CHAPTER 1

Why Study Public Finance?

James Andrews/iStock/Getty Images

Questions to keep in mind

- What is public finance, and what are the key questions that the field addresses?
- What are the key facts about the size and growth of government and the distribution of taxes and spending?
- How does the debate over health care reform in the United States reflect the key questions of public finance?

On January 21, 2020, the United States reported its first official case of what would eventually become known as the Covid-19 virus. Within two months, the number of cases had risen to 50,342 with 701 deaths tied to the virus. Growth thereafter was explosive, with the number of new cases reaching a peak of 34,844 per day on April 9. By May 21, just four months later, 1,640,915 had been reported infected, and 97,895 had died.[1]

The pain was not just felt by the millions of individuals becoming ill and the tens of thousands who were dying; it was felt by the economy as well. Economic activity slowed dramatically beginning in March and fell further with state and local "shutdowns" that restricted commercial activity. The result was a historic collapse in both economic activity and employment. U.S. GDP fell by 31.2% in the second quarter of 2020, and the unemployment rate rose astronomically from 3.5% in January 2020 to 14.8% in April 2020, over four percentage points higher than the Great Recession in 2009.[2]

While the virus raged, so did debates over the proper role of the U.S. government in making a response to it at all levels. One key topic of disagreement was how aggressive the government should be in restricting economic activity in order to slow the spread of the virus. At the federal level, the Trump administration was opposed to virtually all restrictions on activity. President Trump dismissed the threat of Covid-19 despite warnings from the Center for Disease Control and U.S. health officials. On February 27, 2020, President Trump stated in a press conference that "[Covid-19 is] going to disappear. One day, it's like a miracle, it will disappear."[3] In early March, he followed up, "This was unexpected. … And it hit the world. And we're prepared, and we're doing a great job with it. And it will go away. Just stay calm. It will go away."[4]

Reactions to the virus at the state level varied widely. On March 19, California was the first state to issue a stay-at-home order. In the following six weeks, 38 other states issued stay-at-home orders, while 6 states issued only advisory orders and 5 states did not issue any order.[5] In New York, Governor Andrew Cuomo issued a ten-point executive policy, forcing all nonessential businesses to close and banning all nonessential gatherings.[6] In Florida, Governor Ron DeSantis at first refused to impose a strict stay-at-home order, instead suggesting that people "just chill out and stay around the house as much as you can," despite the pleas of health officials for a more encompassing order.[7]

There was also the question of the proper role of government in addressing the enormous economic dislocation created by the virus. The initial response by the federal government was rapid and sizeable. President Trump signed into law the $2.2 trillion CARES Act on March 27, 2020.[8] This law paved the way for a vast array of enormous interventions. These included provisions to provide relief for struggling American households through direct payments as well as assistance for rent and utilities. In an attempt to combat unemployment, these interventions also incorporated measures that bailed out hard-hit industries such as the aviation

industry, provided payroll protection for small businesses, and increased unemployment benefits.[9]

By early summer, the coronavirus that causes Covid-19 had slowed its progression. In early June, new cases reported had decreased to around 22,000 per day and the U.S. economy was recovering rapidly.[10] Economic growth expanded and employment increased, with the GDP increasing by 33.8% and the employment rate dropping to 11.1% by late June.[11]

But states had started to loosen their restrictions on interpersonal interactions and economic activity, and cases started to rise again. By late June, Covid-19 cases and deaths per day exceeded their previous peak, and by the end of 2020, more than 3,000 Americans were dying from Covid every single day. On January 13, the daily death toll peaked at 4,169 reported deaths; more Americans died that day from Covid-19 than were killed in the 9/11 attacks.[12]

Meanwhile, economic growth had once again stalled, and Americans were suffering. At the end of the year, 10.7 million Americans remained unemployed.[13] States and localities, which are mandated to balance their budgets, were unable to offset these economic shocks and faced looming layoffs of their own. Yet Congress and the Trump administration were unable to agree on a follow-up stimulus package to the CARES Act. Democrats sought to increase unemployment benefits and payments to households while targeting underrepresented groups; as Speaker of the House Nancy Pelosi said, "A building is on fire, and they are deciding how much water they want to have in the bucket … millions of people could have fallen into poverty without this $600."[14]

Republicans, however, sought to limit these benefits, citing fears that increased unemployment benefits would deter many Americans from returning to work. As Senate Majority Leader Mitch McConnell stated, "We need to discontinue paying people, in effect, a bonus not to work."[15] Instead, Republicans pushed for liability protection for businesses and health care providers against Covid-19 lawsuits. Democrats largely opposed these measures, asserting that more funds should be allocated to state and local governments so that they could continue to provide essential services.[16]

Discussions stalled throughout the second half of 2020, despite rising caseloads and ongoing economic distress. Finally, on December 27, a $900 billion relief package passed Congress and was signed by President Trump. The new stimulus package

contained direct payments to individuals, increased unemployment insurance payments, extended loans for small businesses, and provided support for education. The new bill also incorporated measures that targeted minorities and low-income households, such as funding for broadband infrastructure, renter's assistance, and food security. However, both Democrats and Republicans were forced to make major concessions. While the new bill did revive jobless benefits, the amount was considerably less than $600 per week included in the spring package. The new bill also did not contain liability protection for businesses, much to the dismay of the Republicans.[17]

The Covid-19 crisis is probably the single most important economic and public health event of the lifetimes of everyone reading this book. And it revealed the fundamental fault lines that underlie the view of government in the United States. While there was initial agreement of the central role for government fiscal relief, that agreement fell apart quickly. Meanwhile, there has never been a consensus on the proper government role for regulating individual behaviors and business activity in the face of this pandemic.

The controversies over the proper role of the government in dealing with social and economic challenges such as pandemics raised the fundamental questions addressed by the branch of economics known as *public finance*. The goal of public finance is to *understand the proper role of the government in the economy.* On the expenditure side of public finance, we ask: Why should the government be spending hundreds of billions of dollars to provide extra unemployment insurance benefits during the pandemic (to cite just one example)? More generally, why is the government the primary provider of goods and services such as highways, education, and transfers to the unemployed, while the provision of goods and services such as clothing, entertainment, and property insurance is generally left to the private sector? On the revenue side of public finance, we ask: How much should the government tax its citizens, and how should that amount be related to the economic circumstances of those individuals? What kinds of activities should be taxed or be given tax relief in difficult times? What effect do taxes have on the functioning of the economy?

Republicans, however, sought to limit these benefits, citing fears that increased unemployment benefits would deter many Americans from returning to work. As Senate Majority Leader Mitch McConnell stated, "We need to discontinue paying people, in effect, a bonus not to work."[15] Instead, Republicans pushed for liability protection for businesses and health care providers against Covid-19 lawsuits. Democrats largely opposed these measures, asserting that more funds should be allocated to state and local governments so that they could continue to provide essential services.[16]

Discussions stalled throughout the second half of 2020, despite rising caseloads and ongoing economic distress. Finally, on December 27, a $900 billion relief package passed Congress and was signed by President Trump. The new stimulus package contained direct payments to individuals, increased unemployment insurance payments, extended loans for small businesses, and provided support for education. The new bill also incorporated measures that targeted minorities and low-income households, such as funding for broadband infrastructure, renter's assistance, and food security. However, both Democrats and Republicans were forced to make major concessions. While the new bill did revive jobless benefits, the amount was considerably less than $600 per week included in the spring package. The new bill also did not contain liability protection for businesses, much to the dismay of the Republicans.[17]

The Covid-19 crisis is probably the single most important economic and public health event of the lifetimes of everyone reading this book. And it revealed the fundamental fault lines that underlie the view of government in the United States. While there was initial agreement of the central role for government fiscal relief, that agreement fell apart quickly. Meanwhile, there has never been a consensus on the proper government role for regulating individual behaviors and business activity in the face of this pandemic.

The controversies over the proper role of the government in dealing with social and economic challenges such as pandemics raised the fundamental questions addressed by the branch of economics known as *public finance*. The goal of public finance is to *understand the proper role of the government in the economy*. On the expenditure side of public finance, we ask: Why should the government be spending hundreds of billions of dollars to provide extra unemployment insurance benefits during the pandemic (to cite just one example)? More generally, why is the government the primary provider of goods and services such as highways, education, and transfers to the unemployed, while the provision of goods and

services such as clothing, entertainment, and property insurance is generally left to the private sector? On the revenue side of public finance, we ask: How much should the government tax its citizens, and how should that amount be related to the economic circumstances of those individuals? What kinds of activities should be taxed or be given tax relief in difficult times? What effect do taxes have on the functioning of the economy?

1.1 The Four Questions of Public Finance

In the simplest terms, **public finance** is the study of the role of the government in the economy. This is a very broad definition. This study involves answering the **four questions of public finance**:

- *When* should the government intervene in the economy?
- *How* might the government intervene?
- *What* is the effect of those interventions on economic outcomes?
- *Why* do governments choose to intervene in the way that they do?

public finance
The study of the role of the government in the economy.
four questions of public finance
When should the government intervene in the economy? How might the government intervene? What is the effect of those interventions on economic outcomes? Why do governments choose to intervene in the way that they do?

In this section, we explore these four questions within the context of a specific example: the market for *health insurance,* in which individuals pay a monthly premium to insurance companies, in return for which insurance companies pay the individuals' medical bills if they are ill. This is only one of many markets in which the government is involved, but it is a particularly useful example because health care spending is the single largest and fastest-growing part of the U.S. government's budget.

When Should the Government Intervene in the Economy?

To understand the reason for government intervention, think of the economy as a series of trades between producers (firms) and consumers. A trade is *efficient* if it makes at least one party better off without making the other party worse off. The total efficiency of the economy is maximized when as many efficient trades as possible are made.

The fundamental lesson of basic microeconomics is that, in most cases, the *competitive market equilibrium is the most efficient outcome for society*—that is, it is the outcome that maximizes the gains from efficient trades. As discussed in much more detail in Chapter 2, the free adjustment of prices guarantees that, in competitive market equilibrium, supply equals demand. When supply equals demand, all trades

that are valued by both producers and consumers are being made. Any good that consumers value above its cost of production will be produced and consumed; goods that consumers value at less than their cost of production will not be produced or consumed.

If the competitive market equilibrium is the most efficient outcome for society, why do governments intervene in the operation of some of these markets? There are two reasons governments may want to intervene in market economies: market failures and redistribution.

Market Failures

The first motivation for government involvement in the economy is the existence of **market failures**, problems that cause a market economy to deliver an outcome that does not maximize efficiency. Throughout this book, we discuss a host of market failures that impede the operation of the market forces you learned about in basic microeconomics. Here, we briefly explore a failure in the health insurance market that may cause its equilibrium outcome to be inefficient.

market failure
A problem that causes the market economy to deliver an outcome that does not maximize efficiency.

At first glance, the market for health insurance seems to be a standard textbook competitive market. Health insurance is supplied by a large number of insurance companies and demanded by a large number of households. In the market equilibrium where supply equals demand, social efficiency should be maximized: anyone who values health insurance above its cost of production is able to buy insurance.

In 2010, before the Affordable Care Act, there were 49 million people without health insurance in the United States, or 18.5% of the non-elderly population (as we'll discuss in Chapter 15, the elderly are provided universal health coverage in the United States under the Medicare program).[18] The existence of such a large number of uninsured does not, however, imply that the market doesn't work. After all, there are many more Americans who don't have a large-screen TV, or a new car, or a home of their own. That a small minority of the population is uninsured does not by itself prove that there is a problem in the market; it just implies that those without insurance don't value it enough to buy it at existing prices.

To answer this question, we turn in Chapter 9 to the tools of **political economy**, the theory of how governments make public policy decisions. Governments face enormous challenges in figuring out what the public wants and how to choose policies that match those wants. In addition, governments may be motivated by much more than simply correcting market failures or redistributing income. Just as there are a host of market failures that can interfere with the welfare-maximizing outcome from the private market, there are a host of *government failures* that can lead to inappropriate government interventions. Politicians must consider a wide variety of viewpoints and pressures, only two of which are the desire to design policies that maximize economic efficiency and redistribute resources in a socially preferred manner.

political economy
The theory of how the political process produces decisions that affect individuals and the economy.

One only needs to look at the wide variety of health insurance policies in very similar countries to see that governments may have more in mind than efficiency or redistribution. Why does the United States rely primarily on private health insurance, while Canada, a similar country bordering the United States, relies on national public health insurance? Why does Germany mandate private health insurance coverage, while the United Kingdom provides free national health care? Coming back to the first question (When should the government intervene?), then, we have an additional concern that must be addressed before recommending government intervention: In practice, will the government actually reduce or solve the problem? Or will government failures cause the problem to grow worse?

1.2 Why Study Public Finance? Facts on Government in the United States and Around the World

Thus far, we have clarified what public finance is. But it still may not be clear why you should spend your precious time on this topic. What makes public finance so compelling is the dominant role that governments play in our everyday lives. In this section, we detail that role by walking you through the key facts about government in the United States and other developed nations. In addition, to motivate the study of public finance, we propose some interesting questions that arise from these facts.

The Size and Growth of Government

Figure 1-1 shows the growth in federal government spending in the United States over the twentieth century. In 1930, the federal government's activity accounted for only about 3.4% of gross domestic product (GDP). Government spending expanded during the Great Depression and grew even more dramatically during World War II, hitting a peak of almost half of GDP in 1944. From the 1950s through the present, the size of government has averaged around 20% of GDP, although it grows during recessions such as those in the early 1980s, the early 1990s, and the early and late 2000s, reaching nearly one-quarter of GDP in 2009 before settling back down at 20.3% by 2014, remaining approximately stable since then.

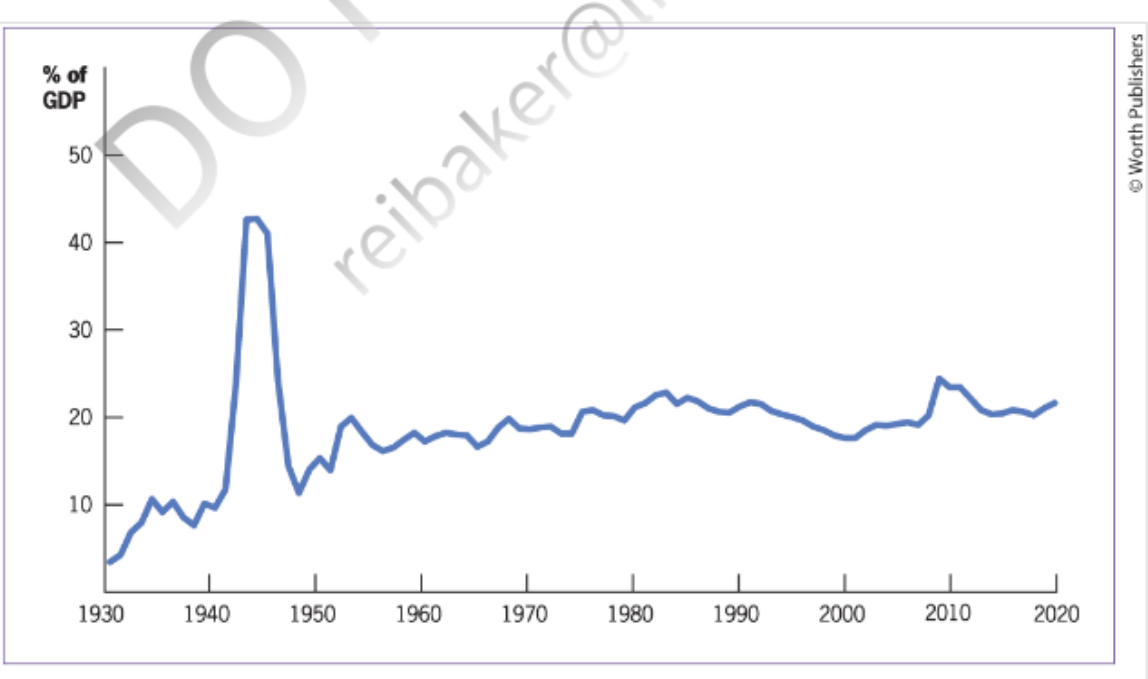

FIGURE 1-1 Federal Government Spending as a Percentage of GDP, 1930–2019 • Over the past 90 years, federal government spending (as a share of GDP) has grown from 3 to 21%. The huge spike in spending over the

1941–1948 period was due to the massive increase in defense spending.

Data from: Office of Management and Budget (2021), Table 1.2.

This growth is mirrored in other developed nations, as seen in Figure 1-2. This figure shows the growth of government spending since 1960 in the United States, Sweden, and Greece and the average for the industrialized nations that are part of the Organization for Economic Cooperation and Development (OECD). The patterns are quite interesting. In 1960, the United States was squarely in line with the average of the OECD in terms of the government spending share of GDP.[38] Yet, government growth was much faster in other OECD nations in the 1960s and 1970s than in the United States, so that by 1980, the U.S. government share was much smaller. Greece started with a government share well below that of the United States in 1960, but government tripled as a share of Greece's GDP so that today its share is much larger than the U.S. government's share.[39] In 1960, Sweden's government's share of GDP was similar to other nations', but this share grew enormously so that by the early 1990s, government spending was about two-thirds of Sweden's GDP. Since then, Sweden's government share has fallen rapidly and now accounts for slightly greater than half of GDP, just barely more than Greece.[40]

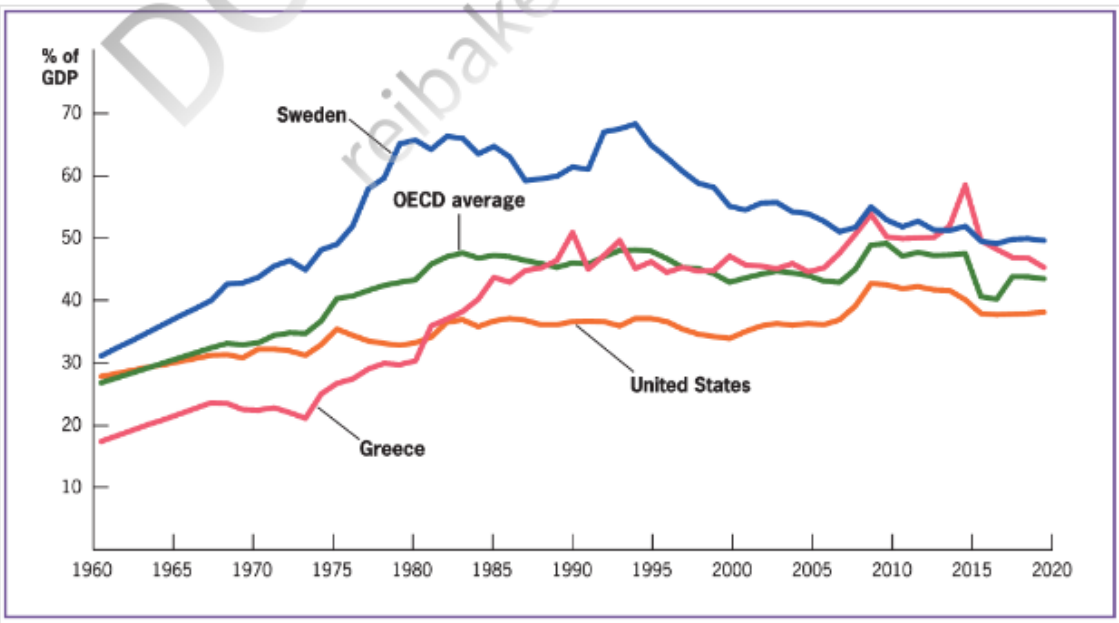

FIGURE 1-2 Total Government Spending Across Developed Nations, 1960–2019 • Government spending as a share of GDP has grown throughout the developed world, but the pace of growth has varied. The United States has seen a modest growth in its government share over this period, while government spending in Greece has more than tripled as a share of the economy.

Data from: Organization for Economic Cooperation and Development (2021), Table 29.

- What explains the growth in government spending over the twentieth century?

Decentralization

A key feature of governments is the degree of *centralization* across local and national government units—that is, the extent to which spending is concentrated at higher (federal) levels or lower (state and local) levels. Figure 1-3 shows government spending in the United States divided into the share of spending by the federal government and the share of spending by other levels of government: state, county, and local governments. The federal government provides the majority of government spending in the United States, but other government spending is quite large as well, amounting to well over 40% of total government spending and more than 17% of GDP. The level of centralization (the share of spending done by the federal government) varies widely across nations, sometimes rising to almost 100% in countries where the federal government does almost all of the government spending.

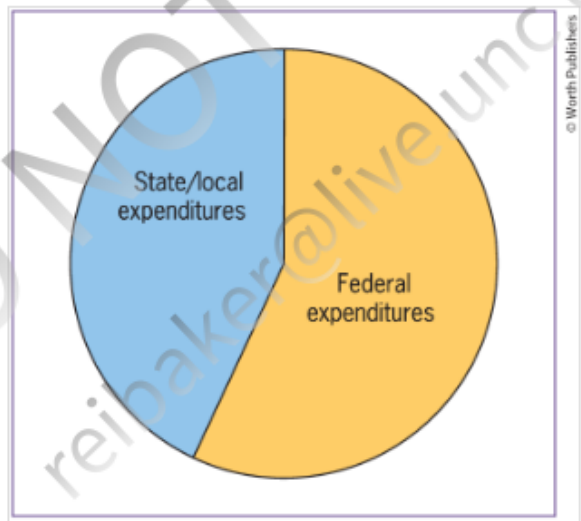

FIGURE 1-3 Federal vs. State/Local Government Spending, 2019 • State and local spending today amounts to almost 40% of total government spending in the United States, at almost 15% of GDP.

Data from: Bureau of Economic Analysis, Tables 3.2, 3.20, 3.21 (2021)

- What is the appropriate extent of centralization and decentralization in government activity?

Spending, Taxes, Deficits, and Debts

Spending, Taxes, Deficits, and Debts

When you run a household, you live on a budget. Outflows of cash for groceries, rent, clothing, entertainment, and other uses must be financed by inflows of cash from work or other sources. Any excess of income over spending is a *cash flow surplus* that can be saved to finance your own spending in future periods or passed on to others after you die. Any shortfall of income below spending is a *cash flow deficit* and must be financed by past savings or by borrowing from others. Any borrowing results in the buildup of some household *debt,* which must ultimately be repaid from future inflows of cash.

