Economics for the Greater Good
Economics for the Greater Good

An Introduction to Economic Thinking for Public Policy

Caroline Krafft
## CONTENTS

<table>
<thead>
<tr>
<th>Acknowledgements</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Part I. Main Body</strong></td>
<td></td>
</tr>
<tr>
<td>1. Production: Can we end hunger?</td>
<td>1</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>2. Supply and Demand: Who gets food, housing, and work?</td>
<td>23</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>3. Trade: An Increasingly Connected World</td>
<td>47</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>4. The Economics of Crime</td>
<td>64</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>5. The Economics of Poverty</td>
<td>81</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>6. The Economics of Discrimination</td>
<td>105</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>7. Who Learns What and Why: The Economics of Education</td>
<td>125</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>8. Why is there pollution and what can we do about it?</td>
<td>155</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
<tr>
<td>9. The Economics of Social Safety Nets: Here to Catch Us When We Fall?</td>
<td>171</td>
</tr>
<tr>
<td>Caroline Krafft</td>
<td></td>
</tr>
</tbody>
</table>
I am grateful to all the wonderful students who tested out draft chapters and gave so much helpful feedback. Support from APDC at St. Catherine University got this book started. My department, including Nasrin Jewell, Deep Shikha, and Kristine West (credit for the title) encouraged me to continue.
PRODUCTION: CAN WE END HUNGER?

Caroline Krafft

Can we end hunger?

In Yemen, 17 million people are food insecure, which means they do not have enough to eat.\(^1\) There are 3.5 million children and women in Yemen who are so malnourished they require urgent treatment. Two-thirds of children under the age of five in Yemen are so chronically malnourished that their growth has been stunted.\(^2\) This rate is an increase in malnutrition from previous decades, when around half of children were stunted.\(^3\) In the long term, these malnourished children will learn less (if they attend school), earn less when they work, and have poorer health throughout their lives.\(^4\) Conflict in Yemen has moved the country closer to famine.\(^5\)

Food security exists when all people, at all times, have physical and economic access to sufficient safe and nutritious food that meets their dietary needs and food preferences for an active and health lifestyle.

–1996 World Food Summit

In the United States, there are 14 million food insecure households (10 percent of households).\(^6\) There are 5 million children in the United States who are in households where food insecurity is affecting children.\(^7\) Despite programs designed to prevent hunger, such as food stamps, children and families are still suffering

1. As of July 2023. World Food Programme 2023
2. World Food Programme 2017
3. El-Kogali and Krafft 2015
4. Dewey and Begum 2011
5. Moyer et al. 2019
6. As of 2021. Coleman-Jensen et al. 2022
7. As of 2021. Coleman-Jensen et al. 2022
from hunger in the United States. These children perform worse in school near the end of the month—when their food stamps have run out.  

Although the United States is a much wealthier nation than Yemen, and hunger is a less common issue, both countries still fail to meet the most basic needs of all of their citizens—starting with food. This chapter introduces the economics of production—how countries create and distribute resources—to help us think about what it would take to end hunger.

How can economics help us end hunger?

Economics is the study of how societies allocate their scarce resources. In both Yemen and the United States—and around the world—understanding the allocation of scarce resources is critical to addressing hunger. Food itself is a scarce resource when families go hungry. Producing food depends on a number of different inputs, such as land, water, and seeds, which are also scarce. Economics can help us think about how to allocate food to people so we can reduce or eliminate hunger. One of the ways economics can help end hunger is by helping us allocate not just food, but the inputs that produce food.

Economists think about allocation decisions on two levels. The first is *microeconomics*, thinking about the behavior of individuals or firms. When thinking about hunger, a microeconomic focus might be on how farmers choose between different crops or decide how much water to allocate to each crop. The second type of economics is *macroeconomics*, the study of the economy of a nation (or a region or the world) as a whole. Macroeconomics would shift the focus to the output of an entire nation or state. This text covers both macroeconomic and microeconomic topics.

How do we produce food?

Whether we are thinking about a single farm (microeconomics) or a nation (macroeconomics), how much food (output) it can produce depends on inputs (seeds, land, farmers). Economists divide inputs into two types—labor and capital. Labor consists of the people (human inputs) that can potentially produce goods and services. When we are thinking about a family farm, labor would be the number of hours family members could work on the farm. Since there are only a fixed number of hours in the day, labor is a scarce resource. When we’re thinking about the entire economy, labor might be measured in millions of workers.

Economists typically measure labor in hours or number of workers. However, this assumes workers are all able to produce the same amount of food, which is not accurate. How much workers can produce depends on a variety of factors—such as their skills or their health. Economists call these aspects of labor human capital.

---

8. Cotti, Gordanier, and Ozturk 2018
Human capital is the information, skills, knowledge, education, and health of workers. It shapes their ability to translate their time into production. For example, an experiment in Guatemala demonstrated that children who received a nutrition supplement during childhood were more productive as adults, earning higher wages, because of the investment in their health—their human capital. Likewise, when students go to college, they expect to be more productive and earn more as adults, because they have invested in their human capital. One way to think about human capital is that labor tends to refer to the quantity of labor, while human capital is one way for economists to think about the quality of labor.

In order to produce food, farms need not only labor and human capital, but also physical capital. Physical capital consists of the physical inputs, such as tractors and land, that help produce outputs. Physical capital also includes infrastructure—roads, electricity, water, and other public works that facilitate production. Natural resources (such as minerals or oil) also are a form of physical capital. Land, water, seeds, and fertilizer are examples of important physical capital inputs that contribute to the production of food.

Technology is the final element of production. Technology refers to the inventions or innovations, as well as new ways of organizing work, that allow us to produce more even with the same amount of inputs. For example, in West and Central Africa, the International Institute of Tropical Agriculture and the International Maize and Wheat Improvement Center work on improving the varieties of maize (a key food crop) available. Essentially, they invent new varieties of maize with enhanced capabilities, such as drought resistance. The emerging technology of new maize varieties allows farmers to produce more food with the same amount of land and labor.

It can be confusing to try to distinguish between technology and capital. Is a new irrigation system technology or is it capital? The technology (idea) of drip irrigation to water crops has been around for decades. However, as Box 1.1 demonstrates, not everyone adopts, or utilizes, new technology. Only when technology is adopted can it be classified as physical capital. Some farmers do not invest in the capital of drip irrigation. The farmers who do invest in drip irrigation have it as part of their capital.

---

**Box 1.1. Economists in action: Tavneet Suri on agricultural technology adoption**

Tavneet Suri is a professor at MIT and an economist who works in development economics (how and why

---

9. Hoddinott et al. 2008
10. Alene et al. 2009
11. Shrestha and Gopalakrishnan 1993
12. Suri 2011
countries grow and develop) and agricultural economics (the economics of growing food). Born and raised in Kenya, she researches technology adoption in sub-Saharan Africa. Agricultural yields (production per acre) have been stagnant or falling in many sub-Saharan African countries, contributing to food insecurity. This decrease in yields is occurring despite the existence of technology—such as hybrid crops and fertilizers—that can increase yields. Adoption of these technologies has been persistently low in many countries. Suri studied farmers’ technology adoption decisions in Kenya and identified problems in access to seed and fertilizer as important barriers to adopting technology that would improve yields for farmers.

The combination of inputs is also important. A simple example of this would be that, on a farm operated by tractors, you would need a driver (labor) for each tractor (physical capital). Any tractor that sits idle due to a lack of driver will not add anything to production. Thus, labor, human capital, and physical capital work best together. Some forms of physical capital, such as hybrid seeds, are more likely to be adopted and can be used more effectively by educated individuals.¹³

**How can we produce more food?**

Now that we have thought through how food is produced, we can turn to a central question for reducing hunger: how can we produce more food? A straightforward answer would seem to be more inputs—and more technology. However, where do these inputs and technology come from? We have to invest in them in the long run (over a long period of time). In the short run (for example, this year), our inputs and technology are fixed. If we want to produce more food, we are going to have to use our inputs differently. This section examines how to produce more food, first in the short run, and then in the long run.

**Producing more food in the short run**

The scarcity of resources is a key economic insight. If we want more workers on the farm, we will have to accept fewer workers in schools or factories. This is the central economic concept of tradeoffs. The saying “there is no free lunch” illustrates this idea. You have to give up something to get something. Economists often use models to think about these tradeoffs. Models are simplifications of the real world that help us think about economic issues—like the tradeoffs involved in producing more food.

The specific model for thinking about how a household or a country might produce more food is called the production possibilities frontier. A production possibilities frontier shows what can currently be produced,
given current resources (physical capital, labor, human capital, and technology). With this model, we simplify to a world where there are two goods. In real life, there are millions of different possible goods and services we can produce, but focusing on just two helps us think through critical tradeoffs. The insights gleaned from this simple model can then be applied to any number of goods.

Figure 1.1 is a production possibilities frontier for food and housing. Here we are thinking on country scale, in terms of millions of tons of food, and millions of housing units (houses). Food is measured on one axis (the x-axis) and housing is on the other axis (the y-axis). Let’s say that this country is Yemen, where there is widespread food insecurity and hunger. Yemen faces a tough choice and important tradeoffs, symbolized by the production possibilities frontier. Yemen (currently, in the short run) has a fixed amount of resources (physical capital, labor, human capital, and technology). The curve—called the frontier—shows the different combinations of goods (food and housing) that Yemen can produce given its current resources. The frontier is the maximum possible. Any point on the frontier represents a particular combination of food and housing. To produce more of one good, Yemen has to give up some of the other good. Point A is one extreme: all food, 30 million tons. However, this means that Yemen builds no housing. Point B is the opposite extreme: 15 million housing units and no food. The remaining points are a mix of goods. Point C is more food (26 million tons) and less housing (7.5 million units). Point D is more housing (14.6 million units) and less food (7.3 million tons). Point E is a mix in the middle, 21.5 million tons of food and 10.4 million units of housing.
Figure 1.1. Production possibilities frontier for food and housing

Point E illustrates an important feature: the production possibilities frontier is curved. We can have more of a mix of goods (31.9 million goods total) than just one alone (15 million housing units or 30 million tons of food). This is because inputs tend to be specialized. **Specialization** means that an input is better suited for producing one output, such as food, than another, such as housing. For labor, this would mean that a carpenter is much better at producing housing than food. If we want to produce all food, that means our carpenter will be farming—and she will not produce very much. Likewise, physical inputs are specialized. Yemen’s resources of seeds will not produce much housing. Some inputs may be quite useful for both goals—land would be useful for both housing and producing food, although the quality of the land might mean some land, for instance with good soil, is better for food, while other land, for instance with sandy soil, is better for housing.

Point F, which is not on the curve, (5 million of each good) is inefficient; we could increase both food and housing at the same time. We are not using all of our resources at point F; that is why it is inefficient. Point G,
19 million housing units and 35 million tons of food, is currently impossible (not feasible). With our current resources we cannot produce so much housing and food.

So how does Yemen decide what to produce, given its current resources? It depends on tradeoffs and people’s preferences. Perhaps it is currently at point E, producing a mix of 21.6 million tons of food and 10.4 million housing units. If Yemen wants to move to point C, which has more food (26.0 million tons), to gain that additional 4.4 million tons of food and reduce hunger, it will have to give up 3.1 million housing units (go from 10.4 million to 7.3 million). This tradeoff is referred to in economics as the opportunity cost. It is not a dollar cost Yemen has to pay, but the opportunities (for housing) that Yemen has to give up to get more food. The opportunity cost, at point E, for 4.4 million tons of food is 3.1 million housing units.

In this example, we’ve been thinking on a large, national scale, and in terms of changes in the millions. A similar principle applies to busy college students facing time constraints during finals. One more hour spent studying for economics means one less hour writing your political science paper. You have a finite amount of resources, including time. The opportunity cost of gaining ten more points on your economics final may be one less page in your political science paper. These decisions about one page or one hour show another aspect common to economics thinking: thinking at the margin. Although it can be helpful to think about what it would take to get to 4.4 million tons more food, most decisions are made like those of the student during finals, thinking about tradeoffs in terms of one more hour. This “one more” thinking is referred to in economics as thinking on the margin. Thinking on the margin refers to thinking about what one additional hour of paper-writing costs, or one additional ton of food takes.

**Producing more food in the long run**

The production possibilities frontier is a model for what we can produce right now, in the short run, given our current resources and technology. We can also use this model to think about the long run. In the long run, our whole production possibilities frontier can shift. In the long run, producers (countries, firms, or family farms) can change their inputs and their technology. When these changes in inputs and technology lead to more production being possible, that means the economy is growing. When these changes in inputs and technology lead to less production being possible, that means the economy is shrinking.

Figure 1.2 shows the economy growing over time. Point J and its frontier were the maximum production options possible in 2019. However, by 2029 point K and its frontier are the most possible, because of better technology and inputs. While point J was 21.6 million tons of food and 10.4 million units of housing, point K is 25.9 million tons of food and 12.5 million units of housing—more of both. Point K was not feasible (possible) in 2019; in 2029 it is now feasible. The shift from the production possibilities frontier with point J to the production possibilities frontier with point K represents economic growth. The key difference between the short run and the long run is that, while in the short run a country can only move along the production possibilities frontier, in the long run it can change the production possibilities frontier, shifting it outward (growing the economy) or inward (shrinking the economy).
In the long-run, growth may occur more in some outputs than others. For example, a house requires a relatively fixed amount of land, so there may not be much growth in the production possibilities for housing over time if there is a limited amount of land inputs in the country. However, improvements in agricultural technology have allowed farmers to have increased yields (output of food) from the same amount of land. When there’s growth in the long-run in one product (but not so much in the other) this is referred to as **specific growth**, as opposed to **general growth**. Figure 1.2 showed general growth, increases in the PPF for both food and housing. Figure 1.3, in contrast, shows a case of food-specific growth; more food is possible than used to be the case. Even though the maximum amount of housing is the same, it may be the case that more housing is produced nonetheless, because some resources (such as land) are shifted from food to housing as agriculture becomes more productive. This is visible in the shift from point L (24.3 million tons of food and 8.8 million
housing units) in 2019 to point M (28.5 million tons of food and 10.2 million housing units) in 2029. Both food and housing increased, food relatively more so as a result of food-specific growth.

![Production possibilities frontier with food-specific growth](image)

**Figure 1.3.** Production possibilities frontier with food-specific growth

What can we do to stimulate growth and produce more food? Policies to promote growth depend on the same factors that determine production in the short run: quantity of labor, quality of labor (human capital), physical capital, and technology. While these were fixed in the short run, in the long run they can be changed. When these inputs grow in number or quality, they lead to economic growth. When these inputs decrease in number or quality (for example, a town’s production shrinking because it has fewer and fewer people), the economy shrinks.

In terms of the quantity of labor, countries facing shortages of labor can either encourage immigration or try to encourage increases in the birth rate. An example of this would be generous parental leave or child
Improving human capital requires upgrading the health, skills, or knowledge of the workforce. There are a variety of different ways to increase skills; increasing education is one that we will examine in more detail in a later chapter. Encouraging apprenticeships, which lead to job-relevant skills, can lead to increases in human capital that pay off economically.\(^\text{15}\)

Improving physical capital can be challenging. Physical capital, whether a sewing machine, new factory, or new road, tends to be a large and expensive investment, requiring savings. We further explore savings in later chapters. Infrastructure investments, such as roads, are one example of an investment in physical capital that can increase production. A study in Nepal demonstrated that expanding roads in rural areas would substantially improve agricultural production, particularly benefiting the rural poor.\(^\text{16}\) Another challenge with infrastructure investments is that they are typically publicly available and used by a large number of people, making them difficult to finance, an issue we return to in later chapters.

Technology has played a particularly critical role in improving well-being over time, including reducing hunger. For example, the “green revolution” in Asia, driven by the introduction of new varieties of wheat and rice in the 1960s and 1970s (new technology), led to large increases in agricultural production which raised farmer’s incomes and well-being substantially.\(^\text{17}\) Other examples of technologies that substantially improved productivity and thus human well-being include the discovery of antibiotics, the invention of the computer, the industrial revolution, and the creation of assembly lines.

Technology is important—but how can we increase our level of technology and thus production? Research and development (R & D) is typically required for intentionally improving technology. The future of technology is very uncertain and difficult to predict, which makes R & D investment challenging and time-consuming. For example, University of Minnesota fruit breeders started work on a new apple variety for Minnesota in the 1990s that only became commercially available in 2018.\(^\text{18}\) A further challenge is that once something is invented, it is typically available to everyone, unless individuals or firms are given exclusive legal rights (such as patents). The broad availability of technology reduces the motivation of individuals or firms to invest in R & D.

Even when seemingly great technology ideas are developed, they may not work well in real life. Box 1 illustrated the challenge of agricultural technology adoption—using new technologies once they are invented. Adoption also depends on the quality of the new technology. In India, cvx rely on traditional stoves that burn firewood, dung, or agricultural waste. In laboratory conditions, improved cookstoves were shown to reduce

\(^\text{14}\) Dion et al. 2015; Rønsen 2004  
\(^\text{15}\) Krafft 2018  
\(^\text{16}\) Jacoby 2000  
\(^\text{17}\) World Bank 2007  
\(^\text{18}\) University of Minnesota 2018
indoor air pollution—which is the leading cause of environmental deaths globally. This technology can particularly benefit women, who are disproportionately responsible for cooking. In development economics, there was a lot of optimism about how this technology could contribute to well-being. In real world conditions, however, the stoves broke, were difficult to maintain, and were costlier to households. Households stopped using them and there were no health improvements. This is just one illustration of the challenges of getting seemingly promising technology ideas to actually work in the real world in order to enhance wellbeing.

Returning to our hybrid seeds example, one of the reasons that farmers may not adopt hybrid seeds is that the real-world quality of such inputs is low. For example, in Uganda, less than half of hybrid maize is authentic (more than half if fake), and thus, generates lower yields. Developing technologies to increase growth is not an easy process.

**Who gets the food?**

The PPF is a very useful model for thinking about how to increase production. It can be applied to questions such as increasing food or housing, education or health care, and manufacturing or services. So far, we’ve been primarily focusing on production at a national level—but also noting examples of how the same model might apply to an individual, family farm, or corporation. Since the PPF is a model that simplifies our economy, one issue missing from the model is the allocation of production—who gets the food. The world actually produces enough food to feed everyone and eradicate hunger. The problem of hunger is a problem not just of production, but also of distribution, who gets the food. Why do certain families or countries experience hunger and others do not? The challenge is that food is produced locally—different people own different amounts of inputs, such as land and labor, and therefore different amounts of resulting food. There are inequalities not only between nations in terms of allocation—a family in Yemen is more likely to be hungry than a family in the United States—but also within countries and even within households. For example, research in Burkina Faso has shown that, even within the same household, women’s farm plots receive fewer inputs and produce less, contributing to inequality within households. In this section, we discuss global economic systems that determine the allocation of resources and ultimately production.

20. Bold et al. 2017
21. Gilbert et al. 2010
22. Udry 1996
Economic systems

Economic systems are determined on a national level and can take a variety of different forms. The two endpoints for the spectrum of economic systems are a command economy or a market economy. When individuals (or firms) own resources and their output, a country is operating under a market economy, also known as a capitalist economy. A market economy gives individuals (and firms) private ownership of resources, referred to as private property rights. Decisions about what to produce are determined by individuals and firms—decision making is decentralized and coordinated through markets (a concept we will return to in the next chapter). Governments have no role in this system, or only enforce private property rights. A command economy involves collective, government ownership of resources, and is also known as a communist or planned economy. Resource and production decisions are planned and made by the command of the central government.

Note that market economies and command economies are the endpoints of a spectrum of potential economic systems. No country is a pure command economy or a pure market economy. Countries mix elements of these different systems, to varying degrees. Their prevalence globally has also changed substantially over time. For example, in 1917, what would become the Soviet Union was formed as a result of the Russian revolution and represented a shift from a market to planned economy. In 1991, the Soviet Union dissolved, and began a shift towards a more market-based system. China likewise shifted to a command economy in the mid-20th century but has, since 1978, undertaken a series of market-oriented reforms. The countries, including the Soviet Union and China, that were the closest to command economies were intensively planned. For instance, countries had five-year plans for what should be produced, and how the benefits of that production would be shared.

