. Historically, in Canada, during the time of the fur trade, books were kept in "currency" but actual currency never changed hands in the barter of trade goods for furs.

Unit of account: the standard in which prices are quoted and accounts are kept.

To serve as a medium of exchange, money *must also be* a store of value. Money works as a time machine allowing people to separate the timing of their expenditures from the timing of their incomes. Money is accepted today with confidence that it can be used some time in the future to make payments when buying goods and services. You would not accept money today that you thought would be unacceptable when you offered it in payment at some later date.

Store of value: an asset that carries purchasing power forward in time for future purchases.

Money is not a unique store of value. Other assets including real estate, financial assets like corporate and government bonds, fine art and antiques all serve as stores of value. These may be better ways to store value, but people still choose to hold some of their wealth as money. This choice to hold money balances is very important to the effects money balances have on financial markets and aggregate expenditure. Chapter 9 examines it in detail.

Money provides a standard for deferred payments. If you take out a student loan the amounts you will repay in the future are measured in dollars. Similarly, servicing and retiring a mortgage on a property or a loan on a car calls for future payments specified in dollars. Domestic money is not essential for this function. Individuals, businesses and governments often borrow or lend in the money of other countries. In those cases the currency in which the loan transaction takes place is usually the standard for payments to settle the debt. The essential attribute of money is its general acceptance as a means of payment. For this money must also be a store of value. This works well when money is also a unit of account and a standard of deferred payments.

Standard of deferred payments: the units in which future financial obligations are measured.

The development of money

The money we use today is the product of a long and continuing evolution in the financial services industry. It is a testament to the ingenuity of people and society seeking to reduce the costs and increase the volume of trade in goods and services.

Historically, there were no banks. Money was a commodity. Gold and silver bullion are two commodities that came to be used extensively because of their relative scarcity and durability. Concerns about the purity of these metals and the inconvenience of weighing out small quantities to make payments led to coinage. The minting of gold and silver coins by heads of state offered a solution to these problems. The 'monarch' certified the purity and quantity of the metal in the coin by having his or her likeness stamped into the metal.

Unfortunately, coinage did not completely solve the concerns about the quantity and quality of gold and silver money. The quantity of gold in a coin could be reduced by clipping its edges, or by rubbing the surfaces of the coin to wear some of the metal away. "Sweating" coins by placing them in a leather bag and shaking them was one technique used to remove some of their precious metal. The edge designs, millings, or facets that we still see on coins today were introduced to combat clipping, and wear on the heads and tails stamped into coins provided evidence of sweating. Coins that were worn or clipped were not accepted at full value in exchange for goods and services.

A second difficulty with precious metal coinage came from the sovereign who controlled the minting process. Adding a little base metal to the gold or silver being minted resulted in coins with less precious metal content than their face value certified. A little of the precious metal was withheld and could be used to mint more coin, which was, in effect, free money for the sovereign. This "debasing" of the coinage was a serious problem at times and, like clipping and sweating, reduced the acceptability of precious metal coinage as money.

The emergence of banks and paper money was a response to the problems with gold and silver commodity money. The first banks were goldsmiths who used gold in the production of jewelry and ornaments. They were accustomed to buying and selling gold bullion, and holding stocks of gold bullion. It was a natural extension of their business to offer to hold gold for safekeeping. Those who stored their gold with goldsmiths were given paper claims or receipts (IOUs), which were convertible back into gold on demand.

When people began to use gold receipts to make payments, gold receipts became a means of payment. They were token money, representing a fixed amount of the precious metal.

Token money: convertible claims on commodity money.

Goldsmiths became bankers when they realized that not all their customers would show up at the same time and want their gold back. The convertibility of gold receipts made them acceptable as a medium of exchange. Gold merchants could make loans by issuing more gold receipts than they had gold in their storage vaults. They only needed gold holdings equal to a fraction of the gold receipts they had issued, as long as people used the receipts as a medium of exchange.

Banks as we know them grew out of this acceptance by society of credit (IOU) money as a medium of exchange. Banks began to accept customer deposits of token money and to issue their own bank notes (credits) as receipts. People liked the convenience and safety of storing some of their wealth with banks. As society became more comfortable with banks and confident in the safety of banks, bank deposits, which could be transferred by cheque, became widely accepted as the medium of exchange. Bank notes and deposits were no longer convertible into gold or commodity money, but they were convertible into legal tender. Governments established central banks to control the supply of legal tender, bank notes, or cash. Bank notes now serve as both a medium of exchange and as the reserves banks hold to ensure the convertibility of their customers' deposits.

Legal tender: money that by law must be accepted as a means of payment.

Bank reserves: cash (legal tender) held by banks to meet possible withdrawals by depositors.

Unlike other financial institutions, such as pension funds and insurance companies, the key aspect of banks is that some of their liabilities are used as the medium of exchange; cheques and debit cards allow their deposits to be used as money to make payments. Bank deposits are credit money.

In Canada today, as in most industrial countries, we use a combination of fiat money and credit money. Fiat money, in contrast to commodity or token money is money that the government has declared to be legal tender. Coins and paper bank notes are fiat money in Canada. If you look carefully at a $5, $10, or $20 Bank of Canada bill you will find the statement: "This note is legal tender." By law it must be accepted as a means of payment for goods and services bought or debts repaid.

Fiat money: money the government has declared as legal tender.

Credit money: the debt of a private business or individual.

Our fiat money is supplemented by credit money. A bank deposit is credit money, and is redeemable in fiat money on demand, or in the case of some savings and time deposits, after a short waiting period. Furthermore, the bank is obliged to pay when your cheque is presented, or when you use your debit card. Bank deposits are a medium of exchange because they are generally accepted as a means of payment, even though they are not legal tender. The sum of bank deposits and fiat money in circulation outside the banks at any time is the stock of medium of exchange and the economy's money supply.

Money supply: the means of payment in the economy, namely currency (notes and coin) in circulation outside the banks and bank deposits.

9.2 Measures of the Canadian Money Supply

The money supply is traditionally defined as cash in circulation outside the banks, plus bank deposits. But as the banking and financial system evolved so did the types of deposits issued to the non-bank public. Now there are questions about the measurement of money supply.

Money supply: notes and coin in circulation outside banks plus bank deposits.

In the early days of banking there were demand deposits on which cheques could be written and savings deposits which often required a period of notice before funds could be withdrawn. Today banks offer a much wider spectrum of deposits to customers from demand to savings deposits that may or may not be chequable, pay interest under different terms, and some which can only be accessed online. Not all deposits serve as means of payment. For these the balance must be transferred to another account before it is available to make a payment. Which deposits should be counted in the money supply?

The structural evolution of the financial system raises further questions. What is a 'bank'? Today banks compete vigorously for deposits with other businesses, including trust companies and credit unions whose deposits are widely accepted as means of payment. There is no longer a reason to exclude those deposits from measures of the money supply. Different measures of the money supply illustrate the importance of different financial institutions in the industry.

The Bank of Canada now publishes data on the monetary base in response to continuing changes in the types of bank deposits available to households and businesses. Advances in technology, financial deregulation, and competition in the financial services sector, which have led to more types of financial assets, make it easy for customers to substitute between those assets we include in narrow definitions of money supply and other assets. But once we leave the monetary base as a measure of money supply, there is no single measure of money that is clearly the means of payment. There is, however, only one type of money that is legal tender; namely, notes and coin.

Monetary base: legal tender comprising notes and coins in circulation plus the cash held by the banks.

Table 8.1 shows the size of the money supply in Canada based on different definitions and money aggregates. These data illustrate the range of choice involved in the selection of a specific measure of the money supply. But one thing is clear: Bank deposits are the major component of money supply by any measure other than the currency component of the monetary base.
Table 8.1 The money supply in Canada in January 2017 (billions $)

Monetary base (MB)	84.6
Currency in circulation (CR)	78.8
M1B=currency+chequable chartered bank deposits	814.8
M2=M1B+notice and savings deposits in the banks	1,510.5
M2+=M2+deposits at other financial institutions	1,905.5

Source: Statistics Canada, CANSIM Table 176-0025

Currency in circulation is only about 5 percent of M2. Deposits account for the remaining 95 percent. The importance of bank deposits as money means that understanding the operations of banks as sources of loans and deposits is the key to understanding the money supply function in the economy.
Example Box 8.1 E-Payments and E-money and Fintech

Payment methods, monies and financial services continue to evolve as illustrated by the development of e-payments and e-money and more broadly by fintech.

E-payments are now familiar and widely used. Debit cards, telephone banking, internet banking and mobile banking provide access to manage deposit accounts and make payments, but they do not involve any creation of money. The central bank still controls money supply.

E-monies fall into two categories. Centralized e-money, for example a multi-purpose pre-paid payment card, denominated in the currency of the controller of the currency, has a money value based on funds received by the issuer. It is accepted as a means of payment in transactions with parties other than the issuer. Stored-value cards that use the Visa or MasterCard systems are widely accepted in payment. In Canada, transactions using these and similar cards are in Canadian dollars.

Decentralized e-money has no centralized issuer and is not denominated in any national currency. It is a *cryptocurrency*. The bitcoin is an example. It is completely decentralized and does not represent a claim on the issuer in the way that a Canadian dollar bank deposit represents a claim on a bank. Furthermore, no bank or institution is an intermediary in a transaction between payor and payee. A cryptocurrency is decentralized over a peer-to peer computer network that directly links users, but no one user in control of the network. It uses 'blockchain technology' which is a shared, continually reconciled data base. However, because cryptocurrencies are not redeemable in national currencies they generally trade online and their values are highly volatile in terms of national currencies.

Fintech describes this emerging and evolving range of technology based financial services. E-payments and e-money are part of this process. But it extends more widely to include technology based loaning, stock trading, robo-advising, digital wallets and all-in-one money management tools, to name a few examples. The growth in fintech is based on the possibility of a wider range of improved services available at lower cost to banks, businesses and consumers. Banks continue to play a major role in financial markets and the payments system and continually adapt to the new technology and competition in other aspects of their operations.

9.2 Money Created By Banks

Banks *create money* when they increase their *deposit liabilities* to pay for the loans they make to customers, or for the financial securities they buy. The public uses the deposit liabilities of the banks as money to make payments or to hold as a store of wealth. There are *four key conditions* that give banks the ability to create money:

1. The non-bank public has confidence in banks and is willing to hold and use bank deposits as money.
2. The non-bank public is willing to borrow from the banks to finance expenditure or asset purchases.
3. The banks are willing to operate with cash reserves equal to some small fraction of their deposit liabilities.
4. The banks are willing to accept the risks involved in lending to the non-bank public.

If any of these is absent, the banks cannot create money, although they may provide safekeeping services.

The first two conditions underlie the demand for banking services. Banks acquire cash by providing customers deposit services and bank customers use bank loans as a source for funds to pay for purchases of goods, services and financial assets like equities and bonds. If the non-bank public is unwilling to use bank services there is no banking industry.

The third condition required for the banks to create money is a bank reserve ratio that is less than one. The reserve ratio (rr) is the ratio of cash on hand to deposit liabilities that banks choose to hold.

rr=reserve assetsdeposit liabilities	(8.1)

Reserve ratio (rr): the ratio of cash reserves to deposit liabilities held by banks.

Cash holdings are reserve assets. If banks choose to hold reserves equal to their deposit liabilities, rr=1 and the banks cannot create deposits. They are simple safety deposit boxes.

A simplified case shows how banks can and do create deposits. Assume banks use a reserve ratio of 10 percent (rr=0.10). Suppose initially the non-bank public has wealth of $1,000 held in cash, before they decide to switch to bank deposit money. This cash is a private sector asset. It is a liability of the central bank or government, which issued it, but not a liability of the private banks. The 'Initial position' in Table 8.4 uses a simple balance sheet to show this cash as an asset of the non-bank private sector.
Table 8.4 How the banking system creates money

Banks				**Non-bank public**			
Assets		**Liabilities**		**Assets**		**Liabilities**	
Cash	0	Deposits	0	Cash	1,000	Bank loans	0
Cash	1,000	Deposits	1,000	Cash	0	Bank loans	0
				Deposits	1,000		
Cash	1,000	Deposits	10,000	Cash	0	Bank loans	9,000
Loans	9,000			Deposits	10,000		

Then in the second part of the table people deposit this $1,000 of cash into the banks by opening bank accounts. Banks get assets of $1,000 in cash, distributed among individual banks by their customers and issue total deposit liabilities of $1,000. These deposits are money the banks owe to their depositors. If banks were simply safety deposit boxes or storerooms, they would hold cash assets equal to their deposit liabilities. Their reserve ratio would be 100 percent of deposits, making rr=1.0. Table 8.4 would end with part 2.

However, if the public uses bank deposits as money, the banks don't need all deposits to be fully covered by cash reserves. It is unlikely that all depositors will show up at the same time and demand cash for their deposits. Recognizing this, the banks decide that reserves equal to 10 percent (rr=0.10) of deposits will cover all *net* customer demands for cash. In this case, the banks have excess reserves which in total equal 90 percent of their deposit liabilities or, initially, $900.

The banks use their excess reserves to expand their lending. Each bank makes new loans equal to its excess reserves. It pays for those loans by *creating an equal amount of deposits*. If you were to borrow from bank your personal deposit would be increased by the amount of the loan. The same thing happens to other people who borrow from their banks.

In our example, all banks combined can create $9,000 of loans based on $1,000 in new cash reserves. In part 3 of Table 8.4, we see loans of $9,000, as assets on the banks' balance sheets, and $9,000 of new deposits to customers, against which they can write cheques or make payments online or by transfers. The newly created deposits of $9,000 are a part of the $10,000 liability on the banks' balance sheets. The public now has bank deposit assets of $10,000 and liabilities, loans owed to the banks, of $9,000. Non-bank public net worth, assets minus liabilities is $1,000, the cash they originally deposited in the banks. Because the public uses bank deposits as money, the banks can buy new loans by creating new deposits.

The reserve ratio is 10 percent in part 3 of Table 8.4 (rr=$1,000 cash/$10,000 deposits =0.10 or 10%). It does not even matter whether the 10 percent reserve ratio is imposed by law or is merely smart profit-maximizing behaviour by the banks that balances risk and reward. The risk is the possibility of being caught short of cash; the reward is the net interest income earned.

Why were the banks able to create money? Originally, there was $1,000 of cash in circulation. That was the money supply. When paid into bank vaults, it went out of circulation as a medium of exchange. But the public got $1,000 of bank deposits against which cheques could be written. The money supply, cash in circulation plus bank deposits, was still $1,000. Then the banks created deposits *not* fully backed by cash reserves. Now the public had $10,000 of deposits against which to write cheques. The money supply rose from $1,000 to $10,000. *The public was willing to use bank deposits as money, willing to borrow from the banks and the banks were willing to lend*. This allowed the banks to create money by making loans based on their fractional reserve ratio.

Alternatively, suppose the public loses confidence in banks and withdraws and holds more currency. The banks are still able to create deposits but the extent of the deposit creation is limited by the public's withdrawal of currency. Bank reserves are reduced. A fall in public confidence in the banks in times of financial problems and bank failures like those in that arose in the autumn of 2008 and even today in some European countries would result in a rise in the currency holdings outside banks. Bank deposits and lending capacity would be reduced as a result, and in extreme cases bank solvency might be at risk without central bank support.

Financial panics

Most people know that banks operate with fractional reserve ratios and are not concerned. But if people begin to suspect that a bank has lent too much, made high risk loans or faces problems in raising funds which would make it difficult to meet depositors' claims for cash, there would be a *run on the bank* and a financial panic. Recognizing the bank cannot repay all depositors immediately, you try to get your money out first while the bank can still pay. Since everyone does the same thing, they ensure that the bank is unable to pay. It holds cash equal to a small percentage of its deposit liabilities and will be unable to liquidate its loans in time to meet the demands for cash.

Financial panic: a loss of confidence in banks and rush to withdraw cash.

Banking problems in Greece in the spring and early summer of 2015 provide an example. Concerns that the Greek government might default on loan payment agreements with IMF and European Union raised the possibility that Greece might leave the euro and return to its earlier national currency: the drachma. Should that happen, all euro deposits in Greek banks would convert to drachmas at an exchange rate that would reduce their real value substantially. Fearing this possibility, depositors in Greek banks tried to withdraw their balances in cash while they were still convertible into euros.

Greek banks, like other banks, operate on a fractional reserve basis. They could not meet this 'run on the bank' without outside support and assistance. The European Central Bank provided emergency cash to the banks but the run continued. In response, limits were placed on the amount of cash a depositor could withdraw at any one time. These measures sustained the banks until a solution to Greek debt crisis was negotiated and immediate concerns about Greek membership in the euro and the value of Greek bank deposits subsided.

However, earlier experience shows how financial crises can arise in other ways. In 2008-2009 the crisis originating in the US mortgage market and real estate sector caused wide spread problems for banks. Many banks had become reliant on large denomination, short-term deposits as sources of funds to support their mortgage lending. Other non-bank financial institutions like insurance companies and pension funds, as well as a relatively small number of individual customers bought these deposits. As the recession and falling property values emerged, the financial community began to worry that home-owners would not be able to pay back their mortgages.

If that happened banks would not be able to pay back depositors money, especially the large denomination short-term deposits. Once non-bank portfolio managers realized that it was difficult if not impossible to evaluate the risks of large denomination deposits, financial institutions that relied on renewing and issuing new deposits to raise funds were in difficulties. The supply of funds to replace expiring deposits dried up and banks could not repay depositors. Several large financial institutions in the United States and in other countries required government rescues or failed. The plight of famous names like *Bear Sterns*, *Countrywide Financial*, *Fannie May*, and *Freddie Mac* became headline news.

Banks in Canada were not immune to the financial difficulties created by the collapse of the large denomination deposit markets. All the major chartered banks were holding some. They were forced to accept that without a market these deposits would no longer be a source of funds. Fortunately, Canadian banks relied more heavily on strong smaller retail depositor bases as sources of funds. The banks remained financially strong and public confidence in the banks did not collapse. No Canadian bank failed or required a government bailout.

Fortunately, financial panics involving depositor runs on the bank are rare, particularly in Canada. A key reason for this, which we discuss in the next chapter, is that the central bank, the Bank of Canada, and other national central banks, will lend cash to banks in temporary difficulties. Furthermore, deposit insurance plans like the Canadian Deposit Insurance Corporation, CDIC, cover individual bank deposits up to $100,000 against default. Knowledge of these institutional arrangements helps prevent a self-fulfilling stampede to withdraw deposits before the bank runs out of cash.

By contrast, the financial crisis and the extended real estate and credit collapse in 2008 created large problems for US banks. Loan and financial asset defaults destroyed bank assets and bank liquidity. Even in the absence of panics and bank runs, many banks became insolvent without sufficient liquid assets to cover their liabilities. Failed bank data illustrates the scale of the problem. The US Federal Deposit Insurance Corporation lists 457 US bank failures over the period January 2008 to September 2012. In the four preceding years, January 2004 to December 2007 there were just 7 US bank failures.

(http://www.fdic.gov/bank/individual/failed/banklist.html)

9.3 Banking in Canada Today

In Canada today, and in other industrial countries, the mainstream banking system is made up of a *central bank* and a number of *commercial banks* and other deposit-taking institutions called *near banks*. Table 8.2 illustrates the structure of the banking industry in Canada. The industry is defined broadly to include deposit-taking institutions, not just those that operate under the federal Bank Act.
Table 8.2 The Canadian banking system in 2017

Banking Institution	**Number**
Central Bank:	
The Bank of Canada	1
Number of Chartered Banks:	
Schedule I domestic banks	26
Schedule II foreign banks subsidiaries	17
Schedule III foreign bank branches	20
Total	**64**

Source: Canadian Bankers' Association. http://www.cba.ca/memberbanks

Banks are financial intermediaries. They borrow money from the public, crediting them with a deposit. The deposit is a liability of the bank. It is money owed to depositors. The money raised from depositors provides the funds to support the bank loans made to businesses, households, and governments.

Financial intermediary: a business that specializes in bringing borrowers and lenders together.

Banks are not the only financial intermediaries. Trust companies, credit unions, caisses populaires, insurance companies, securities dealers, mutual fund companies, and independent financial advisors all play a role in this industry. But banks hold more than 70 percent of the assets in the financial services sector, and the six largest Canadian banks account for over 90 percent of the assets of the banking industry. Trust companies, credit unions, and caisses populaires also accept deposits that are used as money, but those deposits are a small fraction of the total of deposit money. As a result, bank deposits are the focus of our discussion of money in Canada.

The Bank of Canada is Canada's central bank. It is the source of the bank notes used to make payments and held as cash reserves by commercial banks. Established by the government in 1935, it has the responsibility to regulate the national money supply and support the operation of financial markets. The Bank's power to meet these responsibilities comes from its monopoly on the issuance of bank notes.

Bank of Canada: Canada's central bank.