Fundamentally, the finances of the government are no different. Its outflows are government spending and its inflows are tax revenues. If revenues exceed spending, then there is a budget surplus; if revenues fall short of spending, there is a budget deficit. Each dollar of government deficit adds to the stock of government debt. That is, the *deficit* measures the year-to-year shortfall of revenues relative to spending; the *debt* measures the accumulation of past deficits over time. This government debt must be financed by borrowing. Governments can borrow from their own citizens, from citizens of other municipalities, or from other nations.

The three panels of Figure 1-4 show government spending and revenues, the deficit or surplus, and the level of government debt for the U.S. federal government. As shown in panels (a) and (b), with the exception of an enormous increase in spending unmatched by increased taxation during World War II (1941–1945), the federal government's budget was close to balanced until the late 1960s. From the mid-1970s through the mid-1990s, there was a relatively large deficit that rose to about 5% of GDP. This deficit shrank dramatically in the 1990s and actually turned into a sizeable surplus by the end of the decade.

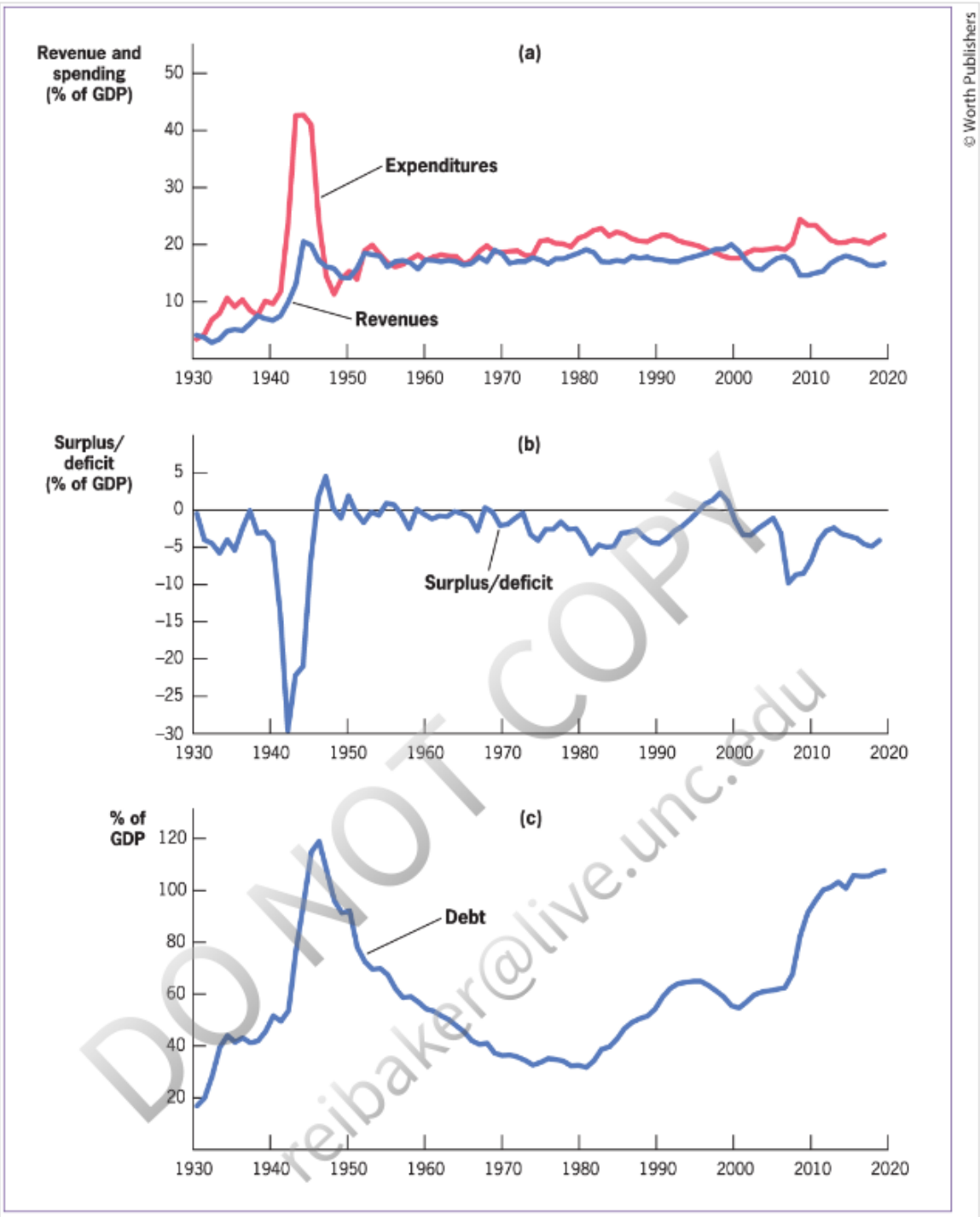

FIGURE 1-4 Federal Revenues and Expenditures, Surplus or Deficit, and Debt, 1930–2019 • For most of the twentieth century, except for the World War II period, federal government receipts kept pace with expenditures, leading to a small national debt. But expenditures have exceeded receipts by several percentage points of GDP on average since the 1970s. The resulting federal debt is now at more than 106% of GDP.

Data from: Office of Management and Budget (2021), Table 1.2, Table 7.1.

But the United States was back in deficit by the early twenty-first century, at levels similar to those in the 1970s. The deficit had become very large in the late 2000s,

reaching levels not seen in the postwar period. By the mid-teens, increases in revenues and decreases in spending brought the deficit back to the level typical of the past 40 years. But expansions in spending and cuts in taxes during the Trump Administration, most recently spurred by the Covid-19 pandemic, have prompted an increase in the deficit.

The resulting implications for the federal debt are shown in panel (c) of Figure 1-4. The stock of debt rose sharply in World War II, then fell steadily until large deficits caused it to rise in the 1980s. The debt has risen considerably since, with a brief pause in the mid- to late 1990s, and now it is more than 106% of GDP. Figure 1-5 compares the level of U.S. debt with the level of debt of other developed nations. The United States has higher debt levels than most other comparable nations, but its load remains well below others.

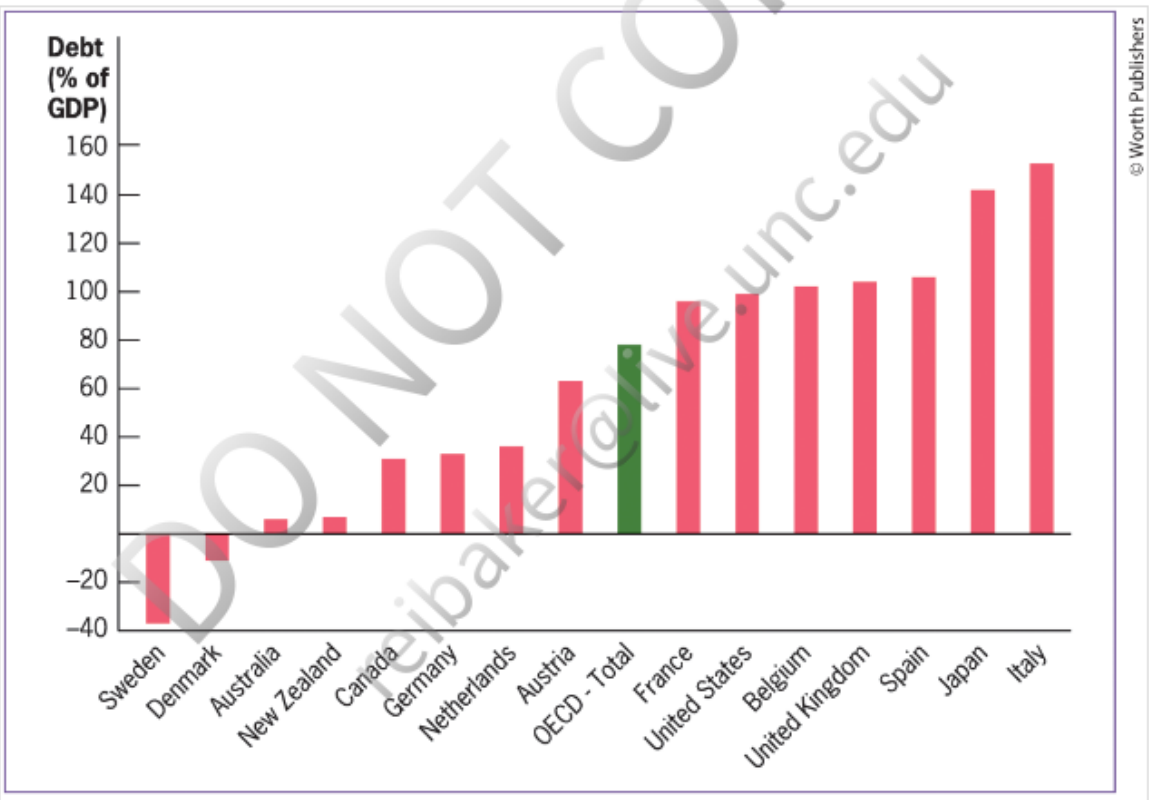

FIGURE 1-5 Debt Levels of OECD Nations in 2020 • The United States has debt levels somewhat above the OECD average, but there is wide variation across developed nations.

Data from: Organization for Economic Cooperation and Development (2021), Table 37.

- What are the costs of having larger deficits and a larger national debt?

Figure 1-6 shows the spending and revenues of state and local governments over time in the United States. Interestingly, unlike the federal government, state and

time in the United States. Interestingly, unlike the federal government, state and local governments' budgets are typically in surplus: there is very little deficit overall across the state and local governments in any year.

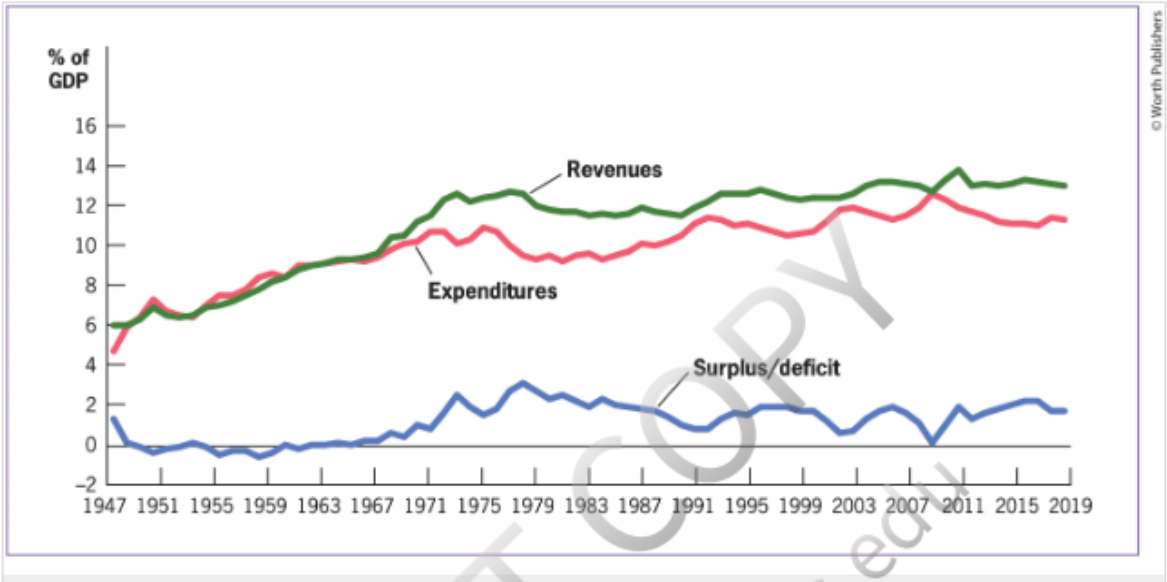

FIGURE 1-6 State and Local Government Receipts, Expenditures, and Surplus, 1947–2019 • State and local revenues almost always exceed expenditures.

Data from: Office of Budget and Management (2021), Table 12.1, Table 14.1, Table 14.3.

- Why are state and local governments able to balance their budgets, while the federal government is not?

Distribution of Spending

Thus far, we have discussed only the sum total of government spending in the United States and not on what these funds are spent. Figure 1-7 shows the distribution of spending across several broad categories for the federal government and state and local governments in 1960 and 2019. Several conclusions are apparent. First, the composition of federal government spending [panel (a)] has changed dramatically over time. In 1960, nearly half of federal government spending was on national defense, military expenditures either at home or abroad.

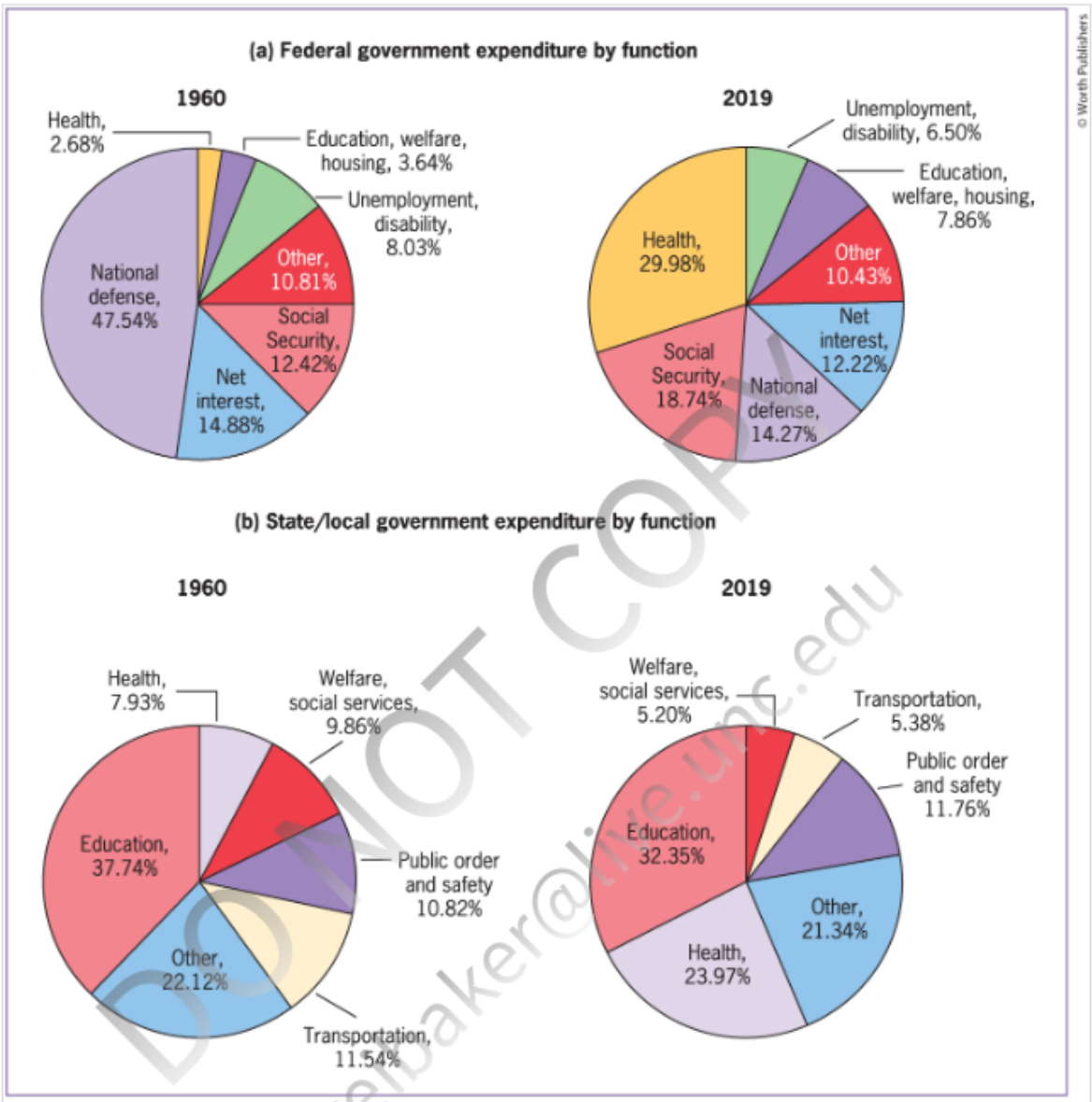

FIGURE 1-7 The Distribution of Federal and State Expenditures, 1960 and 2019 • This figure shows the changing composition of federal and state spending over time, as a share of total spending. (a) For the federal government, defense spending has fallen and Social Security and health spending have risen. (b) For the states, the growth in health care spending has led to a reduction in education spending.

Data from: Bureau of Economic Analysis (2021), Table 3.16.

Defense is a classic example of what economists call a **public good**, a good for which the investment of any one individual benefits a larger group of individuals: if I purchased a missile to protect Boston, that would benefit not just me but all of the residents of the city. As we will discuss at length in Chapter 7, the private sector may underprovide such public goods: if I bear the full cost of buying a missile, but it benefits everyone in town, then I probably won't spend the money on that missile. This makes provision of public goods an important job for the government, as

reflected in the large share of government spending in this area.

public goods
Goods for which the investment of any one individual benefits everyone in a larger group.

Today, however, defense spending has fallen to less than one-sixth of the federal budget. The offsetting spending growth can be found largely in two areas. The first is the Social Security program, which provides income support to the elderly who are retired from their jobs. This is the single largest government program in the United States today, consuming about 18.7% of the entire federal budget. Another large and rapidly growing category is health care programs, a variety of federal government interventions to provide health insurance for the elderly, the poor, and the disabled; these programs consume almost 30% of the budget.[41]

Programs such as Social Security and government health insurance programs are called **social insurance programs**, programs designed to address failures in private insurance markets. As we discussed earlier, private health insurance markets may not provide the appropriate amount of health insurance to the population. This market failure has motivated the government to intervene in health insurance markets; indeed, almost one-half of all health spending in the United States is done by governments. Similarly, the federal government is concerned that individuals may not plan appropriately for the decline in income they will face when they retire, which motivates the existence of the Social Security program.

- Are large government interventions in insurance markets warranted, and do they correct or exacerbate market failures?

social insurance programs
Government provision of insurance against adverse events to address failures in the private insurance market.

The distribution of state and local spending [Figure 1-7, panel (b)] is much different. At the state and local levels, education and public safety account for almost half of spending. Less than 10% of federal spending supports these programs. Likewise, there are no Social Security or defense expenditures at the state or local level. The major development over time at the state and local levels has been the enormous growth in health care spending, which has risen from only 8% of state and local spending to almost 24%. Because this increase appears to have been offset by reductions in spending on education, welfare and social services, transportation, and other categories, it raises important issues of whether the more expensive health care system is reducing (or what we will later refer to as “crowding out”) other state and local services.

other state and local services.

- What is the appropriate type of spending to be done at the federal versus state or local level?

Distribution of Revenue Sources

Figure 1-8 breaks down the sources of federal and state and local revenue over time. The major source of revenue for the federal government [panel (a)] is the individual *income tax,* a tax levied on the income of U.S. residents. This tax provides somewhat less than half of federal revenues and has remained roughly constant as a share of revenues over time. The major shift over time at the federal level has been the rapid shrinking of corporate tax revenues, the funds raised by taxing the incomes of businesses in the United States. While corporate tax revenues once provided almost 25% of federal government revenue, they now provide only about 6%. There has also been a sizeable reduction in *excise taxes,* taxes levied on the consumption of certain goods such as tobacco, alcohol, or gasoline.

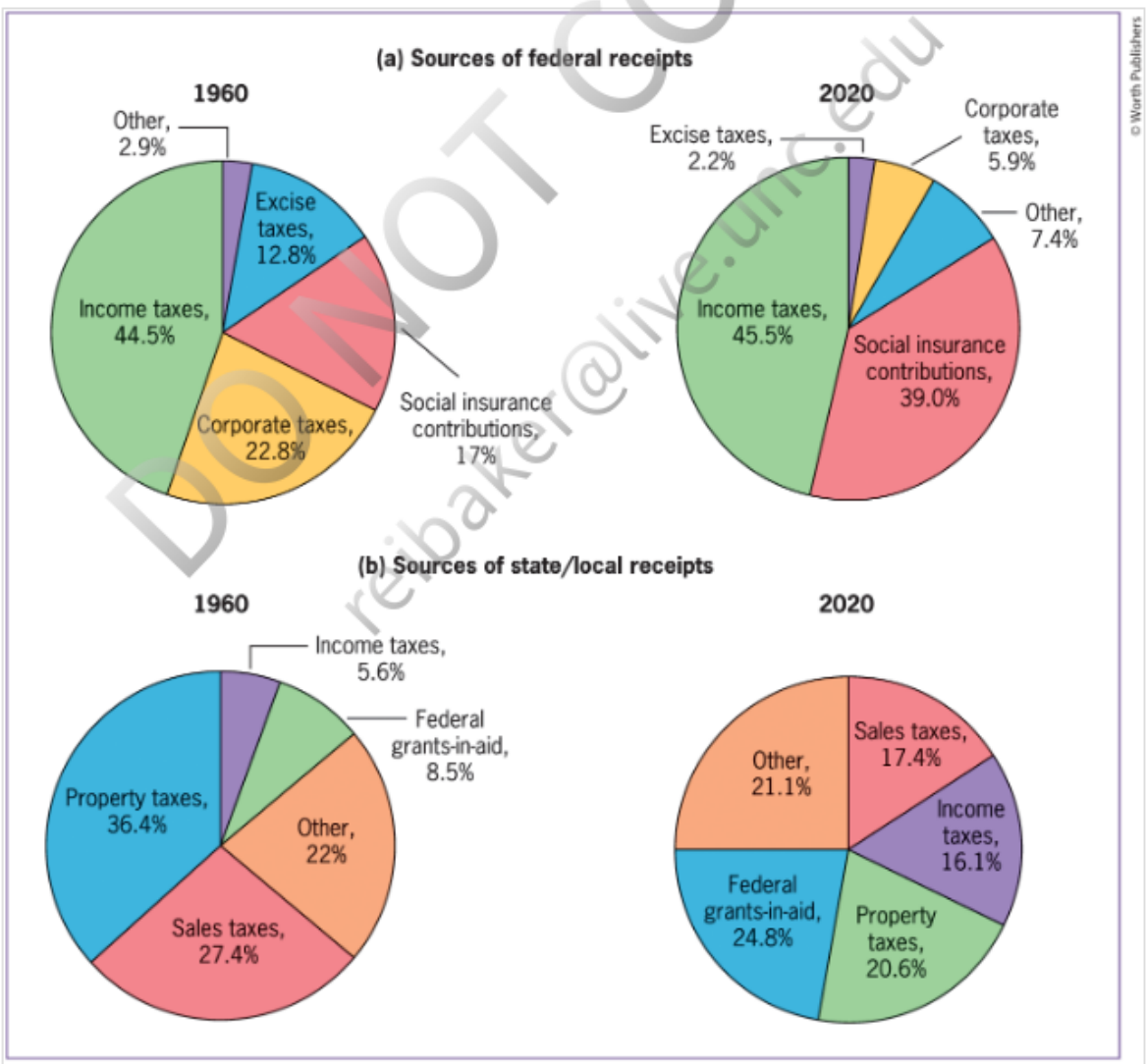

FIGURE 1-8 The Distribution of Federal and State Revenues, 1960 and 2020 • This figure shows the changing composition of federal and state revenue sources over time, as a share of total revenues. (a) At the federal level, there has been a large reduction in corporate and excise tax revenues and a rise in payroll tax revenues. (b) For the states, there has been a decline in property taxes and a rise in income taxes and federal grants.