There are substantial tradeoffs in countries’ decisions around economic systems. A central tension in economics, and particularly economic systems, is the tradeoff that often occurs between equity and efficiency. Market economies tend to be more efficient and lead to more growth (there are a number of exceptions that we will discuss in subsequent chapters), while command economies tend to be more equitable. China’s “Great Leap Forward” reforms starting in 1958 illustrate this tension. China shifted to a much more equitable but less efficient economic model, organizing rural communities into large-scale communes. The resulting loss in efficiency in production lead to the deaths of approximately 23 million people. However, inequality may also reduce growth, as illustrated in Box 1.2. The relationships and tradeoffs between equity and efficiency are complex and hotly debated in economics.

24.Yining 2012
25.Peng 1987
Nancy Birdsall is an economist who has worked at international aid organizations, including the Inter-American Development Bank, the World Bank, and the Carnegie Endowment for International Peace. She was the founding president of the Center for Global Development, which “work[s] to change the policies and practices of rich countries and powerful institutions to reduce global poverty and inequality.”

One important area of Nancy Birdsall’s research has been on the relationship between inequality and growth. Economics typically emphasizes a tradeoff between growth and equity. Birdsall demonstrated that policies that reduce poverty and inequality, such as education, can also increase growth. She demonstrated that low inequality in East Asia contributed to the region’s rapid growth from 1960-1985.

The main reason market economies can be more efficient and lead to more growth than command economies is that the private property rights of market economies create important incentives (motivations). When individuals own resources and the output of those resources, it creates an incentive to be more efficient or work harder. Additional effort and investment on the family farm leads to more food for the family under a market economy. In a command economy, investment and effort by the family will not change how much food they receive. An example of the consequences of poor incentives is that, in China, when family farms had less ownership of land, they were less likely to invest in their land, lowering production.\(^{27}\)

Although there are often tradeoffs in terms of equity and efficiency in economic systems, economics can also help us identify ways to enhance both equity and efficiency (or growth). For example, investing in early childhood education (a form of human capital) has been identified as one of the world’s best investments by economists.\(^{28}\) It has also been identified by economists as the most effective education policy for reducing inequality.\(^{29}\) The importance of early childhood education for both efficiency and equity is why a number of

---

26. Birdsall, Ross, and Sabot 1995; Center for Global Development 2017
27. Jacoby, Li, and Rozelle 2002
28. Temple and Reynolds 2007
29. Checchi and van de Werfhorst 2014
countries that have a larger government role in the economy (a shift towards the command economy side of the spectrum), such as Norway, invest public funds heavily in early childhood education.  

**Measuring our output**

So far, we have been focusing on a two-good world in our production possibilities frontier model. However, in real life, there are not just two goods. How can we measure production with many goods? Economists measure output with the **gross domestic product (GDP)**, which is the market value of all final goods and services produced in a country over a specific time period. Market prices are used to calculate GDP based on what is produced. So, if a (very small) country produced five apples, priced at $2 each, two pears, priced at $3 each, and four bananas, priced at $1 each, its GDP would be $20 (5*2 + 2*3 + 4*1). Typically, GDP is an annual (year) measurement, but many countries also produce quarterly or monthly estimates of GDP.

GDP can measure the level of a nation’s production in current prices. When current prices are used to calculate GDP, this is referred to as **nominal GDP**. Nominal GDP could change over time due to prices rising, even if the country is producing the exact same amount of products. When prices increase for the same good or service, this is referred to as **inflation**. Economists measure changes in prices in order to know if a country is really producing more or not. When accounting for inflation, economists use **price indices**, measures of prices, such as the Consumer Price Index (CPI) to adjust GDP. GDP that has been corrected for inflation using a price index is referred to as **real GDP**, measuring the real (not just price) changes in production. Real GDP tells us whether a country is actually producing more.

Figure 1.4 shows real and nominal GDP in the United States and Yemen over time. Note that GDP for the United States is in trillions, and thus almost 1,000 times larger than that of Yemen, in billions. Real and nominal GDP have both been rising over time in the United States, with a dip around 2009 due to the Great Recession and another dip in 2020 with the COVID-19 pandemic. Nominal GDP has been rising faster than real GDP because prices have been rising over time; the two are the same in 2015, which was the base year used for prices in this case. Yemen had likewise been experiencing both nominal and real GDP growth until 2010. At that point, nominal GDP continued to rise quite rapidly because of rising prices, but real GDP fell, meaning fewer actual goods and services were being produced. Both real and nominal GDP began falling in Yemen starting in 2015, due to the ongoing war. The war has killed a quarter of a million people, destroyed housing and infrastructure, disrupted services, and pushed the country closer to famine.

---

30. Drange and Havnes 2015
31. World Bank 2022
32. Moyer et al. 2019
Two of the most important, headline-grabbing measures of the economy are the GDP growth rate and the inflation rate. These rates measure how much production is changing, which determines what a country can possibly consume, and how much prices are changing. While Figure 1.4 showed the level of output and prices, Figure 1.5 shows how they have changed over time. The (real) GDP growth rate shows how much production has expanded or, as in 2009 and 2020, contracted in the United States. The volatility of GDP growth is referred to as the business cycle, reflecting short run fluctuations. In the U.S., real GDP growth in the long run has tended to be between 2% and 4%. Prices have usually increased between 1% and 3% through 2020, but then inflation rose rapidly in 2021, due in part to supply chain disruptions caused by the pandemic. In contrast, Yemen has experienced much more volatile production and prices, with the real GDP growth rate ranging from a high of 8% to a low of -28%. Inflation has likewise varied from a 55% annual increase in prices down to as low as 2%. These fluctuations represent profound changes in the ability of the country to produce food or other goods, as well as for consumers to afford the prices of different goods, including food.

---

33. World Bank 2022
Another important factor to consider with GDP is how many people are in a country. When we divide the production of the country by its population, we calculate **per capita GDP** (per capita means per person). Per capita GDP can be in real or nominal terms. Although per capita GDP does not account for inequalities and is just an average, it is still a helpful starting point for considering a country’s resources and production. Figure 1.6 illustrates this point with real per capita GDP for the United States and Yemen. As of 2022, real GDP per capita was $62,867 in the United States and $1,018 in Yemen. While real GDP per capita in the United States increased over time, with a dip during the Great Recession and pandemic, real GDP per capita has risen and fallen in Yemen. It reached a high of $2,547 per capita in 2010 before falling to less than half of that level by 2022. Already starting from a low level, this means that the country has, on average, less than half as much production to meet the basic needs of its citizens.

![Figure 1.5. Real GDP growth rates and inflation rates (percentages) in the United States and Yemen](image)

<table>
<thead>
<tr>
<th>United States</th>
<th>Yemen</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. GDP growth rate</td>
<td>Yemen GDP growth rate</td>
</tr>
<tr>
<td>U.S. Inflation rate</td>
<td>Yemen Inflation rate</td>
</tr>
</tbody>
</table>

34. World Bank 2023
There are enormous disparities in the average level of production available to support the needs of Americans and Yemenis. Yet both countries have problems with hunger (although to varying degrees). This is because of inequality within countries and insufficient resources available to certain individuals and households. There are a variety of ways to measure poverty, when people do not have enough resources to meet their basic needs. In a later chapter, we examine poverty in the United States, and anti-poverty policies, including anti-hunger programs, in detail. For now, we will look at one measure of poverty—the percentage of people living on less than $2.15 per day. Taking into account price differences across country, this is considered a measure of extreme poverty. The good news is that extreme poverty has decreased over time in every single region of the globe (Figure 1.735). Despite a rising global population, there are fewer people living in extreme poverty than there

35. Hasell et al. 2023
were 20 years ago.\textsuperscript{36} Although we have not solved world hunger, and no region or country has zero poverty, we have, globally, made major progress.

![Figure 1.7. Extreme poverty rates (percentage of population living on less than $2.15 per day) by region. Notes: in 2017 international dollars, adjusted for price differences across countries and time.]

**Conclusions**

There is reason to hope that we can end hunger. The world already produces enough food to feed everyone.\textsuperscript{37} This production is, however, unequally distributed across countries and across households and individuals within countries. Differences in technology, physical and human capital and labor across countries are a key driver of international inequality. These differences arise for a variety of complex forces, ranging from the impact of colonialism to differences in physical climate.\textsuperscript{38}

\begin{itemize}
\item \textsuperscript{36} Roser and Ortiz-Ospina 2013
\item \textsuperscript{37} Gilbert et al. 2010
\item \textsuperscript{38} Sala-i-Martin, Doppelhofer, and Miller 2004
\end{itemize}
Global development goals (such as the Sustainable Development Goals) include eradicating hunger. Global organizations such as the World Food Program (WFP), United Nations (UN), World Bank, Food and Agriculture Organization (FAO), and International Fund for Agricultural Development (IFAD) work to reduce hunger. They provide a variety of short run help (humanitarian aid). They also provide grants, loans, and technical assistance (advice) to allow for investments in human and physical capital and technology that can increase production and reduce hunger in the long-run. Funded by wealthier countries, these agencies have helped (but not solved) the problem of world hunger; WFP has already cut back the food aid it is giving Yemen due to insufficient funding.\textsuperscript{39}

Although private property and other aspects of market economies create incentives that can spur growth, who owns property (or other forms of wealth) and who receives investments matters. Property ownership and investments are unequally distributed within countries. We will return to this point in the chapter on discrimination. Despite the challenges of inequality, we have seen global progress in reducing hunger and poverty. Investments in different aspects of production—labor, human capital, physical capital, and technology—have all contributed to increases in production. Families can produce more food by investing in these areas. Depending on the role they choose for government, countries can reduce inequality or spur growth through public policies that encourage or create investments in these areas. One day, these investments may be enough to end hunger.

\textbf{List of terms}

- Economics
- Microeconomics
- Macroeconomics
- Labor
- Human Capital
- Physical Capital
- Technology
- Long run
- Short run
- Production Possibilities Frontier
- Specialization
- Opportunity cost
- Margin
- Research and development (R&D)
- Command economy
- Market economy
- Gross Domestic Product (GDP)
- Nominal
- Inflation
- Price index
- Real
- Per capita
- Poverty

\textsuperscript{39} World Food Programme 2023
References


SUPPLY AND DEMAND: WHO GETS FOOD, HOUSING, AND WORK?

Caroline Krafft

Who gets food? The role of markets

Yemen’s food insecurity is an enormous challenge to the wellbeing of Yemenis. In the last chapter, we examined issues of production and ownership of production in thinking about why so many Yemenis were going hungry and what we can do to solve hunger. In this chapter we will focus on how food is allocated within Yemen through markets. The World Food Program (WFP) goes to markets in different areas of Yemen to check on the prices of different foods in local marketplaces. Unaffordable food prices in local markets are an important part of the story of food insecurity in Yemen—and the U.S. In this chapter, first we will learn about the role of markets, including supply and demand. We will then turn to the forces that can shift supply and demand to think about how to make food more affordable. Finally, we will analyze the effect of price controls on markets, focusing on two specific topics: rent control in the housing market, and minimum wages in the labor market.

Demand: What people want

How much food do families want? The answer to this question is going to depend on a number of different factors that we will explore throughout the chapter, but we will start with the price of food. The price is what a buyer pays (and what a seller receives) for a good or service. In Yemen, the currency is the Yemeni rial, so the prices are, for example, 100 rials for a pound of wheat flour. How many pounds of flour a family wants will depend on this price. The relationship between the price and the quantity people want is called demand. It is possible to describe this relationship with a table, an equation, or a graph. In this book, we will focus on graphing demand.

Figure 2.1 shows the demand for wheat flour in Amran governorate (a governorate is like a state) in Yemen.

1. WFP 2017
2. Walker, Keane, and Burke 2010
3. June 2017 prices from WFP 2017
On the x-axis (the horizontal axis) is the quantity, measured as thousands of pounds of wheat flour. On the y-axis (the vertical axis) is the price of a pound of flour, in Yemeni rials. The amount people want at a particular price is called the **quantity demanded**. At a price of 200 rials, the quantity demanded is 10,000 pounds (point A). At point B, when the price is 100 rials, the quantity demanded is 20,000 pounds of wheat flour. If the price rises to 200 rials, families will demand fewer pounds of flour. If the price drops even further, to 50 rials, they will demand more pounds of flour. This pattern occurs for a variety of goods: when the price goes down, people demand more. Likewise, if the price goes up, people demand less. This relationship is sometimes referred to as the “law of demand,” like the law of gravity. Graphically, this relationship is the downward slope of demand. Note that the whole curve is referred to as demand, while a particular point (such as point A) shows a price and quantity demanded.

![Demand for wheat flour in Amran governorate, Yemen](image)

**Figure 2.1 Demand for wheat flour in Amran governorate, Yemen**

**Supply: What people can provide**

Demand is just one side of the story of how much food (or hunger) there is. The other side is **supply**: how
much people or firms are willing to produce and provide to buyers. Like demand, supply depends on price and can be presented as a table, equation, or graph. The amount people want to provide at a particular price is called the **quantity supplied**. Figure 2.2 shows the supply of wheat flour in Amran governorate. At point C, when the price is 250 rials, the quantity supplied is 25,000 pounds. Point D shows that, at a price of 150 rials, the quantity supplied is 15,000 pounds. Suppliers want to supply less as the price decreases, and they want to supply more as the price increases. This relationship is often referred to as the “law of supply.”

Figure 2.2 Supply of wheat flour in Amran governorate, Yemen
Box 2.1. Tricks for graphing supply and demand

When you are first learning supply and demand, it may be tricky to keep track of what the graph should look like. Here are some tricks to help:

- **Demand** is downward sloping (both down and demand start with D)
- **Supply** is upward sloping (both supply and up have an “up”)
- The axes are labeled alphabetically: p goes before q in the alphabet, so when you are drawing your graph, starting at the upper left (just like writing) you’ll label p price first on the y-axis (vertical axis) and q second on the x-axis (horizontal axis)

Markets and equilibrium: Exchanges of goods and services

Supply and demand for a particular good or service—such as wheat flour in Amran governorate—are two sides of a market. Simply put, a market is where supply and demand come together. Sometimes “where” is a physical place, like the vegetable market (souq) in Amran city. Other times “where” is more of a concept. For example, the market for social media apps takes place on the internet.

Together supply and demand determine the price and quantity of a good. In Figure 2.3, supply and demand for wheat flour are shown on the same graph. Where supply and demand meet (cross or intersect) is referred to as equilibrium. In equilibrium, the price is 150 rials and the quantity is 15,000 pounds of flour. At this price, the quantity supplied and quantity demanded are equal. Unless something else changes in the economy to shift the curves (a possibility we will turn to next), price and quantity at equilibrium will stay the same.
Figure 2.3. Supply, demand, and equilibrium for wheat flour in Amran governorate, Yemen

We can understand why markets will stay in equilibrium with a counter-example. What if the price decreased to 100 rials? Then quantity demanded (20,000) would be greater than quantity supplied (10,000). Some people would not be able to buy wheat flour—there would be a shortage (quantity demanded would be greater than quantity supplied) in the market. However, some enterprising shop, since they kept selling out, would surely raise the price of flour, maybe to 110 rials at first, then 120, then 130, then 140. Other sellers would see that they, too, could raise prices. With each increase, although the shortage would be less, they would still sell out, until they reached 150 rials.

At 150 rials, if they raised their prices any further, say to 160 rials, there would be a surplus: quantity supplied would be greater than quantity demanded. A smart shop would lower its price, maybe to 155 rials at first, to make sure that it sold all of its flour. Other shops would follow suit, until they reached 150. Essentially, because quantity supplied equals quantity demanded, there is no motivation to change the price at equilibrium. Markets trend towards equilibrium through buyers and sellers acting in their own self-interest.
The trend towards equilibrium from these unseen forces, without requiring any action on the part of society or government, is often referred to as the “invisible hand.”

What causes changes in supply and demand?

Supply and demand can help us understand not only why prices and quantities are what they are, but also why they change—or how we can change them, for instance, how we can decrease hunger by shifting supply and demand for food. This section is all about what causes changes, called shifts, in supply and demand.

Shifts in demand

What determines demand? There are five main factors that affect demand:

- Preferences
- Income
- Population
- Related goods’ prices
- Expectations about the future

Preferences—what people like—can be shaped by information (including advertising), new experiences, or social trends. For example, if people heard that neighbors got sick after cooking with wheat flour, they might then demand less. Or, if people heard that wheat flour was good for their health, they might demand more.

Income plays a very important role in demand. Demand shows what people want and can pay for—not simply the quantity they want. When income increases, demand for most goods and services increases. One of the reasons for high levels of hunger in Yemen has been long periods of government employees going without pay. This decreased their income, which decreased demand for all sorts of goods and services.

Population also plays an important role in demand. When there are more people, there is more demand for food; when there are fewer people, there is less demand for food. One of the challenges that Yemen is facing is that it has some of the world’s fastest population growth. The fertility rate is 5.3, which means a woman will, on average, have 5.3 children. Yemen’s rapid population growth has increased the demand for food.

The prices of related goods can affect demand, but the effect depends on how the goods are related. Economists distinguish between goods that are complements—goods that go together—and goods that are substitutes—goods that can be used in the place of each other. For wheat flour, yeast might be a complement, since it goes with wheat flour when baking bread. If the price of a complement, such as yeast, increases, then

4. Scommegna 2012
demand for a good, such as wheat flour, will decrease. People will want to bake less bread and need less flour. Corn flour could be a substitute for wheat flour. If the price of a substitute, such as corn flour, rises, then demand for wheat flour will increase. Essentially, when the price of complements goes up, demand for a good goes down. When the price of a substitute goes up, demand for a good goes up. Likewise, when the price of a complement goes down, demand for a good will increase and when the price of a substitute goes down, demand for a good will decrease.

Expectations about tomorrow affect people’s demand today. If people expect the price of wheat flour to increase, they will demand more wheat flour now, in order to save and stock up for the future. Likewise, if they expect the price of wheat flour to go down soon, they will not demand as much wheat flour now.

Shifts in supply

What determines shifts in supply? There are four factors that can cause changes in supply:

- The costs of production
- The prices of related goods in production
- Sellers’ expectations
- The number of sellers

The costs of production are closely related to our last chapter. Although we did not put prices on inputs, we discussed how inputs (resources) and technology can increase or decrease production. The same factors affect costs. If inputs become less expensive, then supply will increase—farmers and shops will supply more food. Likewise, if technology improves, then supply will increase—farmers and shops will supply more food. However, if input prices increase, or technology is lost (uncommon), then supply will decrease. For the example of wheat flour, if seeds become less expensive, that is a decrease in input prices that would increase supply. Similarly, if the faculty of agriculture at Sana’a University in Yemen develops a new type of seed (technology) that doubles the size of the wheat plants that can grow in Yemen, that would increase supply.

The prices of related goods in production depend on the choices that farmers have. For example, in Yemen, sorghum is the most common crop. If the price of sorghum increases, farmers may switch their wheat fields to sorghum, reducing the supply of wheat flour. Likewise, if the price of sorghum decreases, farmers may then decide to supply less sorghum and more wheat flour.

Sellers’ expectations have a similar effect to buyers’ expectations. If farmers expect that the weather will be good for wheat, they will plant more wheat, increasing supply. If they expect that they will not be able to get enough water to irrigate their wheat, they may decrease supply.

5. Alabsi 2006
The number of sellers has a parallel effect to changes in population to demand. When there are more sellers, supply increases. When there are fewer sellers, supply decreases. Competition—how easy it is to enter a market and start a business—plays an important role in the number of sellers.

Analyzing shifts

Now that we understand how and why supply and demand change, we can start analyzing changes in the market and how they will affect the price and quantity of food—and thus hunger. Returning to our market for wheat flour, we can analyze the recent decrease in Somali refugees in Yemen and how it may impact the price and quantity of food in Amran governorate. Although Yemen has a number of its own development challenges, it has a long tradition of accepting refugees from other countries. It is the only country in the Arabian Peninsula to have signed the Refugee Convention. When conflict erupted in Yemen in March 2015, there were 280,000 refugees, 255,000 of them Somalis. As the conflict has continued, some of these refugees have left Yemen and returned to Somalia. Our supply and demand model can help us predict what the impact of the Somali refugees leaving Yemen will be on food prices and quantity.