The Bank of Canada also is the provider of:

- Banking services for the commercial banks in the system
- Banking services for the federal government
- Lender-of-last-resort facilities in times of liquidity crises and reserve shortfalls

Commercial banks hold some of their reserves as deposits in the Bank of Canada, and make payments among themselves using their Bank of Canada deposits. These interbank payments arise from wire transfers, direct deposits, pre-authorized debits, bill payments, point-of-sale debits, and online payments made by bank customers. For example, cheques written by customers at one bank, say *Scotiabank*, but paid to and deposited by customers of the *Royal Bank* result in transfers of deposits between these banks. To settle these transfers, *Scotiabank* must pay the Royal Bank. Funds held by *Scotiabank* on deposit in the Bank of Canada are used for this purpose. They are called "settlement balances." In 2016, *Payments Canada*, which coordinates this clearing of interbank transactions, handled more than 7.4 trillion transactions including 600 million cheques and 4.5 billion point-of-sale debits.

The government holds some deposits in the Bank of Canada. Government receipts, like income taxes paid to the Receiver General, are deposited in government accounts in the Bank of Canada. Government payments like Old Age Security, Employment Insurance benefits, bond interest, and income tax refunds are paid with government cheques or transfers drawn on its Bank of Canada account. Government funds over and above those needed to make regular payments are held on deposit in the commercial banks, and earn interest income for the government.

The key difference between a central bank and the commercial banks in the banking system is the profit motive. Central banks *do not* pursue profits. Their operations focus on the management of the cash reserves available to the public and the banks. The supply of cash reserves affects the behaviour of other banks and financial markets more generally. This is the monetary policy role of the central bank. We will examine it in detail in Chapter 10.

Commercial banks, on the other hand, are profit-oriented businesses. They operate, as we will see shortly, to maximize the profit they earn for their owners. To this end, they offer banking services to the public. Using the notes and deposits issued by the Bank of Canada as reserves, they issue bank deposits to their customers—which are widely used as the medium of exchange—and they make loans to finance purchases made by businesses and households.

To illustrate the business of these banks, Table 8.3 shows the consolidated balance sheet of Canadian chartered banks in December 2016. In the table we see that the banks held small cash balances as reserves against their deposit liabilities. Their other Canadian assets were mainly loans to households and businesses, including mortgage loans, and their holdings of financial securities. Because cash and many of their financial securities have high liquidity, banks can make long-term loans and still have cash and funds available if depositors withdraw their money.

Liquidity: the cost, speed, and certainty with which asset values can be converted into cash.
Table 8.3 Balance sheet of Canadian chartered banks, December 2016

Assets	billions $	Liabilities	billions $
Canadian dollars:		Canadian dollars:	
Cash	33.5	Personal deposits	927.9
Government of Canada securities	275.6	Non-personal deposits	830.3
Corporate securities	193.6	Government deposits	17.2
Personal and business loans	951.0	Advances from Bank of Canada	1.0
Mortgages	1,127.5	Other liabilities	807.0
Foreign currency assets	2,585.7	Foreign currency liabilities	2,638.3
Other assets	54.8		
Total assets	**5,221.7**	**Total liabilities and shareholders' equity**	**5,221.7**

Source: Bank of Canada, Banking and Financial Statistics, December 2016, Tables C3 and C4 and author's calculations. Figures have been rounded to one decimal place.

However, many loans to businesses and households are quite *illiquid*. The bank cannot easily get its money back in a hurry. This is not really a cause for concern when people and businesses have confidence in the banks and make widespread use of bank deposits as money. Payments and receipts are both in bank deposit form, which are cleared quickly and efficiently through the cheque-clearing and transfer facilities. Banks need only small cash balances to cover the net clearings and net public demand for cash. In Table 8.3, the banks are holding only $33.5 billion against deposit liabilities of $1,775.4 billion.

Canadian banks also carry on important international banking operations, as do banks in many other countries. We see this business recorded on the balance sheet as foreign currency assets and liabilities. The foreign currency assets are mainly loans to customers and holdings of foreign financial securities. Foreign currency deposits of customers are the main foreign currency liabilities. These foreign currency operations are similar to the banks' domestic currency operations. The banks provide loan financing to customers needing foreign currency to make payments in other countries, and they provide deposit facilities for customers using foreign currency for international transactions.

Competition and co-operation are important to the efficient operation of the banking system. Banks compete among themselves for customer deposits and customer loans. Some of the competition for deposits is based on the location, convenience, and quality of bank branches, some on the offers of service packages including personal financial advice and wealth management, and some on the interest rates offered on deposit balances. If you watch TV, you are probably aware that some small banks like *President's Choice Financial* and *Tangerine Bank* offer you a relatively high interest rate and will make no service charges if you would put some of your funds on deposit with them. Success in attracting deposits is very important to size and growth of a bank's business.

Credit-worthy customers willing to borrow funds are equally important to a bank's operations. Interest income earned on customer loans is the major source of bank revenue. As a result, banks compete in the personal and business loan markets, using both the terms of loans and the interest rates charged on loans to attract borrowers. The market for mortgage funds is one of the most competitive areas of bank operations. Mortgage rates and terms are advertised widely in the media and in displays in bank offices and even in supermarkets.

Despite this competition for deposits and loans, the banking system depends on the co-operation among banks that makes deposits the medium of exchange. Co-operation in the cheque-clearing system and the debit card *Interac* system are two important examples of banks working jointly to provide the payments system. A cheque book or a debit card is not very useful if it can make payments only to other people or businesses that do business with the same bank you use. Joint interests in *VISA* and *MASTERCARD* are a second important part of inter-bank co-operation that makes these cards widely acceptable as a source of credit.

There are also important areas of bank co-operation on the lending side of their operations. It often happens that businesses and industries have projects that need more financing than any one bank can or wants to provide. However, several banks might agree to provide funding jointly, increasing their lending capacity and spreading the risks associated with the project among them.

These dimensions of competition and co-operation among banks, and their contribution to the efficient functioning of the money and financial sector of the economy, appear regularly in the debate over bank mergers in Canada.

9.5 The Monetary Base and the Money Supply

Table 8.1 showed that bank deposits are the major component of the money supply in Canada, as in most industrial countries. Bank deposits depend in turn on the cash reserves held by banks and the public's willingness to hold bank deposits and borrow from the banks.

To complete our analysis of how the money supply is determined, we need to examine three things:

The source of the cash in the economy.

The amount of that cash that is deposited in the banking system, rather than held as cash balances by the public.

The relationship between the cash supply to the economy and the money supply that results from public and bank behaviour.

Today, in developed countries, central banks are the source of bank reserves. The central bank, the Bank of Canada in Canada, controls the issue of token money in the form of Bank of Canada notes. These are the $5, $10, $20, $50, and $100 bank notes you can withdraw from the bank when you wish to covert some of your bank balance to cash. Bank reserves are mainly the banks' holdings of these central bank notes in their vaults and bank machines. Our bank deposits are now convertible into Bank of Canada notes. The central bank has the responsibility to manage the supply of cash in the economy.

The cash the central bank provides to the economy is called the monetary base (MB) and is sometimes referred to as the stock of high-powered money. It is the legal tender into which bank deposits can be converted. It is the ultimate means of payment in transactions and the settlement of debts. Notes and coins in circulation and held by the banking system are the main part of the money issued by the central bank. As we discussed earlier, the commercial banks hold small settlement balances in the central bank to make inter-bank payments arising from cheque clearings.

Monetary base (MB): legal tender comprising notes and coins in circulation plus the cash held by the banks.

The public's decisions about the use of cash or banks deposits determine how much of the monetary base is held by the banks. The simple example of deposit creation in Table 8.4 assumed the public deposited all its cash with the banks. This was a useful simplification that ignores changes in the cash people hold. We will drop this assumption in what follows.

Our main interest is the relationship between the money supply in the economy, the total of cash in circulation plus bank deposits, and the monetary base created by the central bank. Assuming the public holds just a small fixed amount of cash and using our earlier discussion of the fractional reserve ratio in the banking system, we can define a deposit multiplier. The deposit multiplier provides the link between the monetary base created by the central bank and the money supply in the economy. It also predicts the change in money supply that would result from a change in the monetary base supplied by the central bank.

Deposit multiplier: the change in the bank deposits caused by a change in the monetary base.

Bank deposits= deposit multiplier × bank reserves

Deposit multiplier=Δ deposits / Δ bank reserves

The value of the deposit multiplier depends on rr, the banks' ratio of cash reserves to total deposits. Banks' choice of a ratio of cash reserves to total deposits (rr) determines how much they can expand lending and create bank deposits based on their reserve holdings. The lower the reserve ratio (rr), the more deposits banks can create against given cash reserves, and the larger is the multiplier. This is the relationship illustrated in Table 8.4.

Similarly, the lower the non-bank public's holding of cash, the larger is the share of the monetary base held by the banks. When the banks hold more monetary base, they can create more bank deposits. The lower the non-bank public's currency ratio, the larger are bank holdings of monetary base and the larger the money supply for any given monetary base.

The Money Multiplier

Suppose banks wish to hold cash reserves R equal to a fraction rr of their deposits D. Then:

Bank cash reserves = reserve ratio × deposits,

$\mathbf{R = \mathit{rrD}}$ (8.2)

To keep the example simple assume that we can ignore the small amount of cash held by the non-bank sector. As a result, the monetary base is mainly held as cash in bank vaults and automatic banking machines. This means from Equation 8.2 that:

Monetary base = reserve ratio × deposits,

$\mathbf{MB = \mathit{rrD}}$ (8.3)

and the deposit multiplier, which defines the change in total deposits as a result of a change in the monetary base, is:

Change in deposits = change in monetary base divided by the reserve ratio,

$\mathbf{\Delta D / \Delta MB = 1 / rr}$ (8.4)

which will be greater than 1 as long as rr is less than 1.

If, for example, banks want to hold cash reserves equal to 5 percent, and the non-bank public does not change their holdings of cash, the deposit multiplier will be:

$\Delta D / \Delta MB = 10.05 = 20$

When public cash holdings are constant, the deposit multiplier tells us how much deposits (and therefore the money supply in the economy), notes, and coins in circulation (outside the banks and bank deposits) would change as a result of a change in the monetary base. In this example, a $1 change in the monetary base results in a change in deposits and the money supply equal to $20.

Money supply (M)=cash in circulation+bank deposits (D)

$\mathbf{M = \mathit{cash} + (R/\mathit{rr})}$ (8.5)

We can see from the way we have found the deposit multiplier that it depends on the decisions made by the banks in terms of their reserve holdings. For simplicity the public holds a fixed amount of cash in addition to bank deposits as money. If you experiment with different values for rr, you will see how the deposit and money multiplier would change if the reserve ratio were to change. Furthermore, if the public were to change their cash holdings the cash reserves available to the banks would change and the deposit multiplier would cause a larger change in the money supply.

The importance of bank reserve decisions and public cash holdings decisions is illustrated by recent financial conditions in Europe. As a result of banking crisis and bailouts during and after the financial crisis of 2008, the public had concerns about the safety of bank deposits and decided to hold more cash. At the same time banks found it difficult to evaluate the credit worthiness of potential borrowers and the risks involved in short-term corporate lending or junior government bonds. The public's cash holdings increased and the banks increased their reserve ratios. These shifts in behaviour would reduce the money supply, making credit conditions tighter, unless the central bank provided offsetting increases in the monetary base.

How big is the money multiplier?

Now that we have a formula for the money multiplier, we can ask: What is the size of the multiplier in Canada? Based on data in Table 8.1 above, in January 2017, the monetary base was $84.6 billion, and the money supply defined as M1B was $814.8 billion. These data suggest a bank reserve ratio with respect of M1B =$84.6/$814.8 which is approximately 10.4 percent giving a money supply multiplier of 1/0.104=9.6.

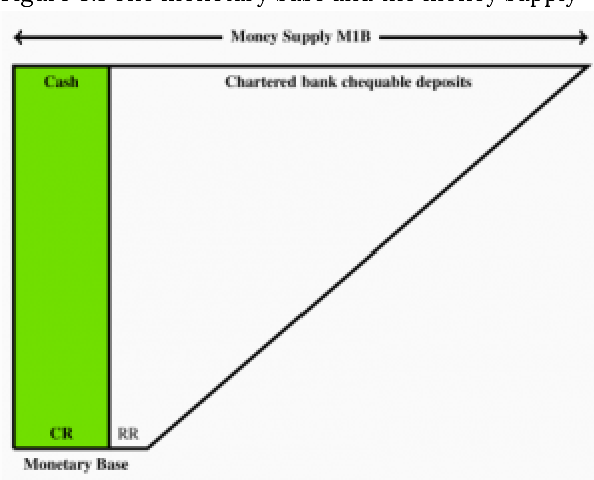
ΔM / ΔMB = $814.8 / $84.6 = 9.6

Each $100 change in monetary base would change the money supply by about **$960.**

However, using a broader definition of money supply such as 'currency outside banks and all chartered bank deposits' gives a Canadian money supply of $1,510.5 and a money multiplier of $1,510.5/$84.6 = 17.85.

Figure 8.1 summarizes the relationship between the monetary base and the money supply. It shows the monetary base used either as cash in circulation or held as cash reserves by the banks. Since banks operate with fractional reserve ratios, the leverage banks have to expand the money supply through their lending and deposits creation based on their reserves RR. We also see that the money supply is heavily dependent on the size of the monetary base, reserve ratios used by the banks and the willingness of the public to hold bank deposits and the willingness of the banks to lend.

Figure 8.1 The monetary base and the money supply



The explanation of banking and the money supply in this chapter provides the money supply function we will use in the next chapter. It is combined there with a demand for money function in the money market to determine the equilibrium rate of interest. That rate of interest integrates money and financial markets with the markets for goods and services in aggregate demand.

A simple money supply function illustrates the determinants of the money supply. The three key variables are:

MB, the monetary base;

the public's holdings of cash; and

rr, the banks' reserve ratio.

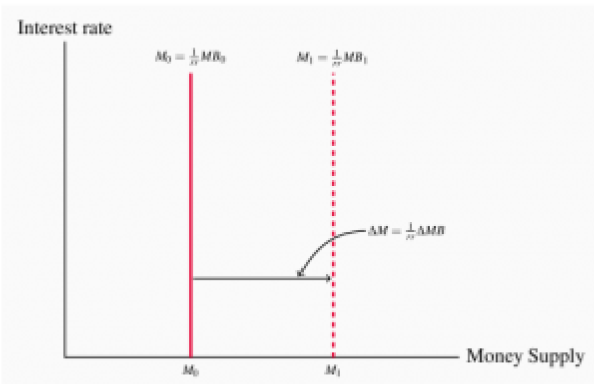
Using Equation 8.5 above, where M is the money supply, we can write:

M = (1/rr) × MB (8.6)

The central bank's control of the monetary base, MB, gives it control of the money supply, M, as long as cash holdings and rr are constant.

Figure 8.2 uses a diagram to illustrate the money supply function and changes in the money supply. The line M0 shows the size of the money supply for a given monetary base MB0 and the money multiplier. The money supply in this diagram is vertical, because we assume cash holdings and the reserve ratio are not affected by the interest rate. M is therefore independent of the nominal interest rate i, which is measured on the vertical axis. This is the supply side of the money market with quantity measured on the horizontal axis and interest rate, which is analogous to price, on the vertical axis.

Figure 8.2 The money supply function



The vertical M1 illustrates the increase in the money supply as a result of an increase in monetary base from MB0 to MB1, working through the money multiplier. A reduction in the monetary base would shift the M line to the left based on the same relationship.

Monetary policy

The central bank conducts monetary policy through its control of the monetary base. The money supply function shows us how, if rr is constant, the central bank's control of the monetary base gives it the power to change money supply and other financial conditions in the economy. If the central bank increases the monetary base, banks have larger cash reserves and increase their lending, offering favourable borrowing rates to attract new loans and create more deposits. In Figure 8.2 the increase in the monetary base to MB1 causes an increase in money supply (ΔM) by the change in MB (ΔMB), multiplied by the money multiplier. The money supply function shifts to the right to M1. A decrease in the monetary base would shift the M function to the left, indicating a fall in the money supply.

9.6 Review and Practice

Summary

In this chapter we investigated the money supply and looked at how it is determined. Money is anything that serves as a medium of exchange. Whatever serves as money also functions as a unit of account and as a store of value. Money may or may not have intrinsic value. In the United States, the total of currency in circulation, traveler's checks, and chequable deposits equals M1. A broader measure of the money supply is M2, which includes M1 plus assets that are highly liquid, but less liquid than those in M1.

Banks create money when they issue loans. The ability of banks to issue loans is controlled by their reserves. Reserves consist of cash in bank vaults and bank deposits with the central bank. Banks operate in a fractional reserve system; that is, they maintain reserves equal to only a small fraction of their deposit liabilities. Banks are heavily regulated to protect individual depositors and to prevent crises of confidence. Deposit insurance protects individual depositors.

A central bank serves as a bank for banks, a regulator of banks, a manager of the money supply, a bank for a nation's government, and a supporter of financial markets generally. In the financial crisis that rocked the United States and much of the world in 2008, the the US central bank, the "Fed" played a central role in keeping bank and nonbank institutions afloat and in keeping credit available. The Federal Reserve System (Fed) is the central bank for the United States. The Fed is governed by a Board of Governors whose members are appointed by the president of the United States, subject to confirmation by the Senate.

The Fed can lend to banks and other institutions through the discount window and other credit facilities, change reserve requirements, and engage in purchases and sales of federal government bonds in the open market. Decisions to buy or sell bonds are made by the Federal Open Market Committee (FOMC); the Fed's open-market operations represent its primary tool for influencing the money supply. Purchases of bonds by the Fed initially increase the reserves of banks. With excess reserves on hand, banks will attempt to increase their loans, and in the process the money supply will change by an amount less than or equal to the deposit multiplier times the change in reserves. Similarly, the Fed can reduce the money supply by selling bonds.

Concept Problems

1. Airlines have "frequent flier" clubs in which customers accumulate miles according to the number of miles they have flown with the airline. Frequent flier miles can then be used to purchase other flights, to rent cars, or to stay in some hotels. Are frequent flier miles money?
2. Debit cards allow an individual to transfer funds directly in a chequable account to a merchant without writing a check. How is this different from the way credit cards work? Are either credit cards or debit cards money? Explain.

3. Many colleges sell special cards that students can use to purchase everything from textbooks or meals in the cafeteria to use of washing machines in the dorm. Students deposit money in their cards; as they use their cards for purchases, electronic scanners remove money from the cards. To replenish a card's money, a student makes a cash deposit that is credited to the card. Would these cards count as part of the money supply?
4. A smart card, also known as an electronic purse, is a plastic card that can be loaded with a monetary value. Its developers argue that, once widely accepted, it could replace the use of currency in vending machines, parking meters, and elsewhere. Suppose smart cards came into widespread use. Present your views on the following issues:
    1. Would you count balances in the purses as part of the money supply? If so, would they be part of M1? M2?
    2. Should any institution be permitted to issue them, or should they be restricted to banks?
    3. Should the issuers be subject to reserve requirements?
    4. Suppose they were issued by banks. How do you think the use of such purses would affect the money supply? Explain your answer carefully.
5. Which of the following items is part of M1? M2?
    1. $0.27 cents that has accumulated under a couch cushion.
    2. Your $2,000 line of credit with your Visa account.
    3. The $210 balance in your chequing account.
    4. $417 in your savings account.
    5. 10 shares of stock your uncle gave you on your 18th birthday, which are now worth $520.
    6. $200 in traveller's checks you have purchased for your spring-break trip.
6. In the Middle Ages, goldsmiths took in customers' deposits (gold coins) and issued receipts that functioned much like checks do today. People used the receipts as a medium of exchange. Goldsmiths also issued loans by writing additional receipts against which they were holding no gold to borrowers. Were goldsmiths engaging in fractional reserve banking? Why do you think that customers turned their gold over to goldsmiths? Who benefited from the goldsmiths' action? Why did such a system generally work? When would it have been likely to fail?
7. A $1,000 deposit in Acme Bank has increased reserves by $1,000. A loan officer at Acme reasons as follows: "The reserve requirement is 10%. That means that the $1,000 in new reserves can back $10,000 in chequable deposits. Therefore I'll loan an additional $10,000." Is there any problem with the loan officer's reasoning? Explain.
8. When the Fed buys and sells bonds through open-market operations, the money supply changes, but there is no effect on the money supply when individuals buy and sell bonds. Explain.

Numerical Problems

1. Consider the following example of bartering:

1 10-ounce T-bone steak can be traded for 5 soft drinks.
1 soft drink can be traded for 10 apples.
100 apples can be traded for a T-shirt.
5 T-shirts can be exchanged for 1 textbook.
It takes 4 textbooks to get 1 VCR.