Data from: Bureau of Economic Analysis (2021), Table 3.2, Table 3.3.

The decrease in revenue from these taxes has been largely replaced by the growth of revenue from *payroll taxes,* the taxes on worker earnings that fund social insurance programs. Payroll taxes differ from the income tax in that the income tax includes all sources of income, such as the return on savings, while payroll taxes apply solely to earnings from work. Payroll taxes have grown from one-sixth of federal revenues to well over one-third.

- What are the implications of moving from taxing businesses and consumption to taxing workers' earnings?

At the state and local levels [Figure 1-8, panel (b)], revenue sources are roughly equally divided between *sales taxes* (including state and local excise taxes on products such as cigarettes and gasoline), federal *grants-in-aid* (redistribution of funds from the federal government to lower levels of government), income taxes, and *property taxes* (taxes on the value of individual properties, mostly homes). Over the past 40 years, the substantial drop in revenue from property taxes has been made up by rising federal grants and income taxes.

- What are the implications of shifting from taxation of property to taxation of income?

Regulatory Role of the Government

The discussion throughout this section has focused on the government as an entity that exerts influence through its powers of taxation and spending. Another critical role the government plays in all nations is that of *regulating economic and social activities.* Consider some examples of how daily existence is affected by the government in the United States:[42]

-

The foods you eat and the medications you take have all been approved by the Food and Drug Administration (FDA), an agency that spends less than 0.025% of the government's budget each year, but whose regulatory powers cover $2.8 trillion worth of goods annually, more than 20% of total consumer expenditures. The FDA regulates the labeling and safety of nearly all food products and bottled water, tests cosmetics to ensure their safety, and approves drugs and medical devices to be sold to the public.

- If you've lost your hearing, got sick from fumes, or had a repetitive stress injury because of your work, you might want to contact the Occupational Safety and Health Administration (OSHA), which is charged with regulating the workplace safety of the 146 million Americans employed at more than 10 million job sites. In 2019, the agency sent its 1,850 inspectors on 33,393 visits to workplaces.
- The radio stations in your car and the channels you watch on cable are overseen by the Federal Communications Commission (FCC), which regulates interstate and international radio, television, wire, satellite, and cable communications. Check any device in your home that emits radiation of communication frequencies (cell phones, remotes, etc.) and you'll find an FCC identification number somewhere on it.
- The air you breathe, the tap water you drink, and the land your home is built upon are all regulated by the Environmental Protection Agency (EPA), which is charged with minimizing dangerous pollutants in the air, water, and food supplies.

1.3 The Questions of Public Finance Are Front and Center During Covid-19

We began this chapter by talking about the Covid-19 pandemic and legislation surrounding Covid-19 relief. The pandemic caused major disruption, and the debates surrounding the government's response to the pandemic reflect exactly the set of key public finance questions that we reviewed earlier.

When Should the Government Intervene?

The Covid-19 pandemic brought about a vast number of government interventions. Some involve questions of personal privacy, such as whether the government should require individuals to wear masks. Others are questions of direct government intervention in business, such as whether states should require businesses to shut down to avoid spread of the disease. And there were enormous debates over how the government should respond to the unprecedented economic dislocation caused by the pandemic. On each of these topics, there were fundamental disagreements across the political spectrum. Republicans generally favored fewer restrictions on individuals and businesses, and a smaller government fiscal response. Democrats favored more active intervention to stop the spread of disease and a more expansive response.

How Should the Government Intervene?

The question of how to intervene was central in the debates in Congress over provisions of the CARES Act and subsequent extensions. Democrats favored mandates for masks, while Republicans proposed looser guidelines. Democrats favored expansions in benefits for unemployed individuals, while Republicans favored loans to businesses. Republicans favored protecting businesses from the legal liabilities of Covid-19, while Democrats instead prioritized funding for state and local government.[43]

What Are the Effects of Interventions?

The $2.2 Trillion CARES Act provided a natural laboratory to understand the effect of interventions to address Covid-19, and economists have jumped in with both feet. The preeminent working paper series for economics is the National Bureau of Economic Research (NBER) working papers. From April 2020 through the end of the

year, there were 327 Covid-related working papers distributed through NBER, representing 23.8% of all NBER working papers.[44]

Economic research provided a vast array of insights into the effects of government interventions. Some studies showed that government activities to limit economic activity had only a modest effect, as most activity was ceasing voluntarily even before the limits were put into place.[45] Other studies showed that the government's Paycheck Protection Program (PPP), which provided forgivable loans to businesses, did not save many jobs.[46] At the same time, studies showed that the unemployment insurance expansions through the CARES Act led to much higher consumption among those receiving payments and did not appear to deter their return to work.[47]

Why Governments Do What They Do

The singular government focus on Covid-19 allows us to also observe important lessons for political economy. Despite the evidence just discussed, for example, when Congress passed a $900 billion relief bill at the end of 2020, it included more money for PPP than it did for extending unemployment insurance, and no money for state and local governments.[48] Given Republican control of Congress, this reflected the Republican prioritization on business funding. Yet as soon as Democrats gained control of the presidency and Congress, they proposed a new approach that placed much more focus on unemployment insurance and funding for governments. The crisis had not changed, yet the government's response had—based on who was in control. This suggests that economics is not the final word in government decision making.

What Are the Effects of Interventions?

The $2.2 Trillion CARES Act provided a natural laboratory to understand the effect of interventions to address Covid-19, and economists have jumped in with both feet. The preeminent working paper series for economics is the National Bureau of Economic Research (NBER) working papers. From April 2020 through the end of the year, there were 327 Covid-related working papers distributed through NBER, representing 23.8% of all NBER working papers.[44]

Economic research provided a vast array of insights into the effects of government interventions. Some studies showed that government activities to limit economic activity had only a modest effect, as most activity was ceasing voluntarily even before the limits were put into place.[45] Other studies showed that the government's Paycheck Protection Program (PPP), which provided forgivable loans to businesses, did not save many jobs.[46] At the same time, studies showed that the unemployment insurance expansions through the CARES Act led to much higher consumption among those receiving payments and did not appear to deter their return to work.[47]

Why Governments Do What They Do

The singular government focus on Covid-19 allows us to also observe important lessons for political economy. Despite the evidence just discussed, for example, when Congress passed a $900 billion relief bill at the end of 2020, it included more money for PPP than it did for extending unemployment insurance, and no money for state and local governments.[48] Given Republican control of Congress, this reflected the Republican prioritization on business funding. Yet as soon as Democrats gained control of the presidency and Congress, they proposed a new approach that placed much more focus on unemployment insurance and funding for governments. The crisis had not changed, yet the government's response had—based on who was in control. This suggests that economics is not the final word in government decision making.

Highlights

- There are four key questions to consider in the study of public finance. The first is: When should the government intervene in the economy? Our baseline presumption is that the competitive equilibrium leads to the outcome that maximizes social efficiency. So government intervention can be justified only on the grounds of market failure (increasing the size of the pie) or redistribution (changing the allocation of the pie).
- Having decided whether to intervene, the government needs to decide how to intervene. There are many policy options that can be pursued to achieve the same goal, such as public provision, mandates for private provision, and subsidies to private provision.
- When deciding how to intervene, the government needs to choose an approach for evaluating the impacts of alternative interventions on the economy. The tools of empirical economics provide one such approach.
- A major question for public finance is: Why do governments choose to pursue the policies that they do? We are particularly concerned about government failure, whereby government intervention can make problems worse, not better.
- Government, which consists of both national (federal) and local units (states, counties, cities, and towns), is large and growing in the United States and throughout the world. The nature of government spending and revenue sources is also evolving over time as governments move away from being providers of traditional public goods (such as defense) to being providers of social insurance (such as Social Security and health insurance).
- Governments also affect our lives through regulatory functions in a wide variety of arenas.

Questions and Problems

1. Many states have language in their constitutions that requires the state to provide for an "adequate" level of education spending. What is the economic rationale for such a requirement?
2. How has the composition of federal, state, and local government spending changed since 1960? What social and economic factors might have contributed to this change in how governments spend their funds?

3. How did the composition of the federal budget change in 2020, in comparison to 2019? What were the most notable changes, and what factors drove those changes?
4. Consider the arguments for and against the passage of a follow-up stimulus package to the CARES Act in response to the Covid-19 pandemic. Specifically, consider arguments for and against additional expanded unemployment benefits and business assistance like the Paycheck Protection Program. Do you think Speaker of the House Nancy Pelosi arguments were more about *efficiency* or *redistribution*? What about the arguments made by Senate Majority Leader Mitch McConnell? Is there an *efficiency* case for expanded unemployment benefits, the PPP, and other Covid-19 relief?
5. Some goods and services are provided directly by the government, while others are funded publicly but provided privately. What is the difference between these two mechanisms of public financing? Why do you think the same government would use one approach sometimes and the other approach at other times?
6. Why does redistribution cause efficiency losses? Why might society choose to redistribute resources from one group to another when doing so reduces the overall size of the economic pie?
7. Consider the four basic questions of public finance listed in the chapter. Which of these questions are positive—that is, questions that can be proved or disproved—and which are normative—that is, questions of opinion? Explain your answer.
8. One rationale for imposing taxes on alcohol consumption is that people who drink alcohol impose negative externalities on the rest of society—for example, through loud and unruly behavior or intoxicated driving. If this rationale is correct, in the absence of governmental taxation, will people tend to consume more alcohol, less alcohol, or same amount of alcohol?
9. What is the role of the Congressional Budget Office (CBO)? Why is independence and impartiality important when conducting empirical analyses?
10. U.S. law prohibits hospitals from denying patients care in an emergency, even if they are unable to pay. Historically, the cost of covering these uninsured patients has been passed on to insured patients and the government. Explain how a law like the Affordable Care Act, which reduces the number of uninsured patients, thus alleviates a market failure.

ADVANCED QUESTIONS

11. In the United States, the federal government pays for a considerably larger share of social welfare spending (i.e., spending on social insurance programs to help low-income, disabled, or elderly people) than it does for K–12 education spending. Similarly, state and local governments provide a larger share of education spending and a smaller share of welfare spending. Is this a coincidence, or can you think of a reason for why this might be so?
12. Americans are split on the subject of school vouchers, which allow parents to use all or part of public funding set aside for the education of their child on private school tuition, as opposed to keeping their child in public school. The controversy especially affects communities with disproportionate numbers of low-income families. Some leaders in these communities strongly support the voucher system and the increased school competition it brings, while others oppose it. Why do you think this split exists?
13. Many states have constitutional requirements that their budgets be in balance (or in surplus) in any given year, but this is not true for the U.S. federal government. Why might it make sense to allow for the federal government to have deficits in some years and surpluses in others?
14. Simple habits like hand-washing and mask-wearing can greatly limit the spread of many diseases, including Covid-19 and the seasonal flu. How might this suggest a role for public interventions? What kinds of public interventions might be possible? Suggest three distinct types of possible intervention.

CHAPTER 2

Theoretical Tools of Public Finance

James Andrews/iStock/Getty Images

Questions to keep in mind

- How do individuals choose how much to consume or how hard to work?
- How do firms choose how much to produce?
- What is the theoretical effect of raising cash welfare benefits on economic efficiency?

Life is going well. After graduating at the top of your college class, you have parlayed your knowledge of public finance into an influential job with your state's Department of Health and Human Services (HHS), which oversees, among other things, the Temporary Assistance for Needy Families (TANF) program. This program provides cash payments to parents whose income is below a specified level.

Your new job thrusts you into the middle of a debate between the state's governor and the head of your department, the HHS secretary. The governor believes that a major problem with the TANF program is that only providing income to parents with very low incomes encourages some parents to stay at home rather than go to work. To provide incentives for these parents to work, the governor wants to cut back on these cash benefits. The secretary of the department disagrees. He thinks that single parents who are home with their children are incapable of finding jobs that pay a wage high enough to encourage them to work. In his view, if the state cuts the cash payments, it will simply penalize those single parents who are staying home to care for their children.

The secretary turns to you to inform this debate by assessing the extent to which cutting cash benefits to low-income single parents will encourage them to work and by evaluating the net welfare implications for the state if these benefits are cut. Such an evaluation will require that you put to work the economics tools that you have learned in your introductory and intermediate courses. These tools come in two flavors.

First are the **theoretical tools**. The primary theoretical tools of economists are graphical and mathematical. The graphical tools, such as supply and demand diagrams and indifference curve/budget constraint graphs, are typically all that you need to understand the key points of theory, but mathematical expositions help to illustrate the subtleties of an argument. In the main body of this book, we rely almost exclusively on graphical analysis, with parallel mathematical analysis presented in some chapter appendices.

theoretical tools
The set of tools designed to understand the mechanics behind economic decision making.

Second, there are **empirical tools**, the set of tools that allows you to analyze data and answer the questions that are raised by theoretical analysis. Most students in this course will have had much less exposure to empirical tools than to theoretical tools. Yet, particularly over the past two decades, empirical tools have become as important as theoretical tools in addressing the problems of public finance, as both the quality of data and our ability to carefully analyze that data have improved dramatically.

empirical tools
The set of tools designed to analyze data and answer questions raised by theoretical analysis.

In the next two chapters, you will be introduced to the key theoretical and empirical tools that you need for this course. In each chapter, we first provide a general background on the concepts and then apply them to our TANF example. The discussion in this chapter is intimately related to the first two of the four questions of public finance we introduced in Chapter 1. The theoretical tools we discuss here are the central means by which economists assess when the government should intervene and how it might intervene.

The remainder of this book relies heavily on the microeconomics concepts that we will be reviewing in this chapter. This chapter does not, however, substitute for an introductory or intermediate microeconomics course. The goal here is to refresh your understanding of the important concepts that you need to undertake theoretical public finance, not to teach them to you for the first time. If the material in this chapter is very unfamiliar, you may want to supplement this text with a more detailed microeconomics text.[1]

Second, there are **empirical tools**, the set of tools that allows you to analyze data and answer the questions that are raised by theoretical analysis. Most students in this course will have had much less exposure to empirical tools than to theoretical tools. Yet, particularly over the past two decades, empirical tools have become as important as theoretical tools in addressing the problems of public finance, as both the quality of data and our ability to carefully analyze that data have improved dramatically.

empirical tools
The set of tools designed to analyze data and answer questions raised by theoretical analysis.

In the next two chapters, you will be introduced to the key theoretical and empirical tools that you need for this course. In each chapter, we first provide a general background on the concepts and then apply them to our TANF example. The discussion in this chapter is intimately related to the first two of the four questions of public finance we introduced in Chapter 1. The theoretical tools we discuss here are the central means by which economists assess when the government should intervene and how it might intervene.

The remainder of this book relies heavily on the microeconomics concepts that we will be reviewing in this chapter. This chapter does not, however, substitute for an introductory or intermediate microeconomics course. The goal here is to refresh your understanding of the important concepts that you need to undertake theoretical public finance, not to teach them to you for the first time. If the material in this chapter is very unfamiliar, you may want to supplement this text with a more detailed microeconomics text.[1]

2.1 Constrained Utility Maximization

The core of theoretical analysis in public finance is the assumption that every individual has a well-defined **utility function**, a mathematical mapping of individual choices over goods into their level of well-being. Economists assume that individuals then undertake **constrained utility maximization**, maximizing their well-being (utility) subject to their available resources. Armed with this assumption, economists proceed to develop **models**—mathematical or graphical representations of reality—to show how constrained utility maximization leads people to make the decisions that they make every day. These models have two key components: the individual's *preferences* over all possible choices of goods and their *budget constraint*, the amount of resources with which they can finance their purchases. The strategy of economic modelers is then to ask: Given a budget constraint, what *bundle of goods* makes a consumer best off?

utility function
A mathematical function representing an individual's set of preferences, which translates their well-being from different consumption bundles into units that can be compared in order to determine choice.
constrained utility maximization
The process of maximizing the well-being (utility) of an individual, subject to their resources (budget constraint).
models
Mathematical or graphical representations of reality.

We can illustrate how consumers are presumed to make choices in four steps. First, we discuss how to model preferences graphically. Second, we show how to take this graphical model of preferences and represent it mathematically with a *utility function*. Third, we model the budget constraints that individuals face. Finally, we show how individuals maximize their utility (make themselves as well off as possible) given their budget constraints.

Preferences and Indifference Curves

In modeling people's preferences, we are not yet imposing any budget constraints; we are simply asking what people prefer, ignoring what they can afford. Later, we impose budget constraints to round out the model.

Much of the power of the preferences models we use in this course derives from one simple assumption: *non-satiation*, or "more is better." Economists assume that more of a good is always better than less. This does not mean that you are equally happy with the tenth pizza as you are with the first; indeed, as we learn later, your

happiness increases less with each additional unit of a good you consume. Non-satiation simply implies that having that tenth pizza is better than not having it.

Armed with this central assumption, we can move on to graphically represent a consumer's preferences across different bundles of goods. Suppose, for example, that Figure 2-1 represents Andrea's preferences between two goods, cakes (with quantity Q_C) and movies (with quantity Q_M). Consider three bundles:

Bundle *A*: 2 cakes and 1 movie
Bundle *B*: 1 cake and 2 movies
Bundle *C*: 2 cakes and 2 movies

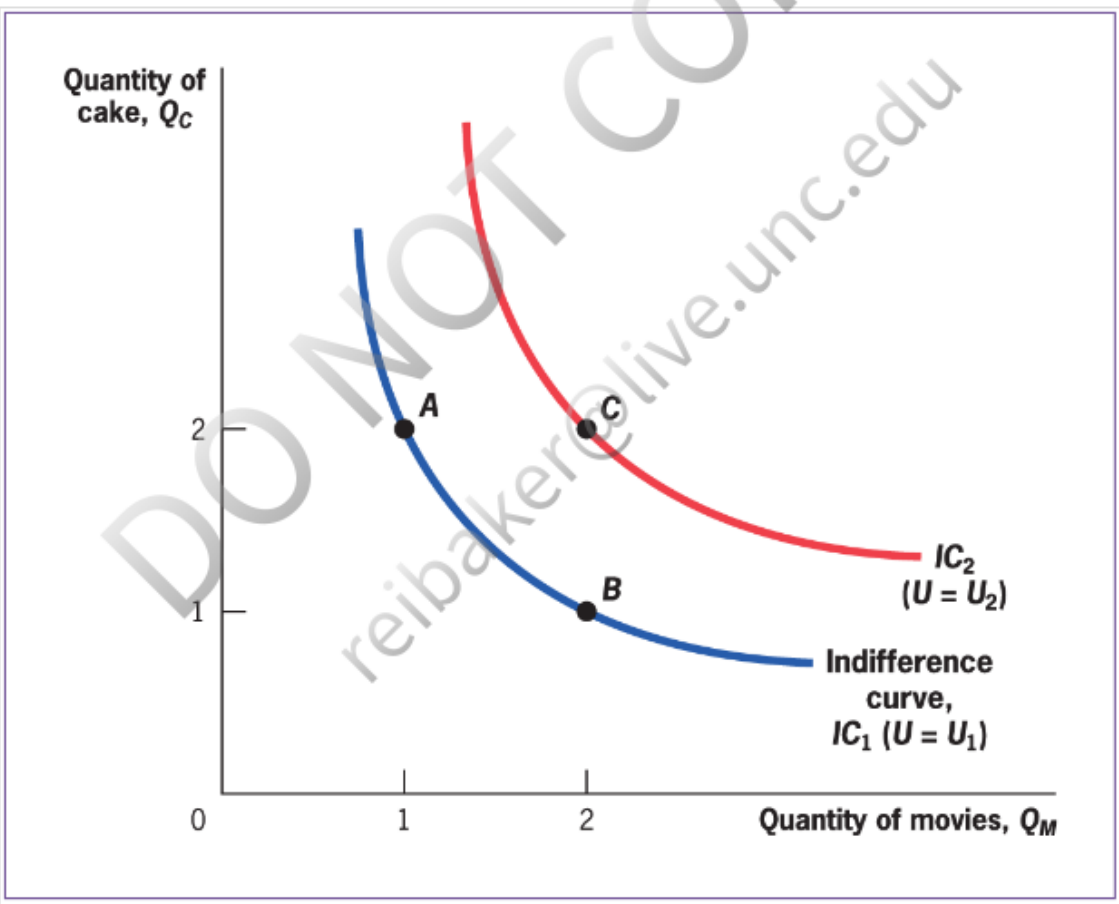

FIGURE 2-1 Indifference Curves for Bundles of Cakes and Movies • Andrea is indifferent between consuming 2 cakes and 1 movie (point *A*) or 1 cake and 2 movies (point *B*), but she prefers 2 cakes and 2 movies (point *C*) to both. Utility is the same along a given indifference curve; indifference curves farther from the origin represent higher utility levels.

Let's assume, for now, that Andrea is indifferent between bundles *A* and *B*, but that she prefers *C* to either; she clearly prefers *C* because of the assumption that more is better. Given this assumption, we can map her preferences across the goods. We do

so using an **indifference curve**, a curve that shows all combinations of consumption that give the individual the same amount of utility (and so among which the individual is indifferent). In this case, Andrea gets the same utility from bundles *A* and *B*, so they lie on the same indifference curve. Because she gets a higher level of utility from consuming bundle *C* instead of either *A* or *B*, bundle *C* is on a higher indifference curve.

indifference curve
A graphical representation of all bundles of goods that make an individual equally well off. Because these bundles have equal utility, an individual is indifferent as to which bundle she consumes.