To figure out the impact of this (and other) changes on a market, we will undertake a four-step process:

1. Draw the market (supply, demand, equilibrium (price and quantity)) before the change.
2. Decide which curve(s) shifted. Supply, demand, or both?
3. Decide what direction the curve(s) shifted. Did they increase or decrease?
4. Draw the new equilibrium and compare the new price and quantity to the old price and quantity.

Let’s apply this process to our market for wheat flour in Amran governorate.

1. Market before the change: Let’s say Figure 2.3 represents the market before the change. The equilibrium price is 150 rials and the quantity is 15,000 pounds.
2. What shifted as a result of Somali refugees leaving Yemen? If we consider the factors that shift supply and demand, this is a change in the population in Yemen. Population affects demand.
3. What direction did the curves shift? The population decreased as a result of the refugees leaving, so this represents a decrease in population, and therefore a decrease in demand (there are fewer people who want food).
4. We can draw this shift and the new equilibrium in Figure 2.4. The old equilibrium is labeled EQ1. The decrease in demand is a shift of the demand curve to the left, from the old D1 to the new D2. The supply curve stays the same. We find EQ2 where the new D2 meets the supply curve. At EQ2, the new
equilibrium, the price is 125 rials and 12,500 pounds of flour. The decrease in demand led to a decrease in price and a decrease in quantity.

When we are thinking about shifts in demand, this will cover everything that is not on the graph changing—so everything except price and quantity. The shifts will then change demand, and thus change the equilibrium quantity and price. Here it is important to note that, while demand shifted, supply did not—there was just a movement along the supply curve that resulted in a different price and quantity.
When you are first learning to shift supply and demand, it may be tricky to keep track of what the shifts should look like. Here are a few tricks to help:

- **Less** (decrease) is a **Left** shift (both less and left start with **L**)
- **More** (increase) is a **Right** shift (both more and right have **R**s)
- You can also check that your shift makes sense graphically. If a curve decreased, you should be able to check that the quantity is less than it used to be at any price. If a curve increased, you should be able to check that the quantity is more at any price.

Now let’s work through another example. Due to the ongoing conflict in Yemen, the main port in Hodeidah was seriously damaged, reducing the ability of firms and aid groups from outside the country to bring in food. Let’s think through how this change will affect the market for wheat flour in Yemen with our four-step process.

1. Market before the change: Let’s say that Figure 2.3 represents the situation in Yemen before the port damage.
2. What curve shifts? Here, the supply curve is shifting, because producers outside the country can no longer supply Yemen (number of suppliers).
3. What direction is the shift? Fewer producers (able to access Yemen) means a decrease in supply.
4. We can draw this shift and the new equilibrium in Figure 2.5. Here we can see the old supply curve, S1, and the new, decreased supply, S2. The equilibrium shifts from EQ1, when the price was 150 rials and the quantity was 15,000, to EQ2, a price of 200 rials and a quantity of 10,000 pounds of flour.

Undoubtedly, one of the factors contributing to hunger and food insecurity in Yemen is the reduction in supply due to the damage to the main port and the inability of many suppliers to access the market.

7. Ghobari 2016
These past two examples have illustrated how we can use the ideas of supply, demand, and markets to understand how current events affect the price and quantity of food—and thus hunger. We can use these same tools to model “what ifs” and make predictions as well. Let’s apply our four-step process to predict what happens if the conflict in Yemen ends. Specifically, let’s consider what happens to the wheat flour market if, simultaneously, Yemeni government employees all get their back pay AND the main port in Hodeidah reopens:

1. Market before the change: Let’s say Figure 2.3 represents the market before the change. The price is 150 rials and the quantity is 15,000 pounds.

2. What shifted as a result of employees getting their back pay and the main port reopening? Employees getting their back pay is about income—that’s demand. The main port reopening will affect supply. So, in this case, both curves shift.

3. What direction did the curves shift? The income increase will increase demand for food. The main port reopening will increase the number of suppliers, thus increasing supply.
4. We can draw this shift and the new equilibrium in Figure 2.6. The old equilibrium is labeled EQ1. The increase in demand is a shift of the demand curve to the right, from the old D1 to the new D2. The supply curve increases from S1 to S2. We find EQ2 where the new D2 meets the new S2 curve. At EQ2, the new equilibrium, the price is 150 rials and 25,000 pounds of flour. The increase in demand and the increase in supply led to an increase in quantity—but the exact same price! This is not always the case, though; the sizes of the shifts of the demand and supply curves determine whether the price changes or not—and the direction of the price change.

An important lesson from this example is that focusing just on price does not tell you whether people are better off. Given the increase in the quantity of food, hunger and food insecurity will decrease. Although people are better off, the price is the same. The WFP carefully monitors prices, and rightly so, because they are an important part of the story—but not the whole story.
Figure 2.6. Both supply and demand increase

The four-step process for analyzing changes in markets can be applied to all sorts of situations, and we will continue to use it throughout future chapters. Next, we turn to two specific markets, and also start looking at policies that affect how markets function. First, we are going to consider the market for housing and rent control. Then we will turn to the market for labor and minimum wages.

**Housing and rent control**

Housing is one of the most basic needs of families. By 2030, three billion people will need new housing.\(^9\) Existing housing suffers from inadequate quality, including poor access to water and sewage, and unsafe

---

\(^9\) Grandolini and Ijjasz-Vasquez 2016
conditions. Housing affordability is also a major challenge for families throughout the globe.\textsuperscript{10} Housing tends to be the single largest purchase of families.\textsuperscript{11} What can countries do about these housing challenges? One common solution to housing affordability is \textit{rent control}, a maximum price set on rents. Local, state, or national governments may enact rent control policies. The goal of rent control is laudable—making sure families can afford housing. However, an economic analysis of rent control demonstrates that the impact is much more complex. Rent control and other \textit{price ceilings}—maximum prices for goods or services—create winners and losers.

Figure 2.7 shows an economic analysis of rent control in the Philippines.\textsuperscript{12} If there were no rent control, the price (in Philippine pesos) and quantity would be determined by supply and demand. The equilibrium would be a price of 14,518 pesos and quantity of 1.02 million apartments. However, in the Philippines, because of rent control, there is a price ceiling. The price of apartments has a maximum at 10,000 pesos. Rents cannot be higher than 10,000 pesos (rents cannot go higher than the ceiling). At a price of 10,000 pesos, the quantity demanded (QD) is 12.5 million apartments; this is where demand intersects the price ceiling. With the price ceiling, the quantity supplied is 7.05 million; this is where supply intersects the price ceiling. The quantity demanded is more (much more) than the quantity supplied. There is a shortage of rental housing of 5.44 million units. The market is not in equilibrium. Because of the price ceiling it cannot reach equilibrium.

\begin{itemize}
  \item \textsuperscript{10} UN-Habitat 2016
  \item \textsuperscript{11} Turner and Kingsley 2008
  \item \textsuperscript{12} Based on Ballesteros, Ramos, and Magtibay 2016
\end{itemize}
As a result of rent control, there are a number of winners and losers in the Philippines’ rental market. Suppliers of apartments all lose—they receive less in rent than they would otherwise. On the demand side, there are a mix of winners and losers. Families who would like to rent but cannot find an apartment because of the shortage lose. Families who do get rent controlled apartments win; they pay less in rent than they would otherwise. Further economic analysis can demonstrate that the losses are larger than the benefits; rent control causes net harm to society. Empirical evidence also shows that, when rent control is implemented, landlords may shift into other markets.\textsuperscript{13} For example, they may stop renting and sell their buildings. Ultimately, this reduces supply and further raises the market price of rentals, the exact opposite of rent control policies intentions.

The implications of most other price ceilings are similar to the impact of rent control. This finding does

\textsuperscript{13} Diamond, McQuade, and Qian 2019
not mean that we have to despair about providing affordable housing—just that rent control is not the right solution. We will learn more about a more promising solution, subsidies, in the chapter on education.

**Labor and minimum wages**

Price ceilings (maximum prices), such as rent control, are one type of price control. The other type of price control is a **price floor**, a minimum price. The most common price floor in the world is a **minimum wage**. To analyze minimum wages, we can study labor markets. In the labor market, the price is called the wage. To keep the numbers easy, we will use hourly wages, but workers can be paid hourly, daily, weekly, or in terms of an annual salary. The quantity of workers can be measured just as the number of workers (we will use that for now) or the total hours people work.

In the labor market, workers are the ones supplying their labor, while firms demand labor. Labor demand from firms depends, in part, on the productivity of workers. Firms only want workers because they will contribute to production. Labor supply by (potential) workers depends, in part, on the population and its characteristics. We will revisit labor supply decisions in the chapter on poverty to better understand how individuals choose how much to work. In the real world there are actually many labor markets, for different types of workers (professors of economics are in a different labor market from professors of political science). For now, we will start with a simple labor market including everyone, like in Figure 2.8, where we see that the labor market equilibrium is 20 million workers being paid $10 an hour.

---

What effect will a minimum wage have on this labor market? That depends on where the minimum wage is relative to equilibrium. If the minimum wage is $7.25 an hour, as it currently is in the United States, then it will actually have no effect on our labor market in Figure 2.8. The equilibrium wage is already above the minimum wage at $10 an hour. A minimum wage of $7.25 has no effect, because firms would choose to pay workers more ($10) to get enough workers for their needs.

Within the United States, there have been a number of states and cities that have passed substantially higher minimum wages. New Jersey, Illinois, Maryland, Massachusetts, and New York have passed $15 minimum wages, which are in various stages of being phased in over time.\(^15\) Minneapolis, Minnesota, and Seattle, Washington, are two examples of cities that have passed $15 minimum wages.\(^16\) What will be the effects of these

\(^{15}\) National Conference of State Legislatures 2019
\(^{16}\) Nelson 2017; Reich, Allegretto, and Godoey 2017
Minimum wages on the labor market? We explore this possibility first in a very simple model, in Figure 2.9. Aside from the minimum wage being imposed, the figure is otherwise the same as Figure 2.8. The equilibrium wage would be $10 and the equilibrium number of workers 20 million. The price floor of the minimum wage now matters, because it is above the equilibrium. At a wage of $15 per hour, labor quantity demanded by firms would be only 15 million, but labor quantity supplied would be 30 million. As a result, 15 million people would be unemployed! Since the unemployment rate is the number of unemployed/labor supply, the unemployment rate would be 50%! This analysis is how minimum wages are often portrayed in introductory economics textbooks. The empirical evidence, however, does not align with this naïve model.

![Figure 2.9. Naïve model of minimum wage](image)

Minimum wage increases do not have large employment effects. The high-quality empirical studies of minimum wage increases typically show no employment effects (or even small increases in employment).\(^{17}\)

---

Why is the empirical evidence so different from the theoretical model? One reason is the shape of the naïve model in Figure 2.9. We have just been drawing easy lines, in the shape of an X, for supply and demand. The slope of the lines—whether they are steep or flatter—has been arbitrary. The shape or slope of the supply and demand curves depends on their elasticity. In general, elasticity refers to how much one economic variable—for example, quantity supplied—responds to changes in another economic variable—such as price. A demand or supply curve is elastic when there is a large change in quantity in response to a change in price. A demand or supply curve is inelastic when there is only a small change in quantity in response to a change in price. Typical introductory economics graphs show labor supply and demand as quite elastic—as in Figure 2.9. However, the empirical evidence suggests that labor supply and demand are inelastic. Labor supply may even bend backwards (which is not normal for supply curves), because if you start to make more money, you may work fewer hours, a point we explore further in later chapters.

A good trick for learning elastic versus inelastic is that inelastic is shaped like an “I,” more straight up and down (showing only small changes in quantity when price changes).

Figure 2.10 presents an alternative model of minimum wages, where labor demand is inelastic and labor supply is backward bending. The equilibrium wage and quantity would be the same as in Figure 2.8 and Figure 2.9. However, when the minimum wage (price floor) is implemented, the effect is very different. Wages rise to $15 per hour. The quantity demanded of labor is 17.5 million and the quantity supplied is 19 million, so there are 1.5 million unemployed—an unemployment rate of 8%. This is much smaller than the 50% in Figure 2.9.

19. Keane 2011
The Figure 2.10 model is likely still too pessimistic in terms of unemployment effects relative to the empirical evidence on minimum wages. Why might minimum wages have, on average, no effect on employment? One reason is that labor is an input into the production of other goods. Recall, from earlier in the chapter, the reasons supply and demand might shift. One is input prices. Labor is an input into the production of almost every other good and service. When wages (labor prices) go up, firms may decrease their supply curves for all their other goods, such that prices increase. This brings up one concern about the effect of minimum wages—and one reason they may not have employment effects—is that increases in minimum wages may just be passed on as price increases. Some studies suggest that minimum wage increases simply become higher prices, while others find no price effects. If minimum wage increases are passed on in terms of prices, then

---

20. Allegretto and Reich 2015
minimum wages may be effective for reducing poverty and inequality, but at the cost of higher prices to everyone.

Another reason that minimum wages may have no effect on employment is that when workers receive higher wages, they also spend more. Recall that one of the factors that affects demand for goods and services is income. When income increases, demand increases. Thus, it may be the case that workers who receive higher minimum wages spend more. This spending increases the demand for goods and services. The increased demand for goods and services in turn increases the demand for labor to produce those goods and services. In the end, there may be no change in employment as a result of the combined effects of minimum wages and higher demand for goods and services.

A third reason that there may be no employment effects from minimum wage increases is that higher wages may make workers more productive. The additional output from workers may more than cover their higher wages. The idea that firms may pay their workers higher wages in order to increase productivity was a key insight of economist Janet Yellen, the U.S. Secretary of Treasury and former chairwoman of the Federal Reserve (the United States central bank, in charge of our money supply). Higher minimum wages have been shown to increase productivity. Reduced turnover as a result of higher wages may also be a substantial benefit to firms, as finding and training new workers decreases productivity. These additional economic explanations demonstrate how overly simplified models (like Figure 2.9) can lead to faulty conclusions at odds with the empirical evidence on policies’ effects.

Conclusions

Markets are central to how all goods and services are allocated. Access to food, housing, and work depends on the functioning of markets. In this chapter, we have learned about supply and demand in markets. The supply and demand model shows how the price and quantity of goods and services are determined. This toolkit lets us analyze policies that change markets’ behavior. For instance, rent control is a price ceiling that lowers the price of housing—but reduces the quantity supplied. Minimum wages are a price floor that has complex consequences on markets. Although overly simple models suggest large unemployment effects from minimum wages, empirically there are, on average, no effects on employment. In future chapters, we will apply and expand our supply and demand models to analyze many other social policies, such as pollution control, free education, and discrimination.

22. Bosch and Manacorda 2010; Dube 2017; West and Reich 2014; Wong 2019
23. Aaronson, Agarwal, and French 2012
24. Yellen 1984
25. Rizov, Croucher, and Lange 2016
26. Boushey and Glynn 2012; Ruffini 2020
### List of terms

- Market
- Supply (curve)
- Demand (curve)
- Quantity supplied
- Quantity demanded
- Price
- Equilibrium
- Shortage
- Surplus
- Substitute
- Complement
- Rent control
- Price ceiling
- Wage
- Price floor
- Minimum wage
- Unemployment
- Elasticity
- Elastic
- Inelastic

### References


Where do goods come from?

Set down your reading for a moment and check the labels on your clothes. Where were your clothes made? My sweater was made in China and my shirt was made in the United States. The U.S. textile and apparel industry (including production of clothes in the U.S.) has a value of $65.2 billion as of 2021. However, as looking at your tags will demonstrate, we do not wear only what is produced in the United States. An enormous amount of what we wear we import, meaning it is produced in other countries and we bring it to the U.S. As of 2019, the U.S. imported $127.7 billion dollars of textiles and apparel. Much more clothing is imported to the U.S. than produced in the U.S. (and you probably have more tags with other country names than the U.S.). At the same time as we import textiles and apparel, the U.S. also exports, meaning goods produced here are sold to other countries. As of 2019, we exported $22.1 billion of textiles and apparel. Why are we both exporting and importing? Why do we buy clothes made mostly abroad? What is the impact of all of this trade on the U.S. and other countries?

This chapter examines the state of global trade, including an understanding of why we trade goods and services with other countries. The chapter also examines some of the controversies around trade and its impact on our country and the world. Trade, trade deals, and trade policies are often controversial, as they can have enormous impacts on lives and livelihoods around the world. One concern with trade deals is that we might be “sending our jobs overseas.” Another concern is the environmental and human impact of trade between countries with unequal environmental and human rights protections. Both of these concerns are reflected in the controversy around the Rana Plaza collapse in Bangladesh. Bangladesh’s garment industry is enormous. It employs somewhere between 3.5-4 million people and contributes 12% of Bangladesh’s GDP. In contrast,

---

3. Ibid.
the U.S. textile and apparel industry employs 0.6 million people. \textsuperscript{5} Among the U.S.’s $127.7 billion in clothing imports, $6.2 billion are from Bangladesh. \textsuperscript{6}

A number of U.S. and European apparel companies order clothes from Bangladeshi factories. One such factory was Rana Plaza, an eight-story building near Dhaka, Bangladesh. The Rana Plaza garment factory collapsed in 2013, killing over 1,100 workers and injuring thousands more. \textsuperscript{7} This collapse, the deadliest disaster to befall the garment industry, was due to poor construction and overloading the building with equipment. The collapse brought attention to the poor and often unsafe conditions garment workers (disproportionately women) face. Would the world be better off without trade? Our economic analysis, below, will ultimately show that, while overall countries gain from trade, there are distinct winners and losers from increasing trade.

**Flows of goods**

How much do we trade? Figure 3.1\textsuperscript{8} shows exports as a percentage of world GDP. Even as early as 1827, exports were substantial (7\%) relative to global production. Declining transport costs, particularly maritime transport costs between countries, as well as strong extractive trade links between colonizers and colonies, played an important role in the continuing growth of trade into the 1910s. The decline thereafter was due to rising transport costs (particularly limiting during the world wars), as well as a shift away from using gold as the basis of currency (the gold standard). \textsuperscript{9} Trade grew again after World War II, with exports at around 24\% of GDP by 2014. Trade is an enormous and increasingly important component of our global economy.

\textsuperscript{5} National Council of Textile Organizations, 2021.
\textsuperscript{6} United States International Trade Commission, 2019.
\textsuperscript{7} Westerman, 2017.
\textsuperscript{8} Ortiz-Ospina, Beltekian, and Roser, 2019.
\textsuperscript{9} Estevadeordal, Frantz, and Taylor, 2017.
The amount of trade varies across countries and goods. Figure 3.2\textsuperscript{10} shows imports to the U.S. from top trading partner countries and exports from the U.S. to the same trading partners, for textiles and apparel. The U.S. imports $43.0 billion of textiles and apparel from China alone. It also exports some $0.7 billion in textiles and apparel to China. Geography clearly plays a role in trade as well. The U.S. exports more to Mexico ($4.1 billion) and Canada ($3.5 billion) than the other countries shown. In fact, the U.S. exports more to Canada than it imports from Canada ($2.1 billion).

\textsuperscript{10} United States International Trade Commission, 2019.
Comparing the exports and imports data for textiles and apparel, it is clear the U.S. imports more textiles and apparel than it exports. For other goods, such as aircraft, the reverse is true, and the U.S. exports more than it imports. When we add up the value of trade across goods and services to estimate total imports and total exports, we can compare the two to see if we are generally exporting more or importing more. This difference (exports-imports) is referred to as the trade balance.

When exports are greater than imports and the trade balance is positive, there is a trade surplus. When exports are less than imports and the trade balance is negative, there is a trade deficit. Although trade deficits are often highly politicized as a problem (we are buying more than we are producing), there is nothing inherently bad (or good) about either a trade deficit or surplus. When there is a trade deficit, this means other countries are buying less from us than we are from them, but are typically investing that difference in assets, including real estate, debt, or bonds. Essentially, the trade deficit allows us to borrow from foreign countries to invest within our country. Although this borrowing can be a problem if it is unsustainable or excessive, a trade deficit is not inherently a problem.