   1. How many 10-ounce T-bone steaks could you exchange for 1 textbook? How many soft drinks? How many apples?
   2. State the price of T-shirts in terms of apples, textbooks, and soft drinks.
   3. Why do you think we use money as a unit of account?
2. Assume that the banking system is loaned up and that any open-market purchase by the Fed directly increases reserves in the banks. If the required reserve ratio is 0.2, by how much could the money supply expand if the Fed purchased $2 billion worth of bonds?
3. Suppose the Fed sells $5 million worth of bonds to Econobank.
   1. What happens to the reserves of the bank?
   2. What happens to the money supply in the economy as a whole if the reserve requirement is 10%, all payments are made by check, and there is no net drain into currency?
   3. How would your answer in part b be affected if you knew that some people involved in the money creation process kept some of their funds as cash?
4. If half the banks in the nation borrow additional reserves totaling $10 million at the Fed discount window, and at the same time the other half of the banks reduce their excess reserves by a total of $10 million, what is likely to happen to the money supply? Explain.
5. Suppose a bank with a 10% reserve requirement has $10 million in reserves and $100 million in chequable deposits, and a major corporation makes a deposit of $1 million.
   1. Explain how the deposit affects the bank's reserves and chequable deposits.
   2. By how much can the bank increase its lending?
6. Suppose a bank with a 25% reserve requirement has $50 million in reserves and $200 million in chequable deposits, and one of the bank's depositors, a major corporation, writes a cheque to another corporation for $5 million. The check is deposited in another bank.

1. Explain how the withdrawal affects the bank's reserves and chequable deposits.
2. By how much will the bank have to reduce its lending?

7. Suppose the bank in problem 6 faces a 20% reserve requirement. The customer writes the same check. How will this affect your answers?
8. Now consider an economy in which the central bank has just purchased $8 billion worth of government bonds from banks in the economy. What would be the effect of this purchase on the money supply in the country, assuming reserve requirements of:
    1. 10%.
    2. 15%.
    3. 20%.
    4. 25%.
9. Now consider the same economy, and the central bank sells $8 billion worth of government bonds to local banks. State the likely effects on the money supply under reserve requirements of:
    1. 10%.
    2. 15%.
    3. 20%.
    4. 25%.
10. How would the purchase of $8 billion of bonds by the central bank from local banks be likely to affect interest rates? How about the effect on interest rates of the sale of $8 billion worth of bonds? Explain your answers carefully.

CHAPTER 10: FINANCIAL MARKETS AND THE ECONOMY

Start Up: Clamping Down on Money Growth

For nearly three decades, Americans have come to expect very low inflation, on the order of 2% to 3% a year. How did this expectation come to be? Was it always so? Absolutely not.

In July 1979, with inflation approaching 14% and interest rates on three-month Treasury bills soaring past 10%, a desperate President Jimmy Carter took action. He appointed Paul Volcker, the president of the New York Federal Reserve Bank, as chairman of the Fed's Board of Governors. Mr. Volcker made clear that his objective as chairman was to bring down the inflation rate—no matter what the consequences for the economy. Mr. Carter gave this effort his full support.

Mr. Volcker wasted no time in putting his policies to work. He slowed the rate of money growth immediately. The economy's response was swift; the United States slipped into a brief recession in 1980, followed by a crushing recession in 1981–1982. In terms of the goal of reducing inflation, Mr. Volcker's monetary policies were a dazzling success. Inflation plunged below a 4% rate within three years; by 1986 the inflation rate had fallen to 1.1%. The tall, bald, cigar-smoking Mr. Volcker emerged as a folk hero in the fight against inflation. Indeed he has returned 20 years later as part of President Obama's economic team to perhaps once again rescue the U.S. economy.

The Fed's seven-year fight against inflation from 1979 to 1986 made the job for Alan Greenspan, Mr. Volcker's successor, that much easier. To see how the decisions of the Federal Reserve affect key macroeconomic variables—real GDP, the price level, and unemployment—in this chapter we will explore how **financial markets**, markets in which funds accumulated by one group are made available to another group, are linked to the economy.

This chapter provides the building blocks for understanding financial markets. Beginning with an overview of bond and foreign exchange markets, we will examine how they are related to the level of real GDP and the price level. The second section completes the model of the money market. We have learned that the Fed can change the amount of reserves in the banking system, and that when it does the money supply changes. Here we explain money demand—the quantity of money people and firms want to hold—which, together with money supply, leads to an equilibrium rate of interest.

The model of aggregate demand and supply shows how changes in the components of aggregate demand affect GDP and the price level. In this chapter, we will learn that changes in the financial markets can affect aggregate demand—and in turn can lead to changes in real GDP and the price level. Showing how the financial markets fit into the model of aggregate demand and aggregate supply we developed earlier provides a more complete picture of how the macroeconomy works.

10.1 Demand, Supply, and Equilibrium in the Money Market

Learning Objectives

1. Explain the motives for holding money and relate them to the interest rate that could be earned from holding alternative assets, such as bonds.
2. Draw a money demand curve and explain how changes in other variables may lead to shifts in the money demand curve.
3. Illustrate and explain the notion of equilibrium in the money market.
4. Use graphs to explain how changes in money demand or money supply are related to changes in the bond market, in interest rates, in aggregate demand, and in real GDP and the price level.

In this section we will explore the link between money markets, bond markets, and interest rates. We first look at the demand for money. The demand curve for money is derived like any other demand curve, by examining the relationship between the "price" of money (which, we will see, is the interest rate) and the quantity demanded, holding all other determinants unchanged. We then link the demand for money to the concept of money supply developed in the last chapter, to determine the equilibrium rate of interest. In turn, we show how changes in interest rates affect the macroeconomy.

The Demand for Money

In deciding how much money to hold, people make a choice about how to hold their wealth. How much wealth shall be held as money and how much as other assets? For a given amount of wealth, the answer to this question will depend on the relative costs and benefits of holding money versus other assets. The demand for money is the relationship between the quantity of money people want to hold and the factors that determine that quantity.

To simplify our analysis, we will assume there are only two ways to hold wealth: as money in a checking account, or as funds in a bond market mutual fund that purchases long-term bonds on behalf of its subscribers. A bond fund is not money. Some money deposits earn interest, but the return on these accounts is generally lower than what could be obtained in a bond fund. The advantage of checking accounts is that they are highly liquid and can thus be spent easily. We will think of the demand for money as a curve that represents the outcomes of choices between the greater liquidity of money deposits and the higher interest rates that can be earned by holding a bond fund. The difference between the interest rates paid on money deposits and the interest return available from bonds is the cost of holding money.

Motives for Holding Money

One reason people hold their assets as money is so that they can purchase goods and services. The money held for the purchase of goods and services may be for everyday transactions such as buying groceries or paying the rent, or it may be kept on hand for contingencies such as having the funds available to pay to have the car fixed or to pay for a trip to the doctor.

The transactions demand for money is money people hold to pay for goods and services they anticipate buying. When you carry money in your purse or wallet to buy a movie ticket or maintain a checking account balance so you can purchase groceries later in the month, you are holding the money as part of your transactions demand for money.

The money people hold for contingencies represents their **precautionary demand for money.** Money held for precautionary purposes may include checking account balances kept for possible home repairs or health-care needs. People do not know precisely when the need for such expenditures will occur, but they can prepare for them by holding money so that they'll have it available when the need arises.

People also hold money for speculative purposes. Bond prices fluctuate constantly. As a result, holders of bonds not only earn interest but experience gains or losses in the value of their assets. Bondholders enjoy gains when bond prices rise and suffer losses when bond prices fall. Because of this, expectations play an important role as a determinant of the demand for bonds. Holding bonds is one alternative to holding money, so these same expectations can affect the demand for money.

John Maynard Keynes, who was an enormously successful speculator in bond markets himself, suggested that bondholders who anticipate a drop in bond prices will try to sell their bonds ahead of the price drop in order to avoid this loss in asset value. Selling a bond means converting it to money. Keynes referred to the speculative demand for money as the money held in response to concern that bond prices and the prices of other financial assets might change.

Of course, money is money. One cannot sort through someone's checking account and locate which funds are held for transactions and which funds are there because the owner of the account is worried about a drop in bond prices or is taking a precaution. We distinguish money held for different motives in order to understand how the quantity of money demanded will be affected by a key determinant of the demand for money: the interest rate.

Interest Rates and the Demand for Money

The quantity of money people hold to pay for transactions and to satisfy precautionary and speculative demand is likely to vary with the interest rates they can earn from alternative assets such as bonds. When interest rates rise relative to the rates that can be earned on money deposits, people hold less money. When interest rates fall, people hold more money. The logic of these conclusions about the money people hold and interest rates depends on the people's motives for holding money.

The quantity of money households want to hold varies according to their income and the interest rate; different average quantities of money held can satisfy their transactions and precautionary demands for money. To see why, suppose a household earns and spends $3,000 per month. It spends an equal amount of money each day. For a month with 30 days, that is $100 per day. One way the household could manage this spending would be to leave the money in a checking account, which we will assume pays zero interest. The household would thus have $3,000 in the checking account when the month begins, $2,900 at the end of the first day, $1,500 halfway through the month, and zero at the end of the last day of the month. Averaging the daily balances, we find that the quantity of money the household demands equals $1,500. This approach to money management, which we will call the "cash approach," has the virtue of simplicity, but the household will earn no interest on its funds.

Consider an alternative money management approach that permits the same pattern of spending. At the beginning of the month, the household deposits $1,000 in its checking account and the other $2,000 in a bond fund. Assume the bond fund pays 1% interest per month, or an annual interest rate of 12.7%. After 10 days, the money in the checking account is exhausted, and the household withdraws another $1,000 from the bond fund for the next 10 days. On the 20th day, the final $1,000 from the bond fund goes into the checking account. With this strategy, the household has an average daily balance of $500, which is the quantity of money it demands. Let us call this money management strategy the "bond fund approach."

Remember that both approaches allow the household to spend $3,000 per month, $100 per day. The cash approach requires a quantity of money demanded of $1,500, while the bond fund approach lowers this quantity to $500.

The bond fund approach generates some interest income. The household has $1,000 in the fund for 10 days (1/3 of a month) and $1,000 for 20 days (2/3 of a month). With an interest rate of 1% per month, the household earns $10 in interest each month ([$1,000 × 0.01 × 1/3] + [$1,000 × 0.01 × 2/3]). The disadvantage of the bond fund, of course, is that it requires more attention—$1,000 must be transferred from the fund twice each month. There may also be fees associated with the transfers.

Of course, the bond fund strategy we have examined here is just one of many. The household could begin each month with $1,500 in the checking account and $1,500 in the bond fund, transferring $1,500 to the checking account midway through the month. This strategy requires one less transfer, but it also generates less interest—$7.50 (= $1,500 × 0.01 × 1/2). With this strategy, the household demands a quantity of money of $750. The household could also maintain a much smaller average quantity of money in its checking account and keep more in its bond fund. For simplicity, we can think of any strategy that involves transferring money in and out of a bond fund or another interest-earning asset as a bond fund strategy.

Which approach should the household use? That is a choice each household must make—it is a question of weighing the interest a bond fund strategy creates against the hassle and possible fees associated with the transfers it requires. Our example does not yield a clear-cut choice for any one household, but we can make some generalizations about its implications.

First, a household is more likely to adopt a bond fund strategy when the interest rate is higher. At low interest rates, a household does not sacrifice much income by pursuing the simpler cash strategy. As the interest rate rises, a bond fund strategy becomes more attractive. That means that the higher the interest rate, the lower the quantity of money demanded.

Second, people are more likely to use a bond fund strategy when the cost of transferring funds is lower. The creation of savings plans, which began in the 1970s and 1980s, that allowed easy transfer of funds between interest-earning assets and chequable deposits tended to reduce the demand for money.

Some money deposits, such as savings accounts and money market deposit accounts, pay interest. In evaluating the choice between holding assets as some form of money or in other forms such as bonds, households will look at the differential between what those funds pay and what they could earn in the bond market. A higher interest rate in the bond market is likely to increase this differential; a lower interest rate will reduce it. An increase in the spread between rates on money deposits and the interest rate in the bond market reduces the quantity of money demanded; a reduction in the spread increases the quantity of money demanded.

Firms, too, must determine how to manage their earnings and expenditures. However, instead of worrying about $3,000 per month, even a relatively small firm may be concerned about $3,000,000 per month. Rather than facing the difference of $10 versus $7.50 in interest earnings used in our household example, this small firm would face a difference of $2,500 per month ($10,000 versus $7,500). For very large firms such as Toyota or AT&T, interest rate differentials among various forms of holding their financial assets translate into millions of dollars per day.

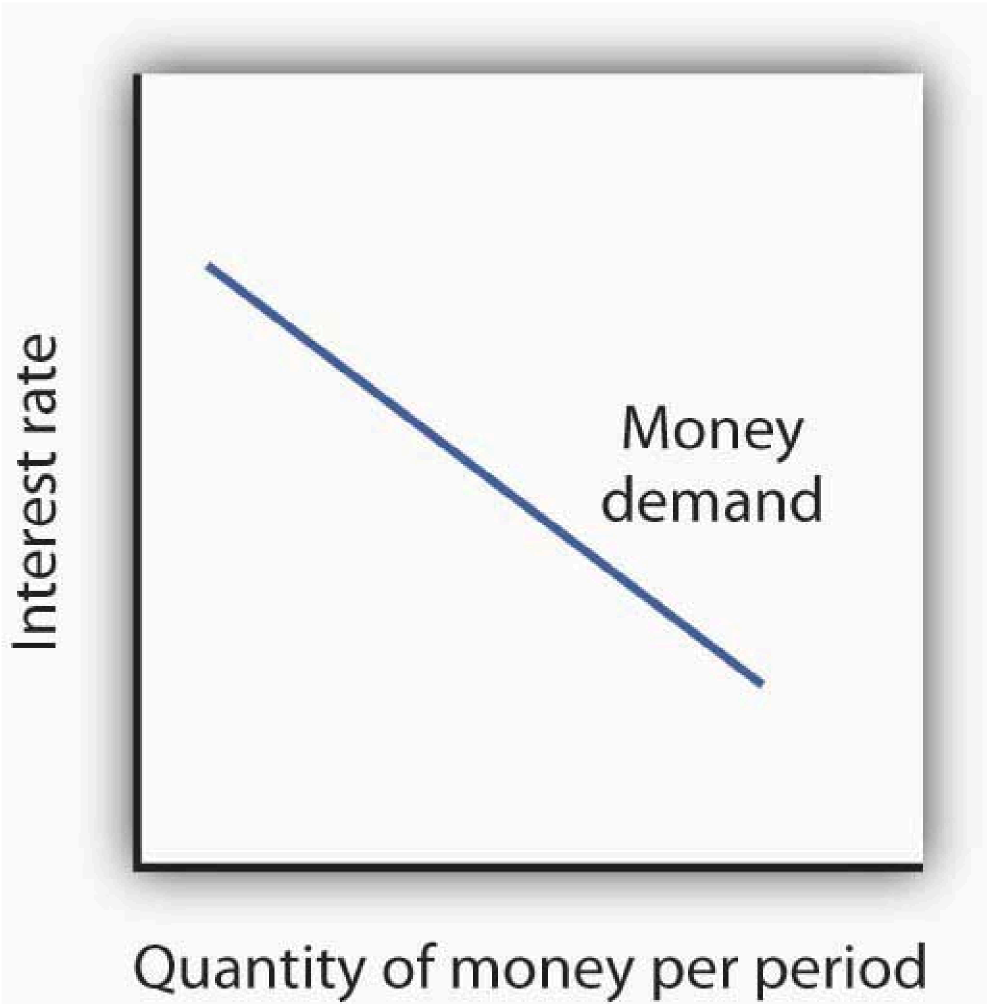
How is the speculative demand for money related to interest rates? When financial investors believe that the prices of bonds and other assets will fall, their speculative demand for money goes up. The speculative demand for money thus depends on expectations about future changes in asset prices. Will this demand also be affected by present interest rates?

If interest rates are low, bond prices are high. It seems likely that if bond prices are high, financial investors will become concerned that bond prices might fall. That suggests that high bond prices—low interest rates—would increase the quantity of money held for speculative purposes. Conversely, if bond prices are already relatively low, it is likely that fewer financial investors will expect them to fall still further. They will hold smaller speculative balances. Economists thus expect that the quantity of money demanded for speculative reasons will vary negatively with the interest rate.

The Demand Curve for Money

We have seen that the transactions, precautionary, and speculative demands for money vary negatively with the interest rate. Putting those three sources of demand together, we can draw a demand curve for money to show how the interest rate affects the total quantity of money people hold. The demand curve for money shows the quantity of money demanded at each interest rate, all other things unchanged. Such a curve is shown in Figure 10.1 "The Demand Curve for Money". An increase in the interest rate reduces the quantity of money demanded. A reduction in the interest rate increases the quantity of money demanded.

Figure 10.1 The Demand Curve for Money



The demand curve for money shows the quantity of money demanded at each interest rate. Its downward slope expresses the negative relationship between the quantity of money demanded and the interest rate.

The relationship between interest rates and the quantity of money demanded is an application of the law of demand. If we think of the alternative to holding money as holding bonds, then the interest rate—or the differential between the interest rate in the bond market and the interest paid on money deposits—represents the price of holding money. As is the case with all goods and services, an increase in price reduces the quantity demanded.

Other Determinants of the Demand for Money

We draw the demand curve for money to show the quantity of money people will hold at each interest rate, all other determinants of money demand unchanged. A change in those "other determinants" will shift the demand for money. Among the most important variables that can shift the demand for money are the level of income and real GDP, the price level, expectations, transfer costs, and preferences.

Real GDP

A household with an income of $10,000 per month is likely to demand a larger quantity of money than a household with an income of $1,000 per month. That relationship suggests that money is a normal good: as income increases, people demand more money at each interest rate, and as income falls, they demand less.

An increase in real GDP increases incomes throughout the economy. The demand for money in the economy is therefore likely to be greater when real GDP is greater.

The Price Level

The higher the price level, the more money is required to purchase a given quantity of goods and services. All other things unchanged, the higher the price level, the greater the demand for money.

Expectations

The speculative demand for money is based on expectations about bond prices. All other things unchanged, if people expect bond prices to fall, they will increase their demand for money. If they expect bond prices to rise, they will reduce their demand for money.

The expectation that bond prices are about to change actually causes bond prices to change. If people expect bond prices to fall, for example, they will sell their bonds, exchanging them for money. That will shift the supply curve for bonds to the right, thus lowering their price. The importance of expectations in moving markets can lead to a self-fulfilling prophecy.

Expectations about future price levels also affect the demand for money. The expectation of a higher price level means that people expect the money they are holding to fall in value. Given that expectation, they are likely to hold less of it in anticipation of a jump in prices.

Expectations about future price levels play a particularly important role during periods of hyperinflation. If prices rise very rapidly and people expect them to continue rising, people are likely to try to reduce the amount of money they hold, knowing that it will fall in value as it sits in their wallets or their bank accounts. Toward the end of the great German hyperinflation of the early 1920s, prices were doubling as often as three times a day. Under those circumstances, people tried not to hold money even for a few minutes—within the space of eight hours money would lose half its value!

Transfer Costs

For a given level of expenditures, reducing the quantity of money demanded requires more frequent transfers between nonmoney and money deposits. As the cost of such transfers rises, some consumers will choose to make fewer of them. They will therefore increase the quantity of money they demand. In general, the demand for money will increase as it becomes more expensive to transfer between money and nonmoney accounts. The demand for money will fall if transfer costs decline. In recent years, transfer costs have fallen, leading to a decrease in money demand.