Indifference curves have two essential properties, both of which follow naturally from the more-is-better assumption:

1. Consumers prefer higher indifference curves. Individuals prefer to consume bundles that are located on indifference curves that are farther out from the origin because they represent bundles that have more of, for example, both cakes *and* movies.
2. Indifference curves always slope downward. Indifference curves cannot slope upward because that would imply that, in this instance, Andrea is indifferent between a given bundle and another bundle that has more of both cakes *and* movies, which violates the more-is-better assumption.

Imagine that the government could somehow insulate Andrea from the utility she loses when prices rise. That is, suppose the government was somehow able to compensate her enough that she could stay on the same indifference curve (IC_1, in our example), even with the new set of prices. Would this mean that the price change will have no effect on her choices? No, it wouldn't because she would still like to choose a different bundle of cakes and movies at this new set of prices. Figure 2-7 also illustrates this point. Despite this price change, the government can hold Andrea's utility constant at these new prices by giving her a budget constraint BC_g, which is parallel to BC_2 but tangent to the same indifference curve IC_1 that corresponds to her original choice. Graphically, the budget constraint has steepened, but Andrea is on the same indifference curve (the same level of utility). Andrea chooses the bundle represented by point *B*: because movies are relatively more expensive, she chooses to consume fewer movies (4.24) and more cakes (4.24). This effect of a price change is called the **substitution effect**: holding utility constant, a relative rise in the price of a good will always cause a consumer to choose less of that good.

substitution effect
Holding utility constant, a relative rise in the price of a good will always cause an individual to choose less of that good.

In the real world, when prices rise, there is no government agency to hold utility constant. This price rise therefore leads to a second effect on demand: Andrea is now effectively poorer because she has to pay higher prices for movies. She is not poorer in an income sense (her income remains at $96), but she is poorer in a real sense because her $96 can buy fewer goods (in particular, fewer movies). This is the **income effect** of a price change: a rise in any price will make the consumer effectively poorer, causing her to choose less of all goods.[2] The quantity demanded falls because Andrea can buy fewer goods with her income.

income effect
A rise in the price of a good will typically cause an individual to choose less of all goods because her income can purchase less than before.

We can measure this income effect by the change from the government-supported budget constraint BC_g to the new budget constraint BC_2. This change represents the restriction in Andrea's opportunity set at the new prices. Because she is poorer, she chooses fewer of all goods, including both movies and cakes, at point *C*. In this case, the income effect reinforces the substitution effect for movies: both cause the

quantity of movies she demands to fall.[3] To sum up, when the price of one good increases relative to another, you choose less of that good for two reasons: because it is relatively more expensive (the substitution effect) and because you are effectively poorer (the income effect).

2.2 Putting the Tools to Work: TANF and Labor Supply Among Single Parents

In your new position with the state government, you have now reviewed the theoretical concepts necessary to address the concerns of the HHS secretary and the governor. Having reviewed these theoretical concepts, let's turn to the question posed at the start of the chapter: Will reducing TANF benefits increase the labor supply of single parents? To answer this question, we can apply the tools of utility maximization to the analysis of the labor supply decision.

The TANF program was created in 1996 by a major overhaul of the *cash transfer* system in the United States. The cash transfer system distributes money from taxpayers to low-income families (as described in much more detail in Chapter 17). TANF provides a monthly support check to families with incomes below a threshold level that is set by each state. In the state of New Jersey, for example, a single parent with two children and no other source of income will receive a monthly check for $454.[4] These checks are largely targeted to single-female-headed households with children because these families are viewed as having the worst prospects for making a living on their own.

Suppose that Joelle is a single mother who spends all of her earnings and TANF benefits on food for herself and her children. By working more hours, she can earn more money for food. But there is a cost to work: she has less time for *leisure*, which includes time spent caring for her kids and managing her household. Suppose that she would prefer time at home to time at work; that is, suppose that *leisure* is a normal good. With these preferences, more work makes Joelle worse off, but it allows her to buy more food.

How does Joelle decide on the optimal amount of labor to supply? To answer this question, we return to the utility maximization framework, but with one twist relative to the decision to purchase cakes and movies. In that case, we were considering two goods. Now, Joelle is considering one good (food consumption) and one "bad" (labor, as time spent at work detracts from time she can spend at home). The trick to modeling this decision is to model the demand for leisure, the good that is the counterpart of labor. That way, we can model the trade-off between two goods using our existing tools and then compute the amount of labor supplied as total work hours minus hours of leisure.

Identifying the Budget Constraint

Suppose that Joelle can work up to 2,000 hours per year at a wage of $10 per hour, that she has no other source of income, and that there is not yet a TANF program in place. By working one less hour in a year, Joelle will lower her consumption by $10 and increase her leisure time by one hour. Thus, the "price" of one hour of leisure time is the hourly wage rate. This fact follows from the principle of opportunity cost: when Joelle opts to take an hour of leisure, her next best alternative activity is to work. Thus, the price of the hour of leisure is $10, the forgone wage she could have earned if working.

The price of food consumption is given directly by the market; let's say that it is $1 per unit of food. This means that Joelle faces a trade-off: each hour of work brings her 10 units of food, and each hour of leisure (time not spent working) costs her 10 units of food. If Joelle can work up to 2,000 hours per year, we can now identify her budget constraint as line *ABC* in Figure 2-8: she can consume a maximum of $20,000 of food per year, a maximum of 2,000 hours of leisure per year, or any combination in between. The slope of the budget constraint is the negative of the ratio of the price of leisure ($10) to the price of consumption ($1), −10.

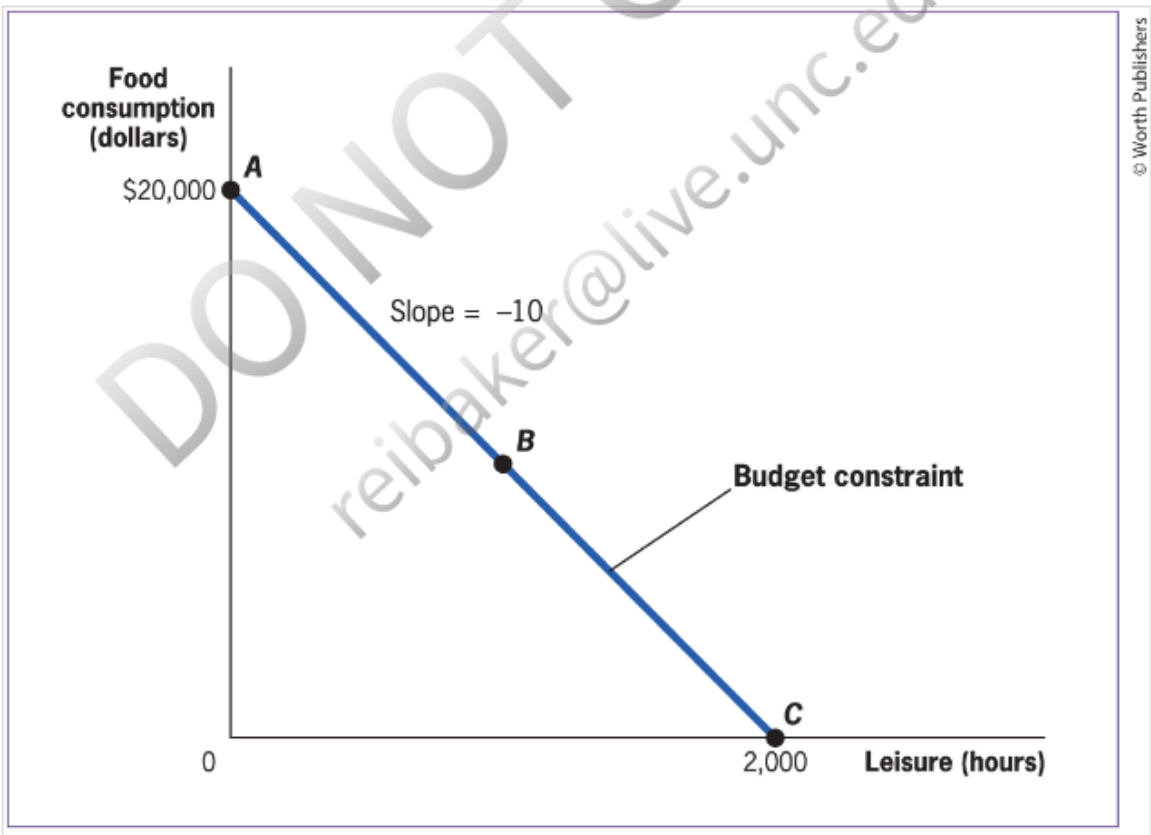

FIGURE 2-8 The Consumption–Leisure Trade-off • Joelle has a choice of taking more leisure and consuming less or taking less leisure (working for more hours) and consuming more. If she takes no leisure, she can have

consumption of $20,000 per year; but if she takes 2,000 hours of leisure, her consumption falls to 0. This is represented by the budget constraint with a slope of −10 the relative price of leisure in terms of food consumption.

The Effect of TANF on the Budget Constraint

Now, let's introduce a TANF program and illustrate what this does to the budget constraint. Programs such as TANF typically have two key features. The first is a *benefit guarantee,* or the baseline amount of money to which recipients are entitled when they enroll in the program. The second is a *benefit reduction rate,* the rate at which the baseline amount is reduced if recipients have other income. For example, a benefit reduction rate of 100% implies that TANF recipients are entitled to the benefit guarantee if they have no other income, but that they lose a dollar of the benefit guarantee for each dollar of other income they earn. A benefit reduction rate of only 50% implies that TANF recipients once again get the full benefit guarantee if they have no other income, but that they lose $0.50 of the benefit guarantee for each $1 they earn. The benefit reduction rate is, in effect, an *implicit tax rate;* it is the rate at which TANF benefits are reduced when recipients earn other income.

We can now add the TANF program to the budget constraint in panel (a) of Figure 2-9. Let's assume the TANF program we're considering has a benefit guarantee of $5,000 and a benefit reduction rate of 50%. The original budget constraint is the line *ABC*. If Joelle chooses 1,000 or fewer hours of leisure, earning $10,000 to $20,000, the budget constraint does not change, remaining as the segment *AB*. This is because, with a benefit guarantee of $5,000 and a benefit reduction rate of 50%, once she earns $10,000 ($5,000/0.5), she is no longer eligible for TANF, so it doesn't affect her budget constraint.

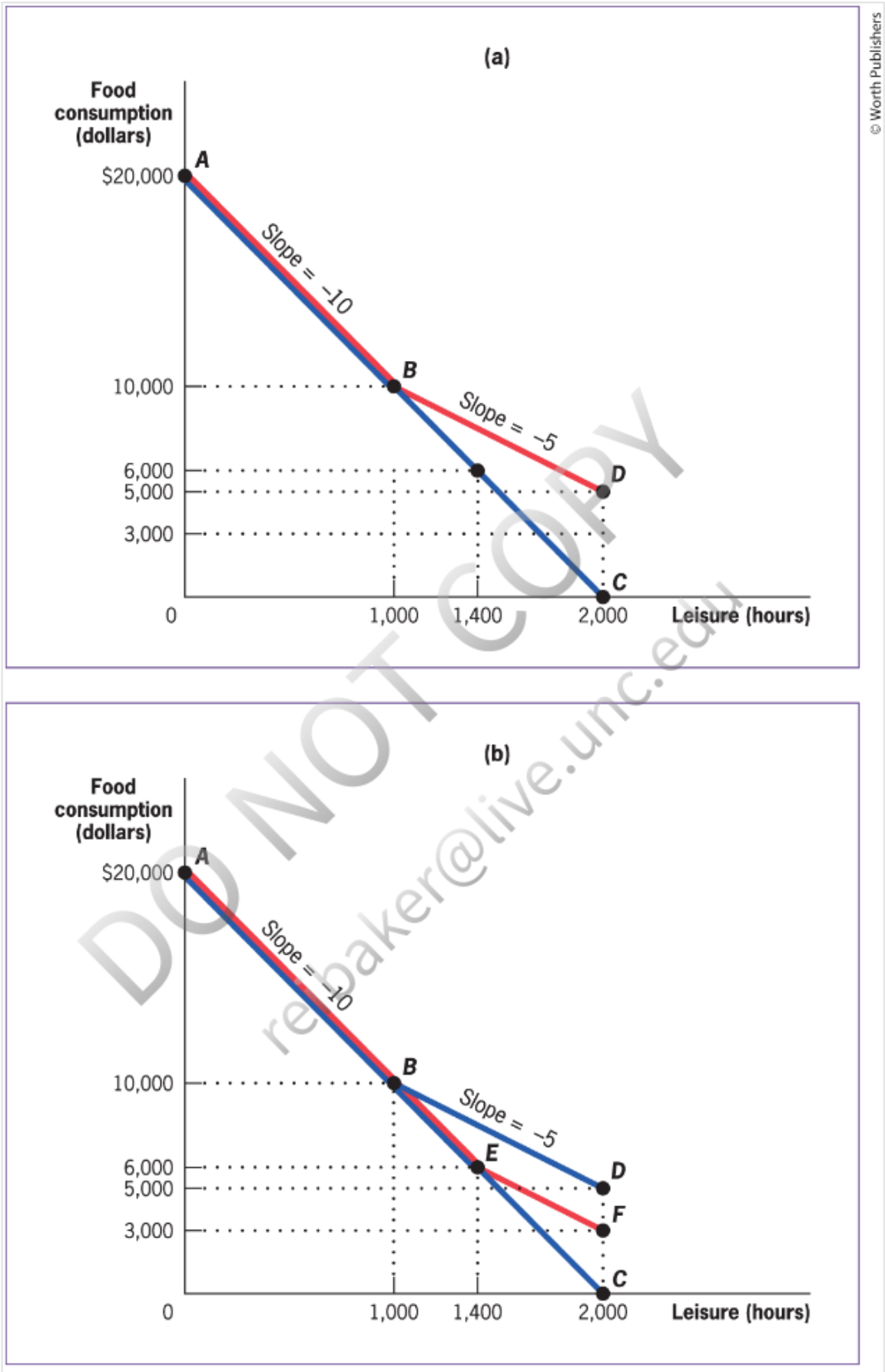

FIGURE 2-9 The Budget Constraint with TANF • Joelle's original budget constraint is *ABC*. With a TANF guarantee of $5,000 and a benefit reduction rate of 50% in panel (a), the budget constraint becomes *ABD*. Once

she has taken more than 1,000 hours of leisure, the budget constraint flattens, and she now can enjoy $5,000 of consumption even with 2,000 hours of leisure at point *D*. When the guarantee falls to $3,000 in panel (b), the budget constraint (*AEF*) doesn't flatten until she takes more than 1,400 hours of leisure; now, with 2,000 hours of leisure, her consumption is only $3,000 at point *F*.

If Joelle chooses to take more than 1,000 hours of leisure, however, the budget constraint becomes flatter. Previously, the price of leisure was $10 per hour because that was the forgone wage. With the 50% benefit reduction rate, however, if Joelle works another hour, she earns $10 in wages but loses $5 in TANF benefits. Under these conditions, the net return to working another hour is now only $5, so the price of leisure falls to $5 per hour. The budget constraint is therefore flatter, with a slope of only −5 rather than −10, because in the range where TANF is available, there is a lower relative price of leisure. Point *D* marks the end of the new budget constraint and provides a new option for Joelle: she can have 2,000 hours of leisure *and* $5,000 in food consumption because of the $5,000 TANF benefit guarantee. Without TANF, if she had chosen to consume 2,000 hours of leisure, she wouldn't have been working at all, and her family would have had no food (point *C*).

Effects of Changes in Benefit Guarantee

Suppose that your state is considering reducing the income guarantee under TANF from $5,000 to $3,000. The effect of this change on the budget constraint is illustrated in panel (b) of Figure 2-9. If Joelle now chooses to take fewer than 1,400 hours of leisure, earning $6,000 to $20,000, the budget constraint does not change, remaining as the segment *AE*. This is because, with the lower benefit guarantee of $3,000 and a benefit reduction rate of 50%, she is now no longer eligible for TANF once she earns $6,000. If she takes more than 1,400 hours of leisure, the budget constraint once again flattens: because she earns $10 in wages but loses $5 in TANF benefits for each hour of work in this range, the slope of the budget constraint along the segment *EF* (the net return to an hour of work) is −5. Point *F* marks the end of the new budget constraint, where Joelle can have 2,000 hours of leisure *and* $3,000 for food consumption because of the $3,000 TANF benefit guarantee. How will single parents react to this policy change?

In answering this question, it is important to return to the concepts of *income and substitution effects* introduced earlier. Suppose, for example, that Joelle earned less than $6,000 before this benefit change. In that case, there is no substitution effect associated with the policy change from a $5,000 benefit guarantee to a $3,000 benefit

guarantee. There is no change in the relative price of leisure, which remains at $5 per hour, so the slope of the budget constraint doesn't change. Whether Joelle gets a $5,000 or a $3,000 check from the government has no impact on the return from working an additional hour ($5 on net), so the price of leisure is unchanged. With relative prices of food and leisure unchanged, there is no desire for substitution across the goods.

There is, however, a clear income effect for Joelle. When the TANF guarantee is reduced, she is poorer. Poorer individuals will reduce their consumption of all normal goods, including leisure. Taking less leisure means working more. In other words, because there is less money available to finance consumption, low-income parents will have to work longer hours. Thus, on net, there is a reduction in leisure from the income effect of reducing the TANF guarantee.

Suppose, instead, that Joelle earned between $6,000 and $10,000 before the benefit change. Once again, this benefit change reduces her income, which will cause her to choose less leisure (and more labor). There is also, however, a change in the price of leisure. In this range of earnings, before the benefits change, an hour of work netted Joelle only $5 per hour, due to the reduction in TANF benefits from additional earnings. Now, because she is no longer eligible for TANF in this income range, an hour of work nets her $10. This relative increase in the price of leisure (taking leisure used to cost $5 but now costs $10 in forgone earnings) will lead to a substitution effect toward less leisure. Thus, in this range, the income and substitution effects work together to reduce leisure.

How Large Will the Labor Supply Response Be?

This example illustrates the power of economic theory. The constrained maximization model implies that a reduction in the benefit guarantee will lead to less leisure and, therefore, more work among single parents. The model does not say, however, how sizeable this response will be. This depends on how much Joelle earned before the benefit change, and the size of the income and substitution effects on her leisure/labor decision.

To illustrate the different possible magnitudes of the response, Figures 2-10 and 2-11 show two different cases. In both cases, we consider utility functions for consumption and leisure, where the utility derived from each is proportional to its *natural logarithm (ln)*. This is a convenient form for utility functions that share most of the properties of the square root utility function we used for cakes and movies,

most notably diminishing marginal utility. As noted earlier, the square root and log forms are just two of many possible forms for utility.

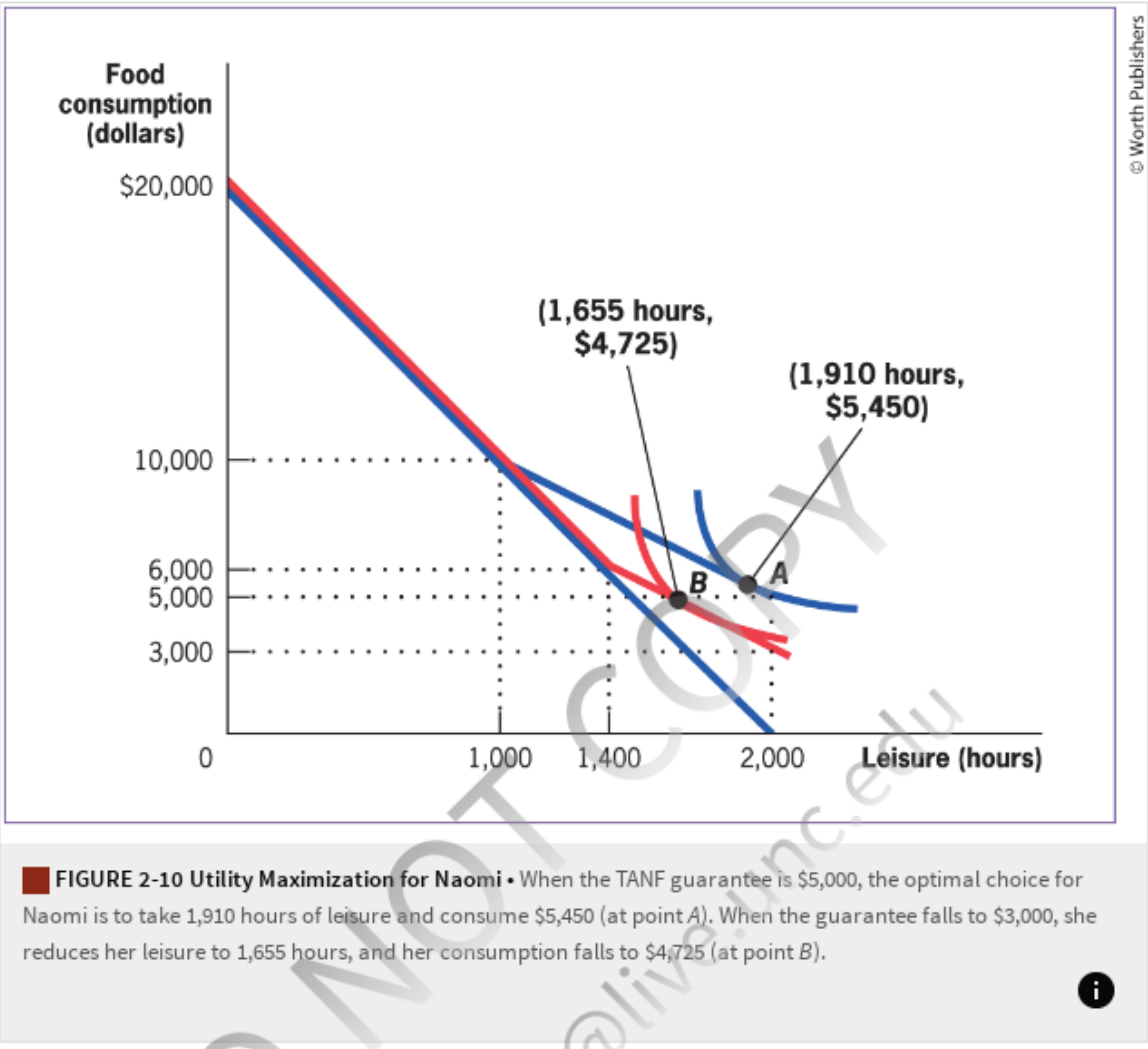

FIGURE 2-10 Utility Maximization for Naomi • When the TANF guarantee is $5,000, the optimal choice for Naomi is to take 1,910 hours of leisure and consume $5,450 (at point *A*). When the guarantee falls to $3,000, she reduces her leisure to 1,655 hours, and her consumption falls to $4,725 (at point *B*).