Figure 3.3 shows the trade balances, in billions of dollars, for the five countries with the biggest trade deficits and the five with the biggest surpluses. It also shows these balances as a percentage of GDP, measuring their size relative to countries’ economies. While the U.S. has the largest trade deficit in absolute terms, $651 billion as of 2020, this deficit is only 3.1% of GDP. The U.S. trade deficit is a smaller share of GDP than some other countries with large trade deficits, such as Egypt (a trade deficit of 7.5% of GDP) or the Philippines (a trade deficit of 7.7% of GDP). The largest global trade surplus in absolute terms is for China, $466 billion, 2.5% of its GDP. Germany is the next largest, at $220 billion, which represents 5.7% of its GDP. Although a lot of political concern focuses on these trade surpluses and deficits, countries with strong economies are on both sides of the spectrum of trade balances.

![Figure 3.3 Trade balances in billions of dollars and as a percentage of GDP, 2020.](image)

Notes: Five largest (in dollar terms) trade deficits and surpluses.

When countries trade goods, they also have to trade currencies. This trade in currencies is referred to as foreign exchange. The exchange rate is how one country’s currency is translated into another country’s currency. For example, the Bangladeshi currency is the taka. One taka is equivalent to $0.012. It takes 84.58 Bangladeshi taka to buy one U.S. dollar. These exchange rates facilitate trade. For example, if a Bangladeshi company wants to buy a computer from Dell, a U.S. company, and the computer costs $500, the Bangladeshi company can exchange 42,290.88 taka for $500. Likewise, when buying clothes from Bangladeshi manufacturers, a U.S.

corporation like Walmart can trade dollars for taka. For a garment costing 800 taka to produce, the U.S. could exchange $9.46. This flow of foreign exchange is how we finance trade.

Why do we trade goods?

Production and consumption

Why do countries trade goods? How do countries decide which goods to make domestically, export, or import? We can answer these questions by returning to our production possibilities frontier model. To make our graphs easier to draw, we’re going to assume the production possibilities frontier is now a straight line (that inputs within a country are not specialized). We are going to ignore currency and prices for the moment and just concentrate on exchanging goods. We’re also going to assume there are only two countries in our model. The intuition of the model will, however, scale up to many countries. As we think about trade, we need to distinguish between production (what countries make) and consumption (what goods or services countries consume). Just like the production possibilities frontier, we can have a consumption possibilities frontier (CPF), the most that people can consume. In Chapter 1, because we were looking only at one country and did not yet consider trade, consumption and production possibilities frontiers were the same.

Figure 3.4 shows a case where production and consumption possibilities are the same and there is no trade between Bangladesh and the U.S. In this case our two goods are garments and milk. In this example, the most milk Bangladesh could produce (and consume) would be 40 million gallons (and no garments). The most garments Bangladesh could produce (and consume) would be 80 million items (and no milk). In contrast, the U.S. could produce 100 million gallons of milk or 100 million garments. Both countries could also produce at points in between, with a mix of milk and garments, along the CPF/PPF.
Figure 3.4 Bangladesh and the U.S. consume only what they produce (no trade)

<table>
<thead>
<tr>
<th>Bangladesh</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk (millions of gallons)</td>
<td>Consumption</td>
</tr>
<tr>
<td>Garments (millions of items)</td>
<td>Production</td>
</tr>
</tbody>
</table>

Why do the countries have different production possibilities? Recall what affects production: quantity of labor, quality of labor (human capital), physical capital (including resources) and technology. In the case of milk versus garments, natural resources are critical. Producing milk requires cows eating grass in a field, or fields of hay or corn feeding cows in a barn. There are areas in the U.S. where resources work well for grazing cattle or growing hay or corn. Bangladesh has an ample water supply, which is good for rice, the country’s staple food crop, but not so good for cows grazing, hay, or corn.

When countries trade, they tend to specialize in certain goods to produce and export, and other goods to not produce and import. In the case of Bangladesh and the U.S., Bangladesh tends to specialize in producing garments to export to the U.S. while the U.S. tends to produce and export dairy to Bangladesh. Why do they specialize in these goods and exchange them? The PPF shows that they have different abilities to produce. First off, the U.S. can produce more of both goods, 100 million gallons of milk compared to Bangladesh’s 40 million, and 100 million garments compared to Bangladesh’s 80 million. This ability to produce more, in absolute terms, is referred to as **absolute advantage**. The U.S. has the absolute advantage over Bangladesh in both milk and garments.

The countries also face different tradeoffs, opportunity costs in changing production. We can see this by comparing the end points of production. If Bangladesh wants to switch from all garments (80 million) to all...
milk (40 million gallons) it has to give up two garments for every gallon of milk. The opportunity cost of one gallon of milk is two garments. Going the other way, if Bangladesh wanted to shift from all milk to all garments, it would have to give up $\frac{1}{2}$ a gallon of milk for every garment. The opportunity cost of a garment is $\frac{1}{2}$ a gallon of milk. Depending on which way you are looking (increasing milk or increasing garments) this holds for all of the (straight line) production possibilities. Wherever production currently is, the tradeoff is 1 gallon of milk for 2 garments. In contrast, the U.S. can produce either 100 million gallons of milk or 100 million garments. Its tradeoffs are 1:1; it has to give up one garment for one gallon of milk. The opportunity cost of a garment is a gallon of milk (and vice versa). These differences in tradeoffs allow us to identify which country has to give up comparatively more to produce a good.

When a country has to give up relatively less of good A to produce good B, a smaller opportunity cost, it is said to have a **comparative advantage**. In the example above, Bangladesh had an opportunity cost of $\frac{1}{2}$ gallon milk for every garment while the U.S. had an opportunity cost of 1 gallon of milk for every garment. Bangladesh’s opportunity cost to produce more garments is lower than the U.S.’s. This means that Bangladesh has a comparative advantage in garments. Although one country—like the U.S.—may have an absolute advantage in both products, it can never have a comparative advantage in both. So while Bangladesh has a comparative advantage in garments, the U.S. has a comparative advantage in milk. It only has to give up one garment to produce one more gallon of milk. Bangladesh has to give up two garments to produce one more gallon of milk.

The benefits when two countries trade

Now that we know that Bangladesh has a comparative advantage in garments and the U.S. has a comparative advantage in milk, we can figure out how the countries should specialize their production and then identify the benefits of trade. Countries should specialize in the goods where they have a comparative advantage. Bangladesh should specialize in garments and the U.S. should specialize in milk. The countries then will determine how many garments to exchange for a gallon of milk. This rate of exchange is referred to as the **terms of trade**. Neither country will accept terms of trade that are worse than what they could do without trade. Thus, the terms of trade must be between 1 and $\frac{1}{2}$ a gallon of milk for a garment. If it had to trade more than 1 gallon of milk for a garment, the U.S. would be better off just producing more garments itself. If it had to trade less than $\frac{1}{2}$ a gallon of milk for a garment, Bangladesh would be better off just producing more milk itself. Where the exchange falls depends on the negotiating skills of the parties involved. Let’s say we end up with an exchange rate of $\frac{3}{4}$ of a gallon of milk for one garment. The U.S. gives Bangladesh $\frac{3}{4}$ of a gallon of milk for every garment it sends. Bangladesh gives the U.S. $\frac{4}{3}$ of a garment for every gallon of milk it sends. Essentially, 3 gallons of milk go from the U.S. for every 4 garments coming from Bangladesh. The U.S. is importing garments and exporting milk; Bangladesh is importing milk and exporting garments. One country’s export is always another’s import.

Having settled on our terms of trade, although the production possibilities will remain the same,
consumption possibilities can change. Since Bangladesh is specializing in garments, still, the most it can get for either production or consumption is 80 million garments. However, if it produces 80 million garments, it can now sell them to the U.S. to get $\frac{3}{4}$ as many gallons of milk: 60 million gallons of milk. The consumption possibilities for milk have increased from 40 million to 60 million (along with various mixes of garments and milk in between shifting out). At the same time, the U.S. will specialize in milk and thus the most it can consume will still be 100 million gallons. However, if it sells 60 million gallons to Bangladesh it can get 80 million garments in return. So if the U.S. produced 60 million gallons of milk and 40 million garments, it could trade for 80 million garments and end up with 120 million garments (and likewise for various other combinations). The basic idea—shown in Figure 3.5—is that countries are always better off when they trade. Their consumption possibilities always increase, which economists refer to as the gains from trade.

<table>
<thead>
<tr>
<th>Bangladesh</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="#" alt="Diagram" /></td>
<td><img src="#" alt="Diagram" /></td>
</tr>
</tbody>
</table>

Figure 3.5 Bangladesh and the U.S. specialize and trade

**Trade agreements**

Now that we know that there are potential gains from trade, we can see why countries often try to make trade agreements, in order to encourage trade. The perspective that trade should be unobstructed and allowed to flow without barriers between countries is referred to as free trade. One recent trade deal is the United States-Canada-Mexico Agreement (USCMA).\(^{15}\) The deal replaced the North American Free Trade Agreement

---

(NAFTA). The new USCMA has a lot of similarities to NAFTA, but also has a number of new regulations that will create winners and losers. For instance, USCMA requires more U.S. parts in autos, likely resulting in higher auto prices.\(^\text{16}\)

Trade agreements are common throughout the world. The European Union (EU) was initially a political union with a common market and easier trade. The EU ultimately worked towards a single integrated market with completely free trade within the EU. Recently, the United Kingdom voted to quit the EU, including leaving the single market.\(^\text{17}\) This shift away from free trade was referred to as “Brexit” (British exit). Another recent (and controversial) trade deal is the Trans-Pacific Partnership (TPP). Initially TPP was a deal between the United States, Canada, Australia, and countries in East Asia (including Japan, Malaysia, and Vietnam). The U.S. withdrew under President Trump, with the 11 remaining countries subsequently negotiating a deal without the U.S.\(^\text{18}\)

USCMA/NAFTA, the EU, and TPP are all examples of trade deals between a few or a group of countries. Countries may also make bilateral (two-way) deals, between just two countries, such as the U.S. and China. There are some definite tradeoffs to making agreements on a bilateral or regional basis, as they can create an uneven playing field for different countries. The organization that aims to help achieve global free trade, as well as deal with trade disputes between countries, is the World Trade Organization (WTO). WTO agreements (rules of trade) have been signed by the majority of the nations of the globe.\(^\text{19}\)

### Concerns around free trade

As USCMA/NAFTA, the EU, and TPP illustrate, trade deals are controversial. Part of that is the nature of deals—countries are negotiating to try to get the best deal they can. They may try to renegotiate if deals are no longer working for them or they think they can get a better deal. As well as renegotiating, some countries leave trade deals, as with Brexit. Why might countries want less free trade? A position of wanting less trade is referred to as protectionism. The idea behind the name protectionism is that countries want to protect their domestic production from foreign competition.

We can better understand protectionism if we return to Figure 3.5. Recall that we said that countries are always better off when they trade, because the consumption possibilities increase. While countries as a whole are better off, trade and trade deals create winners and losers. In the example of the U.S. and Bangladesh, think of what happens when we go from Figure 3.4, with no trade, to Figure 3.5, with trade. Perhaps in the no trade case Bangladesh chose to produce 20 million gallons of milk and 40 million garments. When it trades,

---

16. Ibid.
it specializes in garments, produces zero gallons of milk, and 80 million garments. Perhaps it trades 32 million garments for 24 million gallons of milk.

Consumers are better off. They are left with 48 million garments (more than 40 million without trade) and now have 24 million gallons of milk (more than 20 million without trade). However, there are winners and losers among producers in Bangladesh. Bangladeshi garment producers win—their business rapidly expands. Bangladeshi milk producers lose—their farms shut down. The loss or shrinking of certain industries due to trade motivates protectionism in Bangladesh. Likewise, in the U.S., consumers win as a result of trade, but garment producers lose and milk producers win. In Bangladesh, the milk producers would protest a trade deal, while in the U.S., garment producers would protest. Empirically, there are distinct winners and losers when trade increases. In Brazil, increased trade likely reduced poverty overall, but with effects that varied by state and rural versus urban areas.20

Trade restrictions can take a variety of different forms, as we will discuss below. Globally, when trade restrictions, such as the Multifiber Arrangement, are lifted countries receive differential benefits.21 Lifting trade restrictions can reduce the prices of goods for consumers. Ending the Multifiber Arrangement in 2005 led to a decrease in costs averaging $63 per household.22 This impact highlights an important and challenging aspect of trade deals. The benefits tend to be large in total but diffuse (every American having slightly cheaper clothes) while the costs are less in total, but concentrated (U.S. cities where garment factories were shut down were strongly affected, with some people losing jobs). Although there are additional gains in other sectors (in Figure 3.5, the U.S. dairy sector would be hiring), these jobs are not necessarily in the same locations, nor do they require the same skills, which is why trade agreements also often include provisions for retraining of workers.

Other protectionist arguments have to do with safety. Countries may want to ensure they have domestic production of certain goods and services. For instance, countries may want a certain amount of domestic military or food production, so that in the event of a global disaster or war they can protect their populace. There may also be concerns about the quality or safety of imported products. Another safety issue is the “race to the bottom” illustrated by the Rana Plaza tragedy in Bangladesh. One of the reasons firms may choose to produce in Bangladesh is that it is cheaper to do so. This cheaper production may be because safety conditions and human rights protections are not observed. Likewise, production may move to locations where there are not environmental protections. Environmental protections are costly, and it is cheaper to produce in ways that damage the environment (a point we will return to in later chapters).

Although worker safety, human rights, and the environmental impact of trade are serious concerns, it is not clear to what extent trade is impacting these areas; there is ongoing debate among economists.23 It may also

23. See, for example, the contrasting claims in Managi, Hibiki, and Tsurumi, 2009; Frankel and Rose, 2005.
be possible to build protections, for instance environmental standards, into trade deals. Consumers may be willing to pay more for goods which are ethically produced, for example, fair trade coffee which more directly benefits farmers and their families.

Trade restrictions

If countries decide to go the protectionist route, how might they reduce or prevent trade? There are three main policy options. The first, most radical, and simplest, is an embargo, forbidding any trade at all (or any trade in certain goods). Putting in place an embargo essentially takes a country allowing trade, with the CPF exceeding the PPF (Figure 3.5), and returns it to the CPF equaling the PPF (Figure 3.4). The countries can only consume what they produce. There is a clear loss to consumers. The other common types of trade restrictions are quotas, which are maximum quantities allowed in, and tariffs, which are a tax on imports. In order to understand the impact of quotas and tariffs on the economy, we will return to our supply and demand model.

Figure 3.6 shows the U.S. market for garments from Bangladesh. If there are no trade restrictions, that means our usual supply and demand analysis applies, and equilibrium is a price of $20 and 20 million garments. However, we can add a quota to the market. Here, we set a quota at a maximum of 15 million garments. Similar to the case with price controls, there are times when the quota may have no effect. If the quota (maximum quantity) is 30 million, in this market, it will have no effect. However, with a quota of 15 million Bangladeshi garments, the quantity definitely goes down in this market. At a quantity of 15 million, the price can be found on the demand curve (since that’s how much people are willing to pay). The price is $25. Suppliers would be foolish to lower the price to try to move to equilibrium, because if they lower the price, the quantity cannot increase, and they will just be making less money. As a result of the quota, consumers face higher prices. Suppliers in Bangladesh who succeed in exporting their goods to the U.S. with the quota receive a higher price, but many suppliers in Bangladesh cannot export. We can actually do a bit of math to determine that at $20 with 20 million garments Bangladeshi suppliers would get $400 million, and at $25 with 15 million garments, suppliers would get only $375 million, a net loss. Consumers in the U.S. are losing, suppliers in Bangladesh are losing. So, who is winning?

The winner is someone not on the graph: U.S. garment producers. Since a bunch of folks in the U.S. are no longer buying clothes from Bangladesh, this will lead to an increase in demand in the U.S. garment market. This increase in demand will also increase prices for consumers. Ultimately, economists have shown that the gains to producers are less than the losses to consumers, and so a quota causes a net loss to the U.S. economy (as well as a loss to Bangladesh).

---

Figure 3.6 U.S. supply and demand for garments from Bangladesh with a quota

What happens if, rather than a quota, we use a tariff (tax on imports)? Figure 3.7 shows a $10 tariff on garments from Bangladesh. We show the tariff as a decrease in supply, raising prices exactly $10 along the whole supply curve. This shift is essentially an increase in the cost of production to get the garment into the U.S. It’s easiest to see the tax where supply hits the price axis; there we can see the difference between the original supply and supply with tax is $10. We find the new equilibrium for consumers where supply with tax meets demand. The price (labeled $P_d$) is $25 and quantity 15 million. This price and quantity are the same results we got with a 15 million quota. However, the difference in this case is that firms in Bangladesh do not receive the full higher price. Instead, they receive the price less the tariff: $25-$10=$15 (labeled $P_s$).
Bangladeshi producers are, in this case, all worse off, as are consumers in the U.S. Producers in the U.S. have increased demand for U.S. clothes, but consumers will face higher prices for those as well. There is another winner in this case as well—the government. The government receives the tariff as revenue. Specifically, in this example, it receives $10 for each of 15 million garments: $150 million. However, economic analysis can demonstrate that the benefits to the winners (U.S. producers and government) are less than the harms (increased prices) to consumers. As with quotas, tariffs create a net loss. For example, from the 2018 U.S. tariffs against washing machines, washing machine prices rose almost 12%.\footnote{Flaanen, Hortacsu, and Tintelnot, 2019.} Consumers paid more than 100% of the costs of the tariff increase, because firms increased dryer prices at the same time, even though they did not face
tariffs. The tariffs did create additional jobs in domestic washing machine manufacturing – but at a very high cost to consumers of $820,000 per job created.27

Summary and Conclusions

We live in a globalizing world. Trade has become increasingly important in the global economy. Trade can definitely increase the consumption possibilities for consumers. However, it creates distinct winners and losers among producers, which is one of the reasons that protectionists argue against free trade. Protectionist measures include embargoes, quotas, and tariffs. Our economic analysis demonstrated that all these measures cause specific winners and losers. Consumers lose due to higher prices, but domestic (protected) producers win. However, the country placing trade restrictions, on the net, loses. Additionally, when countries such as the U.S. start to place restrictions on trade, their trading partners are likely to undertake a tit-for-tat retaliation strategy, raising trade restrictions on U.S. goods and harming (different) U.S. producers. Currently, the U.S. is engaged in a trade war with China; China has retaliated with tariffs that are hurting U.S. exporters, whether Minnesota farmers exporting soybeans or Massachusetts cranberry producers.28

Trade deals are one tool to try to address some of the concerns around trade through a package of measures; countries can ensure they each have winning industries. Trade deals can also, potentially, allow for including human rights and environmental protections. Such protections help avoid a “race to the bottom” in production. Deals can also include funds for retraining or relocating displaced workers in order to ensure shifts in production do not lead to large increases in unemployment in certain sectors or locations.

List of terms

- Import
- Export
- Trade balance
- Trade surplus
- Trade deficit
- Foreign exchange (rate)
- Consumption possibilities curve
- Absolute advantage
- Comparative advantage
- Terms of trade
- Gains from trade
- North American Free Trade Agreement (NAFTA)
- United States-Canada-Mexico Agreement (USCMA)
- European Union (EU)

27. Ibid.
References


THE ECONOMICS OF CRIME

Caroline Krafft

What is crime?

Before we can measure crime, understand its causes, or think about policy solutions to prevent and reduce crime, it is important to take a step back and ask a fundamental question: What is crime? In the simplest terms, a crime is an illegal act or action.

Since different countries, states, and even cities have different laws, what is considered a crime varies across localities. The definition of crime has also evolved over time. Context can often determine what is and is not legal. For example, in Maryland it is illegal to keep chickens (and other poultry) without registering the birds. The law is motivated by both economic and public health concerns. Poultry is the state’s number one agricultural product. Backyard birds may be at risk in cases of disease outbreaks, as well as potentially spreading disease to commercial flocks. In other contexts, for instance in Connecticut where poultry production is low, poultry registration is not mandated by law.