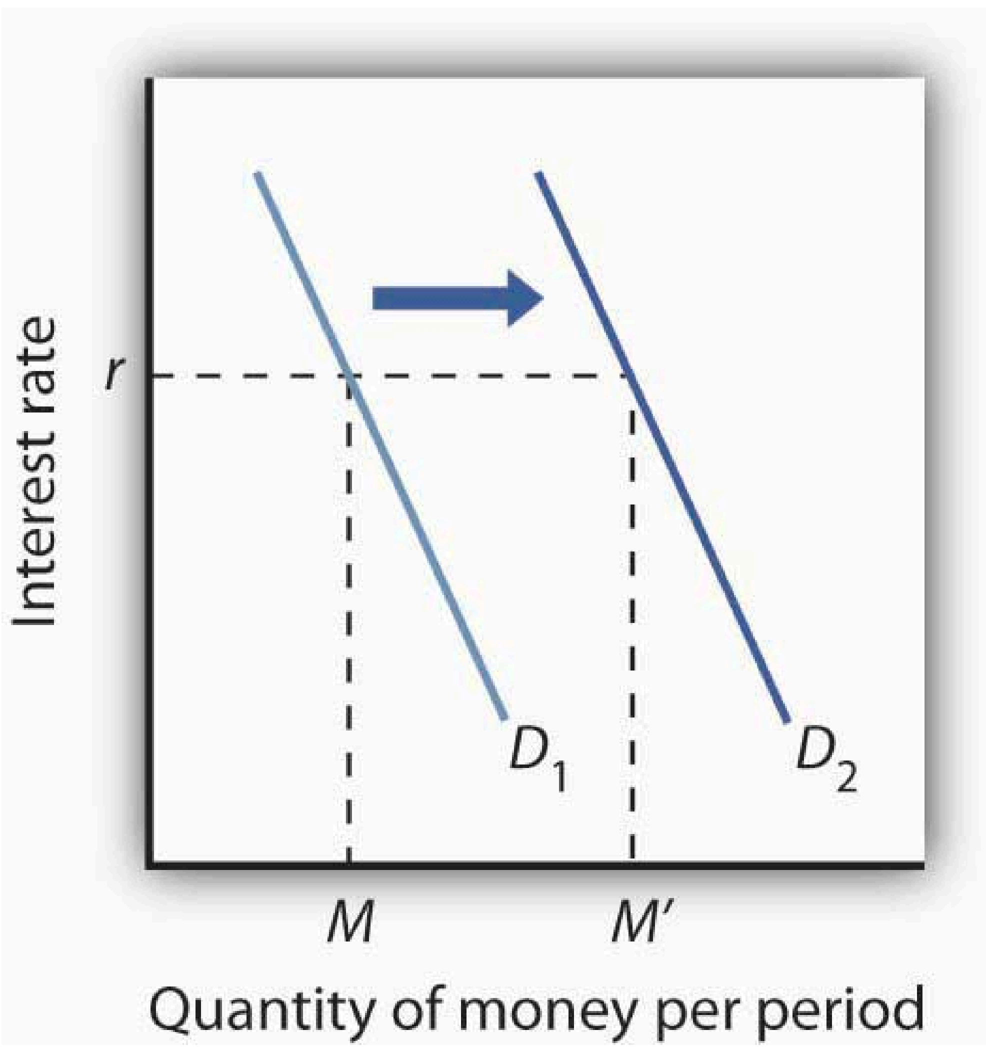
Preferences

Preferences also play a role in determining the demand for money. Some people place a high value on having a considerable amount of money on hand. For others, this may not be important.

Household attitudes toward risk are another aspect of preferences that affect money demand. As we have seen, bonds pay higher interest rates than money deposits, but holding bonds entails a risk that bond prices might fall. There is also a chance that the issuer of a bond will default, that is, will not pay the amount specified on the bond to bondholders; indeed, bond issuers may end up paying nothing at all. A money deposit, such as a savings deposit, might earn a lower yield, but it is a safe yield. People's attitudes about the trade-off between risk and yields affect the degree to which they hold their wealth as money. Heightened concerns about risk in the last half of 2008 led many households to increase their demand for money.

Figure 10.2 "An Increase in Money Demand" shows an increase in the demand for money. Such an increase could result from a higher real GDP, a higher price level, a change in expectations, an increase in transfer costs, or a change in preferences.

Figure 10.2 An Increase in Money Demand

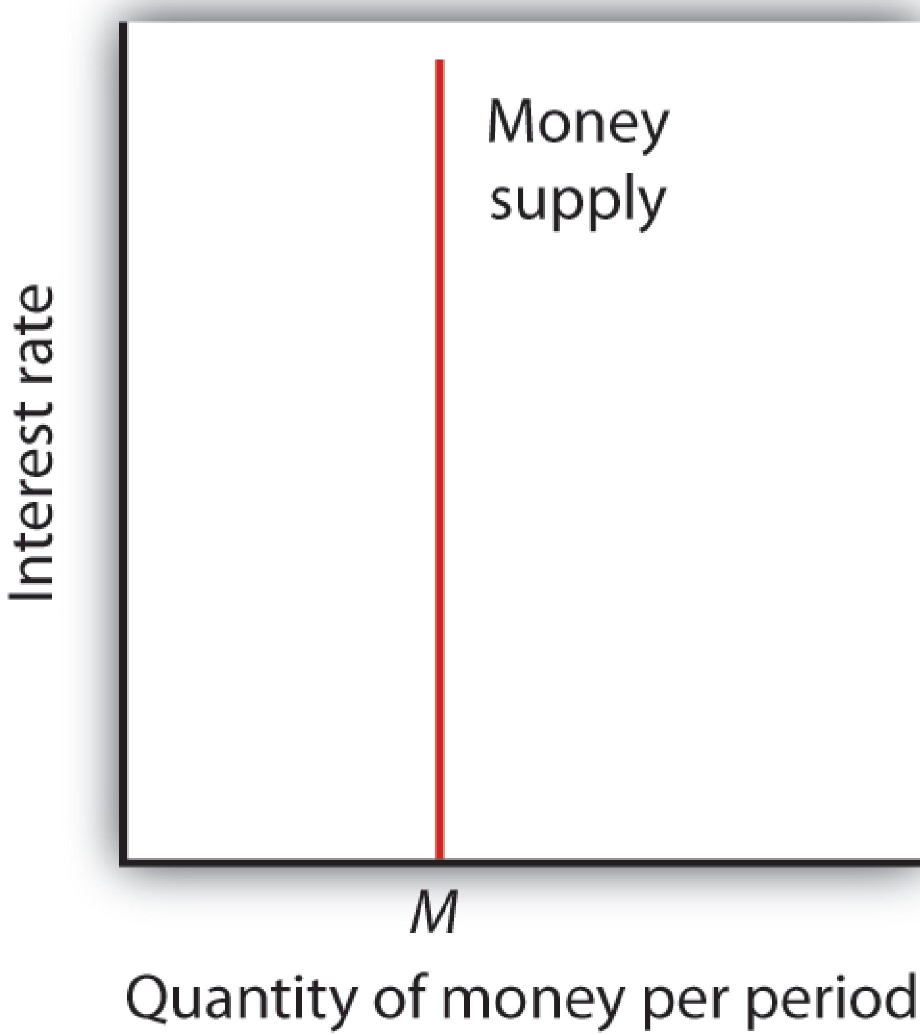


An increase in real GDP, the price level, or transfer costs, for example, will increase the quantity of money demanded at any interest rate *r*, increasing the demand for money from D_1 to D_2. The quantity of money demanded at interest rate *r* rises from *M* to *M′*. The reverse of any such events would reduce the quantity of money demanded at every interest rate, shifting the demand curve to the left.

The Supply of Money

The supply curve of money shows the relationship between the quantity of money supplied and the market interest rate, all other determinants of supply unchanged. We have learned that the Fed, through its open-market operations, determines the total quantity of reserves in the banking system. We shall assume that banks increase the money supply in fixed proportion to their reserves. Because the quantity of reserves is determined by Federal Reserve policy, we draw the supply curve of money in Figure 10.3 "The Supply Curve of Money" as a vertical line, determined by the Fed's monetary policies. In drawing the supply curve of money as a vertical line, we are assuming the money supply does not depend on the interest rate. Changing the quantity of reserves and hence the money supply is an example of monetary policy.

Figure 10.3 The Supply Curve of Money

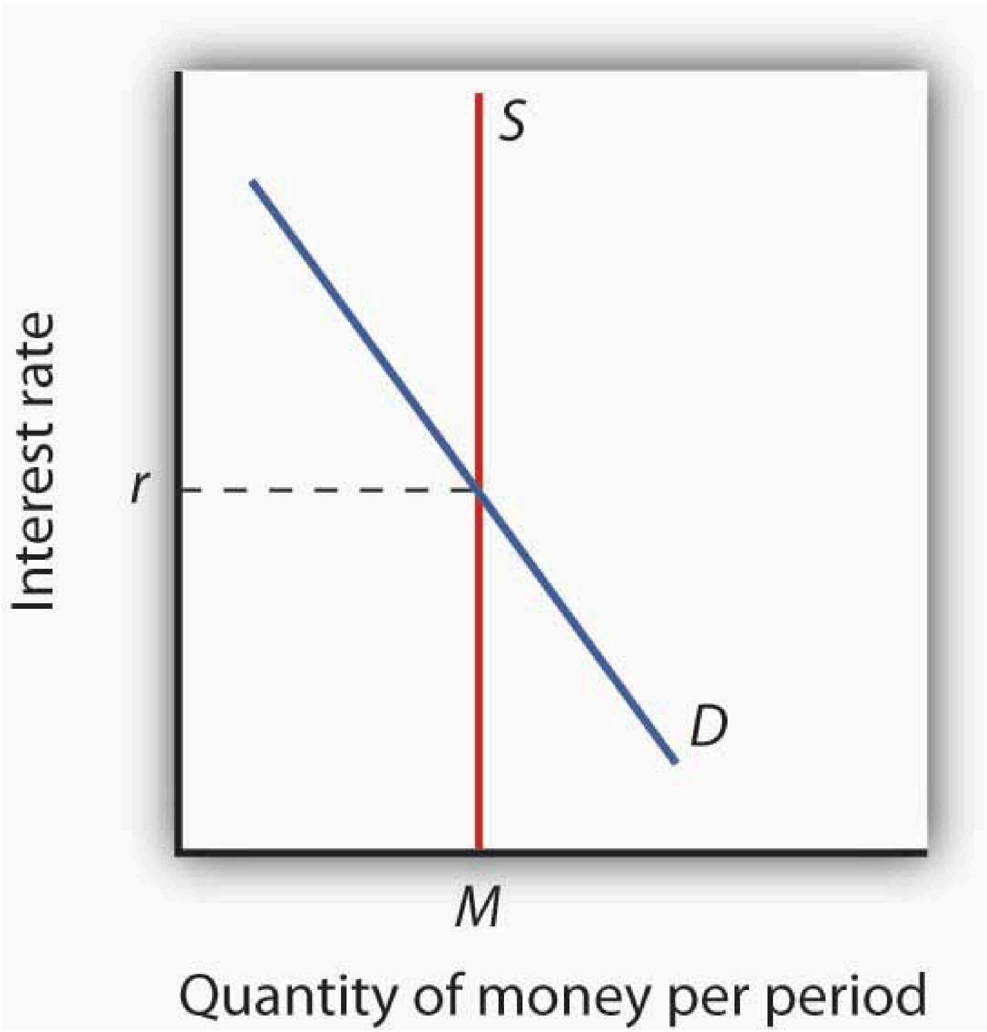


We assume that the quantity of money supplied in the economy is determined as a fixed multiple of the quantity of bank reserves, which is determined by the Fed. The supply curve of money is a vertical line at that quantity.

Equilibrium in the Market for Money

The money market is the interaction among institutions through which money is supplied to individuals, firms, and other institutions that demand money. Money market equilibrium occurs at the interest rate at which the quantity of money demanded is equal to the quantity of money supplied. Figure 10.4 "Money Market Equilibrium" combines demand and supply curves for money to illustrate equilibrium in the market for money. With a stock of money (M), the equilibrium interest rate is *r*.

Figure 10.4 Money Market Equilibrium



The market for money is in equilibrium if the quantity of money demanded is equal to the quantity of money supplied. Here, equilibrium occurs at interest rate *r*.

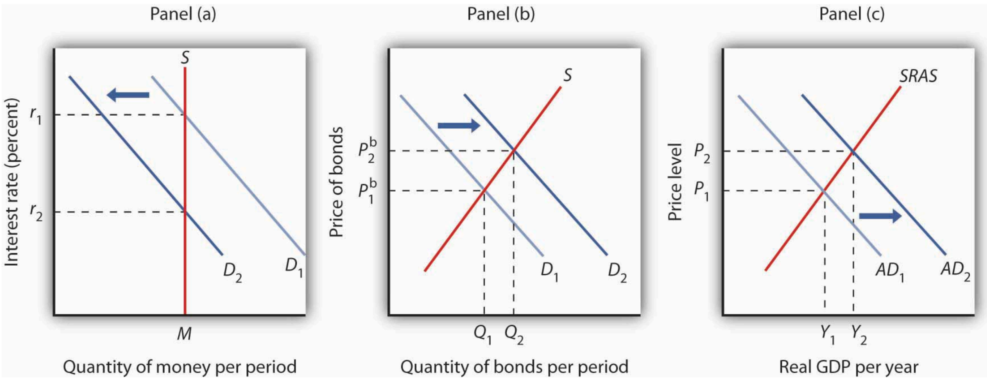
Effects of Changes in the Money Market

A shift in money demand or supply will lead to a change in the equilibrium interest rate. Let's look at the effects of such changes on the economy.

Changes in Money Demand

Suppose that the money market is initially in equilibrium at r_1 with supply curve S and a demand curve D_1 as shown in Panel (a) of [Figure 10.5 "A Decrease in the Demand for Money"](). Now suppose that there is a decrease in money demand, all other things unchanged. A decrease in money demand could result from a decrease in the cost of transferring between money and nonmoney deposits, from a change in expectations, or from a change in preferences[1]. Panel (a) shows that the money demand curve shifts to the left to D_2. We can see that the interest rate will fall to r_2. To see why the interest rate falls, we recall that if people want to hold less money, then they will want to hold more bonds. Thus, Panel (b) shows that the demand for bonds increases. The higher price of bonds means lower interest rates; lower interest rates restore equilibrium in the money market.

Figure 10.5 A Decrease in the Demand for Money



A decrease in the demand for money due to a change in transactions costs, preferences, or expectations, as shown in Panel (a), will be accompanied by an increase in the demand for bonds as shown in Panel (b), and a fall in the interest rate. The fall in the interest rate will cause a rightward shift in the aggregate demand curve from AD_1 to AD_2, as shown in Panel (c). As a result, real GDP and the price level rise.

Lower interest rates in turn increase the quantity of investment. They also stimulate net exports, as lower interest rates lead to a lower exchange rate. The aggregate demand curve shifts to the right as shown in Panel (c) from AD_1 to AD_2. Given the short-run aggregate supply curve SRAS, the economy moves to a higher real GDP and a higher price level.

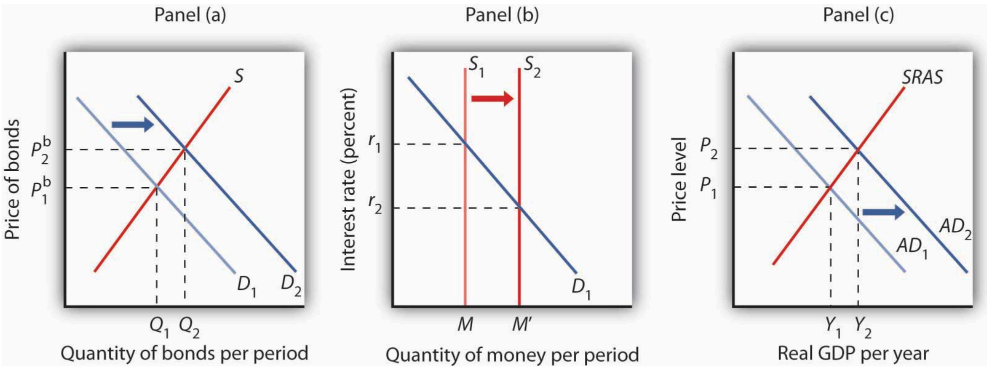
An increase in money demand due to a change in expectations, preferences, or transactions costs that make people want to hold more money at each interest rate will have the opposite effect. The money demand curve will shift to the right and the demand for bonds will shift to the left. The resulting higher interest rate will lead to a lower quantity of investment. Also, higher interest rates will lead to a higher exchange rate and depress net exports. Thus, the aggregate demand curve will shift to the left. All other things unchanged, real GDP and the price level will fall.

Changes in the Money Supply

Now suppose the market for money is in equilibrium and the Fed changes the money supply. All other things unchanged, how will this change in the money supply affect the equilibrium interest rate and aggregate demand, real GDP, and the price level?

Suppose the Fed conducts open-market operations in which it buys bonds. This is an example of expansionary monetary policy. The impact of Fed bond purchases is illustrated in Panel (a) of Figure 10.6 "An Increase in the Money Supply". The Fed's purchase of bonds shifts the demand curve for bonds to the right, raising bond prices to P^b_2. As we learned, when the Fed buys bonds, the supply of money increases. Panel (b) of Figure 10.6 "An Increase in the Money Supply" shows an economy with a money supply of M, which is in equilibrium at an interest rate of r_1. Now suppose the bond purchases by the Fed as shown in Panel (a) result in an increase in the money supply to M'; that policy change shifts the supply curve for money to the right to S_2. At the original interest rate r_1, people do not wish to hold the newly supplied money; they would prefer to hold nonmoney assets. To reestablish equilibrium in the money market, the interest rate must fall to increase the quantity of money demanded. In the economy shown, the interest rate must fall to r_2 to increase the quantity of money demanded to *M*′.

Figure 10.6 An Increase in the Money Supply



The Fed increases the money supply by buying bonds, increasing the demand for bonds in Panel (a) from D_1 to D_2 and the price of bonds to $P^b{}_2$. This corresponds to an increase in the money supply to *M*′ in Panel (b). The interest rate must fall to r_2 to achieve equilibrium. The lower interest rate leads to an increase in investment and net exports, which shifts the aggregate demand curve from AD_1 to AD_2 in Panel (c). Real GDP and the price level rise.

The reduction in interest rates required to restore equilibrium to the market for money after an increase in the money supply is achieved in the bond market. The increase in bond prices lowers interest rates, which will increase the quantity of money people demand. Lower interest rates will stimulate investment and net exports, via changes in the foreign exchange market, and cause the aggregate demand curve to shift to the right, as shown in Panel (c), from AD_1 to AD_2. Given the short-run aggregate supply curve SRAS, the economy moves to a higher real GDP and a higher price level.

Open-market operations in which the Fed sells bonds—that is, a contractionary monetary policy—will have the opposite effect. When the Fed sells bonds, the supply curve of bonds shifts to the right and the price of bonds falls. The bond sales lead to a reduction in the money supply, causing the money supply curve to shift to the left and raising the equilibrium interest rate. Higher interest rates lead to a shift in the aggregate demand curve to the left.

As we have seen in looking at both changes in demand for and in supply of money, the process of achieving equilibrium in the money market works in tandem with the achievement of equilibrium in the bond market. The interest rate determined by money market equilibrium is consistent with the interest rate achieved in the bond market.

Key Takeaways

- People hold money in order to buy goods and services (transactions demand), to have it available for contingencies (precautionary demand), and in order to avoid possible drops in the value of other assets such as bonds (speculative demand).

- The higher the interest rate, the lower the quantities of money demanded for transactions, for precautionary, and for speculative purposes. The lower the interest rate, the higher the quantities of money demanded for these purposes.
- The demand for money will change as a result of a change in real GDP, the price level, transfer costs, expectations, or preferences.
- We assume that the supply of money is determined by the Fed. The supply curve for money is thus a vertical line. Money market equilibrium occurs at the interest rate at which the quantity of money demanded equals the quantity of money supplied.
- All other things unchanged, a shift in money demand or supply will lead to a change in the equilibrium interest rate and therefore to changes in the level of real GDP and the price level.

Try It!

In 2005 the Fed was concerned about the possibility that the United States was moving into an inflationary gap, and it adopted a contractionary monetary policy as a result. Draw a four-panel graph showing this policy and its expected results. In Panel (a), use the model of aggregate demand and aggregate supply to illustrate an economy with an inflationary gap. In Panel (b), show how the Fed's policy will affect the market for bonds. In Panel (c), show how it will affect the demand for and supply of money. In Panel (d), show how it will affect the exchange rate. Finally, return to Panel (a) and incorporate these developments into your analysis of aggregate demand and aggregate supply, and show how the Fed's policy will affect real GDP and the price level in the short run.

Case in Point: Money in Today's World

Figure 25.13

Can Pac Swire – Hong Kong/ Travel Wallet – CC BY-NC 2.0.

The models of the money and bond markets presented in this chapter suggest that the Fed can control the interest rate by deciding on a money supply that would lead to the desired equilibrium interest rate in the money market. Yet, Fed policy announcements typically focus on what it wants the federal funds rate to be with scant attention to the money supply. Whereas throughout the 1990s, the Fed would announce a target federal funds rate and also indicate an expected change in the money supply, in 2000, when legislation requiring it to do so expired, it abandoned the practice of setting money supply targets.

Why the shift? The factors that have made focusing on the money supply as a policy target difficult for the past 25 years are first banking deregulation in the 1980s followed by financial innovations associated with technological changes—in particular the maturation of electronic payment and transfer mechanisms—thereafter.

Before the 1980s, M1 was a fairly reliable measure of the money people held, primarily for transactions. To buy things, one used cash, checks written on demand deposits, or traveler's checks. The Fed could thus use reliable estimates of the money demand curve to predict what the money supply would need to be in order to bring about a certain interest rate in the money market.

Legislation in the early 1980s allowed for money market deposit accounts (MMDAs), which are essentially interest-bearing savings accounts on which checks can be written. MMDAs are part of M2. Shortly after, other forms of payments for transactions developed or became more common. For example, credit and debit card use has mushroomed (from $10.8 billion in 1990 to $30 billion in 2000), and people can pay their credit card bills, electronically or with paper checks, from accounts that are part of either M1 or M2. Another innovation of the last 20 years is the automatic transfer service (ATS) that allows consumers to move money between checking and savings accounts at an ATM machine, or online, or through prearranged agreements with their financial institutions. While we take these methods of payment for granted today, they did not exist before 1980 because of restrictive banking legislation and the lack of technological know-how. Indeed, before 1980, being able to pay bills from accounts that earned interest was unheard of.