Our first case describes Naomi, a single parent who has a utility function of the form $U = 100 \times \ln(C) + 175 \times \ln(L)$, where C is consumption, L is leisure, and ln is the natural logarithmic function. Naomi values both consumption and leisure, but she values leisure somewhat more. Figure 2-10 shows her indifference curves and budget constraint. When the guarantee is $5,000, Naomi chooses to consume 1,910 hours of leisure and work 90 hours per year (point *A*, where her indifference curve is tangent to the budget constraint with a $5,000 guarantee). At that level of labor supply, her wage earnings are $900. Because her TANF guarantee is reduced by $0.50 for each $1 of earnings, however, her total income is the $900 in earnings plus a net TANF benefit of $\$5{,}000 - 0.5 \times \$900 = \$4{,}550$. So her total consumption expenditures are $\$900 + \$4{,}550 = \$5{,}450$. (The mathematics of this example is shown in the appendix to this chapter.)

When the TANF guarantee is reduced to $3,000, Naomi chooses to reduce her leisure because she is now poorer (the income effect), moving to point *B* on the new budget constraint. At that point, she takes only 1,655 hours of leisure per year, works 345 hours, and earns $3,450. For Sarah, the governor is right; the reduction in TANF guarantee has raised her labor supply from 90 hours to 345 hours. Note that because Naomi's TANF benefits are reduced by half her earnings, her TANF benefits are now $\$3{,}000 - 0.5 \times \$3{,}450 = \$1{,}275$. Thus, her total budget is $4,725; her consumption has fallen by $725 from the days of the higher TANF guarantee ($\$5{,}450 - \$4{,}725 = \$725$). Her consumption has not fallen by the full $2,000 cut in the guarantee because she has compensated for the guarantee reduction by working more hours.

Figure 2-11 illustrates the case of a different single parent, Sarah, who has a utility function $U = 75 \times \ln(C) + 300 \times \ln(L)$. Sarah puts a much larger weight on leisure relative to consumption, when compared to Naomi. (Her indifference curves are steeper, indicating that a larger increase in consumption is required to compensate for any reduction in leisure.) For Sarah, the optimal choice when the TANF guarantee is $5,000 is to not work at all; she consumes 2,000 hours of leisure and $5,000 of food (point *A*). When the guarantee is reduced to $3,000, this parent continues not to work and just lets her consumption fall to $3,000. That is, she cares so much more about leisure than about consumption that she won't supplement her TANF guarantee with earnings even at the lower guarantee level. For Sarah, the secretary is right; the reduction in TANF guarantee has had no effect on labor supply, it has simply cut her family's level of food consumption.

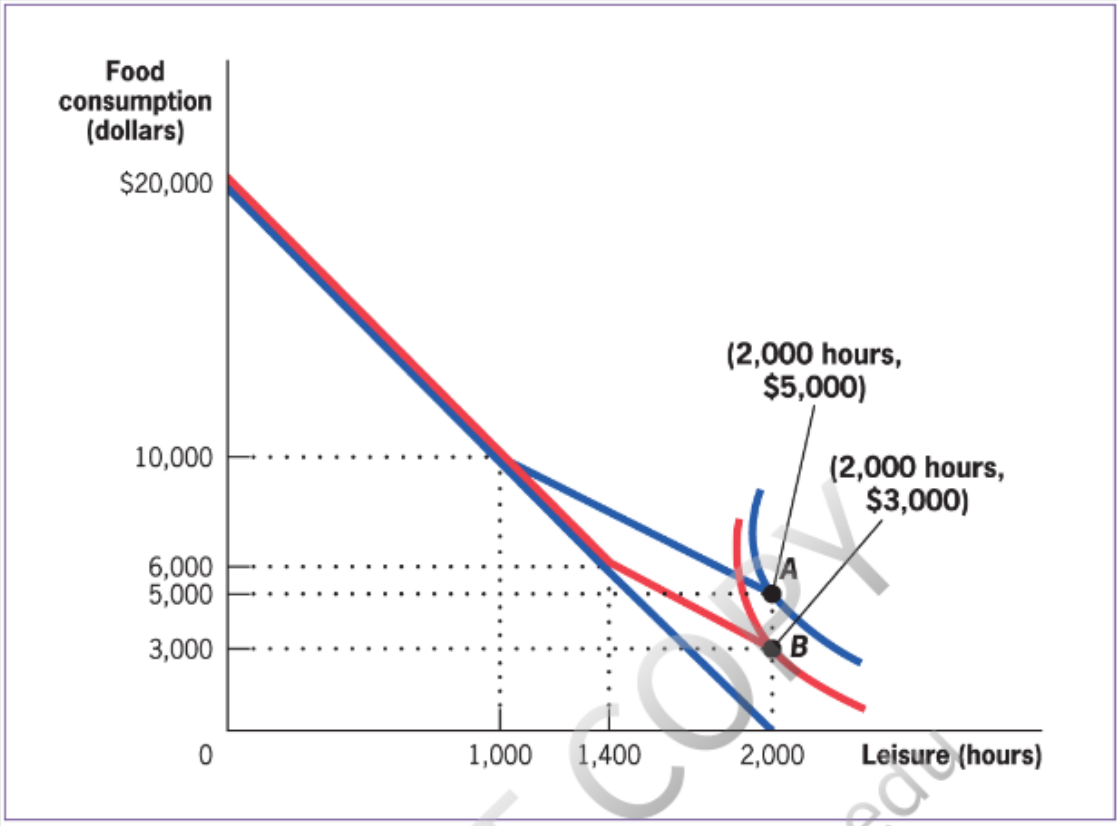

FIGURE 2-11 Utility Maximization for Sarah • Because Sarah values leisure more highly relative to consumption than Naomi in Figure 2-10, she chooses 2,000 hours of leisure regardless of the TANF guarantee. The reduction in guarantee therefore lowers Sarah's consumption from $5,000 (at point *A*) to $3,000 (at point *B*).

Thus, theory alone cannot tell you whether this policy change will increase labor supply or by how much. Theoretically, labor supply could rise, but it might not. To move beyond this uncertainty, you will have to analyze available data on single-parent labor supply, and the next chapter presents the empirical methods for doing so. From these various methods, you will conclude that the governor is right: there is strong evidence that cutting TANF benefits will increase labor supply.

Thus, theory alone cannot tell you whether this policy change will increase labor supply or by how much. Theoretically, labor supply could rise, but it might not. To move beyond this uncertainty, you will have to analyze available data on single-parent labor supply, and the next chapter presents the empirical methods for doing so. From these various methods, you will conclude that the governor is right: there is strong evidence that cutting TANF benefits will increase labor supply.

2.3 Equilibrium and Social Welfare

The disagreement we have been discussing is over whether the labor supply of single parents will rise or not when TANF benefits are cut. As a good public finance economist, however, you know not to stop there. What really should matter to the governor and to the secretary of your department is not a simple fact about whether the labor supply of single parents rises or falls. What should matter is the *normative* question (the analysis of what should be): Does this policy change make society as a whole better off or not?

To address this question, we turn to the tools of normative analysis, **welfare economics.** Welfare economics is the study of the determinants of well-being, or welfare, in society. To avoid confusion, it is important to recall that the term "welfare" is also sometimes used to refer to cash payments (such as those from the TANF program) to low-income single families. Thus, when referring to cash payments in this chapter, we will use the term TANF; our use of the term "welfare" in this chapter refers to the normative concept of well-being.

welfare economics
The study of the determinants of well-being, or welfare, in society.

We discuss the determination of welfare in two steps. First, we discuss the determinants of social efficiency or the size of the economic pie. Social efficiency is determined by the net benefits that consumers and producers receive as a result of their trades of goods and services. We develop the demand and supply curves that measure those net benefits, show how they interact to determine equilibrium, and then discuss why this equilibrium maximizes efficiency. We then turn to a discussion of how to integrate redistribution, or the division of the economic pie, into this analysis so that we can measure the total well-being of society or *social welfare*. In this section, we discuss these concepts with reference to our earlier example of Andrea choosing between movies and cakes; we then apply these lessons to a discussion of the welfare implications of changes in TANF benefits.

Demand Curves

Armed with our understanding of how consumers make choices, we can now turn to understanding how these choices underlie the **demand curve,** the relationship between the price of a good or service and the quantity demanded. Figure 2-12 shows how constrained choice outcomes are translated into the demand curve for movies for Andrea. In panel (a), we vary the price of movies, which changes the

slope of the budget constraint (which is determined by the ratio of movie-to-cake prices). For each new budget constraint, Andrea's optimal choice remains the tangency of that budget constraint with the highest possible indifference curve.

demand curve
A curve showing the quantity of a good demanded by individuals at each price.

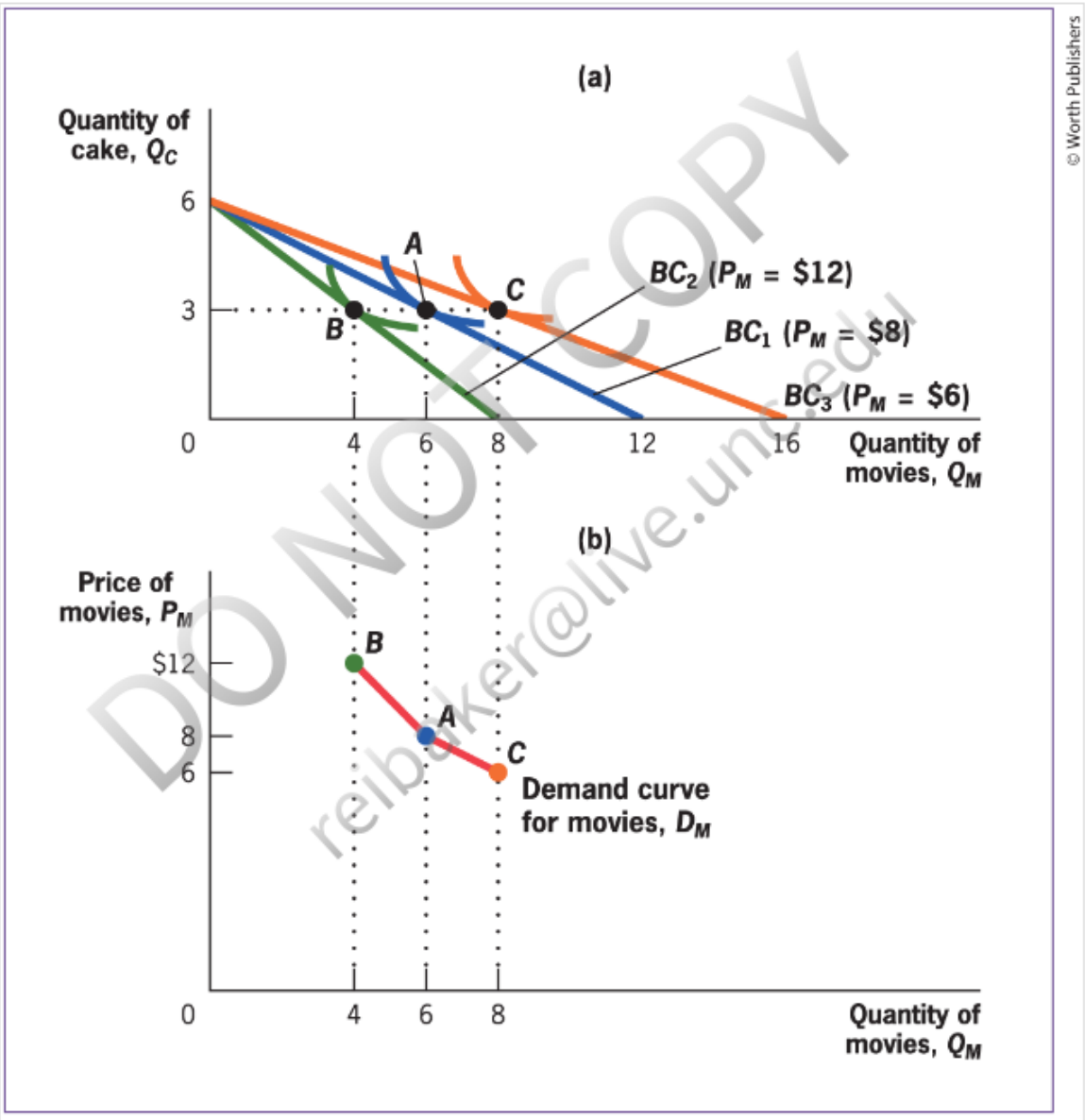

FIGURE 2-12 Deriving the Demand Curve • Changes in the price of movies rotate the budget constraint, changing the number of movies demanded by individuals. When the price of movies rises to $12, then the number of movies demanded falls to 4, and when the price of movies demanded falls to $6, the number of movies demanded rises to 8. We can use this relationship between the price and utility-maximizing choices to trace out the demand curve for movies, D_M, as shown in panel (b).

For example, we have already shown that given her income of $96, at a price of $16 for cakes and $8 for movies, Andrea will choose 6 movies and 3 cakes (point A on BC_1). An increase in the price of movies to $12 will steepen the budget constraint, with the slope rising from $-1/2$ to $-3/4$, as illustrated by BC_2. This increase in price will reduce the quantity of movies demanded so that she chooses 3 cakes and 4 movies (point B on BC_2). A decrease in the price of movies to $6 will flatten the budget constraint, with the slope falling from $-1/2$ to $-3/8$ as illustrated by BC_3. This decrease in price will increase the quantity of movies demanded, and Andrea will now choose to buy 3 cakes and 8 movies (point C on BC_3).

Using this information, we can trace out the demand curve for movies, which shows the quantity of a good or service demanded by individuals at each market price. The demand curve for movies, shown in panel (b), maps the relationship between the price of movies and the quantity of movies demanded.

Elasticity of Demand

A key feature of demand analysis is the **elasticity of demand**, the percentage change in quantity demanded for each percentage change in prices:

$$\varepsilon = \frac{\text{percentage change in quanity demanded}}{\text{percentage change in price}} = \frac{\Delta Q/Q}{\Delta P/P}$$

For example, when the price of movies rises from $8 to $12, the number of movies purchased falls from 6 to 4. So a 50% rise in price leads to a 33% reduction in quantity purchased, for an elasticity of −0.666.

elasticity of demand
The percentage change in the quantity demanded of a good caused by each 1% change in the price of that good.

There are several key points to make about elasticities of demand:

- They are typically negative because quantity demanded typically falls as price rises.
- They are typically not constant along a demand curve. So, in our previous example, the price elasticity of demand is −0.666 when the price of movies rises but is −1.32 when the price of movies falls (a 25% reduction in price from $8 to $6 leads to a 33% increase in demand from 6 to 8 movies).

- A vertical demand curve is one in which the quantity demanded does not change when price rises; in this case, demand is *perfectly inelastic*.
- A horizontal demand curve is one in which quantity demanded changes infinitely for even a very small change in price; in this case, demand is *perfectly elastic*.
- Finally, the example here is a special case in which the demand for cakes doesn't change as the price of movies changes. The effect of one good's price on the demand for another good is the *cross-price elasticity,* and with the particular utility function we are using here, that cross-price elasticity is zero. Typically, however, a change in the price of one good will affect demand for other goods as well.

Supply Curves

The discussion thus far has focused on consumers and the derivation of demand curves. This tells about only one side of the market, however. The other side of the market is represented by the **supply curve**, which shows the quantity supplied of a good or service at each market price. Just as the demand curve is the outcome of utility maximization by individuals, the supply curve is the outcome of *profit maximization* by firms.

supply curve
A curve showing the quantity of a good that firms are willing to supply at each price.

The analysis of firms' profit maximization is similar to that of consumer utility maximization. Just as individuals have a utility function that measures the impact of goods consumption on well-being, firms have a *production function* that measures the impact of firm input use on firm output levels. For ease, we typically assume that firms have only two types of inputs: *labor* (workers) and *capital* (machines, buildings). Consider a firm that produces movies. This firm's production function may take the form $q = \sqrt{K \times L}$, where q is the quantity of movies produced, K is units of capital (such as studio sets), and L is units of labor (such as hours of acting time employed).

The impact of a one-unit change in an input, holding other inputs constant, on the firm's output is the **marginal productivity** of that input. Just as the marginal utility of consumption diminishes with each additional unit of consumption of a good, the marginal productivity of an input diminishes with each additional unit of the input used in production; that is, production generally features *diminishing marginal*

productivity. For this production function, for example, holding K constant, adding additional units of L raises production by less and less, just as with the utility function (of this same form), holding cakes constant, consuming additional movies raised utility by less and less.[5]

marginal productivity
The impact of a one-unit change in any input, holding other inputs constant, on the firm's output.

This production function dictates the cost of producing any given quantity as a function of the prices of inputs and the quantity of inputs used. The total costs of production, *TC*, are determined by $TC = rK + wL$, where r is the price of capital (the rental rate) and w is the price of labor (the wage rate). For day-to-day decisions by the firm, the amount of capital is fixed, while the amount of labor can be varied. Given this assumption, we can define the **marginal cost**, or the incremental cost to producing one more unit, as the wage rate times the amount of labor needed to produce one more unit.

marginal cost
The incremental cost to a firm of producing one more unit of a good.

For example, consider the production function just described, and suppose that the firm is producing movies using 1 unit of capital and 4 units of labor. Now, holding the amount of capital fixed, it wants to produce 3 movies. To do so, it will have to increase its use of labor by 5 units (to 9 total units). If the wage rate is $1 per unit, then the marginal cost of raising production from 2 to 3 movies is $5.

The key point of this discussion is that diminishing marginal productivity generally implies rising marginal costs. To produce a fourth movie would require an increase in labor of 7 units, at a cost of $7; to produce a fifth movie would cost $9. Because each additional unit of production means calling forth labor that is less and less productive, at the same wage rate, the costs of that production are rising.

Recall that the goal of the firm is to maximize its **profit**, the difference between revenues and costs. Profit is maximized when the revenue from the next unit, or the *marginal revenue*, equals the cost of producing that next unit, the *marginal cost*. In a competitive industry, the revenue from any unit is the price the firm obtains in the market. Thus, the firm's profit maximization rule is to produce until price equals marginal cost.

profit

The difference between a firm's revenues and costs, maximized when marginal revenues equal marginal costs.

We can see this through the type of "hill-climbing" exercise proposed in the Quick Hint on p. 34. Suppose the price of movies in the market is $8, the cost of capital is $1 per unit, the cost of labor is $1 per unit, and the firm has 1 unit of capital. Then, if the firm produces 1 movie, it will need to use 1 unit of labor, so that total costs are $2. Because revenues on that first unit are $8, it should clearly produce that first movie. To produce a second movie, the firm will need to use 4 units of labor, or an increase of 3 units of labor. Thus, the marginal cost of that second unit is $3, but the marginal revenue (price) is $8, so the second movie should be produced. For the third movie, the marginal cost is $5, as just noted, which remains below price.

But now imagine the firm is producing 4 movies and is deciding whether to produce a fifth. Producing the fifth movie will require an increase in labor input from 16 to 25 units, or an increase of 9 units. This will cost $9. But the price that the producer gets for this movie is only $8. As a result, producing that fifth unit will be a money loser, and the firm will not do it. Thus, profit maximization dictates that the firm produce until its marginal costs (which are rising by assumption of diminishing marginal productivity) reach the price.

Profit maximization is the source of the supply curve, the relationship between the price and how much producers will supply to the market. At any price, we now know that producers will supply a quantity such that the marginal cost equals that price. Thus, the marginal cost curve is the firm's supply curve, showing the relationship between price and quantity. As quantity rises, and marginal costs rise, the firm will require higher and higher prices to justify producing additional units.

Equilibrium

We have discussed the source of individual demand curves (utility maximization) and firm supply curves (profit maximization). To undertake welfare analysis, we need to translate these concepts to their counterparts at the level of the **market**, the arena in which demanders and suppliers actually interact (such as the supermarket or a website). To do so, we add up the demands of each individual who is demanding goods in this market, and the supplies of each firm that is supplying goods in this market. We *horizontally sum* these curves. That is, at each price, we add up all the quantities available to be purchased at that price by demanders to obtain market-level demand, and all the quantities available to be supplied at that

price by suppliers to obtain market-level supply. The result is the market-level supply and demand curves shown in Figure 2-13.

market
The arena in which demanders and suppliers interact.