Though the term “crime” is often used in moral and justice-based contexts, for example “a crime against humanity,” laws are not inherently moral or just. What is considered moral is often in the hands of lawmakers, and laws often embody the common attitudes of the times. In the United States, laws are typically passed by local, state, and federal legislatures by elected representatives. Thus, the attitudes, including the biases, of voters and policymakers, may become embedded in the law.

Ideally, law and morality are one and the same. However, this has proven not to be true historically. Laws—and definitions of crimes—can be immoral or unjust. For example, historically in the state of Virginia, marriage between individuals of different races was illegal. Such laws are referred to as anti-miscegenation laws. Thus, the 1958 marriage of Mildred Jeter, an African-American woman, and Richard Loving, a white man, was a crime in the state of Virginia. The Lovings were indicted (charged) and sentenced for this crime. Ultimately, the Supreme Court struck down the law (and similar laws in other states) in the 1967 case Loving...

1. Maryland Department of Agriculture
2. AVMA 2006
3. Connecticut Department of Agriculture 2017
v. Virginia. The ruling was based on the principle of individuals having equal protection before the law, prohibiting discrimination based on race.⁴

A more recent example is the laws and punishments for different forms of cocaine. Crack cocaine is produced from powder cocaine, but the amount of crack cocaine that triggers eligibility for a mandatory minimum sentence is 1/100th the amount of powder cocaine. This disparity embodies racial biases, as the majority of individuals (more than 80%) of those convicted for crack cocaine offenses are African American, compared to a third of those convicted for powder cocaine and other drugs. Recent reform, in response to outcry around the bias in sentencing, has changed this ratio to a minimum sentence for crack cocaine at 1/20th the amount of powder cocaine. However, the law still embodies biases that disproportionately harm African-American communities.⁵

Criminal justice systems

The criminal justice system is the name for the agencies and actors who implement policies relating to crime.⁶ While the structure of the criminal justice system varies from country to country, typically it has three main parts: (1) the police, who identify and apprehend those considered to be criminals, (2) the courts, who decide the guilt or innocence of those accused of crimes and sentence them, and (3) corrections institutions, such as prisons, that carry out the penalty determined by the courts. Different levels of government—local, state, or federal levels in the United States—have different, linked roles in the criminal justice system.

Although there are common elements to criminal justice systems throughout the world, countries take different approaches to how crime is addressed. The five basic approaches to controlling crime are: (1) retribution, (2) incapacitation, (3) prevention (4) rehabilitation and (5) restoration. Retribution seeks to punish individuals for their past crimes. Incapacitation tries to prevent future crimes by removing criminals from the community (historically by banishing, more recently by imprisoning). Prevention aims to prevent crimes by changing the costs, benefits, or opportunities of committing a crime. Rehabilitation is an approach focused on transforming an offender into a law-abiding, healthier, and more productive member of society. Restoration (restorative justice) focuses on crime as a form of injury to the victims, community, and offender and aims to repair the injuries and reconcile all parties.⁷ Elements of these models are typically present in all criminal justice systems, but to varying degrees.

Figure 4.1⁸ demonstrates the degree to which countries use imprisonment as a form of criminal justice. The

---

4. U.S. Supreme Court 1967
5. Bjerk 2016; Vagins and McCurdy 2006
6. Neubauer and Fradella 2017
7. Neubauer and Fradella 2017
8. Institute for Criminal Policy Research 2020
United States has by far the largest number of individuals imprisoned, 2.1 million. It is important to note that the United States also has a large population, but even in terms of the rate of imprisonment (prisoners per thousand individuals), the United States has the highest rate of imprisonment of any major nation.9

![Prison populations by country, in thousands](image)

Figure 4.1 Prison populations by country, in thousands

As well as debates about how to approach criminal justice—which are tied, in part, to debates around the causes of crime, discussed below—there are debates about the performance of the different models, and

particularly critiques about disparities in who ends up in jail. Figure 4.2 presents information on the rate of imprisonment (per 100,000 people) by sex and race/ethnicity in the United States. Black men are almost six times as likely to be imprisoned as white men, and Hispanic men and men of other races are more than twice as likely to be imprisoned as white men. Although women are much less likely to be imprisoned in general, there are also substantial disparities among women in rates of imprisonment by race and ethnicity.

![Figure 4.2. Rate of imprisonment (prisoners per 100,000 population), by sex and race/ethnicity](image)

Why do these disparities in imprisonment occur? One reason may be that, for instance, men commit more crimes than women. However, individuals may also receive disparate and unfair treatment in the criminal justice system, as in the case of crack versus powder cocaine, depending on their identities. An important role for economists and statisticians in researching crime is identifying disparities in the criminal justice system, for instance racial profiling by police officers, and designing policies to reduce those disparities.

### Causes of crime

Why do people commit crimes? In order to understand how to effectively prevent and respond to crimes,
we first need to understand the causes of crime. The typical way economists approach crime is to treat it as a rational, logical decision. Individuals weigh the benefits of crime versus the costs of crime. In general, comparing costs and benefits is referred to as **cost-benefit analysis**. When the benefits exceed the costs, individuals commit a crime. So, for example, an individual might decide to rob a bank by comparing the benefits of committing the crime, for instance receiving $10,000 in cash, to the costs of crime. The costs would depend on the chances or probability of getting caught and the consequences of getting caught—the punishment in terms of jail time or a fine. If an individual thinks they have a 10% chance of getting caught, and would expect to have to pay back twice what they stole ($20,000), then it would make rational sense to commit a crime; the benefits ($10,000) exceed the expected costs ($2,000=$20,000 for the cost*0.10 for the 10% chance).

One important cost to consider when assessing crime is the opportunity cost of potential criminals’ time. Engaging in illegal activities, such as growing marijuana, means forgoing alternative, legal uses of time, such as engaging in work. Since there are multiple alternative uses of an individual’s time, he or she will only choose to engage in crime when crime is more lucrative than all alternatives. An individual who owns a truck may choose to use it for the transportation of legal goods, to sell it, or to use it for the transportation of different illegal goods (e.g. guns, drugs, human trafficking). Their decision will be the alternative where the benefits, relative to the costs, are greatest. Opportunity costs are important in determining crime. Improvements in economic conditions, such as lower unemployment and higher wages, play an important role in reducing crime. Changes in opportunity costs affect crime globally; in coastal countries, if fish catches are lower, individuals are more likely to engage in pirate attacks, because the alternative of fishing is less beneficial.

Although economics typically treats criminal behavioral as a rational decision within an individual’s control, this approach may be more appropriate for some situations than others. Consider the cost-benefit analysis for tax evasion. How much money do I get to keep if I do not pay my taxes (benefit)? What are my chances of getting caught if I do not pay my taxes (part of the costs)? How much will the fine be if I do get caught (part of the costs)? This analysis assumes individuals have enough information to assess costs and benefits. Do you know the fine for tax evasion or the chances of being caught?

The cost-benefit analysis model may also be problematic when considering crimes of passion, such as unpremeditated murder; it’s possible to consider this in an economic model where the emotional benefits of crimes of passion are extremely high, but psychological evidence doesn’t always align with this conceptualization. While not a perfect model, cost-benefit analysis can help us think about ways to disincentivize crimes by reducing benefits or increasing costs.

---

12. Corman and Mocan 2005; Edmark 2005
13. Fluckiger and Ludwig 2015
14. Pomeranz 2015
15. Hollander-Blumoff 2012
Cost-benefit analysis leads to important conclusions about the underlying causes of crime, such as poverty, limited educational opportunities, unemployment, and social exclusion. These underlying social problems tend to reduce individuals’ opportunity costs of crime, by reducing the benefits of alternatives to crime. For example, if human trafficking pays $200 a day, a salesperson earning $58 a day is more likely to choose to engage in human trafficking than an economist earning $420 a day. Limited labor market opportunities, for instance low wages or high rates of unemployment, therefore can cause crime.\textsuperscript{16} Low levels of education\textsuperscript{17} can also contribute to crime, due to the role that human capital (education) plays in the opportunity costs, particularly the labor market alternatives, that individuals face.

Box 4.1: Why do people commit acts of terror?\textsuperscript{18}

Economic analysis of crime can also be extended to understanding why people or groups may commit acts of terror. Economist Sarah Brockhoff and her colleagues have examined the link between domestic acts of terror and education. One perspective on the education and domestic terrorism relationship is that education reduces domestic terrorism by increasing economic and social opportunities and improving critical thinking (thus reducing the influence of propaganda). An alternative perspective is that education increases acts of terror by facilitating them, for instance facilitating the ability to organize an attack. The research demonstrates that the relationship between education and acts of terror depends on economic and political conditions. When individuals can translate their education into opportunities, they are less likely to commit acts of terror. However, when individuals cannot translate their education into opportunities, they are more likely to commit acts of terror.

\textsuperscript{16} Witte and Witt 2000  
\textsuperscript{17} Lochner and Moretti 2004  
\textsuperscript{18} Brockhoff, Krieger, and Meierrieks 2014
Reducing Crime

A collective problem

Recognizing crime as a major social problem and with an understanding of the underlying causes of crime, how should societies fight crime? Before we can even answer this, we must answer an important underlying question: who should fight crime? Criminal justice systems are typically both funded and operated by the government—local, state, and national levels of police, courts, and prisons. Often, a private, free market is the optimal configuration for goods and services, where a government just sets the rules of the game. However, in some cases, typically referred to as market failures, the free market does not achieve the social optimum.

The private market will not achieve the optimal level of crime—or crime prevention—because crime prevention is a public good. Public goods can be contrasted with private goods. Private goods have two key features—that they are (1) excludable, which means that other individuals can be prevented from consuming the goods unless they pay for them and (2) rival, which means that the consumption of the good by one individual prevents the consumption of that good by someone else. One example of a private good is a cup of tea. The tea shop can prevent me or anyone else from getting that cup of tea unless we pay; tea is excludable. When I drink my cup of tea, this means that there’s none left for other people; tea is rival.

In contrast, public goods are non-excludable, meaning individuals cannot be prevented from enjoying the goods, and non-rival, meaning that one individual’s use does not prevent the use of another. National defense is a form of crime prevention (preventing invasion) that is non-excludable and non-rival. It is non-excludable because you cannot keep an individual within the borders of the country from enjoying the benefits, even if they do not pay for the service. It is non-rival because one individuals’ protection does not preclude the protection of others.

The nature of public goods leads to the free rider problem—that individuals can benefit from goods or services without paying for them; they can get a free ride. When such goods are privately provided, less than the socially optimal amount of the good will result. Consider your local street. It’s difficult to exclude individuals from a street; barriers could be set up so only those who pay enter (like a toll road), but this is costly, difficult, and slows down traffic. Unless traffic is terrible, one individual using the road does not prevent another from doing so. If you were asked to donate to pay for your local street, you might be tempted to chip in less than your fair share, based on your use, because other users will cover most of the costs—you can get away with still using the street without paying.

Due to their non-excludable and non-rival nature, public goods are, as their name suggests, typically paid for by the public, through taxes, and often publicly provided. National defense and other forms of crime prevention and criminal justice are public goods. This does not mean the private sector has no role in their provision. For national defense, the government typically contracts with private firms. Around $700 billion is spent in the United States on national defense each year. In 2010, $219 billion of that was spent in the form of defense contracting with private firms (not including the cost of civilian employees, military bases, or funding...
for wars in Iraq and Afghanistan). Likewise, although the courts and police are typically public institutions, the corrections (prison) system is a mix of public agencies and private firms. Arguments in favor of private prisons focus on cost-savings generated by the private sector; however, the incentives of private prisons are likely to be problematic. Private prisons may lower the quality of their services (for instance violating human rights or shirking on rehabilitation) to minimize costs and thus maximize profits.

Costs of crime

When thinking about the costs of crime to individuals or society, there are three types of costs that must be considered: direct costs, indirect costs, and intangible costs. Whether something is a cost or a benefit can be confusing to consider; what is a cost to one party can be a benefit to another. While a criminal may benefit from a robbery, this is a cost to society. Direct costs are those that have a clear price. When an individual tags (vandalizes with spray paint) a grocery store, the direct costs to the business (and society) are the cost of a fresh coat of paint and the wages of the painter. The indirect costs are losses of economic value that do not have a clear price tag. In the case of tagging a grocery store, this might be a reduction in the house prices in the neighborhood, as a result of the perception that crime is a problem in the neighborhood. Intangible costs are the costs to victims of crime and society in terms of fear, pain, and suffering. Putting a price on indirect and especially intangible costs is challenging and requires substantial economic research.

As of 2019, in the United States there were 367 violent crimes per 100,000 people (a decrease of 1.0% from 2018) and 2,110 property crimes per 100,000 people (a decrease of 4.5% from 2017). The cost of crime was estimated to be as high as 3.9 trillion dollars per year as of 2012. Homicides (murders) are the most costly offenses, but drug crime and fraud are also a large share of the costs of crime. Globally, the cost of crime is high. For example, organized crime (mafia) in southern Italy is estimated to have an impact of a 16% reduction in GDP per capita.

Cost-benefit analysis of reducing crime

Given the high costs of crime, and our understanding of the underlying causes of crime, the question now becomes how to reduce crime. Assessing the best approach requires applying a cost-benefit analysis. We need

19. Thorpe 2014
20. Witte and Witt 2001
21. McCollister, French, and Fang 2010; Wickramasekera et al. 2015
22. U.S. Department of Justice--Federal Bureau of Investigation 2020
23. Wickramasekera et al. 2015
24. Pinotti 2015
to determine the costs of crime-reduction programs and compare those costs to the benefits of the programs. It is important to note that the benefits of these programs would be measured in the reduced costs to society as a result of reduced crime.

When estimating the benefits of a program, it’s crucial to first estimate its impact. For example, does early childhood education reduce crime? By how much? These impacts then need to be assessed for their economic value, as there are likely to be direct and indirect benefits of programs (recall the terms direct and indirect costs discussed earlier). Programs such as early childhood education may in fact have a variety of benefits, as they can not only reduce crime, but they can also increase wages.  

Programs with benefits greater than their costs should be implemented. When we can only choose one option among many—since, for instance, one person with a drug addiction can only be sent to either (1) prison (2) in-patient drug treatment or (3) out-patient drug treatment—we should select the option where the ratio of benefits to cost (the cost-benefit ratio) is highest. The following sections assess different approaches to reducing crime.

**Approaches to reducing crime**

**Legalization**

One simple approach to reducing crime is to legalize activities that are currently illegal. Clearly, this solution will only work for those crimes where the social costs of the crimes are low. The legalization of drugs, such as marijuana, is one area where cost-benefit analysis may be in favor of legalization as an alternative to prohibition of drugs. Prohibiting drugs leads to major costs not only in terms of enforcement but also in terms of violence, cartels, increased poisonings and overdoses, and increased property crime. Prohibition may have benefits if it reduces drug consumption, since drug consumption has negative health and social effects. Alternatives to prohibition include treating drugs the same way as alcohol or cigarettes (with taxation and regulation), or harm reduction (subsidizing treatment and providing, for instance, clean needles or substitute drugs).  

In the U.S., a number of states have recently legalized marijuana. While it is early to assess the full economic consequences of these decisions, there have been at most small increases in marijuana use, coupled with substantial increases in tax revenue.

---

25. Temple and Reynolds 2007  
27. Dills, Goffard, and Miron 2016
Prevention

It might seem ideal to prevent all crimes. However, the marginal costs of crime prevention increase as there are fewer and fewer crimes. Therefore, the goal of crime prevention should be only to reduce crimes to the point where the benefit equals the cost of additional crime prevention. Effective crime prevention programs tend to raise the costs of crime either through deterrence or increasing the opportunity costs of crime.28

Shifting the opportunity costs of crime occurs throughout the life course. Early childhood interventions that improve cognitive and non-cognitive skills tend to improve labor force outcomes and reduce crime. Programs targeted at children at high risk for poor early outcomes led to a 33%-40% decrease in arrests by age 19.29 Such early childhood programs have a benefit-cost ratio in the range of 4-10 (between four and ten dollars of benefits for each dollar of costs).

Similarly, lead was removed from gasoline in the late 1970s and 1980s, reducing childhood lead exposure. A reduction in crime followed, starting in the 1990s. The social value of the crime reductions greatly exceeded the cost of removing lead from gasoline.30 When youth are taught to consider costs and benefits more carefully, and their behaviors are changed to allow for conscious deliberation, it is possible to reduce crime.31 When tested among economically disadvantaged youth in Chicago, the benefit-cost ratios of such training ranged from 5 up to more than 30. Although difficult to implement as a specific policy intervention, improving education and economic conditions can also achieve this end and prevent crime.32 One specific policy intervention, however, may be de-concentrating poverty. When families were randomly given the opportunity to move from high-poverty to lower-poverty neighborhoods, it reduced violent crimes among teens.33

Deterrence approaches seek to prevent crime by raising the costs or the probability of getting caught. Something as simple as more light can reduce crime.34 Body cameras for police officers can reduce police violence, with $5 of benefits for every $1 of costs.35 In the case of taxes, increasing the paper trail for tax enforcement has been shown to deter firms from committing tax fraud.36 However, other corporate crime deterrence strategies tend to show only weak effects.37 Increases in the probability of being arrested when

28. Corman and Mocan 2005; Mocan and Rees 2005
29. Temple and Reynolds 2007
30. Reyes 2007
32. Corman and Mocan 2005; Lochner and Moretti 2004; Mocan and Rees 2005
34. Doleac and Sanders 2015
35. Williams et al. 2021
36. Pomeranz 2015
37. Simpson et al. 2014
committing a crime have been linked to decreased crime in New York City.\textsuperscript{38} Sentencing laws can create deterrents; when gun crime sentences lengthened, gun crime saw a 5\% reduction.\textsuperscript{39} Deterrence, however, has mixed effectiveness. Scared Straight, a program designed to deter criminal behavior in youths, actually \textit{increases} crime. The program involves visits to prisons by children identified as at risk for future crime. Rigorous evaluations of the program have demonstrated that those exposed to the program were \textit{more}, not less, likely to subsequently commit crimes.\textsuperscript{40}

What role do police tactics have in reducing crime? In general, highly targeted efforts to reduce crime are more effective than generalized increases in policing resources, such as an additional police officer.\textsuperscript{41} Additional funds devoted to community-oriented policing, which involves building ties and trust with communities, can increase police legitimacy and reduce crime.\textsuperscript{42} Intensive foot patrols in crime hotspots can reduce violent crime.\textsuperscript{43} On the community side of policing, neighborhood watch programs can reduce crime, along with improvements in street lighting.\textsuperscript{44} Closed Circuit Televisions (CCTV) can reduce crime but are most effective in targeted locations.\textsuperscript{45} Body cameras on police officers themselves can reduce police misconduct.\textsuperscript{46} While there are some effective strategies for policing, there have not been enough cost-benefit analyses to allow us to identify which ought to be priorities.