Further blurring the lines between M1 and M2 has been the development and growing popularity of what are called retail sweep programs. Since 1994, banks have been using retail-sweeping software to dynamically reclassify balances as either checking account balances (part of M1) or MMDAs (part of M2). They do this to avoid reserve requirements on checking accounts. The software not only moves the funds but also ensures that the bank does not exceed the legal limit of six reclassifications in any month. In the last 10 years these retail sweeps rose from zero to nearly the size of M1 itself!

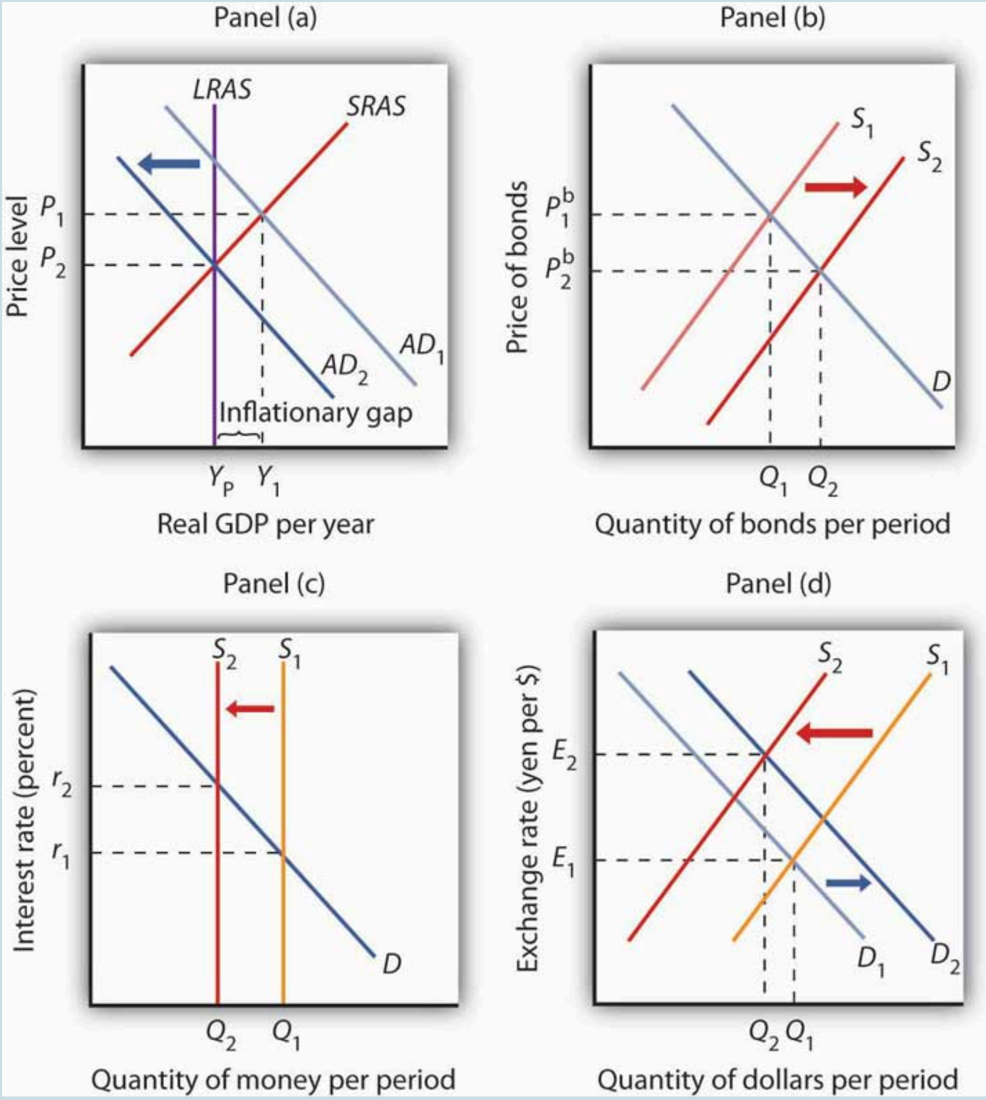
Such changes in the ways people pay for transactions and banks do their business have led economists to think about new definitions of money that would better track what is actually used for the purposes behind the money demand curve. One notion is called MZM, which stands for "money zero maturity." The idea behind MZM is that people can easily use any deposits that do not have specified maturity terms to pay for transactions, as these accounts are quite liquid, regardless of what classification of money they fall into. Some research shows that using MZM allows for a stable picture of the money market. Until more agreement has been reached, though, we should expect the Fed to continue to downplay the role of the money supply in its policy deliberations and to continue to announce its intentions in terms of the federal funds rate.

Source: Pedre Teles and Ruilin Zhou, "A Stable Money Demand: Looking for the Right Monetary Aggregate," *Federal Reserve Bank of Chicago Economic Perspectives* 29 (First Quarter, 2005): 50–59.

Answer to Try It! Problem

In Panel (a), with the aggregate demand curve AD_1, short-run aggregate supply curve *SRAS*, and long-run aggregate supply curve *LRAS*, the economy has an inflationary gap of $Y_1 - Y_P$. The contractionary monetary policy means that the Fed sells bonds—a rightward shift of the bond supply curve in Panel (b), which decreases the money supply—as shown by a leftward shift in the money supply curve in Panel (c). In Panel (b), we see that the price of bonds falls, and in Panel (c) that the interest rate rises. A higher interest rate will reduce the quantity of investment demanded. The higher interest rate also leads to a higher exchange rate, as shown in Panel (d), as the demand for dollars increases and the supply decreases. The higher exchange rate will lead to a decrease in net exports. As a result of these changes in financial markets, the aggregate demand curve shifts to the left to AD_2 in Panel (a). If all goes according to plan (and we will learn in the next chapter that it may not!), the new aggregate demand curve will intersect *SRAS* and *LRAS* at Y_P.

Figure 25.14



[1]In this chapter we are looking only at changes that originate in financial markets to see their impact on aggregate demand and aggregate supply. Changes in the price level and in real GDP also shift the money demand curve, but these changes are the result of changes in aggregate demand or aggregate supply and are considered in more advanced courses in macroeconomics.

10.2 Money Market Equilibrium

The money supply and the demand for money in the financial market determine nominal interest rates. A nominal money supply depends primarily on the monetary base and the money multiplier, namely:

$$M = 1/rr \times MB \quad (10.6)$$

The demand for money is a demand for *real* money balances as determined by real income and interest rates.
$L=kY-hi$

The real money supply is simply the nominal money supply M divided by the price level P, M/P, which measures its purchasing power in terms of goods and services.

Real money supply (M/P): the nominal money supply M divided by the price level P.

The central bank, as the source of the monetary base MB, controls the nominal money supply, *as long as the reserve ratio rr and the public's holdings of cash are constant*. The next chapter explains how the central bank manages the monetary base. If the price level is fixed, the central bank also controls the real money supply. Changes in nominal money tend to lead eventually to changes in prices. However, the central bank can still control the real money supply in the short run—it can change M faster than prices P respond—but in the long run other forces determine real money M/P. For the moment, assume the price level as fixed.

The demand for money as a demand for real money balances is summarized above. The quantity of real money demanded rises when real income rises, but falls when nominal interest rates rise.

Money market equilibrium

Figure 10.1 combines the demand curve for real money balances with the money supply function to give a money market diagram. The demand curve is drawn for a given level of real income, Y_0, and the supply curve for a given monetary base MB_0. With a given price level, the central bank controls the supply of nominal and real money. The supply curve is vertical at M_0/P_0. Equilibrium in the money market is at E. At the interest rate i0, the real money balances people wish to hold just equal the money supplied by the central bank and the banking system.

Figure 10.1 Equilibrium in the money market

To see how this market operates, suppose the interest rate is i_1, lower than the equilibrium level i_0. There is excess demand for money in the amount AB in the diagram. People want to hold money balances equal to B at the interest rate i_0, but only A is available. How does the market adjust to remove this excess demand? The answer lies in the portfolio decisions that distribute wealth between money holdings and bonds.

Consider the interaction between the bond and money markets. When portfolio managers want to restructure their holdings of bonds and money they do so by buying or selling bonds on the bond market. Their actions cannot change the supply of money balances. That is fixed by the monetary base and the money supply multiplier. As a result, bond prices and interest rates change to maintain money market equilibrium.

In Figure 10.1 the excess demand for money at the interest rate i_1 will result in a rise in interest rates. With an excess demand for money, people sell bonds to adjust their money balances. There is an excess supply of bonds. Bond prices fall. Lower bond prices mean higher bond yields and interest rates, as you will recall from our earlier discussion of asset prices and yields. The higher interest rates reduce both the excess supply of bonds and the excess demand for money. The money market adjusts by *moving along* the L curve from B to E, as people want smaller money balances relative to their bond holdings at higher interest rates.

This inverse interest rate – bond price relationship is the key to adjustments in the money market caused by changes in either the demand for or supply of money balances. Those adjustments involve trades in bonds that change bond prices and interest rates to maintain money market equilibrium.

Changes in financial market equilibrium

A shift in either the money supply or money demand changes equilibrium in the money market (and the bond market). Interest rates move to restore equilibrium.

The effect of a change in the money supply

Suppose the central bank lowers the monetary base and the money supply contracts. For a fixed price level, lower nominal money reduces the real money supply. Figure 10.2 shows this leftward shift in the money supply curve from M_0/P_0 to M_1/P_0. The equilibrium interest rate rises from i_0 to i_1 as people sell bonds. A higher interest rate reduces the quantity of real money balances demanded, moving along the demand curve $L(Y_0)$, bringing quantity of balances demanded into line with the reduced supply. Hence, a lower money supply raises equilibrium interest rates. Conversely, a rise in the money supply lowers the equilibrium interest rate.

Figure 10.2 Effect of a fall in the money supply

The effect of a change in real income

Figure 10.3 shows real money demand $L(Y_0)$ for the real income Y_0. A rise in real income increases the quantity of real money balances demanded at each interest rate, shifting the demand for money function from $L(Y_0)$ to $L(Y_1)$. The equilibrium interest rate rises as portfolio managers sell bonds in an attempt to increase their money holdings. The rise in the interest rate lowers the quantity of real balances demanded, moving along the money demand function $L(Y_1)$, and keeps demand for money equal to the unchanged supply. Conversely, a fall in real income would shift the demand for money to the left and reduce the equilibrium interest rate.

Figure 10.3 Effect of a rise in real income

10.3 Review and Practice

Summary

We saw how the money market works. The quantity of money demanded varies negatively with the interest rate. Factors that cause the demand curve for money to shift include changes in real GDP, the price level, expectations, the cost of transferring funds between money and nonmoney accounts, and preferences, especially preferences concerning risk. Equilibrium in the market for money is achieved at the interest rate at which the quantity of money demanded equals the quantity of money supplied. We assumed that the supply of money is determined by the Bank of Canada. An increase in money demand raises the equilibrium interest rate, and a decrease in money demand lowers the equilibrium interest rate. An increase in the money supply lowers the equilibrium interest rate; a reduction in the money supply raises the equilibrium interest rate.

Concept Problems

1. How would each of the following affect the demand for money?
    1. A tax on bonds held by individuals
    2. A forecast by the central bank that interest rates will rise sharply in the next quarter
    3. A wave of crime
    4. An announcement by the Prime Minister that, beginning in the next fiscal year, government spending will be reduced by an amount sufficient to eliminate all future borrowing
2. Some low-income countries do not have a bond market. In such countries, what substitutes for money do you think people would hold?
3. Explain what is meant by the statement that people are holding more money than they want to hold.
4. Explain how the action of the Bank of Canada to sell bonds in the open market would shift the supply curve for money.

Numerical Problems

1. Suppose that the demand and supply schedules for bonds that have a face value of $100 and a maturity date one year hence are as follows:

Price	Quantity Demanded	Quantity Supplied
$100	0	600
95	100	500
90	200	400
85	300	300
80	400	200
75	500	100
70	600	0

1. Draw the demand and supply curves for these bonds, find the equilibrium price, and determine the interest rate.
2. Now suppose the quantity demanded increases by 200 bonds at each price. Draw the new demand curve and find the new equilibrium price. What has happened to the interest rate?
1.

2. Consider the euro-zone of the European Union and Japan. The demand and supply curves for euros are given by the following table (prices for the euro are given in Japanese yen; quantities of euros are in millions):

Price (in euros)	Euros Demanded	Euros Supplied
¥75	0	600
70	100	500
65	200	400
60	300	300
55	400	200
50	500	100
45	600	0

1. Draw the demand and supply curves for euros and state the equilibrium exchange rate (in yen) for the euro. How many euros are required to purchase one yen?
2. Suppose an increase in interest rates in the European Union increases the demand for euros by 100 million at each price. At the same time, it reduces the supply by 100 million at each price. Draw the new demand and supply curves and state the new equilibrium exchange rate for the euro. How many euros are now required to purchase one yen?
3. How will the event in (b) affect net exports in the European Union?
4. How will the event in (b) affect aggregate demand in the European Union?
5. How will the event in (b) affect net exports in Japan?
6. How will the event in (b) affect aggregate demand in Japan?
1.

3. Suppose the quantity demanded of money at an interest rate of 5% is $2 billion per day, at an interest rate of 3% is $3 billion per day, and at an interest rate of 1% is $4 billion per day. Suppose the money supply is $3 billion per day.

1. Draw a graph of the money market and find the equilibrium interest rate.
2. Suppose the quantity of money demanded decreases by $1 billion per day at each interest rate. Graph this situation and find the new equilibrium interest rate. Explain the process of achieving the new equilibrium in the money market.
3. Suppose instead that the money supply decreases by $1 billion per day. Explain the process of achieving the new equilibrium in the money market.

4. We know that the economy faced a recessionary gap in 2008 and that the central bank responded with an expansionary monetary policy. Present the results of the bank's action in a two-panel graph. In Panel (a), show the initial situation, using the model of aggregate demand and aggregate supply. In Panel (b), show how the Bank's policy affects the money market.

CHAPTER 11: GOVERNMENT AND FISCAL POLICY

Start Up: A Massive Stimulus

Shaken by the severity of the recession that began in December 2007, Congress passed a huge $787 billion stimulus package in February 2009. President Obama described the measure as only "the beginning" of what the federal government ultimately would do to right the economy.

Roughly a third of the Recovery and Reinvestment Act is for a variety of tax cuts for individuals and firms. For example, each worker making less than $75,000 a year will receive $400 ($800 for a working couple earning up to $150,000) as a kind of rebate for payroll taxes. That works out to $8 a week. Qualifying college students are eligible for $2,500 tax credits for educational expenses. The other two-thirds is for a variety of government spending programs. The president said that the measure would "ignite spending by businesses and consumers … and make the investment necessary for lasting growth and economic prosperity."Barack Obama, Weekly Address of the President to the Nation, February 14, 2009, available at http://www.whitehouse.gov/blog/09/02/14/A-major-milestone/.

The measure illustrates an important difficulty of using fiscal policy in an effort to stabilize economic activity. It was passed over a year after the recession began. According to an estimate by the Congressional Budget Office (CBO), only about 20% of the spending called for by the legislation will take place in 2009, rising to about two-thirds through the middle of 2010. It is a guess what state the economy will be in then.

A fiscal stimulus package of over $150 billion had already been tried earlier in February 2008. It included $100 billion in tax rebates to households—up to $600 for individuals and $1,200 for couples— and over $50 billion in tax breaks for businesses. The boost to aggregate demand seemed slight—consumers saved much of their rebate money. In November 2008, unemployment insurance benefits were extended for seven additional weeks, in recognition of the growing unemployment problem.

President Obama argued that his proposals for dealing with the economy in the short term would, coincidentally, also promote long-term economic health. Some critics argued for a greater focus on tax cuts while others were concerned about whether the spending would focus on getting the greatest employment increase or be driven by political considerations.

How do government tax and expenditure policies affect real GDP and the price level? Why do economists differ so sharply in assessing the likely impact of such policies? Can fiscal policy be used to stabilize the economy in the short run? What are the long-run effects of government spending and taxing?

We begin with a look at the government's budget to see how it spends the tax revenue it collects. Clearly, the government's budget is not always in balance, so we will also look at government deficits and debt. We will then look at how fiscal policy works to stabilize the economy, distinguishing between built-in stabilization methods and discretionary measures. We will end the chapter with a discussion of why fiscal policy is so controversial.

As in the previous chapter on monetary policy, our primary focus will be U.S. policy. However, the tools available to governments around the world are quite similar, as are the issues surrounding the use of fiscal policy.

11.1 Government and the Economy

Learning Objectives

1. Understand the major components of government spending and sources of government revenues.
2. Define the terms *budget surplus*, *budget deficit*, *balanced budget*, and *national debt*, and discuss their trends over time.

We begin our analysis of fiscal policy with an examination of government purchases, transfer payments, and taxes in the U.S. economy.

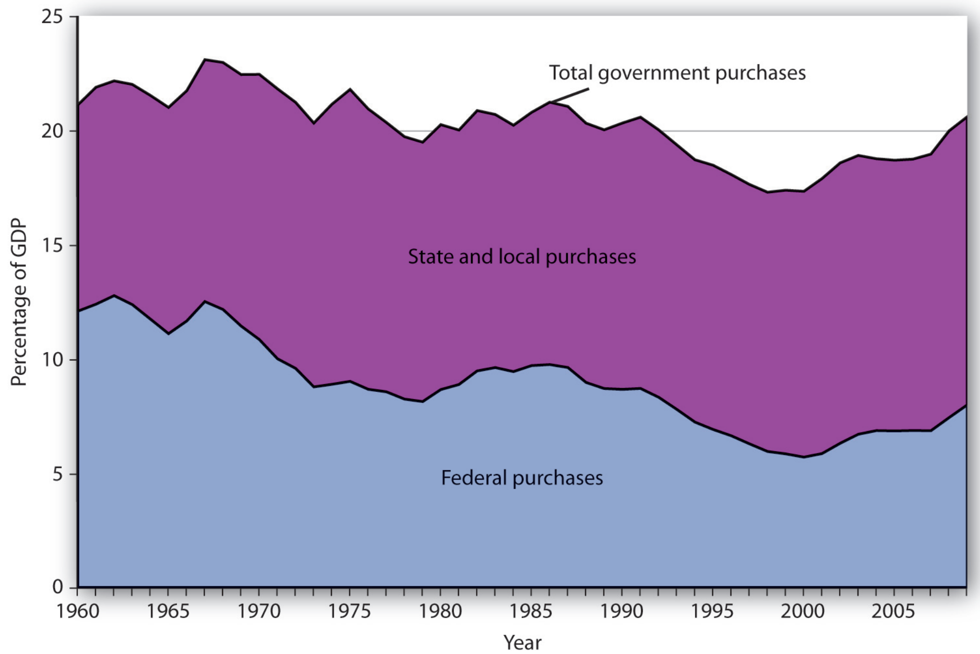
Government Purchases

The government-purchases component of aggregate demand includes all purchases by government agencies of goods and services produced by firms, as well as direct production by government agencies themselves. When the federal government buys staples and staplers, the transaction is part of government purchases. The production of educational and research services by public colleges and universities is also counted in the government-purchases component of GDP.

While government spending has grown over time, government purchases as a share of GDP generally declined from over 20% until the early 1990s to under 18% in 2001. Since then, though, the percentage of government purchases in GDP began to increase back toward 20%, first as military spending picked up and then more recently to over 20% during the 2007-2009 recession.

Figure 11.1 "Federal, State, and Local Purchases Relative to GDP, 1960–2009" shows federal as well as state and local government purchases as a percentage of GDP from 1960 to 2009. Notice the changes that have occurred over this period. In 1960, the federal government accounted for the majority share of total purchases. Since then, however, federal purchases have fallen by almost half relative to GDP, while state and local purchases relative to GDP have risen.

Figure 11.1 Federal, State, and Local Purchases Relative to GDP, 1960–2009



Government purchases were generally above 20% of GDP from 1960 until the early 1990s and then below 20% of GDP until the 2007-2009 recession. The share of government purchases in GDP began rising again in the 21st century.

Source: Bureau of Economic Analysis, NIPA Table 1.1 and 3.1 (November 23, 2010 revision).