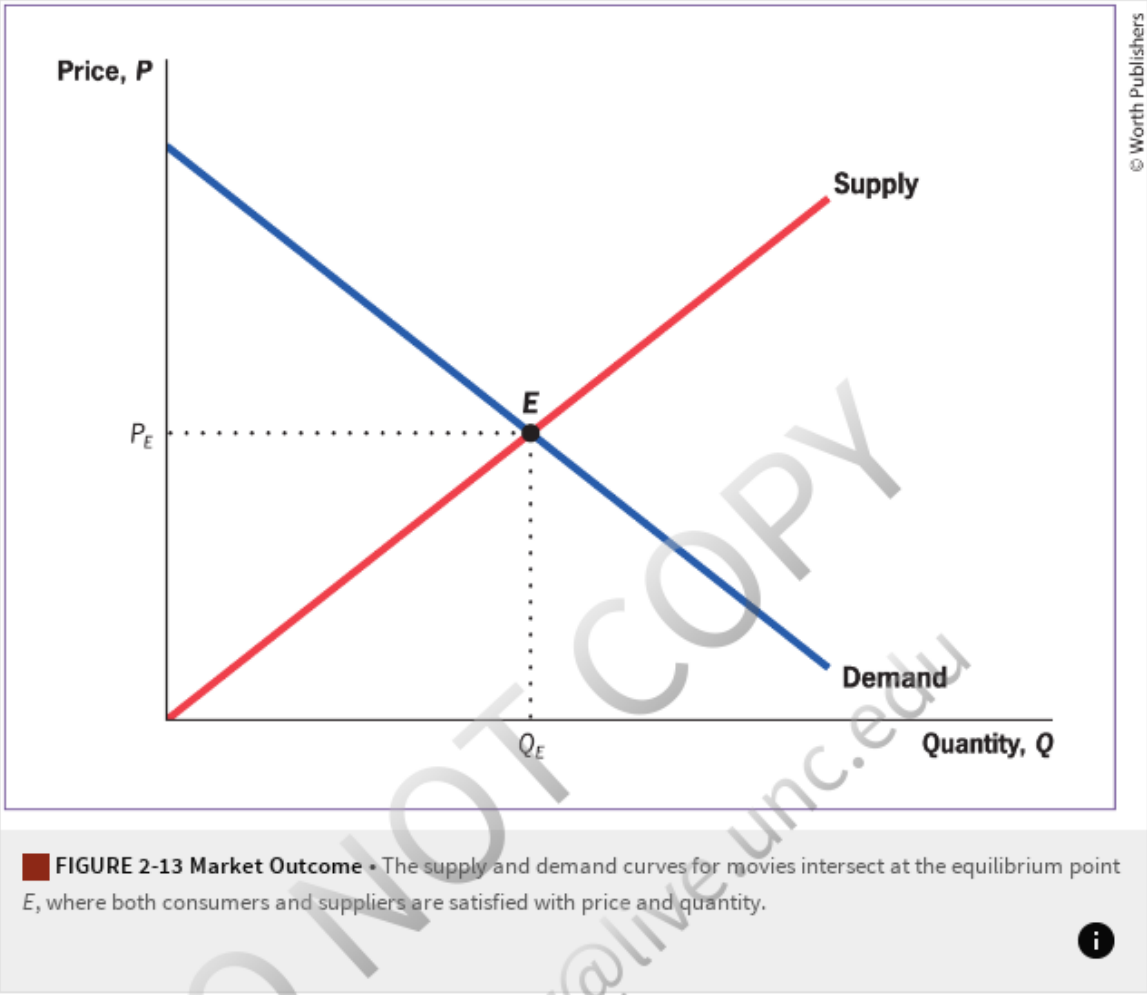

FIGURE 2-13 Market Outcome • The supply and demand curves for movies intersect at the equilibrium point *E*, where both consumers and suppliers are satisfied with price and quantity.

The market-level supply and demand curves interact to determine the **market equilibrium**, the price and quantity pair that will satisfy both demand and supply. This point occurs at the intersection of the supply and demand curves, such as point *E* in Figure 2-13. Given the equilibrium price P_E, demanders will demand the equilibrium quantity, Q_E, and suppliers will be willing to supply that equilibrium quantity. The competitive market equilibrium represents the unique point at which both consumers and suppliers are satisfied with price and quantity.

market equilibrium
The combination of price and quantity that satisfies both demand and supply, determined by the interaction of the supply and demand curves.

Social Efficiency

Social Efficiency

Armed with the analysis thus far, we are now ready to take the final step: to measure *social efficiency,* or the size of the pie. Social efficiency represents the net gains to society from all trades that are made in a particular market, and it consists of two components: consumer and producer surplus.

Consumer Surplus

The gain to consumers from trades in a market for consumer goods is **consumer surplus,** the benefit that consumers derive from consuming a good, above and beyond what they paid for the good. Once we know the demand curve, consumer surplus is easy to measure because each point on a demand curve represents the consumer's willingness to pay for that quantity. It is important to always keep in mind that willingness to pay is dependent on the consumer's resources; willingness to pay is shorthand for "willingness to pay given available resources."

consumer surplus
The benefit that consumers derive from consuming a good, above and beyond the price they paid for the good.

Panel (a) of Figure 2-14 shows a graphical representation of consumer surplus in the movie market: the shaded area below the demand curve and above the equilibrium price P_E (area WZX). This area is consumer surplus because these are units where the willingness to pay (represented by the demand curve) is higher than the amount actually paid, P_E. Consumer surplus is largest on the very first unit because this represents the consumer who most wanted the good (the person who is willing to buy the good at a very high price). For that first unit, consumer surplus equals the distance WX on the vertical axis. Consumer surplus then falls as additional consumers derive less and less marginal utility from the good. Finally, for the consumer whose demand (willingness to pay) equals the price (at point Z), consumer surplus is zero.

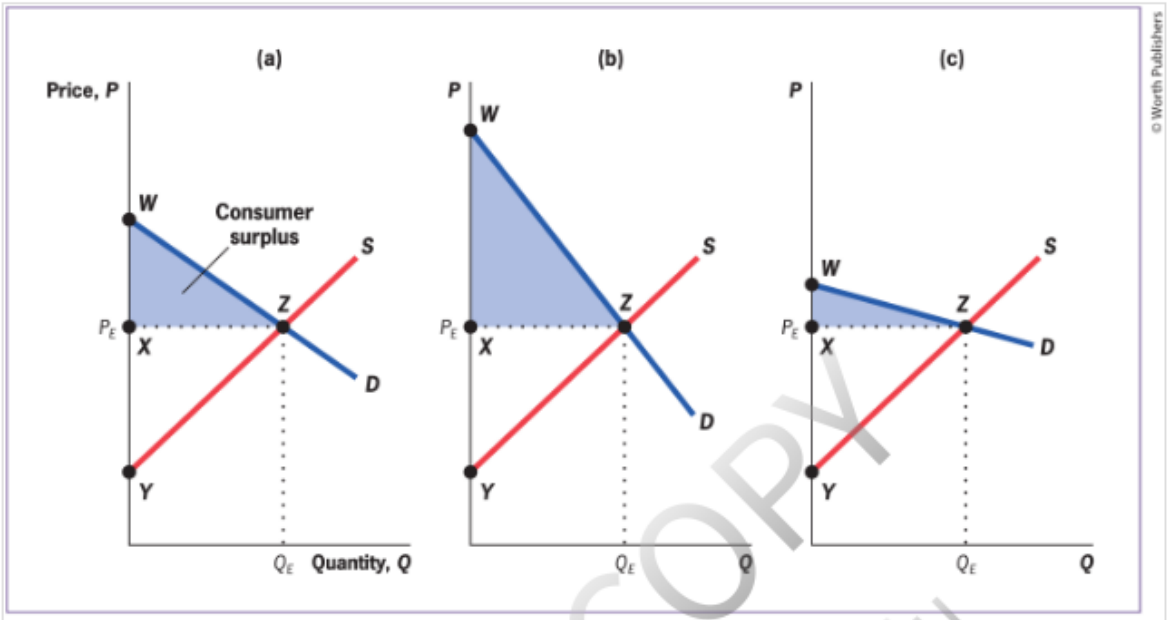

FIGURE 2-14 Consumer Surplus • The consumer surplus is the area below the demand curve and above the equilibrium market price, the shaded area *WZX* in all three panels of this graph. This represents the value to consumers of consuming goods, above and beyond the price paid for those goods. As demand becomes more inelastic, consumer surplus rises; as demand becomes more elastic, consumer surplus falls.

Choosing an Equity Criterion

The form of the social welfare function clearly plays an important role in driving government policy. Yet the *SWF* is not handed down from some higher power but determined in some way by the interplay of politicians and the voting public. The mechanisms through which the *SWF* might evolve through the political process are discussed at length in Chapter 9, but it is important to recognize that there are other criteria besides an *SWF* that might lead to redistributive concerns. For example, some policy makers take the **commodity egalitarianism** view, in which all that matters is that individuals have met a basic level of need for goods such as housing or medical care, and that once they have met this basic level, income distribution is irrelevant. Others argue that all that matters is **equality of opportunity**, whereby individuals are guaranteed an equal chance to succeed, but if some do and others do not, that is not the concern of the government. We discuss these alternative views and their implications for government policy in Chapter 17.

commodity egalitarianism
The principle that society should ensure that individuals meet a set of basic needs, but that beyond that point income distribution is irrelevant.

equality of opportunity
The principle that society should ensure that all individuals have equal opportunities for success but not focus on the outcomes of choices made.

2.4 Welfare Implications of Benefit Reductions: The TANF Example Continued

The equilibrium and social welfare tools developed in Section 2.3 can be applied to evaluate the benefits and costs to society of reducing TANF benefits. The benefits are the improvement in efficiency from removing a barrier to labor supply by single parents, raising single parents' labor supply and raising the size of the social surplus. (Relying on the empirical evidence discussed in the next chapter, we assume that labor supply increases when benefits fall.) The costs are the reductions in equity that arise from reducing income support to one of the lowest-income groups in our society. The job of public finance economists is to measure these efficiency and equity consequences. The job of policy makers is to trade the consequences off to decide on appropriate policy choices.

Efficiency

We can apply the tools of welfare analysis to model the welfare implications of cutting TANF benefits. Figure 2-17 shows the market for labor services by single parents. The price of labor, the wage (W), is on the vertical axis; the amount of hours worked in aggregate in the market (H) is on the horizontal axis.

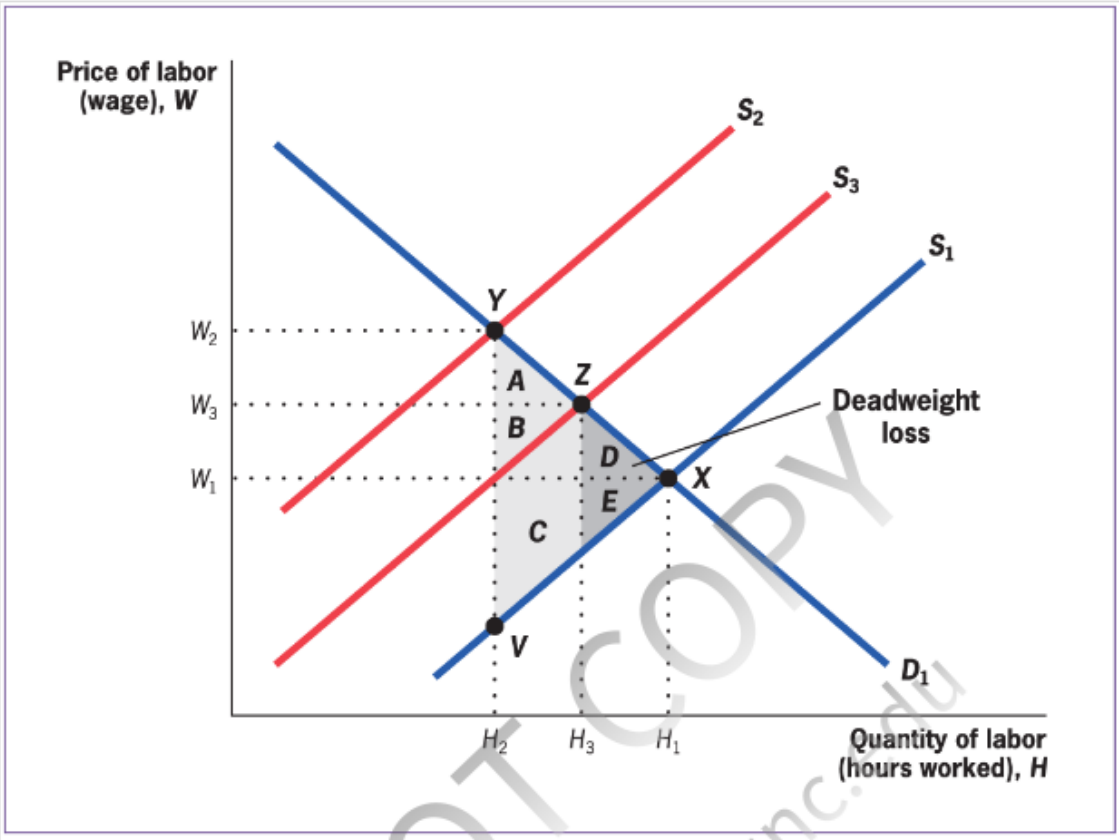

FIGURE 2-17 Welfare Implications of TANF • Without TANF, the labor market is in competitive equilibrium at point *X*, the intersection of S_1 and D_1. When TANF is introduced, labor supply falls to S_2, and the market moves to a new equilibrium at point *Y*, creating a deadweight loss of $A + B + C + D + E$. When TANF benefits are reduced, supply increases to S_3, and social efficiency rises by $A + B + C$.

Unlike Figure 2-13, the demand for the good (the single parent's hours of work) comes from firms, and the supply comes from individuals. Nevertheless, as in Figure 2-13, the demand curve slopes downward (as wages rise, firms demand fewer hours of work) and the supply curve slopes upward (as wages rise, individuals are willing to supply more hours of work—assuming that substitution effects are larger than income effects).

Suppose that, in the absence of the TANF program, there are no other government interventions that affect the labor market. In that case, without TANF, labor supply, S_1, intersects labor demand, D_1, at point *X*, and the market is in competitive equilibrium, maximizing social efficiency at hours of work H_1.

When TANF is introduced, however, single parents work fewer hours, reducing the supply of labor at every wage, so that the supply curve shifts left to S_2. The labor market will reach a new equilibrium at point *Y*. Relative to the original equilibrium,

the number of hours worked has fallen from H_1 to H_2. This reduction in hours worked causes a deadweight loss of the area $A + B + C + D + E$. The difference between H_1 and H_2 represents hours of work that the single parent would willingly provide to the firm, and the firm would happily demand from her, were it not for the TANF program. Social efficiency has thus fallen.

If TANF benefits are cut, the labor supply of single parents increases and the supply curve shifts out to S_3. At the new equilibrium Z, the single parents supply H_3 hours of labor, and the deadweight loss has been reduced to $D + E$. That is, social efficiency has grown by the area $A + B + C$ due to this reduction in TANF benefits.

We can now quantify the social efficiency gain to lower TANF benefits: area $A + B + C$ is gained when single parents increase their supply of labor. If we know the slopes of these demand and supply curves, we can then measure this social efficiency gain. These slopes can be estimated using the types of empirical methods we discuss in Chapter 3.

Equity

Given this large efficiency gain, why not cut TANF benefits? Indeed, why have the TANF program at all? As just noted, governments have programs such as TANF because their citizens care not only about efficiency but also about equity, the fair distribution of resources in society. For many specifications of social welfare, the competitive equilibrium, while being the social efficiency-maximizing point, may not be the social welfare-maximizing point.

Currently, the share of single parents living below the *poverty line*, a measure of the minimal income required to live in the United States, is 27.4%, compared with only 11.6% for all families.[7] Cutting TANF benefits would therefore worsen outcomes for a population that is already one of the worst off in society. Cutting TANF benefits could have dramatic equity costs that offset the efficiency gains.

To consider a simple example, imagine that society has a utilitarian *SWF* and that each individual in society has a utility function of the form $U = \sqrt{C}$, where C = consumption = income. Imagine further that 10% of citizens are single parents who have an initial income of $10,000, and the remaining 90% of citizens have an initial income of $50,000. Suppose that if we cut TANF benefits, the income of single parents falls to $5,000, while the income of everyone else rises to $51,000. Under this policy, the average level of income in society rises from $46,000 to $46,400, so total

social efficiency has risen. Yet social welfare has fallen; the average utility level has fallen from 211.2 to 210.3 (computed by averaging across all citizens the square root of income both before and after this change). This is because we are adding small amounts of income to the high-income majority, who already have a low marginal utility of income, but we are taking large amounts of income away from the low-income minority, who have a very high marginal utility of income. While this policy move raises efficiency, it harms equity even more in the context of this *SWF*.

Measuring empirically the cost to society from this reduced equity is quite difficult. Essentially, the analyst must make some assumption about how society values the well-being of different groups, such as single parents versus other taxpayers.

To consider a simple example, imagine that society has a utilitarian *SWF* and that each individual in society has a utility function of the form $U = \sqrt{C}$, where $C =$ consumption $=$ income. Imagine further that 10% of citizens are single parents who have an initial income of $10,000, and the remaining 90% of citizens have an initial income of $50,000. Suppose that if we cut TANF benefits, the income of single parents falls to $5,000, while the income of everyone else rises to $51,000. Under this policy, the average level of income in society rises from $46,000 to $46,400, so total social efficiency has risen. Yet social welfare has fallen; the average utility level has fallen from 211.2 to 210.3 (computed by averaging across all citizens the square root of income both before and after this change). This is because we are adding small amounts of income to the high-income majority, who already have a low marginal utility of income, but we are taking large amounts of income away from the low-income minority, who have a very high marginal utility of income. While this policy move raises efficiency, it harms equity even more in the context of this *SWF*. Measuring empirically the cost to society from this reduced equity is quite difficult. Essentially, the analyst must make some assumption about how society values the well-being of different groups, such as single parents versus other taxpayers.

To consider a simple example, imagine that society has a utilitarian *SWF* and that each individual in society has a utility function of the form $U = \sqrt{C}$, where $C =$ consumption $=$ income. Imagine further that 10% of citizens are single parents who have an initial income of $10,000, and the remaining 90% of citizens have an initial income of $50,000. Suppose that if we cut TANF benefits, the income of single parents falls to $5,000, while the income of everyone else rises to $51,000. Under this policy, the average level of income in society rises from $46,000 to $46,400, so total social efficiency has risen. Yet social welfare has fallen; the average utility level has fallen from 211.2 to 210.3 (computed by averaging across all citizens the square root of income both before and after this change). This is because we are adding small amounts of income to the high-income majority, who already have a low marginal utility of income, but we are taking large amounts of income away from the low-income minority, who have a very high marginal utility of income. While this policy move raises efficiency, it harms equity even more in the context of this *SWF*. Measuring empirically the cost to society from this reduced equity is quite difficult. Essentially, the analyst must make some assumption about how society values the well-being of different groups, such as single parents versus other taxpayers.

2.5 Conclusion

This chapter has shown both the power and the limitations of the theoretical tools of economics. On the one hand, by making relatively straightforward assumptions about how individuals and firms behave, we are able to address complicated questions such as how TANF benefits affect the labor supply of single parents, and the implications of that response for social welfare. On the other hand, while we have answered these questions in a general sense, we have been very imprecise about the potential size of the changes that occur in response to changes in TANF benefits. That is, theoretical models can help point to the likely impacts of policy changes on individual decisions and social welfare, but they cannot tell us the magnitude of those effects. To do so, we have to turn to empirical economics, which we will do in the next chapter.

Highlights

- Policy debates such as that over the appropriate level of Temporary Assistance for Needy Families (TANF) benefits motivate the need for theoretical modeling of individual and firm decision-making behaviors.
- Modeling the impact of policy changes on individual behavior requires the use of utility-maximization models in which individuals maximize their well-being, subject to market prices and their available resources.
- Individual well-being, or utility, is maximized when individuals choose the bundle of goods that equates the rate at which they want to trade off one good for another (the marginal rate of substitution) with the rate at which the market allows them to trade off one good for another (the price ratio).
- TANF-like programs introduce complicated budget constraints with several possible segments, depending on whether a parent is on or off the program.
- Reducing TANF benefits is likely to increase the labor supply of single parents, but the size of the increase is unclear and depends on the parents' preferences for leisure and consumption.
- Social welfare is determined by considering both social efficiency (the size of the pie) and equity (the distribution of the pie).
- Social efficiency is maximized at the competitive equilibrium, where demand (which is derived from underlying utility maximization) equals supply (which is derived from underlying profit maximization).
- Social welfare is maximized by using a social welfare function to incorporate both efficiency and society's preferences for redistribution into policy making.
- Because reducing TANF benefits moves the labor market closer to the competitive equilibrium, it raises total social efficiency, but at a cost of lowering the incomes of a particularly needy group. The net impact on social welfare is unclear.

APPENDIX TO CHAPTER 2: The Mathematics of Utility Maximization

This appendix develops the mathematics behind the utility-maximization example presented on pp. 41–43 and in Figures 2-10 and 2-11. The utility function that underlies the indifference curves in Figure 2-10 is:

$$U = 100 \times \ln(C) + 175 \times \ln(L)$$

where C is consumption, and L is leisure.

For this utility function, the marginal rate of substitution is:

$$MRS = MU_L / MU_C = (175/L) / (100/C) = 1.75 \times (C/L)$$

Naomi has a market wage of $10 per hour and can work up to 2,000 hours per year. She is also subject to a TANF program that features a benefit guarantee of $5,000 and a benefit reduction rate of 50%. As a result, the budget constraint has two segments:

$C = 5{,}000 + (2{,}000 - L) \times 10 \times 0.5$ if leisure is more than 1,000 hours (TANF segment)

$C = (2{,}000 - L) \times 10$ if leisure is less than 1,000 hours (non-TANF segment)

Given this budget constraint, we can solve for the optimal amount of leisure and consumption for this single parent. We do this by first finding her optimal leisure and consumption bundle on each of the two segments of the budget constraint and then evaluating which of those choices leads to higher total utility.

On the first (TANF) segment of the budget constraint, we solve the problem:

$$\text{Maximize } U = 100 \times \ln(C) + 175 \times \ln(L)$$

$$\text{Subject to } C = 5{,}000 + (2{,}000 - L) \times 10 \times 0.5$$

Substituting from the budget constraint into the utility function, we obtain:

$$\text{Maximize } U = 100 \times \ln(5{,}000 + (2{,}000 - L) \times 10 \times 0.5) + 175 \times \ln(L)$$

We maximize this by taking the differential of utility with respect to leisure and setting it equal to zero:

$$(100 \times -5)/(5{,}000 + (2{,}000 - L) \times 5) + 175/L = 0$$

Solving this equation, we obtain $L = 1{,}910$. At that level of leisure, consumption is 5,450. This implies a utility of $100 \times \ln(5{,}450) + 175 \times \ln(1{,}910) = 2{,}182$.

Now, we can solve the problem again for the second (non-TANF) segment of the budget constraint:

$$\text{Maximize } U = 100 \times \ln(C) + 175 \times \ln(L)$$

$$\text{Subject to } C = (2{,}000 - L) \times 10$$

Once again, substituting from the budget constraint into the utility function, we obtain:

$$\text{Maximize } U = 100 \times \ln((2{,}000 - L) \times 10) + 175 \times \ln(L)$$

Taking the differential of utility with respect to leisure and setting this to zero, we can solve for an optimal L of 1,273 and resulting consumption of 7,270. Plugging

these values back into the utility function, we get a value for utility from this choice of 2,140. This utility value is lower than 2,182, so the individual will choose point A on the first (TANF) segment of the budget constraint.