\textbf{Incarceration}

When crimes have been committed, should criminals be sent to prison? If so, for how long? One of the motivations for sentencing criminals to prison is deterrence, as discussed previously. An additional motivation for incarceration is to prevent a person from committing another crime, which is also called \textit{recidivism}. Whether or how long individuals should be in prison is debatable. For juveniles, incarceration is by far the worst option; compared to administering no consequences, incarceration actually increases future crime. Juvenile incarceration is often ineffective. It reduces the chances of returning to school, which then reduces human capital and therefore the opportunity cost of crime. Diversion programs, including community service, are a more effective alternative than either juvenile incarceration or doing nothing.\textsuperscript{47}

\begin{itemize}
\item \textsuperscript{38} Corman and Mocan 2005
\item \textsuperscript{39} Abrams 2012
\item \textsuperscript{40} Petrosino et al. 2013
\item \textsuperscript{41} Farrington and Welsh 2005; Sherman and Eck 2002
\item \textsuperscript{42} Mazerolle et al. 2013; Zhao, Scheider, and Thurman 2002
\item \textsuperscript{43} Braga, Papachristos, and Hureau 2007
\item \textsuperscript{44} Bennett, Holloway, and Farrington 2008; Welsh and Farrington 2008a
\item \textsuperscript{45} Welsh and Farrington 2008b
\item \textsuperscript{46} Ariel, Farrar, and Sutherland 2015
\item \textsuperscript{47} Aizer and Doyle 2015; Petrosino and Turpin-Petrosino, Carolyn Guckenburg 2010
\end{itemize}
As for adults in prison, evidence demonstrates that continued incarceration does reduce crime, with the benefits approximately equal to the cost (on average). This suggests that early release is not a universally optimal policy. Neither, however, is incarceration. In fact, there is no universally optimal policy. Each type of offender requires a different approach. For example, incarceration is an effective strategy for high-risk offenders, but not for less prolific offenders or for drug offenses. Not prosecuting non-violent misdemeanors actually reduces subsequent crime. In addition to ineffectiveness in terms of crime reduction, incarceration as the default response to all offenders is costly. Incarceration costs $35,347 per year for federal prisoners as of 2019 in the U.S. Other strategies are more cost-effective, such as diversion to drug treatment. Alternatively, early release combined with the threat of a longer sentence for reoffending can reduce the risk of recidivism. Intensive monitoring, however, for example by probation officers, is not cost-effective and may actually increase recidivism. Reducing the intensity of supervision is thus cost-effective.

Rehabilitation

Rehabilitation is an important aspect of preventing crimes. Measures toward rehabilitation can be taken either in place of incarceration, while a person is in jail, or after they’ve been released. Drug courts, which incorporate drug treatment and monitoring into the criminal justice process, reduce recidivism, more so for adults than juveniles. Rehabilitation generally has positive effects, and is more effective than punitive measures in reducing recidivism. Rehabilitation can include elements of therapy or counseling. It can also take the form of educational and vocational programs that seek to shift a person’s options, and thus cost-benefit analysis, away from reoffending.

Providing economic opportunities to those at risk of recidivism can shift the cost-benefit calculation away from reoffending. In Liberia, unemployed ex-fighters engaged in illegal activities are at high risk for returning to violence or crime, such as engaging in mercenary work. A recent economics experiment provided a combination of capital, the promise of future cash, agricultural training, and counseling for these high-risk men. Training alone did not have a significant impact, but current capital and future cash led to improved

48. Levitt 1996
49. McDougall et al. 2008
50. Agan, Doleac, and Harvey 2021
51. Department of Justice Bureau of Prisons 2021
52. Drago, Galbiati, and Vertova 2009
53. Doleac 2018
55. Lipsey and Cullen 2007; Lotti 2022
56. Lipsey and Cullen 2007
legal employment activities and, therefore, less illegal work and mercenary activity. The effectiveness of these economic opportunity programs, however, depends on the context in which they’re implemented. In the U.S., for example, ex-offender employment programs have not had a significant effect on recidivism.

Restorative justice

Restorative justice programs bring together offenders, victims, and communities to determine how the offender can repair the injuries they’ve wrought. While such programs cause only modest reductions in reoffending, they are cost-effective, with a benefit-cost ratio of eight. Such programs also compare favorably to traditional criminal justice prosecution. Although restorative justice is an option only in cases when appropriate and ethical, it is particularly appealing as a lower-cost option to normal criminal justice processes and leads to higher victim satisfaction.

Summary and conclusions

Crimes are defined by the local legal codes. While laws may embody ideals of justice and morality, they can also embody the biases of voters and policymakers and even contribute to social problems. Crimes have the potential to impose enormous direct, indirect, and intangible costs on society. To try to reduce those costs, countries use a variety of different approaches within their criminal justice systems. Criminal justice systems are primarily public, because crime prevention is a public good. Since individuals typically decide to commit crimes when the benefits outweigh the costs, policies to reduce crime often target raising the costs of crime, including the opportunity costs. This may take the form of improving labor market alternatives.

The cost-benefit analysis framework can be applied to evaluating approaches to crime reduction, including legalization, prevention, incarceration, rehabilitation, or restorative justice. Which option is best depends on the relative costs and benefits, as well as the type of crime. While incarceration may be effective for high-risk offenders, preventing crime through better early childhood experiences is more cost-effective. For juveniles, incarceration is the worst option. Doing nothing is an even better option, although diversion is best. Drug treatment and other forms of rehabilitation, as well as restorative justice, also tend to be cost-effective. Ultimately, countries are likely to choose a mix of these policies in response to the unique range of criminal justice challenges they face.

57. Blattman and Annan 2015
58. Visher, Winterfield, and Coggeshall 2005
59. Strang et al. 2013
List of terms

- Crime
- Criminal justice system
- Retribution
- Incapacitation
- Deterrence
- Rehabilitation
- Restoration
- Cost-benefit analysis
- Market failures
- Public good
- Private good
- Excludable
- Rival
- Free rider problem
- Direct costs
- Indirect costs
- Intangible costs
- Cost-benefit ratio
- Recidivism

References


AVMA. 2006. “Maryland Starts Poultry Registration Program.” *JAVMAnews*.


Petrosino, Anthony, Carolyn Turpin-Petrosino, Meghan E. Hollis-Peel, and Julia G. Lavenberg. 2013. “Scared Straight and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency.” Campbell Systematic Reviews No. 5.


What is poverty?

In the United States, poverty is measured relative to a federal poverty line (FPL) and depends on a family’s size. The basis of the poverty line is spending on food, updated for inflation, and multiplied by three. This basis is because, in 1955, when the data that determined the first poverty line were collected, food was one third of a family’s income. Box 5.1 discusses this historical basis of the poverty line. Table 5.1 shows the poverty lines in the U.S. as of 2022. A family of four, in 2022, would be considered “in poverty” if their income was below $27,750. Globally, similar poverty line construction is used in countries such as Egypt, which calculates a variety of poverty lines, also based off of food expenditure.²

<table>
<thead>
<tr>
<th>Persons in family/household</th>
<th>Poverty line</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$13,590</td>
</tr>
<tr>
<td>2</td>
<td>$18,310</td>
</tr>
<tr>
<td>3</td>
<td>$23,030</td>
</tr>
<tr>
<td>4</td>
<td>$27,750</td>
</tr>
<tr>
<td>5</td>
<td>$32,470</td>
</tr>
<tr>
<td>6</td>
<td>$37,190</td>
</tr>
<tr>
<td>7</td>
<td>$41,910</td>
</tr>
<tr>
<td>8</td>
<td>$46,630</td>
</tr>
</tbody>
</table>

For families/households with more than 8 persons, add $4,720 for each additional person.

Table 5.1 U.S. poverty lines by family/household size, 2022

The FPL is used to determine eligibility for a number of different anti-poverty programs in the U.S. There are a variety of critiques in using the FPL either to determine program eligibility or measure wellbeing. Recent estimates indicate that food is one-eighth, not one-third, of budgets, suggesting FPL provides a substantial underestimate of poverty. Some families may have additional expenses—such as child care needs—that are not accounted for in measuring poverty. Local variations in cost of living are not taken into account. The measure does not take into account taxes, which reduce disposable income. The measure also does not take into account non-cash public benefits, such as public housing, that may improve well-being substantially without
changing (cash) income. Although alternative measures have been proposed, none have yet been implemented in the U.S.  

**Box 5.1: Mollie Orshansky and the History of Poverty Lines in the U.S.**

Mollie Orshansky developed the modern system of measuring poverty for the U.S. At the time she developed the poverty lines, Orshansky was working for the Social Security Administration as a social science research analyst. She had training in economics and statistics and studied families’ budgets and spending. Orshansky calculated in 1955 that a family spent one third of their money on food. She then used the amount of money needed for an economy food plan as the basis for the poverty line, multiplying that amount by three to cover non-food expenses. Her initial work on poverty lines was published in a *Social Security Bulletin* in 1963. When President Lyndon Johnson declared a war on poverty in 1964, Orshansky played a key role in designing thresholds that depended on family size and spending a third of the budget on food. Her poverty thresholds were made the official statistical definition of poverty in the U.S. in 1969.

The progress the U.S. has made in fighting poverty over the past 50 years has been mixed. Figure 5.1 shows poverty rates from 1966, at the start of the “War on Poverty,” until 2020. Notably, the poverty rate for seniors (those 65+) has declined from a high of 29.5% in 1967 to just 9.0% in 2020. However, child poverty has fluctuated and is, as of 2020, higher than it was in 1970. Adult poverty has also increased slightly since 1970. Notably, the “Great Recession,” the economic downturn the U.S. (and the global economy) experienced from 2007-2009 led to increases in poverty among children and adults, but not among seniors. A similar pattern of rising poverty for children and adults but not seniors occurred in 2020 compared to 2019, due to the COVID-19 pandemic. The majority of seniors receive most of their income from Social Security, which is not impacted by economic downturns. Social Security also played a key role in the decline in poverty among seniors in the late 1960s.

3. Blank and Greenberg 2008; Tanner 1996
4. Fisher 2008; Orshansky 1965
5. U.S. Census Bureau 2022
Causes and models of poverty

On an economy-wide level, we know from Chapter 1 that a variety of factors, including the nation’s stock of capital, the quality of the labor force, the population, and technology play a critical role in a country’s level of development. When considering individuals or households, there is some overlap in the factors that cause poverty. Human capital and physical capital ownership (wealth) play a critical role in families’ poverty in the long run. There are also intergenerational aspects of poverty that are important to consider, as poverty tends to be transmitted across generations. Combatting poverty in the long run requires improving individuals’ and families’ levels of physical and human capital. In the short run, we can think about the causes of poverty as insufficient current resources. This section discusses models for understanding the causes of poverty in the short run and the long run.
The budget constraint: A short run model

In the short run, we can model poverty using the idea of a budget constraint. The budget constraint shows the tradeoffs that a household faces, given its current resources. As with the production possibilities curve, the budget constraint simplifies the household’s spending choices into a model with two goods, such as housing and food, as in Figure 5.2. The units are the number of goods, determined given the price of different goods and the family’s income. Figure 5.2 is the budget constraint for a four-person household, with two parents and two children. One parent is working full time, forty hours a week, bringing home pay equivalent to the federal minimum wage of $7.25 per hour (we will ignore any taxes or subsidies for now). Over the course of a month, the family’s income is $1,232.50, and over the year, the family’s income is $14,790. They are in poverty (far below the FPL). Each month, the family spends only on meals, which cost $2.43 per meal (based on USDA estimates for a family of four with two young children and moderate costs), and housing, renting an apartment. The cost for each square foot of apartment space is $2 (this cost is on the low end of major metropolitan rents).

The (monthly) budget constraint shows the tradeoffs they face. They can, at point A, eat no meals, and have a 616 square foot (typical one bedroom) apartment. At point B, each person can eat one meal each day for the month, which with four people total, is 120 meals over the course of a 30-day month (this is equivalent to spending $292 on food, leaving $940 for housing, which pays for 470 square feet, a studio apartment). At point C, each person can eat two meals a day, but then the family can only afford 325 sq. ft. of housing—a very small studio apartment. At point D, the family can eat 3 meals each day, but now can only afford 179 sq. ft. of housing—basically the size of a bedroom and a bathroom. At point E, the family has no housing, but can eat 4.2 meals per day. The slope of the budget constraint shows the tradeoffs the family faces. Given that a meal is $2.43 and a square foot of housing is $2, for each meal the family adds, they have to give up 1.2 square feet of housing. Likewise for each square foot of housing the family adds, they have to give up 0.8 meals. No matter what, this family cannot meet their basic needs; they cannot afford three meals a day and even a very small studio apartment at the same time, much less any transportation, health care, or other expenses.

8. USDA 2017a
How will this family decide what to do? What trade-offs will they make? To understand the idea of how individuals and families think about tradeoffs, economists use the idea of an indifference curve. An **indifference curve** shows the combinations of goods that make individuals the same level of “happy,” referred to as utility. So, for instance, the family might be equally happy with only one meal per day but a big apartment (point B) as with 3 meals per day but only renting one bedroom (point D). This combination is shown as indifference curve 1 on Figure 5.3. All the combinations that make a family equally happy are shown on the same indifference curve. Indifference curves that include more of both goods and services (shifted outward) indicate more happiness. People also tend to prefer a mix of goods and services, so point C would be preferred to both B and D (indifference curve 2). People will choose the point that offers them the highest happiness within their budget constraint. In Figure 5.3, this is point C, where indifference curve 2 is just tangent to (touching) the budget constraint.
Production functions: A long run model

In the long run, it may be possible to shift a family’s budget constraint so they earn more. In order to do this, they would need additional sources of income. Human capital is one potential source of income; because higher-quality human capital (for instance, economic analysis skills) increases productivity, workers will be paid a higher wage (remember that labor demand is determined by workers’ productivity). This increase in productivity is usually modeled with a **production function**. For instance, a worker may be producing economic analyses, and one input into the production function is their human capital. Raising that human capital increases production (and thus the worker’s productivity). The higher wage will increase (shift out) the worker’s budget constraint. This shift is shown in Figure 5.4, where the family now has their worker earning $12/hour rather than $7.25. Now it is possible for the family to, for example, eat 3 meals a day and have 583 square feet of housing (a studio apartment).
An additional source of income is physical capital, or, when we think about individuals owning physical capital, wealth. Like human capital, physical capital can act as an input into the production function. For instance, a company may buy new sewing machines for their garment factory. If the new sewing machines are faster, this will increase production, since sewing machines are an important element of the garment production function. Individuals may directly own companies, own shares (stocks) in the companies, or own financial capital that is loaned to companies. When they own companies, the physical capital pays off in the long run in higher profits. When individuals own financial capital (stocks), they receive a financial return, either the payout of a dividend on stocks, or an increase in the cash value of their stocks. When people provide the capital for loans by saving in a bank, they receive a return on their savings (the interest rate they receive). Physical capital can produce a stream of income, which can either be re-invested to generate more, future wealth, in the long run, or can shift out families’ budget constraints. If families who do not have any wealth want to start accruing wealth, they would have to shift their budget constraint inward. They would have to consume less now in order to have a stream of additional income, from their savings becoming wealth and generating a return, in the future.

Differences in wealth have an important role in not only current disparities but also perpetuating those disparities across generations. Figure 5.5 shows net worth (assets minus liabilities, basically what you own minus what you owe) by race and ethnicity. The figure also shows one of the most common forms of wealth, homeownership. Black, non-Hispanic and Hispanic families have much lower median net worth than white families.

Figure 5.4. Increase in the budget constraint

10. Federal Reserve 2021
families. Median net worth is $36,050 for Hispanic families and $24,100 for Black, non-Hispanic families compared to $189,100 for White, non-Hispanic families. Differences in homeownership are an important component of wealth disparities, with a 48% rate of homeownership among Hispanic families, 45% among Black, non-Hispanic families, and 74% among White, non-Hispanic families. Historical policies such as redlining (not offering loans in minority communities) as well as ongoing disparities in bank practices contribute to current disparities.\textsuperscript{11}

<table>
<thead>
<tr>
<th>Net worth</th>
<th>Homeownership</th>
</tr>
</thead>
<tbody>
<tr>
<td>White, non-Hispanic</td>
<td>$189,100</td>
</tr>
<tr>
<td>Black, non-Hispanic</td>
<td>$24,100</td>
</tr>
<tr>
<td>Hispanic</td>
<td>$36,050</td>
</tr>
<tr>
<td>Other</td>
<td>$74,500</td>
</tr>
</tbody>
</table>

Figure 5.5 Family net worth (median) and homeownership rates (percentage) by race/ethnicity, 2019

Not only do these disparities in wealth affect income in the present, they also have an intergenerational component, affecting families’ ability to invest in their children. For instance, the low-income family in Figure 5.3 may struggle to feed their children. This food insecurity will reduce the children’s human capital (health and education). Such disparities can be long lasting and affect multiple generations. One of the ways in which slavery in the U.S. continues to affect Black families today is disparities in human capital.\textsuperscript{12} Children’s human capital, income, and wealth, are closely linked to that of their parents and even grandparents.\textsuperscript{13}

\begin{itemize}
\item \textsuperscript{11} Sullivan et al. 2015
\item \textsuperscript{12} Bertocchi 2016
\item \textsuperscript{13} Sacerdote 2005
\end{itemize}
Programs to combat poverty

Programs to combat poverty work to tackle the short run and long run underlying causes of poverty. As well as working on whether families have adequate resources now, programs often aim to shift the underlying human or physical capital of families so that they can move out of poverty. This section looks at a variety of different approaches to combat poverty, including income supports, which expand the budget constraint in the present, as well as investments that can shift budgets in the long run, by increasing human capital or physical capital.

When examining programs to combat poverty in the short run, we are typically thinking about different forms of income supports—to shift families’ budgets outward and allow them to meet more of their basic needs. Income supports come in a variety of forms, including (1) in-kind aid, such as food, (2) conditional cash transfers (cash transfers where recipients have to undertake certain activities or behaviors, such as attending job training or taking children to the doctor) and (3) unconditional cash transfers (receiving cash without any conditions beyond being below a certain income). These different approaches have different impacts on families’ budget constraints as well as their long run human and physical capital.

In-kind aid

Food assistance is a typical type of in-kind aid, targeted to a specific good (food). In the U.S., two important programs that provide assistance with food are the Supplemental Nutrition Assistance Program (SNAP), also known as food stamps, and the Women, Infants, and Children (WIC) program.

Box 5.2: Eligibility and Benefits for SNAP and WIC

Families are typically eligible for SNAP if they (1) are below 130% of the FPL, and (2) have less than $2,500 in resources (bank account balance, cars (in some states), etc.). People have to meet work requirements (working or in a training program) to qualify for SNAP, with some exceptions (children, seniors, pregnant women). Households can receive a maximum allotment that goes down as their income goes up. For each additional dollar of income, a household’s SNAP benefits are reduced by 30 cents. A household of four has a maximum monthly allotment of $835. SNAP

14. USDA 2017b; 2016; 2021
benefits are issued on an Electronic Benefit Transfer (EBT) card—like a bank debit card. They can be used for eligible food purchases at participating stores.

WIC provides benefits to women who are pregnant and during the year after a child is born, as well as for infants and children up to five years old. States set income eligibility between 100% and 185% of FPL. Individuals must also be determined to be at nutrition risk. The benefits provided by WIC are vouchers, checks, or EBT to purchase specific nutritious foods such as cereal, fruit/vegetable juice, eggs, milk, cheese, peanut butter, beans/peas, and canned fish. For women with infants who do not breastfeed, WIC provides infant formula.

In-kind food programs shift the budget constraint for only one good. A family with $1,132 in monthly income would receive $309 each month in SNAP benefits. This can only be used for food. We can show this in the budget constraint in Figure 5.6, where the family can buy 127 meals per month with SNAP. This creates the flat section of the budget constraint, where they spend all their cash income on housing (for 566 square feet) and then can choose how much of SNAP to use, up to the maximum SNAP benefit of 127 meals. Past 127 meals, they are facing a similar tradeoff to before; they have to reduce square footage to afford additional meals. However, SNAP still allows for 127 more meals; instead of up to 466 meals per month the family can now afford up to 593.

What will the family choose? The answer depends on their preferences (indifference curves). Figure 5.6 shows two possibilities—and an important critique of in-kind income supports. The family in panel (a) likes a mix of both goods; they buy more of both housing and meals with SNAP benefits. They are not restricted by SNAP being only usable for meals because they want to buy more than 127 anyways. This shift in both goods is an example of how benefits are often fungible; for the family in panel (a) SNAP has the same effect as cash would. The family in panel (b) strongly prefers housing. They are restricted by SNAP being only for food; they buy exactly 127 meals with SNAP, but they would actually be happier (on a higher indifference curve) with cash instead. They would buy fewer meals and more housing and be happier. In-kind benefits restrict choice, which will make some families less happy than cash. They also tend to be more administratively burdensome than cash. However, if the goal of the program is to alleviate a particular problem, such as food insecurity, or if there are additional benefits to food consumption (such as accumulation of human capital for children), this may justify the program design.
An additional critique of SNAP and similar income support programs is that they may disincentivize work. If families can meet their basic needs with public benefits, they may choose not to work and instead receive benefits, which is costly to society. SNAP alone would not meet families’ basic needs. However, it may create some work disincentives in how the benefit is determined; for each additional dollar of income families receive 30 cents less in benefits, meaning the net effect of working another hour or getting a raise is only 70% of what it would be without SNAP. We will investigate these sorts of labor disincentives in the next section.