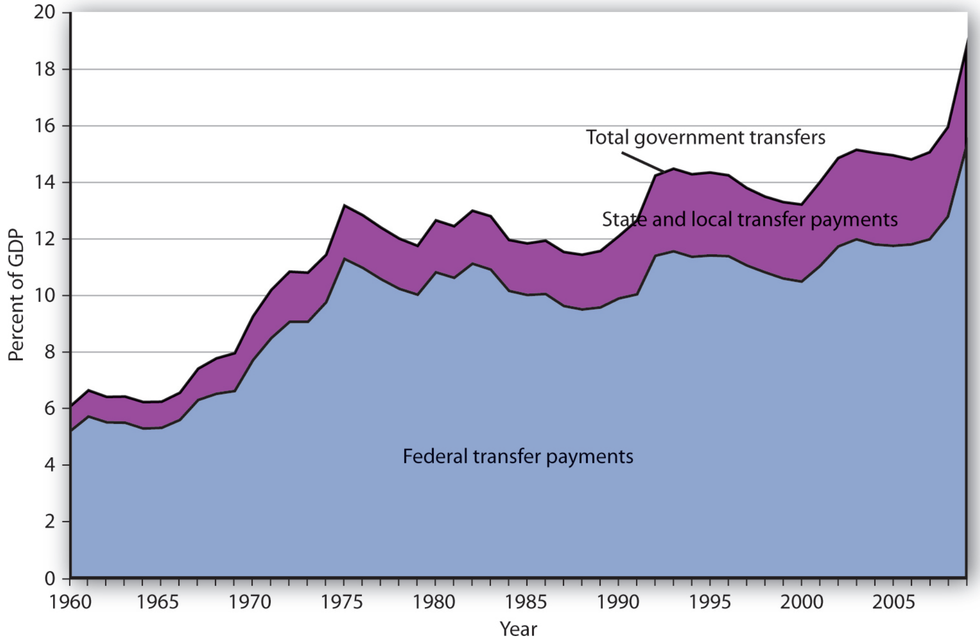
Transfer Payments

A transfer payment is the provision of aid or money to an individual who is not required to provide anything in exchange. Social Security and welfare benefits are examples of transfer payments. During the 2007-2009 recession, transfers again rose.

A number of changes have influenced transfer payments over the past several decades. First, they increased rapidly during the late 1960s and early 1970s. This was the period in which federal programs such as Medicare (health insurance for the elderly) and Medicaid (health insurance for the poor) were created and other programs were expanded.

Figure 11.2 "Federal, State, and Local Transfer Payments as a Percentage of GDP, 1960–2009" shows that transfer payment spending by the federal government and by state and local governments has risen as a percentage of GDP. In 1960, such spending totaled about 6% of GDP; by 2009, it had risen to about 18%. The federal government accounts for the bulk of transfer payment spending in the United States.

Figure 11.2 Federal, State, and Local Transfer Payments as a Percentage of GDP, 1960–2009



The chart shows transfer payment spending as a percentage of GDP from 1960 through 2009. This spending rose dramatically relative to GDP during the late 1960s and the 1970s as federal programs expanded. More recently, sharp increases in health-care costs have driven upward the spending for transfer payment programs such as Medicare and Medicaid. Transfer payments fluctuate with the business cycle, rising in times of recession and falling during times of expansion. As such, they rose sharply during the deep 2007-2009 recession.

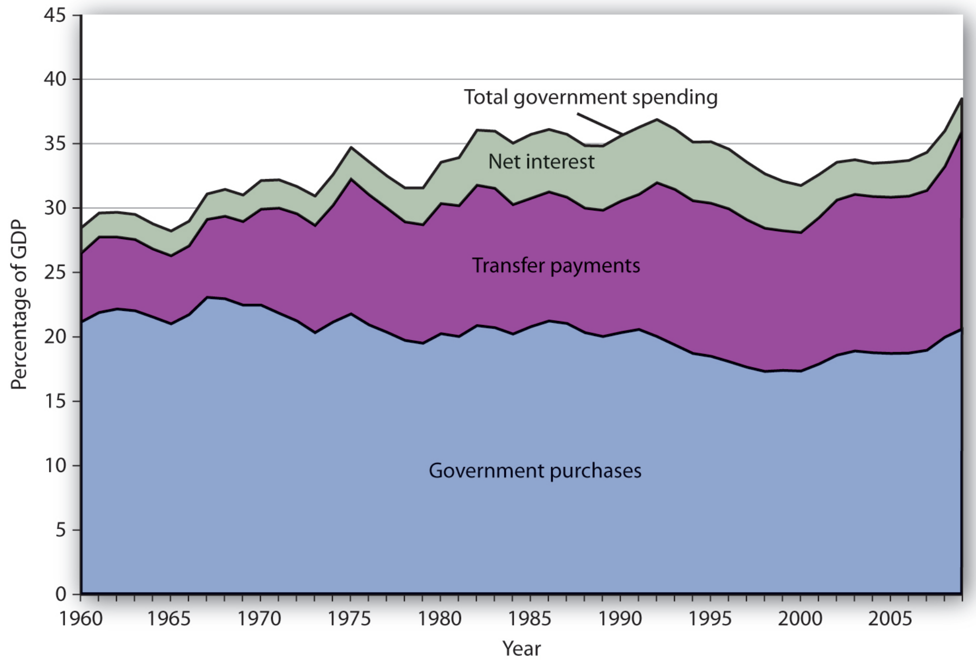
Source: Bureau of Economic Analysis, NIPA Table 1.1, 3.2, and 3.3 (November 23, 2010 revision).

Transfer payment spending relative to GDP tends to fluctuate with the business cycle. Transfer payments fell during the late 1970s, a period of expansion, then rose as the economy slipped into a recessionary gap during the 1979–1982 period. Transfer payments fell during the expansion that began late in 1982, then began rising in 1989 as the expansion began to slow. Transfer payments continued to rise relative to GDP during the recessions of 1990–1991 and 2001–2002 and then fell as the economy entered expansionary phases after each of those recessions. During the 2007—2009 recession, transfers again rose.

When economic activity falls, incomes fall, people lose jobs, and more people qualify for aid. People qualify to receive welfare benefits, such as cash, food stamps, or Medicaid, only if their income falls below a certain level. They qualify for unemployment compensation by losing their jobs. More people qualify for transfer payments during recessions. When the economy expands, incomes and employment rise, and fewer people qualify for welfare or unemployment benefits. Spending for those programs therefore tends to fall during an expansion.

Figure 11.3 "Government Spending as a Percentage of GDP, 1960–2009" summarizes trends in government spending since 1960. It shows three categories of government spending relative to GDP: government purchases, transfer payments, and net interest. Net interest includes payments of interest by governments at all levels on money borrowed, less interest earned on saving.

Figure 27.3 Government Spending as a Percentage of GDP, 1960–2009



This chart shows three major categories of government spending as percentages of GDP: government purchases, transfer payments, and net interest.

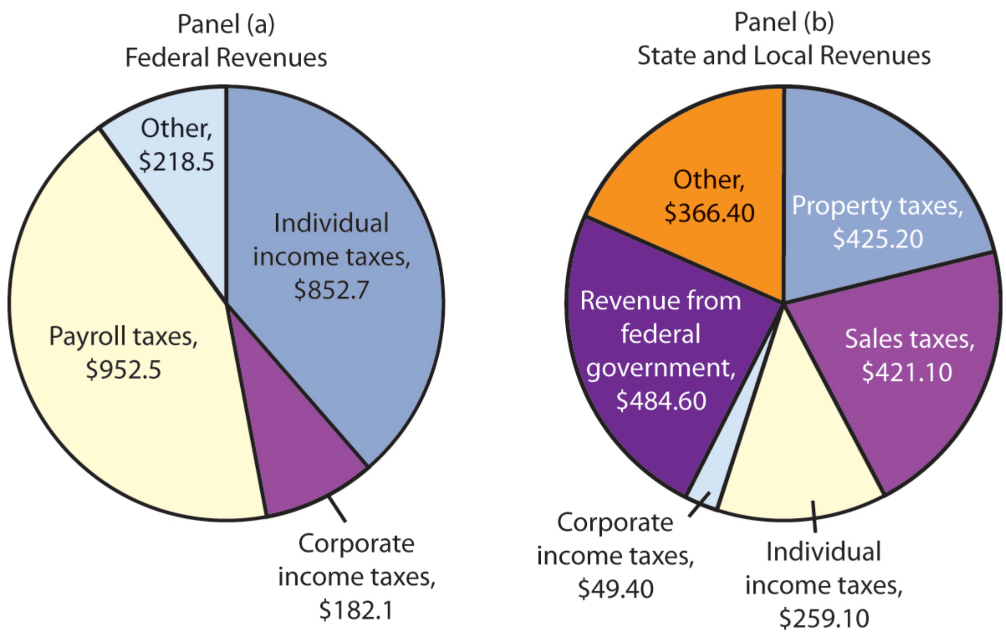
Source: Bureau of Economic Analysis, NIPA Table 1.1 and 3.1 (November 23, 2010 revision).

Taxes

Taxes affect the relationship between real GDP and personal disposable income; they therefore affect consumption. They also influence investment decisions. Taxes imposed on firms affect the profitability of investment decisions and therefore affect the levels of investment firms will choose. Payroll taxes imposed on firms affect the costs of hiring workers; they therefore have an impact on employment and on the real wages earned by workers.

The bulk of federal receipts come from the personal income tax and from payroll taxes. State and local tax receipts are dominated by property taxes and sales taxes. The federal government, as well as state and local governments, also collects taxes imposed on business firms, such as taxes on corporate profits. Figure 11.4 "The Composition of Federal, State, and Local Revenues" shows the composition of federal, state, and local receipts in a recent year.

Figure 11.4 The Composition of Federal, State, and Local Revenues



Federal receipts come primarily from payroll taxes and from personal taxes such as the personal income tax. State and local tax receipts come from a variety of sources; the most important are property taxes, sales taxes, income taxes, and grants from the federal government. Revenue shares are for 2009.

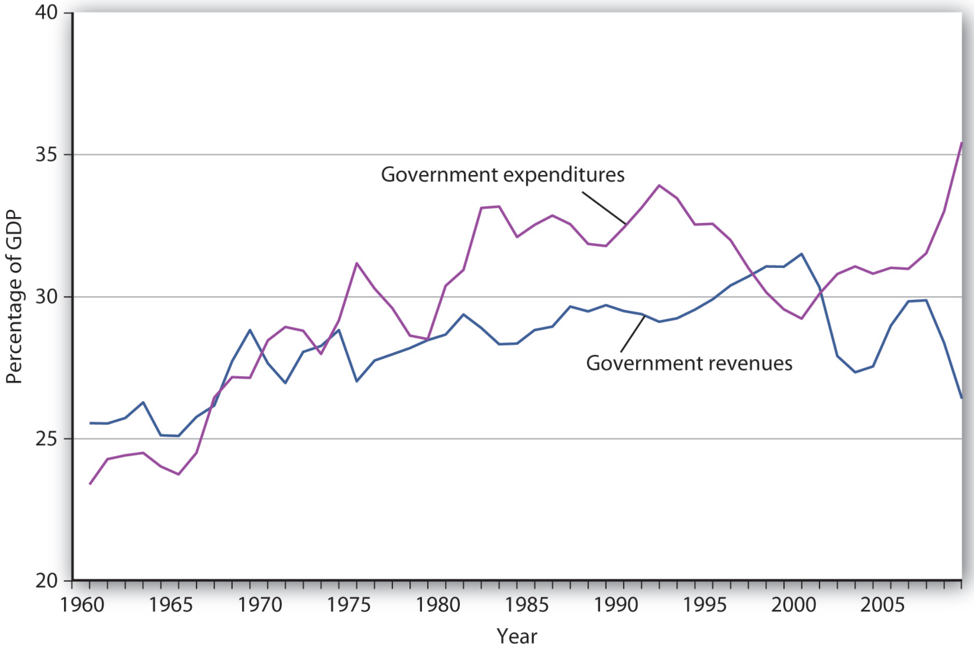
Source: Bureau of Economic Analysis, NIPA Table 3.2 and 3.3 (November 23, 2010 revision).

The Government Budget Balance

The government's budget balance is the difference between the government's revenues and its expenditures. A budget surplus occurs if government revenues exceed expenditures. A budget deficit occurs if government expenditures exceed revenues. The minus sign is often omitted when reporting a deficit. If the budget surplus equals zero, we say the government has a balanced budget.

Figure 11.5 "Government Revenue and Expenditure as a Percentage of GDP, 1960–2009" compares federal, state, and local government revenues to expenditures relative to GDP since 1960. The government's budget was generally in surplus in the 1960s, then mostly in deficit since, except for a brief period between 1998 and 2001. Bear in mind that these data are for all levels of government.

Figure 11.5 Government Revenue and Expenditure as a Percentage of GDP, 1960–2009



The government's budget was generally in surplus in the 1960s, then mostly in deficit since, except for a brief period between 1998 and 2001.

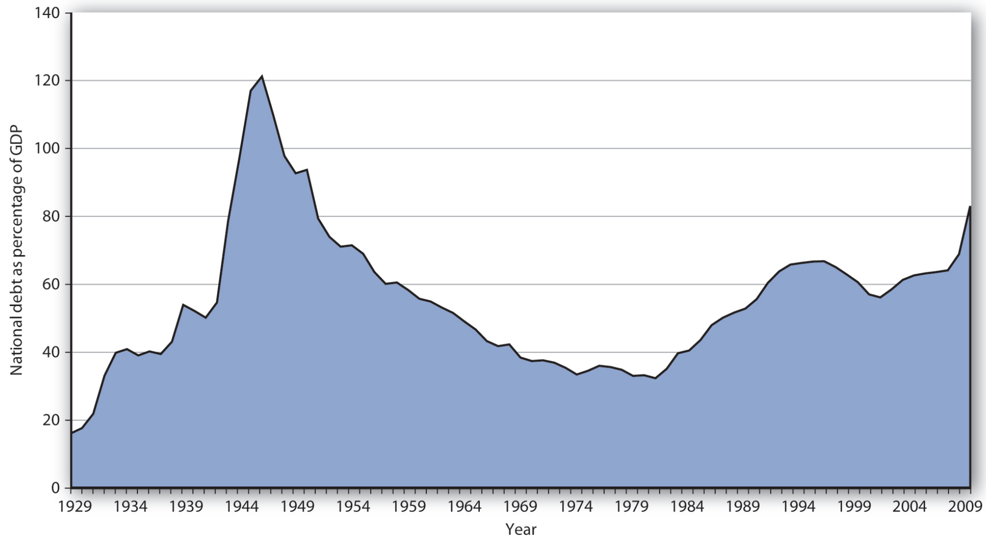
Source: Bureau of Economic Analysis, NIPA Table 1.1 and 3.1 (November 23, 2010 revision).

The administration of George W. Bush saw a large increase in the federal deficit. In part, this is the result of the government's response to the terrorist attacks in 2001. It also results, however, from large increases in federal spending at all levels together with tax cuts in 2001, 2002, and 2003. The federal deficit is projected to be even larger during the administration of Barack Obama.

The National Debt

The national debt is the sum of all past federal deficits, minus any surpluses. Figure 11.6 "The National Debt and the Economy, 1929–2009" shows the national debt as a percentage of GDP. It suggests that, relative to the level of economic activity, the debt is well below the levels reached during World War II. The ratio of debt to GDP rose from 1981 to 1996 and fell in the last years of the 20th century; it began rising again in 2002.

Figure 11.6 The National Debt and the Economy, 1929–2009



The national debt relative to GDP is much smaller today than it was during World War II. The ratio of debt to GDP rose from 1981 to 1996 and fell in the last years of the 20th century; it began rising again in 2002.

Sources: Data for 1929–1938 from *Historical Statistics of the United States, Colonial Times to 1957*—not strictly comparable with later data. Data for remaining years from Office of Management and Budget, *Budget of the United States Government, Fiscal Year 2011, Historical Tables.*

Judged by international standards, the U.S. national debt relative to its GDP is somewhat above average among developed nations. Figure 11.7 "Debts and Deficits for 26 Nations, 2008" shows national debt as a percentage of GDP for 26 countries in 2008. It also shows deficits or surpluses as a percentage of GDP.

Figure 11.7 Debts and Deficits for 26 Nations, 2008

Country	National Debt as percentage of GDP	Deficit (−) or surplus as percentage of GDP
Japan	172.1	−2.7
Italy	114.4	−2.7
Greece	102.6	−7.8
Iceland	96.3	−13.6
Belgium	93.5	−1.2
Hungary	77.0	−3.7
France	75.7	−3.4
Portugal	75.2	−2.8
United States	70.0	−6.5
Canada	69.7	0.1
Germany	68.8	−0.2
Austria	66.2	−0.5
Netherlands	65.8	0.7
United Kingdom	56.8	−5.3
Norway	56.0	18.8
Poland	54.0	−3.7
Sweden	47.1	2.5
Spain	47.0	−4.1
Switzerland	44.0	1.6
Finland	40.7	4.4
Denmark	39.8	3.4
Slovak Republic	30.8	−2.3
Korea	26.8	3.3
New Zealand	25.3	3.1
Luxembourg	16.3	2.5
Australia	14.3	1.0
Average	63.3	

The chart shows national debt as a percentage of GDP and deficits or surpluses as a percentage of GDP in 2008. The national debt of the United States relative to its GDP was somewhat above average among these nations.

Source: Organisation for Economic Co-operation and Development (OECD). *Factbook* 2010. OECD Publishing, May 25, 2010.

Key Takeaways

- Over the last 50 years, government purchases fell from about 20% of U.S. GDP to below 20%, but have been rising over the last decade.
- Transfer payment spending has risen sharply, both in absolute terms and as a percentage of real GDP since 1960.

- The bulk of federal revenues comes from income and payroll taxes. State and local revenues consist primarily of sales and property taxes.
- The government budget balance is the difference between government revenues and government expenditures.
- The national debt is the sum of all past federal deficits minus any surpluses.

Try It!

What happens to the national debt when there is a budget surplus? What happens to it when there is a budget deficit? What happens to the national debt if there is a decrease in a surplus? What happens to it if the deficit falls?

Case in Point: Generational Accounting

Figure 27.8

Kevin Krejci – Our National Debt – CC BY 2.0.

One method of assessing the degree to which current fiscal policies affect future generations is through a device introduced in the early 1990s called generational accounting. It measures the impact of current fiscal policies on different generations in the economy, including future generations. Generational accounting is now practiced by governments in many countries, including the United States and the European Union.

As populations age, the burden of current fiscal policy is increasingly borne by younger people in the population. In most countries, economists computing generational accounts have found that people age 40 or below will pay more in taxes than they receive in transfer payments, while those age 60 or above will receive more in transfers than they pay in taxes. The differences are huge. According to a recent study by Jagadeesh Gokhale, summarized in the table below, in 2004 in the United States, a male age 30 could expect to pay $201,300 more than he receives in government transfers during his lifetime, while another male age 75 could expect to receive $171,100 more in transfers than he paid in taxes during his lifetime. That is a difference of $372,400! For future generations, those born after the year 2004, the difference is even more staggering. A male born after the year 2005 can expect to pay $332,200 more in taxes than he will receive in transfer payments. For a woman, the differences are also large but not as great. A woman age 30 in 2004 could expect to pay $30,200 more in taxes than she will receive in transfers during her lifetime, while a woman age 75 could expect to receive transfers of $184,100 in excess of her lifetime tax burden.

The table below gives generational accounting estimates for the United States for the year 2004 for males and females. Notice that the net burden on females is much lower than for males. That is because women live longer than men and thus receive Social Security and Medicare benefits over a longer period of time. Women also have lower labor force participation rates and earn less than men, and pay lower taxes as a result.

Table 27.1 Generational Accounts for the United States (thousands of 2004 dollars)

Year of birth	Age in 2004	Male	Female
2005 (future born)	−1	333.2	26.0
2004 (newborn)	0	104.3	8.1
1989	15	185.7	42.0
1974	30	201.3	30.2
1959	45	67.8	−54.1
1944	60	−162.6	−189.4
1929	75	−171.1	−184.1
1914	90	−65.0	−69.2

Generational accounting has its critics—for example, the table above only measures direct taxes and transfers but omits benefits from government spending on public goods and services. In addition, government spending programs can be modified, which would alter the impact on future generations. Nonetheless, it does help to focus attention on the sustainability of current fiscal policies. Can future generations pay for Social Security, Medicare, and other retirement and health care spending as currently configured? Should they be asked to do so?

Sources: Jagadeesh Gokhale, "Generational Accounting," *The New Palgrave Dictionary of Economics*, 2nd ed. (forthcoming).

Answer to Try It! Problem

A budget surplus leads to a decline in national debt; a budget deficit causes the national debt to grow. If there is a decrease in a budget surplus, national debt still declines but by less than it would have had the surplus not gotten smaller. If there is a decrease in the budget deficit, the national debt still grows, but by less than it would have if the deficit had not gotten smaller.

11.2 The Use of Fiscal Policy to Stabilize the Economy

Learning Objectives

1. Define automatic stabilizers and explain how they work.
2. Explain and illustrate graphically how discretionary fiscal policy works and compare the changes in aggregate demand that result from changes in government purchases, income taxes, and transfer payments.

Fiscal policy—the use of government expenditures and taxes to influence the level of economic activity—is the government counterpart to monetary policy. Like monetary policy, it can be used in an effort to close a recessionary or an inflationary gap.