What happens when we lower the TANF guarantee to $3,000? We can solve the same problem, but now with the lower guarantee level. Doing so, we find that the single parent would still choose to be on the TANF segment of the budget constraint, with leisure of 1,655 hours (and work of 345 hours).

The utility function that underlies the indifference curves in Figure 2-11 is:

$$U = 75 \times \ln(C) + 300 \times \ln(L)$$

For this utility function, the marginal rate of substitution is:

$$MRS = MU_L / MU_C = (300/L)/(75/C) = 4 \times (C/L)$$

Sarah's budget constraint is the same as that of Naomi's:

$C = 5{,}000 + (2{,}000 - L) \times 10 \times 0.5$ if leisure is more than 1,000 hours (TANF segment)

$C = (2{,}000 - L) \times 10$ if leisure is less than 1,000 hours (non-TANF segment)

On the first (TANF) segment of the budget constraint, we solve the problem:

$$\text{Maximize } U = 75 \times \ln(C) + 300 \times \ln(L)$$

$$\text{Subject to } C = 5{,}000 + (2{,}000 - L) \times 10 \times 0.5$$

Doing so, we obtain an optimal value of leisure of 3,200. This value exceeds the maximum possible level of leisure, 2,000. So the parent chooses to take that maximum value, with leisure of 2,000 and consumption of 5,000, for a utility level of 2,919.

This parent will be worse off on the non-TANF segment of the budget constraint because she wants so much leisure. Likewise, solving the problem for the $3,000 guarantee, we once again find that she chooses the "corner" solution of 2,000 hours of leisure and 3,000 units of consumption.

CHAPTER 3

Empirical Tools of Public Finance

James Andrews/iStock/Getty Images

Questions to keep in mind

- How do you estimate the effect of policy changes on individual behavior?
- What are "quasi-experiments," and how can they be used to obtain behavioral estimates?
- How do changes in TANF benefits affect single-parent labor supply?

Once again, we return to your days as an employee of your state's Department of Health and Human Services. After doing the careful theoretical analysis outlined in Chapter 2, you are somewhat closer to making a meaningful contribution to the debate between the governor and the secretary of Health and Human Services. You can tell the governor and the secretary that a reduction in TANF benefits is likely, but not certain, to raise labor supply among single parents and that the implications of this response depend on their concerns about equity versus efficiency. Yet these politicians don't just want to know that TANF reductions might raise labor supply,

nor are they interested in the graphical calculations of the social welfare effects of lower benefits. They want numbers.

To provide these numbers, you now turn to the **empirical tools of public finance**, the use of data and statistical methodologies to measure the impact of government policy on individuals and markets. Many of these tools were developed more recently than the classical analyses of utility maximization and market equilibrium that we worked with in Chapter 2. As a result, they are also more imperfect, and there are lively debates about the best way to approach problems such as estimating the labor-supply response of single parents to TANF benefit changes.

empirical tools of public finance
The use of data and statistical methods to measure the impact of government policy on individuals and markets.

In this chapter, we review these empirical methods and encounter the fundamental issue faced by those doing empirical work in economics: disentangling causality from correlation. We say that two economic variables are **correlated** if they move together. But this relationship is **causal** only if one of the variables *causes* the movement in the other. If, instead, there is a third factor that causes both to move together, the correlation is not causal.

correlated
Two economic variables are correlated if they move together.
causal
Two economic variables are causally related if the movement of one causes movement of the other.

This chapter begins with a review of this fundamental problem. We then turn to a discussion of the "gold standard" for measuring the causal effect of an intervention, a *randomized trial,* in which individuals are randomly assigned to receive or not receive the intervention. Randomized trials, which are much more common in medicine than in public finance, provide a benchmark against which other empirical methods can be evaluated. We then discuss the range of other empirical methods used by public finance economists to answer questions such as the causal impact of TANF benefit changes on the labor supply of single parents. Throughout, we use this TANF example, using real-world data on benefit levels and the single-parent labor supply, to assess the questions raised by the theoretical analysis in Chapter 2.

This chapter begins with a review of this fundamental problem. We then turn to a discussion of the "gold standard" for measuring the causal effect of an intervention, a *randomized trial,* in which individuals are randomly assigned to receive or not receive the intervention. Randomized trials, which are much more common in medicine than in public finance, provide a benchmark against which other empirical methods can be evaluated. We then discuss the range of other empirical methods used by public finance economists to answer questions such as the causal impact of TANF benefit changes on the labor supply of single parents. Throughout, we use this TANF example, using real-world data on benefit levels and the single-parent labor supply, to assess the questions raised by the theoretical analysis in Chapter 2.

3.1 The Important Distinction Between Correlation and Causality

In the early nineteenth century, a cholera epidemic swept through Russia. The government, in an effort to stem the disease, sent doctors to the worst-affected areas. The peasants of a particular province observed a very high correlation between the number of doctors in a given area and the incidence of cholera in that area. Relying on this fact, they banded together and murdered their doctors.[1]

The fundamental problem in this example is that the peasants in this town clearly confused *correlation* with *causality*. They correctly observed that there was a positive association between physician presence and the incidence of illness. But they took that as evidence that the presence of physicians *caused* illness to be more prevalent. What they missed, of course, was that the link actually ran the other way: it was a higher incidence of illness that caused more physicians to be present. In statistics, this is called the *identification problem*: Given that sets of data on two variables are correlated, how do you identify whether one variable is causing the pattern seen in the other?

This is not just a problem from ancient times. In September 2014, a group of aid workers trying to assist communities in Guinea in dealing with an outbreak of Ebola were attacked, and eight were killed. The reason, according to a local leader: "Wherever those people have passed, the communities have been hit by illness."[2]

Foreign nations are not the only ones plagued by this problem. In 1988, a Harvard University dean conducted a series of interviews with Harvard freshmen and found that those who had taken SAT preparation courses scored, on average, 63 points lower (out of 1,600, the then-maximum number of points) than those who hadn't. The dean concluded that SAT preparation courses were unhelpful and that "the coaching industry is playing on parental anxiety."[3] This conclusion is another excellent

©Amy Hwang/CartoonStock.com

"First, we eat. Then, we'll quickly find data that correlates free donuts with increased productivity."

"First, we eat. Then, we'll quickly find data that correlates free donuts with increased productivity."

example of confusing correlation with causation. While SAT preparation courses are ubiquitous today, in 1998 they were much less common. So who was most likely to take SAT preparation courses? Those students who needed the most help with the exam! So all this study found was that students who needed the most help with the SAT did the worst on the exam. The courses did not cause students to do worse on the SATs; rather, students who would likely do worse on the SATs were the ones who took the courses.

Another example comes from the medical evaluation of the benefits of breastfeeding infants. Child-feeding recommendations typically include breastfeeding beyond 12 months, but some medical researchers have documented increased rates of malnutrition in breastfed toddlers. This has led them to conclude that breastfeeding for too long is nutritionally detrimental. But the misleading nature of this conclusion was illustrated by a study of toddlers in Peru that showed that it was those babies who were already underweight or malnourished who were breastfed the longest.[4] Increased breastfeeding did not lead to poor growth; children's poor growth and health led to increased breastfeeding.

The Problem

In all of these examples, the analysis suffered from a common problem: the attempt to interpret a correlation as a causal relationship without sufficient thought to the underlying process that generated the data. Noting that those who take SAT preparation courses do worse on SATs, or that those infants who breastfeed longest are the least healthy, is only the first stage in the research process—that of documenting the correlation. Once the data are available on any two measures, it is easy to see whether or not they move together, a characteristic we call being *correlated*.

What is harder to assess is whether the movements in one measure are *causing* the movements in the other. For any correlation between two variables A and B, there are three possible explanations, one or more of which could result in the correlation:

- A causes B.
- B causes A.
- Some third factor causes both.

Consider the previous SAT preparation example. For this sample of Harvard students, those who took an SAT prep course performed worse on their SATs. The

interpretation drawn by the Harvard administrator was one of only many possible interpretations, such as:

- SAT prep courses worsen preparation for SATs.
- Those who are of lower test-taking ability take preparation courses to try to catch up.
- Those who are generally nervous people like to take prep courses, and being nervous is associated with doing worse on standardized exams.

The Harvard administrator drew the first conclusion, but the others may be equally valid. Together, these three interpretations show that this correlation cannot be interpreted as a causal effect of test preparation on test scores without more information or additional assumptions.

Similarly, consider the breastfeeding interpretation. Once again, there are many possible interpretations, including:

- Longer breastfeeding is bad for children's health.
- Those infants who are in the worst health get breastfed the longest.
- The mothers with the lowest incomes breastfeed longer because this is the cheapest form of nutrition for children, and low income is associated with poor infant health.

Once again, all of these explanations are consistent with the observed correlation. But, once again, the studies that argued for the negative effect of breastfeeding on health *assumed* the first interpretation while ignoring the others.

The general problem that empirical economists face in trying to use existing data to assess the causal influence of one factor on another is that one cannot immediately go from correlation to causation. This is a problem for policy purposes because what matters most is causation. Policy makers typically want to use the results of empirical studies as a basis for predicting how government interventions will affect behaviors. Knowing that two factors are correlated provides no predictive power; prediction requires understanding the causal links between the factors. For example, the government shouldn't make policy based on the fact that breastfeeding infants are less healthy. Rather, it should assess the true causal effect of breastfeeding on infant health and use that as a basis for making government policy. The next section begins to explore the answer to one of the most important questions in empirical research: How can you draw causal conclusions about the relationships between correlated variables?

3.2 Measuring Causation with Data We'd Like to Have: Randomized Trials

One of the most important empirical issues facing society today is understanding how new medical treatments affect the patient health. An excellent example of this issue is the case of estrogen replacement therapy (ERT), a popular treatment for middle-aged and older women who have gone through menopause (the end of menstruation).[5] Menopause is associated with many negative side effects, such as rapid changes in body temperature ("hot flashes"), difficulty sleeping, and higher risk of urinary tract infection. ERT reduces those side effects by mimicking the estrogen produced by the woman's body before the onset of menopause.

Although ERT helped ameliorate the negative side effects of menopause, there was also a concern about ERT. Anecdotal evidence suggested that ERT might raise the risk of heart disease and, in turn, the risk of heart attacks or strokes. A series of studies beginning in the early 1980s investigated this issue by comparing women who did and did not receive ERT after menopause. These studies concluded that those who received ERT were at no higher risk of heart disease than those who did not; indeed, there was some suggestion that ERT actually lowered the risk of heart disease.

There was concern, however, that such a comparison did not truly reflect the causal impact of ERT on heart disease because women who underwent ERT were more likely to be under a doctor's care, to lead a healthier lifestyle, and to have higher incomes—all of which are associated with a lower chance of heart disease (an example of the third general explanation discussed earlier: some third factor is correlated with both ERT and heart disease). So it is possible that ERT might have raised the risk of heart disease but that this increase was masked because the women taking the drug were in better health otherwise.

Randomized Trials as a Solution

How can researchers address this problem? The best solution is through the gold standard of testing for causality: **randomized trials**. Randomized trials involve taking a group of volunteers and randomly assigning them to either a **treatment group**, the set of individuals who are subject to an intervention (such as a medical treatment or government benefit) being studied, or a **control group**, who are not given the intervention. Effectively, volunteers are assigned to treatment or control by the flip of a coin.

randomized trial
The ideal type of experiment designed to test causality, whereby a group of individuals is randomly divided into a treatment group, which receives the treatment of interest, and a control group, which does not.
treatment group
The set of individuals who are subject to an intervention being studied.
control group
The set of individuals comparable to the treatment group who are not subject to the intervention being studied.

To see why randomized trials solve our problem, consider what researchers would ideally do in this context: take one set of older women, clone them, and place the originals and the clones in parallel universes. Everything would be the same in these parallel universes except for the use of ERT. Then, researchers could observe the differences in the incidence of heart disease between these two groups of women. Because the women would be precisely the same, we would know by definition that any differences would be causal. Because both sets of women are the same except for the treatment, there would be only one possible reason why the set of women assigned ERT would have higher rates of heart disease.

Unfortunately, we live in the real world and not in some science-fiction story, so we can't do this parallel universe experiment. But, amazingly, we can approximate this alternative reality through the randomized trial. This is because of the definition of *randomization:* assignment to treatment groups and control groups is not determined by anything about the subjects but by the flip of a coin. As a result, the treatment group is identical to the control group in every facet but one: the treatment group gets the treatment (in this case, the ERT).

The Problem of Bias

We can rephrase all of the studies discussed so far in this chapter in the treatment/control framework. In the SAT example, students who took preparatory classes were the treatment group, and students who did not take the classes were the control group. In the breastfeeding example, the infants who breastfed for more than a year were the treatment group, and the infants who did not were the control group. In the ERT studies that occurred before randomized trials, those who received ERT were the treatment group, and those who did not were the control group. Even in the Russian doctor example, the areas where the doctors were sent were the treatment group, and the areas where the doctors were not sent were the control group. Virtually any empirical problem we discuss in this course can be thought of as a comparison between treatment and control groups.

We can, therefore, always start our analysis of an empirical methodology with a simple question: Do the treatment and control groups differ for any reason other than the treatment? All of the earlier examples involve cases in which the treatment groups differ in consistent ways from those in the control groups: those taking SAT prep courses may be of lower test-taking ability than those not taking the courses; those breastfed longest may be in worse health than those not breastfed as long; and those taking ERT may be in better health than those not taking ERT. These non–treatment-related differences between treatment and control groups are the fundamental problem in assigning causal interpretations to correlations.

We call these differences **bias**, a term that represents any source of difference between treatment and control groups that is correlated with the treatment but is *not due* to the treatment. The estimates of the impact of SAT prep courses on SAT scores, for example, are *biased* by the fact that those who take the prep courses are likely to do worse on the SATs for other reasons. Similarly, the estimates of the impact of breastfeeding past one year on health are biased by the fact that those infants in the worst health are the ones likely to be breastfed the longest. The estimates of the impact of ERT on heart disease are biased by the fact that those who take ERT are likely in better health than those who do not. Whenever treatment and control groups consistently differ in a manner that is correlated with, but not due to, the treatment, there can be bias.

bias
Any source of difference between treatment and control groups that is correlated with the treatment but is not due to the treatment.

Moreover, even the gold standard of randomized trials has some potential problems.[10] First, the results are only valid for the sample of individuals who volunteer to be either treatments or controls, and this sample may be different from the population at large. For example, those in a randomized trial sample may be less averse to risk, or they may be more desperately ill. Thus, the answer we obtain from a randomized trial, while correct for this sample, may not be valid for the average person in the population.

A second problem with randomized trials is that of **attrition**: individuals may leave the experiment before it is complete. This is not a problem if individuals leave randomly because the sample will remain random. Suppose, however, that the experiment had positive effects on half the treatment group and negative effects on the other half, and, as a result, the half with negative effects left the experiment before it was done. If we focused only on the remaining half, we would wrongly conclude that the treatment had overall positive impacts.

attrition
Reduction in the size of samples over time, which, if not random, can lead to biased estimates.

In the remainder of this chapter, we discuss several approaches taken by economists to try to assess causal relationships in empirical research. We will do so through the use of the TANF example. The general lesson from this discussion is that there is no way to consistently achieve the ideal of the randomized trial; bias is a pervasive problem that is not easily remedied. There are, however, methods available that can allow us to approach the gold standard of randomized trials.

3.3 Estimating Causation with Data We Actually Get: Observational Data

In the previous section, we showed how a randomized trial can be used to measure the impacts of an intervention such as ERT or lower TANF benefits on outcomes such as heart attacks or labor supply. As we highlighted, however, data from such randomized trials are not always available when important empirical questions need to be answered. Typically, the analyst has **observational data** instead, which are data generated from individual behavior observed in the real world. For example, instead of information on a randomized trial of a new medicine, we may simply have data on who took the medicine and what their outcomes were (the source of the original conclusions on ERT). Several well-developed methods can be used by analysts to address the problem of bias with observational data, and the use of these tools often closely approximates the gold standard of randomized trials.

observational data
Data generated by individual behavior observed in the real world, not in the context of deliberately designed experiments.

This section explores how researchers can use observational data to estimate causal effects instead of just correlations. We do so within the context of the TANF example. It is useful throughout to refer to the empirical framework established in the previous section: those with higher TANF benefits are the control group, those with lower TANF benefits are the treatment group, and our concern is to remove any sources of bias between the two groups (i.e., any differences between them that might affect their labor supply, other than TANF benefits differences). Thus, the major concern throughout this section is how to overcome any potential bias so that we can measure the causal relationship (if there is one) between TANF benefits and labor supply.

Time Series Analysis

One common approach to measuring causal effects with observational data is **time series analysis**, documenting the correlation between the variables of interest over time. In the context of TANF, for example, we can gather data over time on the benefit guarantee in each year and compare these data with the amount of labor supply delivered by eligible single parents in those same years.

time series analysis
Analysis of the comovement of two series over time.

Figure 3-1 shows such a time series analysis of data for the period 1975 through 2019. On the horizontal axis are the years. The left-hand vertical axis charts the average real monthly benefit guarantee for a family of four (controlled for inflation by expressing income in constant 2020 dollars) available in the United States over this period. Benefits declined dramatically from $1,099 in 1975 to $661 in 2019, nearly falling by half in real terms because benefit levels have not kept up with inflation. The right-hand vertical axis charts the average hours of work per year for single mothers (including zeros for those mothers who do not work). The hours worked have risen, from below 1,000 hours per year in 1975 to over 1,450 hours per year in 2019. Thus, there appears to be a negative relationship between benefit guarantees and labor supply: falling benefit guarantees are associated with higher levels of labor supply by single mothers.

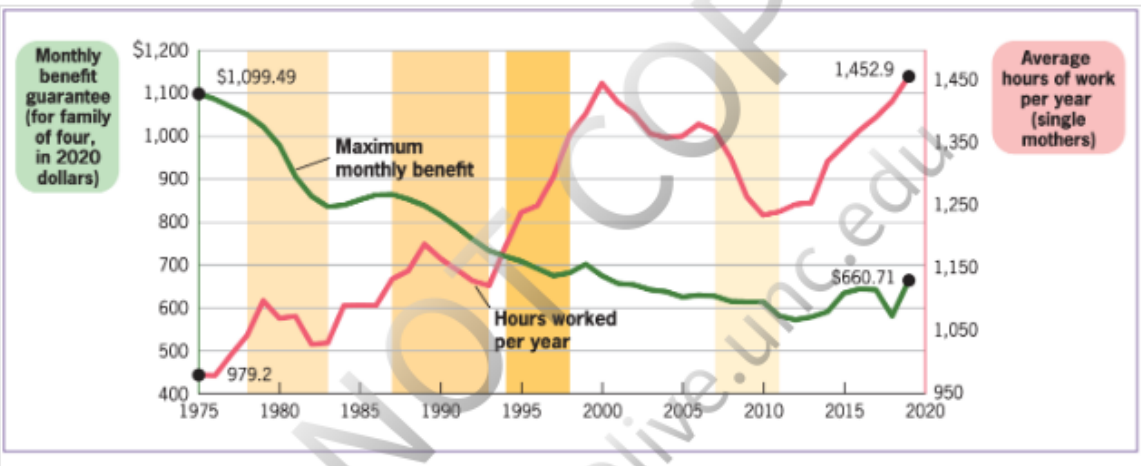

FIGURE 3-1 Average Benefit Guarantee and Single Mother Labor Supply, 1975–2019 • The left-hand vertical axis shows the monthly benefit guarantee under cash welfare, which falls from $1,099 in 1975 to $661 in 2019. The right-hand vertical axis shows average hours of work per year for single mothers, which rises from 979 in 1975 to 1,453 in 2019. Over this entire 45-year period, there is a strong negative correlation between the average benefit guarantee and the level of labor supply of single mothers, but there is not a very strong relationship within subperiods of this overall time span.

Data from: Calculations based on data from the Current Population Survey's annual March supplements (https://www.census.gov/programs-surveys/cps.html).

Problems with Time Series Analysis

Although this time series correlation is striking, it does not necessarily demonstrate a causal effect of TANF benefits on labor supply. When there is a slow-moving trend in one variable through time, as is true for the general decline in income guarantees over this period, it is very difficult to infer its causal effects on another variable. There could be many reasons why single mothers work more now than they did in 1975: greater acceptance of women in the workplace; better and more options for

child care; and even more social pressures on mothers to work. The simple fact that labor supply is higher today than it was 40 years ago does not prove that this increase has been caused by the steep decline in income guarantees.

This problem is highlighted by examining subperiods of this overall time span. From 1978 through 1983, the period of steepest benefits decline, benefits fell by over 20% in real terms (from $1,051 to $836 per month), yet hours of work fell by 1% (from 1,042 hours to 1,030 hours), whereas a causal effect of benefits would imply a rise in hours of work. From 1987 through 1993, benefits fell by about 15% (from $864 to $733 per month), yet labor supply first increased and then decreased, with a total decrease over this period of 1%. The subperiods, therefore, give a very different impression of the relationship between benefits and labor supply than does the overall time series.

A particularly instructive example about the limitations of time series analysis is the experience of the 1993–1998 period. In this subperiod, there is both a fall in benefits (falling by about 7%, from $733 to $681 per month) and a sharp rise in labor supply of single mothers (rising by about 21%, from 1,121 hours per year to 1,361 hours per year). The data from this subperiod seem to support the notion that lower benefits cause rising labor supply. Yet during this period, the economy was experiencing dramatic growth, with the general unemployment rate falling from 7.3% in January 1993 to 4.4% in December 1998. It was also a period that saw an enormous expansion in the Earned Income Tax Credit (EITC), a federal wage subsidy that has been shown to be effective in increasing the labor supply of single mothers. Contrast that period with the time from 2007 to 2011. Benefits were falling at roughly the same rate in both periods, yet from 1993 to 1998 (during an economic boom), hours of work rose whereas during 2007–2011 (worse economic times), hours of work fell.

Therefore, it is likely that other factors, not falling benefits, caused the increased labor supply of single mothers. So once again, other factors get in the way of a causal interpretation of this correlation over time; factors such as economic growth and a more generous EITC can cause bias in this time series analysis because they are also correlated with the outcome of interest.

When Is Time Series Analysis Useful?