### Income support through taxation

In the U.S. and worldwide, income support is often provided in part through taxation. In this section we will modify our budget constraint model to understand the impact of different tax policies on families’ budgets, and also on their labor supply. In Figure 5.7 we shift our budget constraint to model income and leisure rather than food and housing. It’s important to keep in mind that the opposite of leisure is labor. Assuming that folks have to sleep at least 64 hours, the most someone can have work (or leisure) in a week is 80 hours. If they have 80 hours of leisure, they will always have income of zero. For each additional hour that they work, they will have one less hour of leisure, but will have their wage in income (which makes them happy because they will spend that income on their preferred mix of goods and services). The slope of the budget constraint is determined by the wage rate; if the wage is $10, each additional hour of work increases income by $10. Although when the budget constraint increased this usually represented additional amounts of both goods and services, when the...
wage rate increases, this pivots the budget constraint, because it is only when individuals work another hour that they earn more. Mapping indifference curves, we end up with a backward bending labor supply curve for a worker.

![Budget Constraint](image)

**Figure 5.7. Budget constraint for income and leisure and labor supply**

Tax policies can shift the budget constraint. The U.S. and many other countries have progressive tax systems, where individuals with higher incomes pay a higher marginal tax rate. Regressive taxes are taxes where those with lower income pay (relatively) more. The marginal tax rate means that, when an individual moves into a higher tax bracket, he or she does not pay any more taxes on the preceding income, just on additional income. There were seven 2022 tax rates: 10%, 12%, 22%, 24%, 32%, 35%, and 37%. What the brackets encompass in terms of income depends on how someone files, for instance if they are single or married. For a single person earning $15,000 a year, the first $9,875 would be taxed at 10% ($1,027.50) and the income from $10,275 to $15,000 would be taxed at 12% ($567), for a total tax of $1,594.50.

These varying marginal tax rates can be represented in the budget constraint. Figure 5.8 shows, for someone who earns $20 per hour, the different marginal tax rates that shape the budget constraint. As the tax rate increases, the slope of the budget constraint decreases. Although the take-home for each additional hour decreases as the tax rate increases, we know that this is going to have complex effects on labor supply depending on preferences and the shape of our backward bending labor supply curve.
One of the key anti-poverty programs in the U.S., the **Earned Income Tax Credit** (EITC) uses negative tax rates to try to reduce poverty while incentivizing work. EITC depends on the number of children in the household and filing status. It provides a tax credit (the government pays families) for families depending on their work. The tax credit increases with income up until a maximum point. The tax credit increases with income up until a maximum point. Figure 5.9 shows the tax credit phasing in for a single mother with one child who makes $12/hour; the credit increases until $10,980 of pre-tax income, when it is $3,733. The credit then remains the same amount until $20,130 of pre-tax income. The credit then phases out (goes to zero) by $43,492 of pre-tax income. This shifts out the budget constraint, particularly for the range of 20-40 hours per week, which may not only alleviate poverty but also incentivize work. However, the EITC structure creates an
effectively quite high marginal tax rate between $20,130 and $43,492, which may disincentivize work past a certain point. Empirically, EITC has been shown to increase labor supply, especially for single mothers.\textsuperscript{15}

![Figure 5.9 Budget constraint with and without EITC, 2022](image)

Women, who are disproportionately caregivers, face a particular challenge in needing child care for each additional hour they work; the need to pay for child care acts like a tax relative to their budget constraint. Child care will typically be a higher share of their income for women with the lowest incomes, acting like a regressive tax on work. Paying for child care affects both single parents and two-parent households trying to decide whether the second parent should stay at home or work. Child care and other carework costs are a global challenge to economic participation. The loss of women’s domestic labor—child care and other carework

\textsuperscript{15} Eissa and Liebman 1996; Meyer and Rosenbaum 2001
along with domestic chores—if they work outside the home plays an important role in low female labor force participation in Egypt. After deducting all their costs, women often find that work is not worthwhile.\textsuperscript{16} In both the U.S. and globally, preschool and kindergarten programs, along with child care subsidies are policies that governments worldwide implement to help alleviate this additional challenge, and also increase the chances of work for women.

**Cash aid programs**

In the United States, a central program for alleviating poverty for families with children is *Temporary Assistance for Needy Families* (TANF). This program is what individuals typically mean when they refer to “welfare,” as it provides cash payments to low-income families. The program is for families with children only. TANF, like SNAP and other programs, has both asset and income requirements. Since the program is administered by states, there are different specific policies and benefits in each state. In general, parents are expected to work. TANF provides employment counseling, including training, to aid in finding work. Child care assistance is typically provided with TANF benefits. Benefits are also capped at 60 months (5 years) total for a family. The benefits provided are cash assistance (and coupled with that, SNAP benefits) that depend on the size and income of the family. In Minnesota, the TANF program is called Minnesota Family Investment Program (MFIP) and for a single parent with two children provides up to $1,189 per month, $14,268 a year in benefits.\textsuperscript{17} This benefit is not enough to lift families out of poverty; the poverty line for a family of three in 2022 was $23,030.\textsuperscript{18}

One common concern with programs such as TANF is that they disincentivize work. In some states, beyond a certain (very low) threshold, any additional income reduces the size of the TANF grant. Other states phase-out the benefits, with families keeping 50 cents of each dollar of additional income up to a certain point. Figure 5.10 shows the budget constraint with TANF where additional income reduces the size of the grant, compared to the current tax structure. In panel (a) the parent prefers to work 50 hours a week at a wage of $9.50 an hour with or without TANF being in place. In panel (b) there is a work disincentive effect; the parent chooses to work 10 hours per week without TANF and no hours with TANF. The parent in (b) is better off (happier, on a higher indifference curve), but no longer working.

---

\textsuperscript{16} Hoodfar 1997

\textsuperscript{17} Minnesota Department of Human Services 2022

\textsuperscript{18} U.S. Department of Health & Human Services--Office of the Assistant Secretary for Planning and Evaluation 2022
While TANF and welfare have the potential to disincentivize work, there are also work requirements built into the program. These work requirements were part of welfare reform undertaken by President Bill Clinton in 1996. While caseloads (participation) declined after reform, measures of independence, avoiding poverty, and avoiding hardship show less success. The employment effects of reforms are empirically small (for instance, a two percentage point increase in employment by one estimate), but women, especially the poorest women, are financially worse off. Substantial debates remain around the future direction of welfare reform, particularly given that TANF is not effective for families with substantial work barriers, such as temporary disability or domestic abuse.

19. Danziger et al. 2016a
20. Cancian and Meyer 2004
22. Danziger et al. 2016b
Box 5.3: Universal Basic Income

One way to alleviate poverty, which is currently being hotly debated around the globe, is to provide a universal basic income (UBI). This is essentially an unconditional guarantee, with no work or any other requirements, of a certain level of income. If the UBI is set above the poverty line, such a policy could instantly alleviate poverty. However, as we have discussed for other anti-poverty programs, UBI could have substantial additional economic effects, including on labor supply. Switzerland recently held a referendum on UBI, as did Finland, The Netherlands, India, and Kenya are all running experiments to study the impact of UBI. One experiment that was undertaken in the 1970s in Manitoba suggests that, while UBI did reduce women’s labor supply, it also improved health. Additional, more recent experiments will be helpful in understanding the impact of UBI in different contexts.

Human Capital Investments

In the long run, poverty alleviation requires shifting families’ human or physical capital to increase their productivity and income. This section discusses some of the evidence on human capital investments as a poverty alleviation strategy. One common global policy that combines elements of cash aid and human capital investment is a conditional cash transfer (CCT). CCTs use means tests, assessing a families’ assets to identify families in poverty or otherwise at risk for poor human capital accumulation. Means tests are used because cash income is often a difficult measure to implement in developing countries. CCTs provide cash aid conditional on families’ behavior; in developing countries where health, malnutrition, and school drop outs are common problems, CCTs often condition on families’ taking children for medical visits, nutrition evaluations, or school enrollment. The conditionality is in part because families might not make these investments otherwise, because such investments may not pay off to the parents, but rather to their children, or because there are market failures preventing them from making such investments. CCTs improve education and nutrition outcomes.

23. BBC News 2016; England 2017; Forget 2011; Hanna and Olken 2018; Lowrey 2017; Prescott, Swidinsky, and Wilton 1986
usually (but not always) to a greater extent than unconditional programs.\textsuperscript{24} CCTs also tend to provide transfers to women, because they usually (but not always) will invest more in the next generation.\textsuperscript{25}

A variety of poverty alleviation programs target the human capital of children at risk of poverty directly, instead of through conditions. Nurse visitation programs, which provide prenatal and infancy home visits to promote maternal and child health and development, improve various aspects of both parents’ and children’s development, including reducing the use of other poverty alleviation programs.\textsuperscript{26} High quality early childhood programs improve education and earnings outcomes, decreasing use of welfare and other social assistance programs.\textsuperscript{27} School meal (breakfast and lunch) programs have both short-term nutrition effects and long-run educational effects.\textsuperscript{28} All of these are effective programs for improving human capital early in life.

For adults, alleviating poverty through human capital interventions is more challenging, in part because much of human capital is shaped during the early years of life. Training programs, which can confer additional human capital, tend to be ineffective in the short term, but may have medium term impacts.\textsuperscript{29} Currently there are a large number of job training programs in the U.S. targeting the unemployed and low-income workers. For training programs to be effective, they need to confer the skills in demand in the labor market, which can be challenging for public programs and requires substantial information about the skills employers seek.\textsuperscript{30}

### Physical Capital: Wealth and Neighborhoods

In order for families to acquire wealth, and from that a stream of income, they must be able to save. Programs to increase saving have particularly targeted low-income countries, where microcredit programs provide small loans, typically to start a business, and often targeted to women who face particular barriers to finance. Microsavings programs, including group savings, may also help low-income families accumulate capital. The balance of evidence is that microfinance programs have at most modest positive effects on households.\textsuperscript{31} Microsavings programs, with savings products designed or subsidized for low-income families, do have large savings and wealth impacts.\textsuperscript{32}

Neighborhoods can play a role in poverty; whether a family lives in an area with a high concentration of

\begin{thebibliography}{10}
\bibitem{24} Baird, McIntosh, and Ozler 2011; Benhassine et al. 2015
\bibitem{25} Benhassine et al. 2015; Duflo 2003; Gitter and Barham 2008
\bibitem{26} Eckenrode et al. 2010; Olds 2007; Olds et al. 2004
\bibitem{27} Temple and Reynolds 2007
\bibitem{28} Bhattacharya, Currie, and Haider 2006; Hinrichs 2010
\bibitem{29} Card, Kluve, and Weber 2010
\bibitem{30} Almeida, Behrman, and Robalino 2012
\bibitem{31} Angelucci, Karlan, and Zinman 2015; Augsburg et al. 2015; Banerjee et al. 2015; Crépon et al. 2015; Tarozzi, Desai, and Johnson 2015
\bibitem{32} Karlan, Ratan, and Zinman 2014
\end{thebibliography}
poverty can affect their outcomes, in part through the services they can access and in part due to their social networks (social capital). Moving families to areas with less concentrated poverty has been suggested as one intervention that can help reduce poverty. Moving families can improve outcomes in some contexts, with positive impacts observed in the U.S.\textsuperscript{33} A program that randomly relocated families in India from slums to improved housing on the edge of the city did not have positive long-term outcomes, in part due to the fact that it disconnected families from their social networks.\textsuperscript{34}

**Conclusions**

What can we do to alleviate poverty for families around the world? In the short run, we can shift out families’ budget constraints by providing them with more resources. In-kind aid, such as food stamps, can improve their well-being, but often has the same effect as cash. Income supports that provide cash have potential work disincentive effects, but programs such as the EITC can actually incentivize work through tax credits. In either case, with both TANF (welfare) and EITC, the labor supply effects of programs tend to be relatively small.

In the long-run, increasing human capital, through investments in children’s health or adults’ skills, plays an important role in increasing individuals’ productivity, ultimately alleviating poverty. Likewise, increases in wealth (physical capital), through savings or loans, as well as geographic mobility can play an important role in long-run poverty alleviation. Although there has been substantial progress in reducing global poverty, in the U.S. poverty has been a persistent problem for adults and children for decades. The elderly have, however, through Social Security, experienced major decreases in poverty since the 1960s. Identifying programs, such as basic incomes, that can achieve similar results for children and families is an important agenda for economic research in the near future.

**List of terms**

- Federal poverty line (FPL)
- Budget constraint
- Production function

\textsuperscript{33} Kling, Ludwig, and Katz 2013; Ludwig, Duncan, and Hirschfield 2001

\textsuperscript{34} Barnhardt, Field, and Pande 2017
• In-kind aid
• Conditional cash transfers
• Unconditional cash transfers
• Supplemental Nutrition Assistance Program (SNAP)
• Women, Infants, and Children Program (WIC)
• Indifference curve
• Progressive taxes
• Regressive taxes
• Earned Income Tax Credit (EITC)
• Temporary Assistance for Needy Families (TANF)
• Microcredit
• Microsavings

References


Fiszbein, Ariel, Norbert Schady, Francisco H. G. Ferreira, Margaret Grosh, Nial Kelleher, Pedro Olinto,


Minnesota Department of Human Services. 2022. “Minnesota Family Investment Program (MFIP).”


6. THE ECONOMICS OF DISCRIMINATION

Caroline Krafft

What is discrimination?

**Discrimination** is the unjust or unequal treatment of an individual or group based on a specific characteristic, such as their race, age, or gender identity. In the United States, a number of laws forbid discrimination on the basis of age, disability, national origin, pregnancy, race/color, religion, or sex. Globally, the United Nations (UN) has passed conventions on eliminating all forms of racial discrimination and discrimination against women.

Although the definition seems straightforward, identifying when an individual or entity is discriminating in practice is quite challenging. This difficulty is because **disparities** (differences in outcomes) may be the result of current or past discrimination. The fact that women earn 81 cents for every dollar men earn may reflect employers’ discrimination in setting wages, but may also reflect the fact that women choose different majors, or are more likely to take time out of the workforce to care for children. Of course, that women choose different majors may also reflect discrimination in human capital accumulation.

Likewise, the fact that Black men have a one in three lifetime likelihood of imprisonment, while white men have a one in seventeen chance may be due to a variety of factors, such as historical housing discrimination and poor local labor market opportunities, as well as discrimination in the criminal justice system. Identifying the source of disparities—for instance in the case of imprisonment, whether disparities are due to unequal and discriminatory outcomes around education, employment, or poverty as factors in committing crimes, or in unequal chances of arrests, convictions, or sentences—is critical to addressing and reducing these disparities.

---

Causes of discrimination

Discriminatory “tastes”

Economists have two main theories concerning the causes of discrimination. The first theory is that individuals have “tastes” or preferences for discrimination. This taste-driven discrimination theory suggests that factors such as social and physical distance and relative socioeconomic status contribute to tastes for discrimination. Contact with a minority group and the size of a “minority” group matter as well (the minority in this case could actually be a majority that has historically been disempowered, e.g. women). Tastes for discrimination mean that individuals are effectively willing to forfeit income to avoid certain transactions or interactions. For instance, landlords may prefer to rent only to individuals of a certain race or religion, even though they could charge higher rents if they opened up to a broader market.

Statistical discrimination

The second theory is statistical discrimination, which assumes discrimination is essentially an information problem. For instance, in the labor market, employers may have imperfect information about the productivity of individual workers. Consider the case of an employer hiring a new carpenter for a construction company. The employer has information from applicants’ resumes on their education, training and past work experience. She can even administer a test to prospective employees to measure their skills, perhaps building a stair rail. The resume information and the test are, however, imperfect signals of the employee’s productivity. This uncertainty and imperfect information cause the employer to take into account another factor: she believes that women are, on average, less productive carpenters than men. This potentially erroneous statistic, combined with the inability of the resume and skills test to fully signal productivity, will lead the employer to conclude that a man is likely to be more productive than an equally qualified woman. The job is then offered to the man instead of the woman.

A “taste” for discrimination and statistical discrimination are often framed as competing theories. However, as we will see in discussing the empirical studies on discrimination below, there is substantial evidence for each theory. One way of reconciling the theories may be to think of taste-driven discrimination as a potential source of assumptions about individuals’ productivity in the face of incomplete information. Additionally, some individuals may change their assumptions in the face of additional evidence, a case which supports the existence of statistical discrimination. Others with more deeply ingrained prejudices would not reevaluate their

assumptions, which lends credence to taste-driven discrimination. Different assumptions regarding individuals and groups may be more or less swayed by information.

**Box 6.1: Economists in Action: Lisa Cook Studies Competition and Discrimination**

Dr. Lisa D. Cook is a Professor of Economics and International Relations, currently serving on the Board of Governors for the Federal Reserve. She researches economic growth and development, along with financial institutions, innovation, and economic history. She was a Senior Economist for the Council of Economic Advisors, serving in the White House, and was elected to the board of the American Economic Association (AEA). She directed the AEA Summer Training Program, which increases diversity in the field of economics by preparing undergraduate students for graduate degrees in economics.

One important area of Dr. Cook’s research is creating and analyzing data on discrimination, including gathering data on Jim Crow era firms that were friendly towards African Americans, as well as the creation of a national lynching database. In one of her papers Dr. Cook examines the determinants of firms’ discrimination towards potential consumers during the Jim Crow era and prior to the Civil Rights Act. She shows that firm owners segregated and discriminated against African Americans based on white consumers’ discrimination, a case of taste-based discrimination. Reductions in the number of white consumers reduced discrimination and activism among African Americans also helped.

**Labor market models of discrimination and its consequences**

Regardless of whether discrimination is taste-driven or caused by statistical discrimination, it can essentially be modeled the same way. The approaches to reducing discrimination will be quite different, but the models for the impact of discrimination will be nearly identical. Consider discrimination for the case of the labor market. Recall that employers’ demand for labor is based on productivity. We are now going to name that productivity the **marginal revenue product** of workers, how much revenue they create for their employers through their

---

7. Cook et al., 2020; Cook, 2011; Cook, 2019; Board of Governors of the Federal Reserve, 2022.
work. Operating under statistical discrimination, productivity is assumed to be lower for certain groups.\(^8\) In the case of gender discrimination, employers may assume the marginal revenue product of women is lower because they disproportionately undertake caregiving responsibilities (a case of statistical discrimination). Alternatively, with taste-based discrimination, hiring a less-preferred group imposes a “cost” on employers, effectively modeled as a decrease in the marginal revenue product.

Figure 1 shows discrimination in the labor market for men and women. To simplify, we assume the same labor supply for men and women. However, the demand, which is equal to the marginal revenue product (MRP) is assumed to be lower for women than for men (discrimination). As a result, the equilibrium outcome is that fewer women are employed and women are earning lower wages than men. Although we focus on models of the labor market here, similar ideas apply to other markets. Housing is another example. A landlord might discriminate in supplying housing to individuals with a different religion than his own. This would shift the supply of housing to differentiate between religious groups, with a higher cost (reduced supply) for those of a different religion.

In all cases, there is a substantial challenge when it comes to proving discrimination. It is difficult to measure the MRP of different workers. This then makes it difficult to distinguish between cases where individuals actually are differentially productive on average, such as workers with more training or experience, and when discrimination is occurring.

\(^{8}\) Hellerstein, Neumark, and Troske, 2002.
Evidence on Discrimination

There are a variety of forms of discrimination and different groups that are discriminated against globally. This section presents just some of the evidence on discrimination, primarily from the U.S., but from other global contexts as well. Economists rely on a host of different techniques to gather evidence on discrimination. One is multiple regression, also called multivariate regression, a statistical technique where economists try to account for differences in observable characteristics. When checking for wage discrimination, for example, these characteristics would include occupation and education. The remainder of the differences in outcomes would then be attributed to discrimination.