Some tax and expenditure programs change automatically with the level of economic activity. We will examine these first. Then we will look at how discretionary fiscal policies work. Four examples of discretionary fiscal policy choices were the tax cuts introduced by the Kennedy, Reagan, and George W. Bush administrations and the increase in government purchases proposed by President Clinton in 1993. The 2009 fiscal stimulus bill passed in the first months of the administration of Barack Obama included both tax cuts and spending increases. All were designed to stimulate aggregate demand and close recessionary gaps.

Automatic Stabilizers

Certain government expenditure and taxation policies tend to insulate individuals from the impact of shocks to the economy. Transfer payments have this effect. Because more people become eligible for income supplements when income is falling, transfer payments reduce the effect of a change in real GDP on disposable personal income and thus help to insulate households from the impact of the change. Income taxes also have this effect. As incomes fall, people pay less in income taxes.

Any government program that tends to reduce fluctuations in GDP automatically is called an automatic stabilizer. Automatic stabilizers tend to increase GDP when it is falling and reduce GDP when it is rising.

To see how automatic stabilizers work, consider the decline in real GDP that occurred during the recession of 1990–1991. Real GDP fell 1.6% from the peak to the trough of that recession. The reduction in economic activity automatically reduced tax payments, reducing the impact of the downturn on disposable personal income. Furthermore, the reduction in incomes increased transfer payment spending, boosting disposable personal income further. Real disposable personal income thus fell by only 0.9% during the 1990–1991 recession, a much smaller percentage than the reduction in real GDP. Rising transfer payments and falling tax collections helped cushion households from the impact of the recession and kept real GDP from falling as much as it would have otherwise.

Automatic stabilizers have emerged as key elements of fiscal policy. Increases in income tax rates and unemployment benefits have enhanced their importance as automatic stabilizers. The introduction in the 1960s and 1970s of means-tested federal transfer payments, in which individuals qualify depending on their income, added to the nation's arsenal of automatic stabilizers. The advantage of automatic stabilizers is suggested by their name. As soon as income starts to change, they go to work. Because they affect disposable personal income directly, and because changes in disposable personal income are closely linked to changes in consumption, automatic stabilizers act swiftly to reduce the degree of changes in real GDP.

It is important to note that changes in expenditures and taxes that occur through automatic stabilizers do not shift the aggregate demand curve. Because they are automatic, their operation is already incorporated in the curve itself.

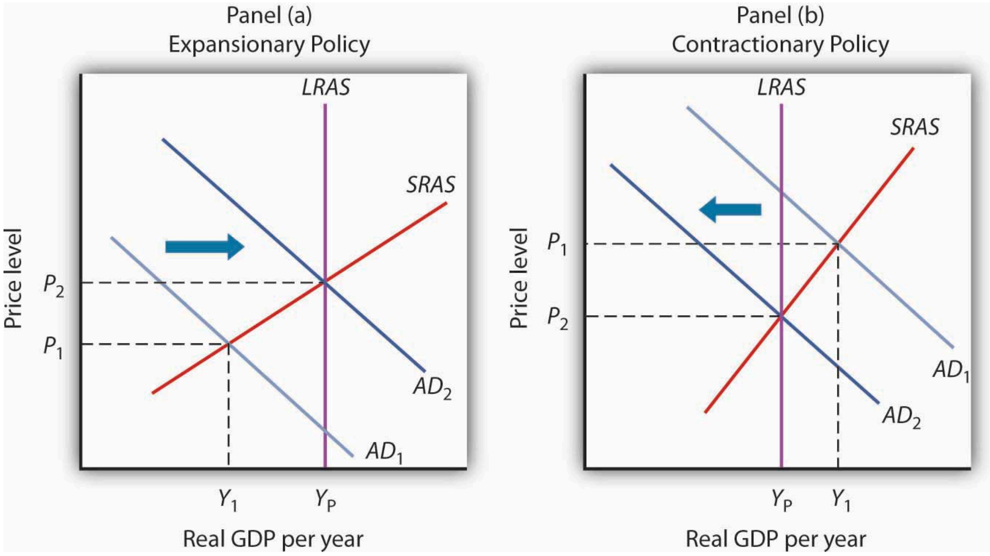
Discretionary Fiscal Policy Tools

As we begin to look at deliberate government efforts to stabilize the economy through fiscal policy choices, we note that most of the government's taxing and spending is for purposes other than economic stabilization. For example, the increase in defense spending in the early 1980s under President Ronald Reagan and in the administration of George W. Bush were undertaken primarily to promote national security. That the increased spending affected real GDP and employment was a by-product. The effect of such changes on real GDP and the price level is secondary, but it cannot be ignored. Our focus here, however, is on discretionary fiscal policy that is undertaken with the intention of stabilizing the economy. As we have seen, the tax cuts introduced by the Bush administration were justified as expansionary measures.

Discretionary government spending and tax policies can be used to shift aggregate demand. Expansionary fiscal policy might consist of an increase in government purchases or transfer payments, a reduction in taxes, or a combination of these tools to shift the aggregate demand curve to the right. A contractionary fiscal policy might involve a reduction in government purchases or transfer payments, an increase in taxes, or a mix of all three to shift the aggregate demand curve to the left.

Figure 11.9 "Expansionary and Contractionary Fiscal Policies to Shift Aggregate Demand" illustrates the use of fiscal policy to shift aggregate demand in response to a recessionary gap and an inflationary gap. In Panel (a), the economy produces a real GDP of Y_1, which is below its potential level of Y_p. An expansionary fiscal policy seeks to shift aggregate demand to AD_2 in order to close the gap. In Panel (b), the economy initially has an inflationary gap at Y_1. A contractionary fiscal policy seeks to reduce aggregate demand to AD_2 and close the gap. Now we shall look at how specific fiscal policy options work. In our preliminary analysis of the effects of fiscal policy on the economy, we will assume that at a given price level these policies do not affect interest rates or exchange rates. We will relax that assumption later in the chapter.

Figure 11.9 Expansionary and Contractionary Fiscal Policies to Shift Aggregate Demand



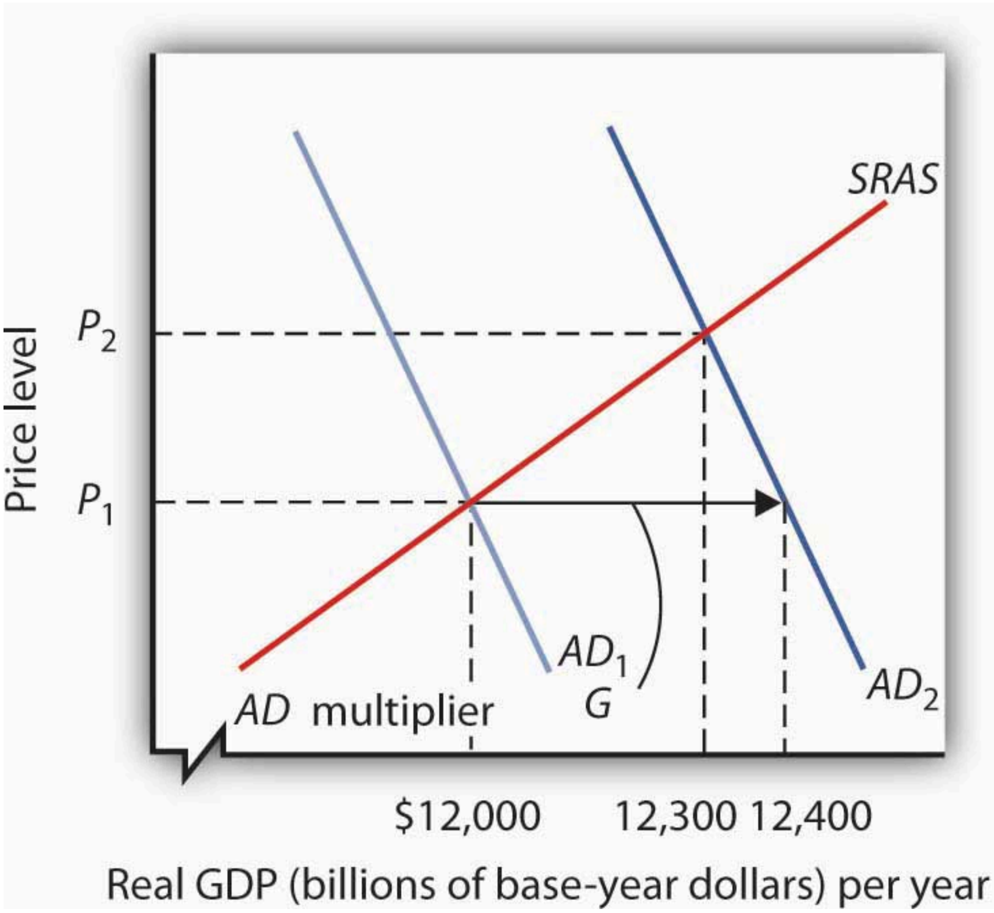
In Panel (a), the economy faces a recessionary gap ($Y_P - Y_1$). An expansionary fiscal policy seeks to shift aggregate demand to AD_2 to close the gap. In Panel (b), the economy faces an inflationary gap ($Y_1 - Y_P$). A contractionary fiscal policy seeks to reduce aggregate demand to AD_2 to close the gap.

Changes in Government Purchases

One policy through which the government could seek to shift the aggregate demand curve is a change in government purchases. We learned that the aggregate demand curve shifts to the right by an amount equal to the initial change in government purchases times the multiplier. This multiplied effect of a change in government purchases occurs because the increase in government purchases increases income, which in turn increases consumption. Then, part of the impact of the increase in aggregate demand is absorbed by higher prices, preventing the full increase in real GDP that would have occurred if the price level did not rise.

Figure 11.10 "An Increase in Government Purchases" shows the effect of an increase in government purchases of $200 billion. The initial price level is P_1 and the initial equilibrium real GDP is $12,000 billion. Suppose the multiplier is 2. The $200 billion increase in government purchases increases the total quantity of goods and services demanded, at a price level of P_1, by $400 billion (the $200 billion increase in government purchases times the multiplier) to $12,400 billion. The aggregate demand thus shifts to the right by that amount to AD_2. The equilibrium level of real GDP rises to $12,300 billion, and the price level rises to P_2.

Figure 11.10 An Increase in Government Purchases



The economy shown here is initially in equilibrium at a real GDP of $12,000 billion and a price level of P_1. An increase of $200 billion in the level of government purchases (ΔG) shifts the aggregate demand curve to the right by $400 billion to AD_2. The equilibrium level of real GDP rises to $12,300 billion, while the price level rises to P_2.

A reduction in government purchases would have the opposite effect. The aggregate demand curve would shift to the left by an amount equal to the initial change in government purchases times the multiplier. Real GDP and the price level would fall.

Changes in Business Taxes

One of the first fiscal policy measures undertaken by the Kennedy administration in the 1960s was an investment tax credit. An investment tax credit allows a firm to reduce its tax liability by a percentage of the investment it undertakes during a particular period. With an investment tax credit of 10%, for example, a firm that engaged in $1 million worth of investment during a year could reduce its tax liability for that year by $100,000. The investment tax credit introduced by the Kennedy administration was later repealed. It was reintroduced during the Reagan administration in 1981, then abolished by the Tax Reform Act of 1986. President Clinton called for a new investment tax credit in 1993 as part of his job stimulus proposal, but that proposal was rejected by Congress. The Bush administration reinstated the investment tax credit as part of its tax cut package.

An investment tax credit is intended, of course, to stimulate additional private sector investment. A reduction in the tax rate on corporate profits would be likely to have a similar effect. Conversely, an increase in the corporate income tax rate or a reduction in an investment tax credit could be expected to reduce investment.

A change in investment affects the aggregate demand curve in precisely the same manner as a change in government purchases. It shifts the aggregate demand curve by an amount equal to the initial change in investment times the multiplier.

An increase in the investment tax credit, or a reduction in corporate income tax rates, will increase investment and shift the aggregate demand curve to the right. Real GDP and the price level will rise. A reduction in the investment tax credit, or an increase in corporate income tax rates, will reduce investment and shift the aggregate demand curve to the left. Real GDP and the price level will fall[1].

Changes in Income Taxes

Income taxes affect the consumption component of aggregate demand. An increase in income taxes reduces disposable personal income and thus reduces consumption (but by less than the change in disposable personal income). That shifts the aggregate demand curve leftward by an amount equal to the initial change in consumption that the change in income taxes produces times the multiplier[2]. A reduction in income taxes increases disposable personal income, increases consumption (but by less than the change in disposable personal income), and increases aggregate demand.

Suppose, for example, that income taxes are reduced by $200 billion. Only some of the increase in disposable personal income will be used for consumption and the rest will be saved. Suppose the initial increase in consumption is $180 billion. Then the shift in the aggregate demand curve will be a multiple of $180 billion; if the multiplier is 2, aggregate demand will shift to the right by $360 billion. Thus, as compared to the $200-billion increase in government purchases that we saw in Figure 11.10 "An Increase in Government Purchases", the shift in the aggregate demand curve due to an income tax cut is somewhat less, as is the effect on real GDP and the price level.

Changes in Transfer Payments

Changes in transfer payments, like changes in income taxes, alter the disposable personal income of households and thus affect their consumption, which is a component of aggregate demand. A change in transfer payments will thus shift the aggregate demand curve because it will affect consumption. Because consumption will change by less than the change in disposable personal income, a change in transfer payments of some amount will result in a smaller change in real GDP than would a change in government purchases of the same amount. As with income taxes, a $200-billion increase in transfer payments will shift the aggregate demand curve to the right by less than the $200-billion increase in government purchases that we saw in Figure 11.10 "An Increase in Government Purchases".

Table 11.2 "Fiscal Policy in the United States Since 1964" summarizes U.S. fiscal policies undertaken to shift aggregate demand since the 1964 tax cuts. We see that expansionary policies have been chosen in response to recessionary gaps and that contractionary policies have been chosen in response to inflationary gaps. Changes in government purchases and in taxes have been the primary tools of fiscal policy in the United States.

Table 11.2 Fiscal Policy in the United States Since 1964

Year	Situation	Policy response
1968	Inflationary gap	A temporary tax increase, first recommended by President Johnson's Council of Economic Advisers in 1965, goes into effect. This one-time surcharge of 10% is added to individual income tax liabilities.
1969	Inflationary gap	President Nixon, facing a continued inflationary gap, orders cuts in government purchases.
1975	Recessionary gap	President Ford, facing a recession induced by an OPEC oil-price increase, proposes a temporary 10% tax cut. It is passed almost immediately and goes into effect within two months.
1981	Recessionary gap	President Reagan had campaigned on a platform of increased defense spending and a sharp cut in income taxes. The tax cuts are approved in 1981 and are implemented over a period of three years. The increased defense spending begins in 1981. While the Reagan administration rejects the use of fiscal policy as a stabilization tool, its policies tend to increase aggregate demand early in the 1980s.
1992	Recessionary gap	President Bush had rejected the use of expansionary fiscal policy during the recession of 1990–1991. Indeed, he agreed late in 1990 to a cut in government purchases and a tax increase. In a campaign year, however, he orders a cut in withholding rates designed to increase disposable personal income in 1992 and to boost consumption.
1993	Recessionary gap	President Clinton calls for a $16-billion jobs package consisting of increased government purchases and tax cuts aimed at stimulating investment. The president says the plan will create 500,000 new jobs. The measure is rejected by Congress.
2001	Recessionary gap	President Bush campaigned to reduce taxes in order to reduce the size of government and encourage long-term growth. When he took office in 2001, the economy was weak and the $1.35-billion tax cut was aimed at both long-term tax relief and at stimulating the economy in the short term. It included, for example, a personal income tax rebate of $300 to $600 per household. With unemployment still high a couple of years into the expansion, another tax cut was passed in 2003.
2008	Recessionary gap	Fiscal stimulus package of $150 billion to spur economy. It included $100 billion in tax rebates and $50 in tax cuts for businesses.
2009	Recessionary gap	Fiscal stimulus package of $787 billion included tax cuts and increased government spending passed in early days of President Obama's administration.

Key Takeaways

- Discretionary fiscal policy may be either expansionary or contractionary.
- A change in government purchases shifts the aggregate demand curve at a given price level by an amount equal to the initial change in government purchases times the multiplier. The change in real GDP, however, will be reduced by the fact that the price level will change.
- A change in income taxes or government transfer payments shifts the aggregate demand curve by a multiple of the initial change in consumption (which is less than the change in personal disposable income) that the change in income taxes or transfer payments causes. Then, the change in real GDP will be reduced by the fact that the price level will change.
- A change in government purchases has a larger impact on the aggregate demand curve than does an equal change in income taxes or transfers.
- Changes in business tax rates, including an investment tax credit, can be used to influence the level of investment and thus the level of aggregate demand.

Try It!

Suppose the economy has an inflationary gap. What fiscal policies might be used to close the gap? Using the model of aggregate demand and aggregate supply, illustrate the effect of these policies.

Case in Point: Post–World War II Experiences with Fiscal Policy in the United States

Figure 27.11

Commonwealth Club – Christina Romer – CC BY 2.0.

Christina Romer, tapped by Barack Obama to head the Council of Economic Advisers, has a long history of writing on economic history. Much of her work focuses on the macroeconomic performance of the United States economy over the past 100-plus years and hence also involves painstaking work to construct historical data series.

One such study titled "Changes in Business Cycles: Evidence and Explanations" draws on a number of her research efforts to compare economic fluctuations before World War I to those after World War II in order to see if the advent of macroeconomic stabilization policy has affected macroeconomic performance. After first showing that macroeconomic performance has not improved as markedly as we might think (excluding the interwar period when "all hell broke loose in the American economy"), she does conclude that monetary and fiscal policies to influence aggregate demand since World War II have "served to dampen many recessions and counteract some shocks entirely."

She notes that before World War I, changes in macroeconomic policy could not have affected economic performance, because the government was simply too small, with, for example, government spending as a percent of GNP averaging between 1.5% and 2.5% between 1901 and 1916. During that period, the government did operate under specified monetary standards and banking regulations, but the Federal Reserve was not created until 1914, so there was no monetary institution to respond to macroeconomic instability. Thus, macroeconomic policy can truly be seen as a post–World War II phenomenon.

Germaine to the focus on fiscal policy in this chapter, Romer found that discretionary fiscal policy after World War II contributed 0.5 percentage points to the rate of growth of real GDP in years following the troughs of recessions, while automatic stabilizers contributed 0.85 percentage points. Adding in the average contribution of monetary policy of 1.5 percentage points, macroeconomic policy in total contributed 2.85 percentage points to the average actual growth of GDP in the years following troughs of 4.6%. She also concluded that macroeconomic policies likely prevented some recessions or near-recessions. For example, automatic stabilizers muted fluctuations in years of extreme changes in GDP, up or down, by 1 to 2 percentage points in absolute value and fluctuations in years of moderate changes in GDP by about 0.5 percentage points in absolute value.

Especially in light of the active use of both monetary and fiscal policies to counter the recession that began in December 2007, she also found that there has been a rise in policy-induced recessions and that this phenomenon explains both why output and other macroeconomic variables have not been more stable in the past half-century and why post–World War II business cycles have been in the moderate range. In particular, she argues that the Fed has generally been too expansionary when the economy was growing, which has led to inflation. Then the Fed has used contractionary policy to reduce inflation. She concludes, "In essence, we have replaced the prewar boom-bust cycle driven by animal spirits and financial panics with the postwar boom-bust cycle driven by policy."

Source: Christina Romer, "Changes in Business Cycles: Evidence and Explanations," *Journal of Economic Perspectives* 13, no. 2 (Spring 1999): 23–44.

Answer to Try It! Problem

Fiscal policies that could be used to close an inflationary gap include reductions in government purchases and transfer payments and increases in taxes. As shown in Panel (b) of Figure 27.9 "Expansionary and Contractionary Fiscal Policies to Shift Aggregate Demand", the goal would be to shift the aggregate demand curve to the left so that it will intersect the short-run aggregate supply curve at Y_P.

[1]Investment also affects the long-run aggregate supply curve, since a change in the capital stock changes the potential level of real GDP. We examined this earlier in the chapter on economic growth.

[2] A change in tax rates will change the value of the multiplier. The reason is explained in another chapter.

11.3 Issues in Fiscal Policy

Learning Objectives

1. Explain how the various kinds of lags influence the effectiveness of discretionary fiscal policy.
2. Explain and illustrate graphically how crowding out (and its reverse) influences the impact of expansionary or contractionary fiscal policy.
3. Discuss the controversy concerning which types of fiscal policies to use, including the arguments from supply-side economics.

The discussion in the previous section about the use of fiscal policy to close gaps suggests that economies can be easily stabilized by government actions to shift the aggregate demand curve. However, as we discovered with monetary policy in the previous chapter, government attempts at stabilization are fraught with difficulties.