All time series analysis is not necessarily useless, however. In some cases, there may be sharp breaks in the time series that are not related to third factors that can cause bias. A classic example is shown in Figure 3-2. This figure shows the price of a pack

of cigarettes (in constant 2020 dollars) on the left vertical axis and the *youth smoking rate*, the percentage of high school seniors who smoke at least once a month, on the right vertical axis. These data are shown for the time period from 1975 to 2018. From 1980 to 1992, there was a steady increase in the real price of cigarettes (from $1.97 to $3.33 per pack) and a steady decline in the youth smoking rate (from 30.5 to 27.8%). As previously noted, these changes over time need not be causally related. Smoking was falling for all groups over this time period due to an increased appreciation of the health risks of smoking, and prices may simply have been rising due to rising costs of tobacco production.

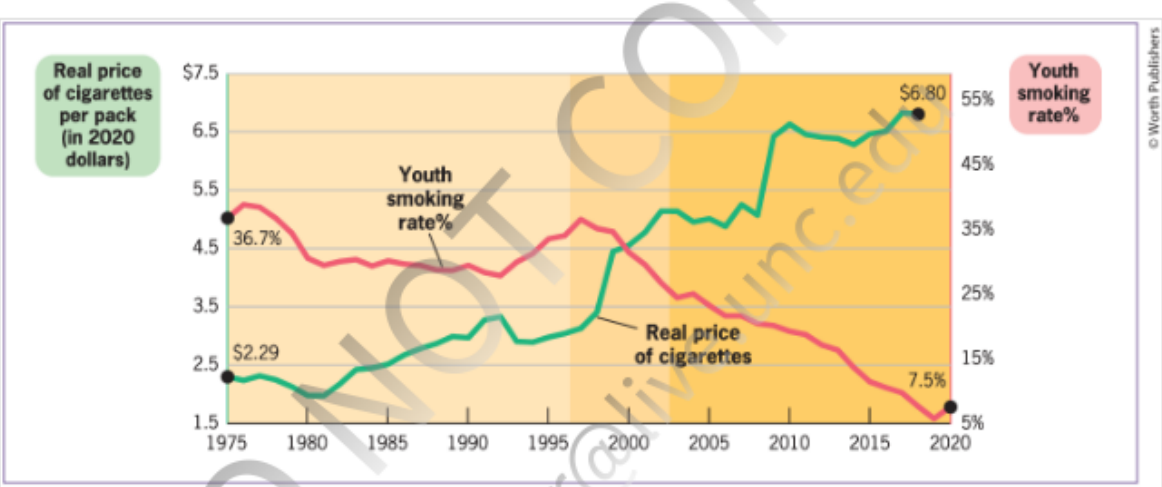

FIGURE 3-2 Real Cigarette Prices and Youth Smoking, 1975–2018 • The left-hand vertical axis shows the real price of cigarettes per pack, which rises from $2.29 in 1975 to $6.80 in 2018. The right-hand vertical axis shows the youth smoking rate (the share of high school seniors who smoke at least once a month), which fell from 1980 to 1992, rose sharply to 1997, and has fallen thereafter. There is a striking negative correspondence between price and youth smoking within subperiods of this era.

Data from: Calculations based on Monitoring the Future (2020).

Then, in April 1993, a "price war" in the tobacco industry, led to a sharp drop in real cigarette prices from $3.33 to $2.90 per pack.[11] At that exact time, youth smoking began to rise. This striking simultaneous reversal in both series is more compelling evidence of a causal relationship than is the long, slow-moving correlation over the 1980–1992 period. But it doesn't *prove* a causal relationship because other things were changing in 1993 as well. It was, for example, the beginning of an important period of economic growth, which could have led to more youth smoking. Moreover, the rise in youth smoking seems too large to be explained solely by the price decrease.

Fortunately, in this case, there is another abrupt change in this time series. In 1998 and thereafter, prices rose steeply when the tobacco industry settled a series of expensive lawsuits with many states (and some private parties) and passed the costs on to cigarette consumers. At that exact time, youth smoking began to fall again. This type of pattern seems to strongly suggest a causal effect, even given the limitations of time series data. That is, it seems unlikely that there is a factor correlated with youth smoking that moved up until 1992, then down until 1997, and then back up again, as did price. That youth smoking follows the opposite pattern as cigarette prices suggests that price is causing these movements. Thus, while time series correlations are not very useful when there are long-moving trends in the data, they are more useful when there are sharp breaks in trends over a relatively narrow period of time.

That said, patterns since 2014 reveal the same type of problem that plagued the previous time series example: outside factors that drive time series trends. From 2014 to 2018, prices rose modestly, from $6.28 to $6.80 per pack (8.7%). But youth smoking fell precipitously, from 13.6 to 7.6% (a 29% decline). This dramatic drop seems unlikely to reflect an enormous price elasticity, but instead the dramatic shift from traditional to e-cigarettes among youth. Between 2017 and 2018 alone, the share of 12th graders who smoked e-cigarettes jumped from 11 to 20.9%, and in 2019, nearly 26% of high school seniors had smoked e-cigarettes in the past month.[12] Thus, while some of the fall in youth smoking may reflect rising prices, it is unlikely to be the primary cause.

Cross-Sectional Regression Analysis

A second approach to identifying causal effects is **cross-sectional regression analysis**, a statistical method for assessing the relationship between two variables while holding other factors constant. *Cross-sectional* means comparing many individuals at one point in time, rather than comparing outcomes over time as in a time series analysis.

cross-sectional regression analysis
Statistical analysis of the relationship between two or more variables exhibited by many individuals at one point in time.

In its simplest form, called a *bivariate regression,* cross-sectional regression analysis is a means of formalizing correlation analysis—that is, of quantifying the extent to which two series are correlated. Returning to the example in Chapter 2, suppose that there are two types of single parents, with preferences over leisure and food

that there are two types of single parents, with preferences over leisure and food consumption represented by Figures 2-10 and 2-11. Before there is any change in TANF benefits, the parent who has a lower preference for leisure (Naomi in Figure 2-10) has both lower TANF benefits and a higher labor supply than the parent who has a greater preference for leisure (Sarah in Figure 2-11). If we took these two parents and correlated TANF benefits to labor supply, we would find that higher TANF benefits are associated with lower labor supply.

This correlation is illustrated graphically in Figure 3-3. We graph the two data points when the benefit guarantee is $5,000. One data point, point *A*, corresponds to Sarah from Figure 2-11 and represents labor supply of 0 hours and an income guarantee of $5,000. The other data point, point *B*, corresponds to Naomi in Figure 2-10 and represents a labor supply of 90 hours per year and TANF benefits of $4,550. The downward-sloping line makes clear the *negative correlation* between TANF benefits and labor supply; the recipient with lower TANF benefits has a higher labor supply.

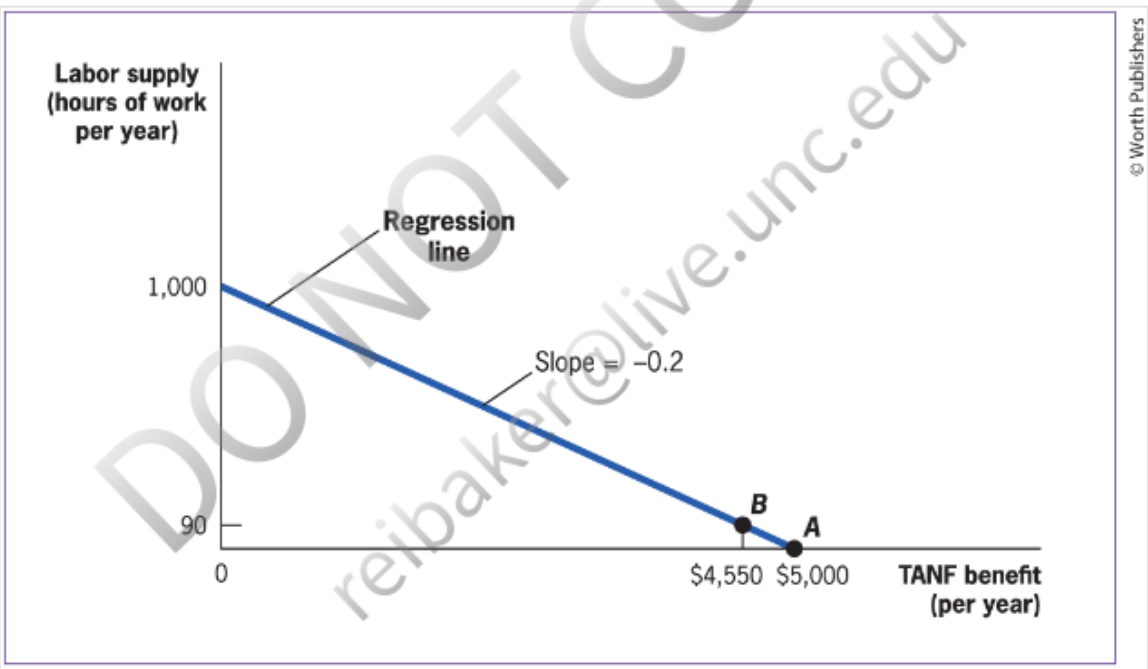

FIGURE 3-3 TANF Benefits and Labor Supply in Theoretical Example • If we plot the data from the theoretical example of Chapter 2, we find a modest negative relationship between TANF benefits and the labor supply of single mothers.

Regression analysis takes this correlation one step further by quantifying the relationship between TANF benefits and labor supply. Regression analysis does so by finding the line that best fits this relationship and then measuring the slope of that line.[13] This is illustrated in Figure 3-3. The line that connects these two points has a slope of −0.2. That is, this bivariate regression indicates that each $1

has a slope of −0.2. That is, this bivariate regression indicates that each $1 reduction in TANF benefits per month leads to a 0.2-hour-per-year increase in labor supply. Regression analysis describes the relationship between the variable that you would like to explain (the *dependent variable,* which is labor supply in this example) and the set of variables that you think might do the explaining (the *independent variables,* in this example, the TANF benefit).

Example with Real-World Data

The example in Figure 3-3 is made up, but we can replicate this exercise using real data from one of the most popular sources of cross-sectional data for those doing applied research in public finance: the Current Population Survey (CPS).

3.4 Conclusion

The central issue for any policy question is establishing a causal relationship between the policy in question and the outcome of interest. Do lower welfare benefits *cause* higher labor supply among single parents? Does more pollution in the air *cause* worse health outcomes? Do larger benefits for unemployment insurance *cause* individuals to stay unemployed longer? These are the types of questions that we address in this book using the empirical methods described here.

In this chapter, we discussed several approaches to distinguish causality from correlation. The gold standard for doing so is the randomized trial, which removes bias through randomly assigning treatment and control groups. Unfortunately, however, such trials are not available for every question we want to address in empirical public finance. As a result, we turn to alternative methods such as time series analysis, cross-sectional regression analysis, and quasi-experimental analysis. Each of these alternatives has weaknesses, but careful consideration of the problem at hand can often lead to a sensible solution to the bias problem that plagues empirical analysis.

3.4 Conclusion

The central issue for any policy question is establishing a causal relationship between the policy in question and the outcome of interest. Do lower welfare benefits *cause* higher labor supply among single parents? Does more pollution in the air *cause* worse health outcomes? Do larger benefits for unemployment insurance *cause* individuals to stay unemployed longer? These are the types of questions that we address in this book using the empirical methods described here.

In this chapter, we discussed several approaches to distinguish causality from correlation. The gold standard for doing so is the randomized trial, which removes bias through randomly assigning treatment and control groups. Unfortunately, however, such trials are not available for every question we want to address in empirical public finance. As a result, we turn to alternative methods such as time series analysis, cross-sectional regression analysis, and quasi-experimental analysis. Each of these alternatives has weaknesses, but careful consideration of the problem at hand can often lead to a sensible solution to the bias problem that plagues empirical analysis.

Highlights

- A primary goal of empirical work is to document the causal effects of one economic factor on another, for example, the causal effect of raising TANF benefits on the labor supply of recipients.
- The difficulty with this goal is that it requires treatment groups (those who are affected by policy) and control groups (those not affected) who are identical except for the policy intervention.
- If these groups are not identical, there can be bias—that is, other consistent differences across treatment/control groups that are correlated with, but not due to, the treatment itself.
- Randomized trials are the gold standard to surmount this problem. Treatments and controls are identical by definition, so, ideally, there is no bias, and any differences across the groups are a causal effect.
- Time series analysis is unlikely to provide a convincing estimate of causal effects because so many other factors change through time.
- Cross-sectional regression analysis also suffers from bias problems because similar people make different choices for reasons that can't be observed, leading once again to bias. Including control variables offers the potential to address this bias.
- Quasi-experimental methods have the potential to approximate randomized trials, but control groups must be selected carefully to avoid biased comparisons.

Questions and Problems

1. Suppose you are running a randomized experiment and you randomly assign study participants into control and treatment groups. After making the assignments, you study the characteristics of the two groups and find that the treatment group has a lower average age than the control group. How could this arise?
2. Why is a randomized trial the "gold standard" for solving the identification problem?
3. What do we mean when we say that correlation does not imply causality? What are some of the ways in which an empirical analyst attempts to disentangle the two?
4. A researcher conducted a cross-sectional analysis of children and found that the average test performance of children with divorced parents was

lower than the average test performance of children of intact families. This researcher then concluded that divorce is bad for children's test outcomes. What is wrong with this analysis?

5. A study in the *Annals of Improbable Research* once reported that counties with large numbers of mobile-home parks had higher rates of tornadoes than did the rest of the population. The authors conclude that mobile home parks cause tornado occurrences. What is an alternative explanation for this fact?
6. What are some of the concerns with conducting randomized trials? How can quasi-experiments potentially help here?
7. You are hired by the government to evaluate the impact of a policy change that affects one group of individuals but not another. Suppose that before the policy change, members of a group affected by the policy averaged $17,000 in earnings, and members of a group unaffected by the policy averaged $16,400. After the policy change, members of the affected group averaged $18,200 in earnings, while members of the unaffected group averaged $17,700 in earnings.
   a. How can you estimate the impact of the policy change? What is the name for this type of estimation?
   b. What are the assumptions you have to make for this to be a valid estimate of the impact of the policy change?
8. Consider the example presented in the appendix to this chapter. Which coefficient estimates would be considered "statistically significant" or distinct from zero?
9. A researcher wants to investigate the effects of education spending on housing prices, but they only have cross-sectional data. When performing regression analysis, they control for average January and July temperatures. Why are they doing this? What other variables would you control for, and why?
10. It is commonly taught in introductory microeconomics courses that minimum wages cause unemployment. In 2020, the federally mandated minimum wage was $7.25, but more than half the states had higher state-mandated minimum wages. Why can't you test the "minimum wages cause unemployment" theory by simply comparing unemployment rates across states with different minimum wages? Can you think of a better way to test it?

ADVANCED QUESTIONS

11. Suppose that your friend Oscar has collected data and determined that towns with newly constructed high schools tend to have higher SAT scores than other towns. He tells you that he has proved that new high schools cause higher SAT scores. When you object that "correlation does not imply causation," he is ready with more data. He shows you convincing evidence that SAT scores tend to increase shortly after towns build new high schools, but that there is no tendency for new high schools to be built in towns that have recently seen large increases in SAT scores. Is this enough evidence to prove that new high schools cause higher SAT scores, or can you think of an alternative explanation for Oscar's data?
12. Researchers often use *panel data* (multiple observations over time of the same people) to conduct regression analysis. With these data, researchers are able to compare the same person over time in order to assess the impacts of policies on individual behavior. How could this provide an improvement over cross-sectional regression analysis of the type described in the text?
13. Suppose that your state announced that it would provide free tuition to high-achieving students graduating from high school starting in 2022. You decide to see whether this new program induces families with high-achieving children graduating in 2022 or later to purchase new cars. To test your findings, you use a "falsification exercise": you observe the new-car-purchasing behavior of families with children graduating in 2021. Why is this a useful exercise?
14. Your state introduced a tax cut in the year 2021. You are interested in seeing whether this tax cut has led to increases in personal consumption within the state. You observe the following information:

Year	Consumption in Your State
2016	330
2018	350
2020	370
2022	430

a.

APPENDIX TO CHAPTER 3: Cross-Sectional Regression Analysis

In the text, we presented a cursory discussion of cross-sectional regression analysis and the role of control variables. In this appendix, we provide a more detailed presentation of this approach within our TANF example.

Data

For this analysis, we use data from the March 2019 Current Population Survey (CPS). From that survey, we selected all women who reported that they were unmarried and had a child younger than age 19. The total sample is 6,779 single mothers.

For this sample, we have gathered data on the following variables for each woman:

- *TANF:* Total income from welfare in the previous year (in thousands of dollars).
- *Hours:* Total hours of work in the previous year, computed as reported weeks of work times usual hours per week.
- *Race:* We divide reported race into White, Black, and other.
- *Age:* Age in years.
- *Education:* We use reported education to divide individuals into four groups: high school dropouts; high school graduates with no college; those with some college; and college graduates.
- *Urbanicity:* We use information on residential location to divide individuals into four groups: central city; other urban; rural; and unclear (the CPS doesn't identify location for some mothers for survey confidentiality reasons).

Regression

Using these data, we can estimate a regression of the impact of welfare on hours of work of the form:

$$(1)\ \text{HOURS}_i = \alpha + \beta \text{TANF}_i + \epsilon_i$$

where there is one observation for each mother i. This is the counterpart of the regression analysis shown in Figure 3-4, but now we are using each individual data point, rather than grouping the data into categories for convenience.

In this regression, α, the constant term, represents the estimated number of hours worked if welfare benefits are zero. β is the slope coefficient, which represents the change in hours worked per dollar of welfare benefits. ε is the error term, which represents the difference for each observation between its actual value and its predicted value based on the model.

The results of estimating this regression model are presented in the first column of the appendix table. The first row shows the constant term a, which is 1,437: this measures the predicted hours of labor supply delivered at zero welfare benefits. The second row shows the coefficient b, which is −163: each $1,000 of welfare benefits lowers hours worked by 163. This is slightly higher than the estimate from the grouped data of –157 discussed in the text. Thus, for a mother with no welfare benefits, predicted hours of work are 1,437; for a mother with $5,000 in welfare benefits, predicted hours of work are $1{,}437 - 5 \times 163 = 622$.

Underneath this estimate in parentheses is the estimate's *standard error*. This figure captures the precision with which these coefficients are estimated and reminds us that we have here only a statistical representation of the relationship between welfare benefits and hours worked. Roughly speaking, we cannot statistically distinguish values of b that are two standard errors below or above the estimated coefficient. In our context, with a standard error of 11.3 hours, the results show that our best estimate is that each thousand dollars of welfare lowers hours worked by 163, but we can't rule out that the effect is only $140.4(163 - 2 \times 11.3)$ or that it is $185.6(163 - 2 \times 11.3)$.

In the context of empirical economics, this is a *very* precise estimate. Typically, as long as the estimate is more than twice the size of its standard error, we say that it is *statistically significant*.

APPENDIX 3 TABLE Cross-Sectional Regression Analysis

	Equation (1)	Equation (2)

Constant	1,437	1,396
	(11)	(64)
TANF benefits	−163	−147
	(11)	(11)
White		30
		(39)
Black		29
		(44)
No high school diploma		−692
		(36)
High school graduate		−389
		(27)
Some college		−234
		(30)
Age		6
		(1)
Central city		89
		(34)
Other urban		79
		(32)
Rural		−39
		(37)
R^2	0.030	0.101

The final row of the table shows the R^2 of the regression. This is a measure of how well the statistical regression model is fitting the underlying data. An R^2 of 1 would mean that the data are perfectly explained by the model so that all data points lie directly on the regression line; an R^2 of 0 means that the data are not at all explained. The value of 0.054 here says that less than 5% of the variation in the data is explained by this regression model.

As discussed in the text, however, this regression model suffers from serious bias problems since those mothers who have a high preference for leisure will have both low hours of work and high welfare payments. One approach to addressing this problem suggested in the text was to include control variables. We don't have the ideal control variable, which is preference for leisure. We do, however, have other variables that might be correlated with preference for leisure or other factors that determine labor supply: race, education, age, and urbanicity. So we can estimate regression models of the form:

$$(2)\ \text{HOURS}_i = \alpha + \beta \text{TANF}_i + \delta \text{CONTROL}_i + \epsilon_i$$

where CONTROL is the set of control variables for individual i.

In the second column of the appendix table, we show the impact of including these other variables. When we have a categorical variable such as race (categorized into White, Black, and other), we include *indicator variables* that take on a value of 1 if the individual is of that race, and 0 otherwise. Note that when we have N categories for any variable (e.g., 3 categories for race), we only include $N - 1$ indicator variables, so that all estimates are relative to the excluded category (e.g., the coefficient on the indicator for "Black" shows the impact of being Black on welfare income, relative to the omitted group of Hispanics).

Adding these control variables does indeed lower the estimated impact of TANF benefits on labor supply. The coefficient falls to –147 but remains highly significant; interestingly, once we add controls, we get an estimate that is almost identical to the estimate from using the grouped data in Figure 3-4. The R^2 more than triples but still indicates that we are explaining less than 15% of the variation in the data.

The control variables are themselves also of interest:

- *Race:* White workers are estimated to work 30 hours per year more than all other workers (the omitted group), but this estimate is smaller than its standard error, so we do not call this a statistically significant difference. Black workers are estimated to work 21 hours per year more than the omitted group.
- *Education:* Hours of work clearly rise with education. People without high school diplomas work 692 fewer hours per year than do college graduates (the omitted group); high school graduates work 389 fewer hours per year; and those with some college work 234 fewer hours per year than those who

graduate from college. All of these estimates are very precise (the coefficients are very large relative to the standard errors beneath them in parentheses).

- *Age:* Hours worked decline with age, with each year of age leading to 6 fewer hours of work; this is a very precise estimate as well.
- *Location:* Relative to those with unidentified urbanicity, people in cities and suburbs work more, by a statistically significant amount. People in rural areas, on the other hand, work less than those with unidentified urbanicity, but these results are not precise.

Do these control variables eliminate bias in the estimated relationship between TANF benefits and labor supply? There is no way to know for sure, but it seems unlikely. The fact that this large set of controls explains only 7 percentage points more of the variation in labor supply across individuals suggests that it is unlikely to capture all of the factors correlated with both labor supply and TANF benefits.