Lags

Discretionary fiscal policy is subject to the same lags that we discussed for monetary policy. It takes some time for policy makers to realize that a recessionary or an inflationary gap exists—the *recognition lag*. Recognition lags stem largely from the difficulty of collecting economic data in a timely and accurate fashion. The current recession was not identified until October 2008, when the Business Cycle Dating Committee of the National Bureau of Economic Research announced that it had begun in December 2007. Then, more time elapses before a fiscal policy, such as a change in government purchases or a change in taxes, is agreed to and put into effect—the *implementation lag*. Finally, still more time goes by before the policy has its full effect on aggregate demand—the *impact lag*.

Changes in fiscal policy are likely to involve a particularly long implementation lag. A tax cut was proposed to presidential candidate John F. Kennedy in 1960 as a means of ending the recession that year. He recommended it to Congress in 1962. It was not passed until 1964, three years after the recession had ended. Some economists have concluded that the long implementation lag for discretionary fiscal policy makes this stabilization tool ineffective. Fortunately, automatic stabilizers respond automatically to changes in the economy. They thus avoid not only the implementation lag but also the recognition lag.

The implementation lag results partly from the nature of bureaucracy itself. The CBO estimate that only a portion of the spending for the stimulus plan passed in 2009 will be spent in the next two years is an example of the implementation lag. Government spending requires bureaucratic approval of that spending. For example, a portion of the stimulus plan must go through the Department of Energy. One division of the department focuses on approving loan guarantees for energy-saving industrial projects. It was created early in 2007 as part of another effort to stimulate economic activity. A Minnesota company, Sage Electrochromics, has developed a process for producing windows that can be darkened or lightened on demand to reduce energy use in buildings. Sage applied two years ago for a guarantee on a loan of $66 million to build a plant that would employ 250 workers. Its application has not been approved. In fact, the loan approval division, which will be crucial for projects in the stimulus plan, has never approved any application made to it in its two years in existence!

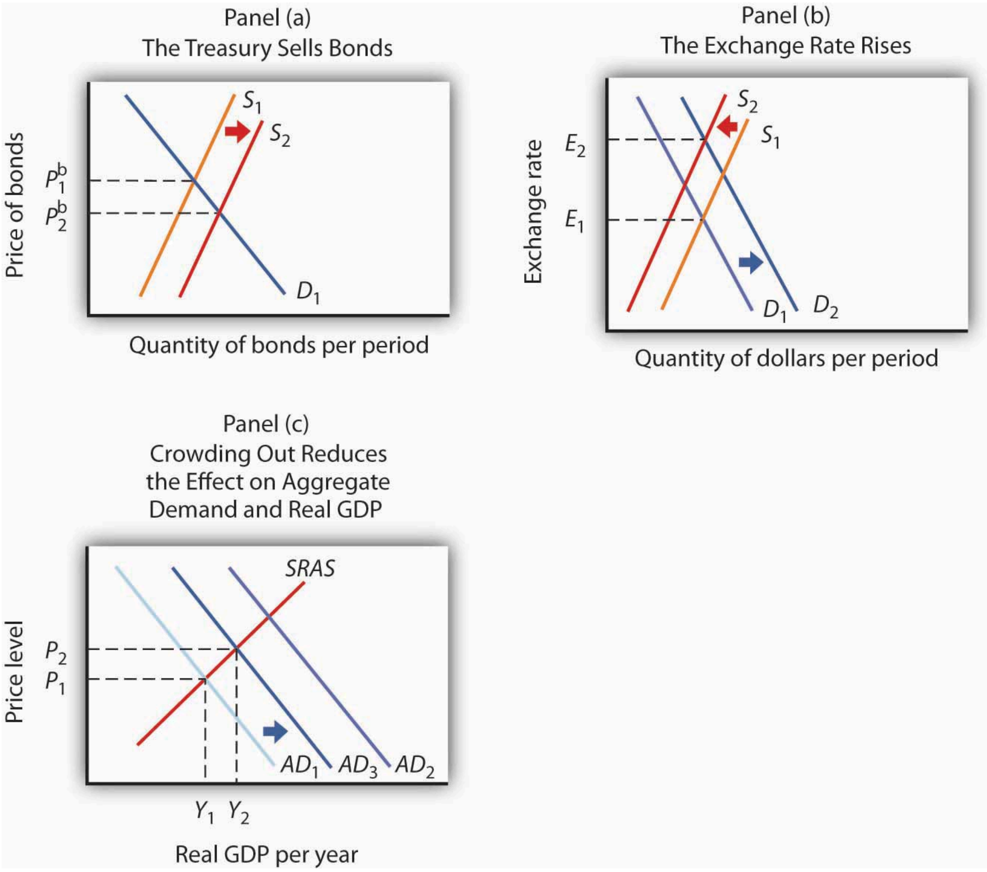
Energy Secretary Steven Chu, a Nobel Prize-winning physicist, recognizes the urgency of the problem. In an interview with the *Wall Street Journal*, Dr. Chu said that his agency would have to do better. “Otherwise, it’s just going to be a bust,” he said (Power & King Jr., 2009).

Crowding Out

Because an expansionary fiscal policy either increases government spending or reduces revenues, it increases the government budget deficit or reduces the surplus. A contractionary policy is likely to reduce a deficit or increase a surplus. In either case, fiscal policy thus affects the bond market. Our analysis of monetary policy showed that developments in the bond market can affect investment and net exports. We shall find in this section that the same is true for fiscal policy.

Figure 11.12 “An Expansionary Fiscal Policy and Crowding Out” shows the impact of an expansionary fiscal policy: an increase in government purchases. The increase in government purchases increases the deficit or reduces the surplus. In either case, the Treasury will sell more bonds than it would have otherwise, shifting the supply curve for bonds to the right in Panel (a). That reduces the price of bonds, raising the interest rate. The increase in the interest rate reduces the quantity of private investment demanded. The higher interest rate increases the demand for and reduces the supply of dollars in the foreign exchange market, raising the exchange rate in Panel (b). A higher exchange rate reduces net exports. Panel (c) shows the effects of all these changes on the aggregate demand curve. Before the change in government purchases, the economy is in equilibrium at a real GDP of Y_1, determined by the intersection of AD_1 and the short-run aggregate supply curve. The increase in government expenditures would shift the curve outward to AD_2 if there were no adverse impact on investment and net exports. But the reduction in investment and net exports partially offsets this increase. Taking the reduction in investment and net exports into account means that the aggregate demand curve shifts only to AD_3. The tendency for an expansionary fiscal policy to reduce other components of aggregate demand is called crowding out. In the short run, this policy leads to an increase in real GDP to Y_2 and a higher price level, P_2.

Figure 11.12 An Expansionary Fiscal Policy and Crowding Out



In Panel (a), increased government purchases are financed through the sale of bonds, lowering their price to $P^b{}_2$. In Panel (b), the higher interest rate causes the exchange rate to rise, reducing net exports. Increased government purchases would shift the aggregate demand curve to AD_2 in Panel (c) if there were no crowding out. Crowding out of investment and net exports, however, causes the aggregate demand curve to shift only to AD_3. Then a higher price level means that GDP rises only to Y_2.

Crowding out reduces the effectiveness of any expansionary fiscal policy, whether it be an increase in government purchases, an increase in transfer payments, or a reduction in income taxes. Each of these policies increases the deficit and thus increases government borrowing. The supply of bonds increases, interest rates rise, investment falls, the exchange rate rises, and net exports fall.

Note, however, that it is private investment that is crowded out. The expansionary fiscal policy could take the form of an increase in the investment component of government purchases. As we have learned, some government purchases are for goods, such as office supplies, and services. But the government can also purchase investment items, such as roads and schools. In that case, government investment may be crowding out private investment.

The reverse of crowding out occurs with a contractionary fiscal policy–a cut in government purchases or transfer payments, or an increase in taxes. Such policies reduce the deficit (or increase the surplus) and thus reduce government borrowing, shifting the supply curve for bonds to the left. Interest rates drop, inducing a greater quantity of investment. Lower interest rates also reduce the demand for and increase the supply of dollars, lowering the exchange rate and boosting net exports. This phenomenon is known as "crowding in."

Crowding out and crowding in clearly weaken the impact of fiscal policy. An expansionary fiscal policy has less punch; a contractionary policy puts less of a damper on economic activity. Some economists argue that these forces are so powerful that a change in fiscal policy will have no effect on aggregate demand. Because empirical studies have been inconclusive, the extent of crowding out (and its reverse) remains a very controversial area of study.

Also, the fact that government deficits today may reduce the capital stock that would otherwise be available to future generations does not imply that such deficits are wrong. If, for example, the deficits are used to finance public sector investment, then the reduction in private capital provided to the future is offset by the increased provision of public sector capital. Future generations may have fewer office buildings but more schools.

Choice of Policy

Suppose Congress and the president agree that something needs to be done to close a recessionary gap. We have learned that fiscal policies that increase government purchases, reduce taxes, or increase transfer payments—or do a combination of these—all have the potential, theoretically, to raise real GDP. The government must decide which kind of fiscal policy to employ. Because the decision makers who determine fiscal policy are all elected politicians, the choice among the policy options available is an intensely political matter, often reflecting the ideology of the politicians.

For example, those who believe that government is too big would argue for tax cuts to close recessionary gaps and for spending cuts to close inflationary gaps. Those who believe that the private sector has failed to provide adequately a host of services that would benefit society, such as better education or public transportation systems, tend to advocate increases in government purchases to close recessionary gaps and tax increases to close inflationary gaps.

Another area of contention comes from those who believe that fiscal policy should be constructed primarily so as to promote long-term growth. Supply-side economics is the school of thought that promotes the use of fiscal policy to stimulate long-run aggregate supply. Supply-side economists advocate reducing tax rates in order to encourage people to work more or more individuals to work and providing investment tax credits to stimulate capital formation.

While there is considerable debate over how strong the supply-side effects are in relation to the demand-side effects, such considerations may affect the choice of policies. Supply-siders tend to favor tax cuts over increases in government purchases or increases in transfer payments. President Reagan advocated tax cuts in 1981 on the basis of their supply-side effects. Coupled with increased defense spending in the early 1980s, fiscal policy under Mr. Reagan clearly stimulated aggregate demand by increasing both consumption and investment. Falling inflation and accelerated growth are signs that supply-side factors may also have been at work during that period. President George W. Bush's chief economic adviser, N. Gregory Mankiw, argued that the Bush tax cuts would encourage economic growth, a supply-side argument. Mr. Bush's next chief economic adviser, Ben Bernanke, who became the next chairman of the Federal Reserve Board in 2006, made a similar argument and urged that the Bush tax cuts be made permanent.

Finally, even when there is agreement to stimulate the economy, say through increasing government expenditures on highways, the *how* question remains. How should the expenditures be allocated? Specifically, which states should the highways run through? Each member of Congress has a political stake in the outcome. These types of considerations make the implementation lag particularly long for fiscal policy.

Key Takeaways

- Discretionary fiscal policy involves the same kind of lags as monetary policy. However, the

implementation lag in fiscal policy is likely to be more pronounced, while the impact lag is likely to be less pronounced.
- Expansionary fiscal policy may result in the crowding out of private investment and net exports, reducing the impact of the policy. Similarly, contractionary policy may "crowd in" additional investment and net exports, reducing the contractionary impact of the policy.
- Supply-side economics stresses the use of fiscal policy to stimulate economic growth. Advocates of supply-side economics generally favor tax cuts to stimulate economic growth.

Try It!

Do the following hypothetical situations tend to enhance or make more difficult the use of fiscal policy as a stabilization tool?

1. Better and more speedily available data on the state of the economy
2. A finding that private sector investment spending is not much affected by interest rate changes
3. A finding that the supply-side effects of a tax cut are substantial

Case in Point: Crowding Out in Canada

Figure 27.13

Moni cah – [canadian dollars]() – CC BY-NC-ND 2.0.

In an intriguing study, economist Baotai Wang examined the degree of crowding out of Canadian private investment as a result of government expenditures from 1961–2000. What made Professor Wang's analysis unusual was that he divided Canadian government expenditures into five categories: expenditures for health and education, expenditures for capital and infrastructure, expenditures for the protection of persons and property (which included defense spending), expenditures for debt services, and expenditures for government and social services.

Mr. Wang found that only government expenditures for capital and infrastructure crowded out private investment. While these expenditures reduced private investment, they represented increased public sector investment for things such as highways and ports.

Expenditures for health and education actually "crowded in" private sector investment. These expenditures, Mr. Wang argued, represented increases in human capital. Such increases complement returns on private sector investment and therefore increase it.

Mr. Wang found that Canadian government expenditures for debt service, the protection of persons and property, and for government and social services had no effect on private sector investment. He argued that expenditures for protection of persons and property may involve some crowding out, but that they also stimulated private investment by firms winning government contracts for defense purchases. The same explanation could be applied to government expenditures for government and social services. These also include an element of investment in human capital.

His results suggest that crowding out depends on the nature of spending done by the government. Some kinds of spending clearly did not crowd out private sector investment in Canada.

Source: Baotai Wang, "Effects of Government Expenditure on Private Investment: Canadian Empirical Evidence," *Empirical Economics* 30, no. 2 (September 2005): 493–504.

Answers to Try It! Problems

1. Data on the economy that are more accurate and more speedily available should enhance the use of fiscal policy by reducing the length of the recognition lag.
2. If private sector investment does not respond much to interest rate changes, then there will be less crowding out when expansionary policies are undertaken. That is, the rising interest rates that accompany expansionary fiscal policy will not reduce investment spending much, making the shift in the aggregate demand curve to the right greater than it would be otherwise. Also, the use of contractionary fiscal policy would be more effective, since the fall in interest rates would "invite in" less investment spending, making the shift in the aggregate demand curve to the left greater than it would otherwise be.
3. Large supply-side effects enhance the impact of tax cuts. For a given expansionary policy, without the supply-side effects, GDP would advance only to the point where the aggregate demand curve intersects the short-run aggregate supply curve. With the supply-side effects, both the short-run and long-run aggregate supply curves shift to the right. The intersection of the AD curve with the now increased short-run aggregate supply curve will be farther to the right than it would have been in the absence of the supply-side effects. The potential level of real GDP will also increase.

References

Power, S., and Neil King, Jr., "Next Challenge on Stimulus: Spending All That Money," *Wall Street Journal*, February 13, 2009, p. A1.

11.4 Review and Practice

Summary

The government sector plays a major role in the economy. The spending, tax, and transfer policies of local, state, and federal agencies affect aggregate demand and aggregate supply and thus affect the level of real GDP and the price level. An expansionary policy tends to increase real GDP. Such a policy could be used to close a recessionary gap. A contractionary fiscal policy tends to reduce real GDP. A contractionary policy could be used to close an inflationary gap.

Government purchases of goods and services have a direct impact on aggregate demand. An increase in government purchases shifts the aggregate demand curve by the amount of the initial change in government purchases times the multiplier. Changes in personal income taxes or in the level of transfer payments affect disposable personal income. They change consumption, though initially by less than the amount of the change in taxes or transfers. They thus cause somewhat smaller shifts in the aggregate demand curve than do equal changes in government purchases.

There are several issues in the use of fiscal policies for stabilization purposes. They include lags associated with fiscal policy, crowding out, the choice of which fiscal policy tool to use, and the possible burdens of accumulating national debt.

Concept Problems

1. What is the difference between government expenditures and government purchases? How do the two variables differ in terms of their effect on GDP?
2. Federally funded student aid programs generally reduce benefits by $1 for every $1 that recipients earn. Do such programs represent government purchases or transfer payments? Are they automatic stabilizers?
3. Crowding out reduces the degree to which a change in government purchases influences the level of economic activity. Is it a form of automatic stabilizer?
4. The Case in Point on fiscal policy before World War I and after World War II mentions the idea of policy-induced recessions. Explain how this notion relates to the difficulties discussed in the text of using discretionary policy to stabilize the economy.
5. Suppose an economy has an inflationary gap. How does the government's actual budget deficit or surplus compare to the deficit or surplus it would have at potential output?
6. Suppose the president were given the authority to increase or decrease federal spending by as much as $100 billion in order to stabilize economic activity. Do you think this would tend to make the economy more or less stable?
7. Suppose the government increases purchases in an economy with a recessionary gap. How would this

policy affect bond prices, interest rates, investment, net exports, real GDP, and the price level? Show your results graphically.

8. Suppose the government cuts transfer payments in an economy with an inflationary gap. How would this policy affect bond prices, interest rates, investment, the exchange rate, net exports, real GDP, and the price level? Show your results graphically.
9. Suppose that at the same time the government undertakes expansionary fiscal policy, such as a cut in taxes, the Fed undertakes contractionary monetary policy. How would this policy affect bond prices, interest rates, investment, net exports, real GDP, and the price level? Show your results graphically.
10. Given the nature of the implementation lag discussed in the text, discuss possible measures that might reduce the lag.

Numerical Problems

1. Look up the table on Federal Receipts and Outlays, by Major Category, in the most recent *Economic Report of the President* available in your library or on the Internet.

    1. Complete the following table:

Category	**Total outlays**	**Percentage of total outlays**
National defense		
International affairs		
Health		
Medicare		
Income security		
Social Security		
Net interest		
Other		

    2. Construct a pie chart showing the percentages of spending for each category in the total.

2. Look up the table on ownership of U.S. Treasury securities in the most recent *Economic Report of the President* available in your library or on the Internet.

    1. Make a pie chart showing the percentage owned by various groups in the earliest year shown in the table.
    2. Make a pie chart showing the percentage owned by various groups in the most recent year shown in the table.
    3. What are some of the major changes in ownership of U.S. government debt over the period?

3. Suppose a country has a national debt of $5,000 billion, a GDP of $10,000 billion, and a budget deficit of $100 billion.
    1. How much will its new national debt be?
    2. Compute its debt-GDP ratio.
    3. Suppose its GDP grows by 1% in the next year and the budget deficit is again $100 billion. Compute its new level of national debt and its new debt-GDP ratio.
4. Suppose a country's debt rises by 10% and its GDP rises by 12%.
    1. What happens to the debt-GDP ratio?
    2. Does the relative level of the initial values affect your answer?
5. The data below show a country's national debt and its prime lending rate.

Year	National debt (billions of $)	Lending rate (%)
1992	4,064	6.0
1993	4,411	6.0
1994	4,692	8.5
1995	4,973	8.7
1996	5,224	8.3
1997	5,413	8.5

    1. Plot the relationship between national debt and the lending rate.
    2. Based on your graph, does crowding out appear to be a problem?
6. Suppose a country increases government purchases by $100 billion. Suppose the multiplier is 1.5 and the economy's real GDP is $5,000 billion.
    1. In which direction will the aggregate demand curve shift and by how much?
    2. Explain using a graph why the change in real GDP is likely to be smaller than the shift in the aggregate demand curve.
7. Suppose a country decreases government purchases by $100 billion. Suppose the multiplier is 1.5 and the economy's real GDP is $5,000 billion.
    1. In which direction will the aggregate demand curve shift and by how much?
    2. Explain using a graph why the change in real GDP is likely to be smaller than the shift in the aggregate demand curve.
8. Suppose a country decreases income taxes by $100 billion, and this leads to an increase in consumption spending of $90 billion. Suppose the multiplier is 1.5 and the economy's real GDP is $5,000 billion.
    1. In which direction will the aggregate demand curve shift and by how much?
    2. Explain using a graph why the change in real GDP is likely to be smaller than the shift in the aggregate demand curve.

9. Suppose a country increases income taxes by $100 billion, and this leads to a decrease in consumption spending of $90 billion. Suppose the multiplier is 1.5 and the economy's real GDP is $5,000 billion.
    1. In which direction will the aggregate demand curve shift and by how much?
    2. Explain using a graph why the change in real GDP is likely to be smaller than the shift in the aggregate demand curve.
10. Suppose a country institutes an investment tax credit, and this leads to an increase in investment spending of $100 billion. Suppose the multiplier is 1.5 and the economy's real GDP is $5,000 billion.
    1. In which direction will the aggregate demand curve shift and by how much?
    2. Explain using a graph why the change in real GDP is likely to be smaller than the shift in the aggregate demand curve.
11. Suppose a country repeals an investment tax credit, and this leads to a decrease in investment spending of $100 billion. Suppose the multiplier is 1.5 and the economy's real GDP is $5,000 billion.
    1. In which direction will the aggregate demand curve shift and by how much?
    2. Explain using a graph why the change in real GDP is likely to be smaller than the shift in the aggregate demand curve.
12. Explain why the shifts in the aggregate demand curves in questions 7 through 11 above are the same or different in absolute value.

Please share your supplementary material!

Have you created any supplementary materials for use with *Principles of Economics* such as test items, a question bank, or presentation slides? If so, please consider sharing your materials through this online textbook!

The University of Minnesota Libraries have begun to share supplementary materials through our online textbooks, using Google Drive to provide access to faculty and staff that may need to use those materials. Please see *Understanding Media and Culture* as an example:

https://open.lib.umn.edu/mediaandculture/

If you have created any supplementary material for *Principles of Economics*, please contact the University of Minnesota Libraries eLearning Support Initiative at learnlib@umn.edu for more information on how you can share your materials and make them available to faculty and students worldwide.