Conducting experiments is another method to assess discrimination. Economists can randomize resumes with different characteristics to apply to jobs or randomize the economic equivalent of “mystery shoppers”
with different characteristics to apply for housing. These experiments are typically referred to as **audit studies**. Experiments are the most effective for being certain about cause and effect, but can be challenging to implement and much more expensive than analyzing existing data with multiple regression. This section presents evidence from both multiple regression models and audit studies on discrimination in education, housing, the labor market, and the criminal justice system.

### In education

Discrimination in the education system leads to disparate human capital outcomes that also contribute to labor market disparities. Teachers play a key role in education and their discriminatory attitudes can affect students in a variety of ways. For instance, one experiment demonstrated that teachers gave worse grades and lower secondary school recommendations when assignments (essays) had minority (Turkish) names. Teachers also have lower expectations and negative attitudes that affect their behavior towards minority students, which may in turn affect those students’ performance. Gender bias may be particularly important for Science, Technology, Engineering, and Math (STEM) fields. Science faculty presented with otherwise identical student resumes bearing either a female or male name rated women as less competent than men. Faculty were less likely to hire women, offered a lower salary, and were less likely to mentor women.

### In housing

Housing was one of the areas where discrimination in the United States was first measured effectively. Fair housing audits were developed by housing organizations to identify racial discrimination in housing opportunities after the passage of the Civil Rights Act. For example, in assessing Black-white housing disparities, an audit will send two auditors to a housing agent, one white and one Black, for a random sample of advertised housing units. When individuals receive differential treatment, specifically in different offers of housing, and this treatment depends on their race, the audit indicates discrimination. Historically, as of 1981, Black housing seekers were told about 30% fewer available housing units than whites.

More recent studies have taken advantage of the power of the internet; an experiment in the U.S. rental apartment market varied first names, using those commonly associated with whites and African Americans.

---

14. Ibid.
In some cases, it also included information about credit history and smoking. African-American sounding names had a 9.3 percentage point lower positive response rate than applicants with white-sounding names, indicating discrimination. The additional information on credit history and smoking did differentially affect the gap in response rates, indicating that information and statistical discrimination contributed to disparities.\textsuperscript{15} In India, a study using India’s largest real estate website showed that, while an upper-caste Hindu had a 35% chance of a response to a housing application, this was only 22% for a Muslim applicant.\textsuperscript{16} An experiment in Sweden varied distinctive ethnic and gender names in applying for rental housing. Arabic/Muslim names received fewer responses than the Swedish male names, and Swedish female names had an easier time accessing housing than Swedish male names.\textsuperscript{17} In addition to long-term rentals, these disparities extend to short-term rentals, such as Airbnb vacation rentals. Applications from guests with African-American names were 16% less likely to be accepted relative to otherwise identical guests with distinctively white names.\textsuperscript{18} Discrimination also occurs against Airbnb ethnic minority hosts.\textsuperscript{19}

\section*{In the labor market}

Discrimination in the labor market manifests in substantial hiring disparities by race, ethnicity, gender, and disability status. A study sending fake resumes to help-wanted ads in Boston and Chicago found that white names received 50% more callbacks for interviews than African-American names.\textsuperscript{20} A similar study in New York City found that Black applicants were half as likely to receive a callback or job offer than white applicants.\textsuperscript{21} In interviews for waitstaff jobs in Philadelphia, job applications from women had a 40 percentage point lower chance of receiving a job offer from high-price (and high earning) restaurants than men, in part embodying customer discrimination.\textsuperscript{22} An experiment that randomized disclosure of disability status found disability halved the chances of a callback.\textsuperscript{23}

In Toronto, a study demonstrated that individuals with foreign experience or with Indian, Pakistani, Chinese, and Greek names were less likely to be hired than those with English names.\textsuperscript{24} In Germany, which has a substantial number of Muslim migrants, especially from Turkey, it is common for applicants to send photos

\begin{thebibliography}{24}
\bibitem{15} Ewens, Tomlin, and Wang, 2014.
\bibitem{16} Datta and Pathania, 2016.
\bibitem{17} Ahmed and Hammarstedt, 2008.
\bibitem{18} Edelman, Luca, and Dan, 2017.
\bibitem{19} Laouénan and Rathelot, 2022.
\bibitem{20} Bertrand and Mullainathan, 2004.
\bibitem{21} Pager, Western, and Bonikowsi, 2009.
\bibitem{22} Neumark, Bank, and Van Nort, 1996.
\bibitem{23} Bellemare et al., 2018.
\bibitem{24} Oreopoulos, 2011.
\end{thebibliography}
with resumes. A study of female applicants that randomized German names, Turkish names, and whether the migrant was wearing a headscarf found significant discrimination against Turkish names and more so against those wearing a headscarf. This discrimination is so pronounced that a female applicant who wears a headscarf and who has a Turkish name would have to send 4.5 times as many applications to receive the same number of callbacks as a woman with a German name and no headscarf.\textsuperscript{25}

In the United States, women’s pay is, as of 2020, 82% of men’s pay.\textsuperscript{26} Notably, women and men are approaching convergence in their pay at the start of their careers. Figure 6.2\textsuperscript{27} shows median weekly earnings by sex, as well as the ratio of women’s wages to men’s. Early on in the life course, women’s wages are 95% (ages 16-24) that of men’s. However, pay diverges over the lifespan, with a major expansion in the gender gap. At ages 25-34, women’s pay is 89% of men’s pay. By ages 35-44 women’s pay is only 81% of men’s, dropping to 78% at ages 45-64. Two key drivers for the gap expansion are differences in career interruptions and differences in weekly hours—both largely associated with motherhood.\textsuperscript{28} In contrast, when men become parents, they tend to receive a premium, an increase in pay, rather than a penalty.\textsuperscript{29} As of 2010, differences in human capital contributed little to the gender wage gap. However, differences in occupations were still important, as women tended to be in traditionally female occupations that are generally lower paying, such as nursing and teaching.\textsuperscript{30}

\textsuperscript{25}. Weichselbaumer, 2019.
\textsuperscript{27}. Ibid.
\textsuperscript{28}. Bertrand, Goldin, and Katz, 2010; Bailey et al., 2019.
\textsuperscript{29}. Lundberg and Rose, 2002.
\textsuperscript{30}. Blau and Kahn, 2016.
Gender pay gaps can be compounded by racial disparities. Figure 6.3\textsuperscript{31} shows median weekly earnings among workers as a percentage of white men’s earnings. People of color tend to earn less than whites, with disparities further exacerbated by the gender pay gap. For instance, Black men earn 75\% of what white men earn, while Hispanic men earn 72\% of what white men earn. Asian men earn more than white men, at 130\% and Asian women earn 103\% of what white men earn. In contrast, other groups of women earn less on average. Black women earn 69\% of white men and Hispanic women 64\%. Relative to white women, who earn 82\% of what white men earn, Asian women are better off but still at a disadvantage when compared to the relatively higher earnings of Asian men.

In studying pay gaps by race, what are referred to as pre-market factors, such as human capital, explain an important share of pay gaps. However, an important share of gaps are also discrimination in the labor market—estimated to be at least one-third of the Black-white wage gap.\textsuperscript{32} Discrimination feeding into pay gaps can occur in complex ways. For example, when Black job-seekers attempt to negotiate for a higher salary, they are penalized in terms of their salary outcomes.\textsuperscript{33} Likewise, women tend to be perceived more negatively than men when they try to negotiate, in part due to gender stereotypes around being “nice.”\textsuperscript{34}

\textbf{In the criminal justice system}

Discrimination is a challenge throughout the criminal justice system and contributes to the large disparities in incarceration by race and gender that were discussed in the crime chapter. Racial disparities in drug arrests are

\textsuperscript{32} Fryer Jr., Pager, and Spenkuch, 2013.
\textsuperscript{33} Hernandez et al., 2019.
\textsuperscript{34} Bowles, Babcock, and Lai, 2007.
not due to differential drug or nondrug offending, nor residing in areas with a police focus on drug offenses; there is strong evidence of discrimination and disparities in police practices driving disparities.\textsuperscript{35} Likewise, studies using the differential ability to tell driver race in the daytime versus the nighttime have demonstrated racial bias in traffic stops in some localities, but not others.\textsuperscript{36} Once arrested, individuals may be discriminated against in terms of the process from pre-trial processing (for instance, setting bail) through setting their sentences.\textsuperscript{37} Offenders who are Black, male, less educated, and lower income receive longer sentences.\textsuperscript{38}

**Policies to reduce discrimination**

**Competition**

The idea of taste-based discrimination has, historically, been linked with the idea that competition may play a key role in reducing discrimination. Consider a case where all workers are equally productive, but some employers have discriminatory tastes. It would follow that the non-discriminating employers would be able to make a greater profit by hiring individuals who tend to be discriminated against but are equally productive. This idea would suggest that the solution to discrimination in any market is simply competition. However, empirical evidence suggests that, while competitive markets deter discrimination, firms that have market power exist and do discriminate.\textsuperscript{39} Simply “waiting out” discrimination will not be effective. Other interventions are required.

**Changing the available information**

An important set of interventions to reduce discrimination focus on changing the available information about individuals. Interventions can remove markers of protected categories, such as gender and race, from the set of available information to reduce discrimination. For example, when symphony orchestras adopted blind auditions—where the candidate plays music behind a screen and is not visible to the hiring committee—this approach led to gender equity in hiring, increasing the proportion of women in symphony orchestras.\textsuperscript{40} However, policies to remove all potentially revealing information are challenging to design, and employers may

\textsuperscript{35} Mitchell and Caudy, 2015; Welty et al., 2016.
\textsuperscript{36} Ritter and Bael, 2009; Antonovics and Knight, 2009.
\textsuperscript{37} Schlesinger, 2007; Starr and Rehavi, 2013; Bielen, Marneffe, and Mocan, 2018.
\textsuperscript{38} Mustard, 2001; Cook et al., 2020.
\textsuperscript{39} Hellerstein, Neumark, and Troske, 2002.
\textsuperscript{40} Goldin and Rouse, 2000.
be resistant to their implementation. For instance, orchestras have to lay down carpet, to muffle the sounds of heeled shoes that are associated with women, or ask women to take off their shoes.

Removing names from the available information may reduce discrimination in a variety of areas. This approach can be particularly effective for reducing discrimination in models like Uber and Lyft or Airbnb where such information could be readily removed without interrupting transactions. Other approaches to removing potential markers of protected categories include anonymizing resumes and using skills-based tests (like the orchestra auditions) for other jobs as well. A number of European countries have experimented with anonymizing applications. Doing so can reduce disparities and equalize the probability of receiving an interview. However, the process still allows for discrimination in hiring after the interview and precludes affirmative actions for otherwise equivalent applicants. In France, anonymous resumes ultimately led to a lower probability of interviewing and hiring minority candidates.

Depending on the nature of discrimination, there may be cases where removing information could be potentially harmful and adding information may be more helpful. One of the studies that identified discrimination in Airbnb determined that discrimination against African-American names disappeared when there was a positive public review. Essentially, positive information about individuals helped reduce discrimination. Having to report gender-disaggregated information about pay has been shown to reduce the gender pay gap. However, having individuals disclose their past salaries when applying to new jobs can perpetuate discrimination, as new employers will use those as a basis for salary offers. This problem has led some states to ban asking applicants about their salary history.

Box 6.2: The Lilly Ledbetter Fair Pay Act

Lilly Ledbetter was an employee of Goodyear from 1979 until 1998. Initially Ledbetter was paid the

41. Ge et al., 2016.
45. Cui, Li, and Zhang, 2016.
46. Bennedsen et al., 2018.
same as the men in the same position. By 1997, Ledbetter was paid $3,727 per month. Male managers were paid between $4,286 and $5,236 per month. In part because Goodyear kept pay information confidential (as is common practice), Ledbetter did not find out about the pay disparity until long after the disparity had occurred. When she sued, under Title VII of the Civil Rights Act, a 5-4 decision in the case before the Supreme Court determined that she had not filed within the statute of limitations—the legal time frame for filing after discrimination occurs. The case treated the discrimination as the decision about her salary by her supervisor, some time ago, not the ongoing disparate paychecks, because the paychecks themselves did not have discriminatory intent, which is required under Title VII. The problem that faced Lilly Ledbetter, that she learned about discrimination long after it occurred and the statute of limitations expired, led to the Lilly Ledbetter Fair Pay Act. The Act, passed in 2009, broadened the definition of discriminatory practice to include, for instance, each disparate paycheck. The case and subsequent act illustrate some of the challenges in identifying and remedying discriminatory practices.

The “Ban the box” campaign is an example of an information removal effort that appears to have achieved the opposite of its goal. We learned in the chapter on crime that ex-offender rehabilitation depends in part on employment opportunities and holding a legitimate job. Yet employers tend to discriminate against those with a criminal record. Employers commonly ask about past criminal convictions on initial job applications. “Ban the box” campaigns forbid asking at the initial job application stage but allow for the question in interviews and with conditional job offers. The goal was that ex-offenders would have better job opportunities. An additional goal was to reduce racial disparities and discrimination in employment, given racial disparities in the criminal justice system. Although well intentioned, “ban the box” laws appear to be counterproductive in reducing discrimination. Employers, without information on criminal history, operate under statistical discrimination and are less likely to interview young, low-skilled Black and Hispanic men.

**Affirmative action**

**Affirmative action** is a “set of procedures designed to eliminate unlawful discrimination between applicants, remedy the results of such prior discrimination, and prevent such discrimination in the future. Applicants may

---

be seeking admission to an educational program or looking for professional employment.” Affirmative action in the United States came about as a 1961 executive order by President John F. Kennedy, with a requirement mandating affirmative action among government contractors. Affirmative action subsequently expanded to other areas, such as education.

Those in favor of affirmative action argue that it equalizes opportunities, benefits qualified women and minorities, and that it is beneficial to society as a whole. Proponents also suggest that affirmative action improves equity and either improves efficiency or has at most minor reductions in efficiency through the reallocation of jobs. Opponents suggest that there are efficiency losses and that the policy itself is inherently racist.

Historically, affirmative action has helped promote the employment of minorities and women. The magnitude of the effects is generally fairly small, although they can cause substantial relative shifts for minority groups. While minorities who benefit on the labor market may have poorer credentials, they have equal performance, suggesting that efficiency concerns have relatively little merit. Further, white males face costs, but they are relatively small. Affirmative action has also increased the probability that under-represented minority groups graduate from selective institutions. However, affirmative action or some approaches to affirmative action have been banned in making university admissions decisions.

Reducing bias in individuals

Individuals’ biases, for example their gender biases, are key drivers of discrimination. Legal changes can potentially change individuals’ attitudes and behavior. For example, the passage of same-sex marriage reforms in U.S. states reduced individuals’ discrimination against sexual minorities. This reduced discrimination in turn contributed to improvements in labor market outcomes for same-sex couples.

Individuals may not be aware of their biases, in which case they are referred to as implicit biases. Training can also reduce implicit biases, particularly those that may be caused by lack of exposure or familiarity with other races. Treating implicit bias like a habit that can be combated through awareness, concern about its effects, and the use of strategies to reduce bias is helpful, particularly for people who are concerned about

57. E.g. Moss-Racusin et al., October 9, 2012.
59. Lebrecht et al., 2009.
discrimination in the first place. 60 This suggests that training to reduce bias in individuals requires some commitment on their part to change their thinking and behaviors, and therefore is likely to work better for some individuals and biases than others.

Professionalizing human resources functions may also help reduce bias in the hiring process. Research in Canada demonstrated that employers discriminated against those with Asian-sounding names. Asian applicants had a 20% disadvantage for large employers but double the disadvantage, 40%, for small employers. Larger organizations may devote more resources to recruitment, have professional human resource strategies, and also have more experience with diverse staff. 61 This professionalism may reduce (although not necessarily eliminate) discrimination.

It may even be possible to reduce the role of biased human decision making in areas such as sentencing. Risk assessments are a potential, but controversial, approach to reducing bias in sentencing, parole, and rehabilitation. 62 Risk assessment instruments model the probability of reoffending based on a number of factors, including criminal history. In part because of different criminal histories, the policy can have disparate impact across racial groups. For example, Black offenders receive higher risk assessments, on average, than white offenders. 63 Especially with disparities in the criminal justice system, such instruments may perpetuate disparities. However, improvements in computing, such as machine learning algorithms, have the potential to reduce jail populations and crime rates, including reducing the percentage of minorities in jail. 64 Yet machine learning and artificial intelligence can also pick up and replicate existing biases. 65

Conclusions

Discrimination occurs in education, employment, housing, and the criminal justice system, as well as many other dimensions of individuals’ lives. Economists tend to understand discrimination through one of two models—based on prejudicial “tastes” for discrimination or based on incomplete information leading to statistical discrimination. Both theories of discrimination show how discrimination contributes to disparate outcomes, such as different wages and employment rates for men and women. Although discrimination is pervasive, the good news is that progress is being made, and (some) disparities have decreased over time as a result of effective policies. Designing effective policies is, however, extremely challenging, as the efforts to “ban the box” illustrate. The challenges of designing effective policies underline an important role for economists

60. Devine et al., 2012.
64. Kleinberg et al., 2017.
and statisticians in the fight against discrimination: carefully evaluating the impact of different policy and program attempts to reduce discrimination.

### List of terms

- Discrimination
- Disparities
- Taste-driven discrimination
- Statistical discrimination
- Marginal revenue product
- Multiple regression
- Audit studies
- Affirmative action
- Implicit biases

### References


Global importance of education

One of the great successes of recent human history has been the expansion of education. Throughout the globe, we have seen enormous progress in terms of human knowledge and skills. These increases in human capital, which in turn increase productivity, have played a critical role in poverty reduction and human progress. However, education does not just affect how productive workers are and their own earnings. Education contributes to economic growth and poverty reduction. Education has an important role in reducing inequality, particularly across generations. Education can help reduce crime, increase political participation, improve health and child health, as well as reduce child marriage, childbearing, and mortality. Education is also instrumental to individuals’ (and especially women’s) empowerment.

Education’s vast potential impact has an important role in how education markets function. Globally, the public sector has a central role in providing education. In this chapter, we will explore why leaving education to the private sector alone will not be economically efficient. We will also examine how the public sector performs in comparison to the private sector in providing education, including issues of quality, efficiency and equity in schooling. The chapter frames debates about education within current global development goals.

Global progress in expanding education

Historically, very few individuals could read or write. Literacy is the ability to read or write. Illiteracy is when

2. Assaad and Saleh, 2018; Celhay and Gallegos, 2015.
5. Glewwe, 1999; Chou et al., 2012; Currie and Moretti, 2003.
someone cannot read or write. Individuals living in the 1800s were more likely to be illiterate than literate (Figure 7.1). In 1820, just 12% of the world’s population was literate. Illiteracy was due, in large part, to limited access to schooling. In 1820, just 17% of people had at least some formal basic education (had been to school). As Figure 1 shows, there has been an enormous shift in schooling and literacy. By 2015, 86% of the world’s population had been to school and as of 2016, 86% of the world’s population could read and write. Some countries made enormous improvements in the span of a single generation. For example, Algeria’s elderly (ages 65) have a literacy rate of 28% compared to 97% among youth (ages 15-24).  

Improvements in literacy and attending school have been uneven across regions. Figure 7.2 shows gross enrollment ratio by region. Gross enrollment ratios (GERs) compare the number of students enrolled at a school in a level (such as elementary, also known as primary schooling) to the number of students who

---


should be in that level. Because the number of students that should be in a level assumes on time progression and no grade repetition, the GER can be greater than 100% if students are delayed in their schooling. Figure 7.2\(^{10}\) shows GERs for pre-primary (kindergarten), primary (elementary education), secondary (middle (lower secondary) and high school (upper secondary)), as well as tertiary (higher) education for four regions of the world that encompass primarily developing countries. No region has universal pre-primary, although Latin America & the Caribbean and East Asia & the Pacific come close. Sub-Saharan Africa and the Middle East and North Africa lag far behind, with GERs for pre-primary of 31-33% in 2018.

At the primary level, regions have converged towards GERs at or above 100%, although Sub-Saharan Africa still falls a bit short of 100% as of 2018. Secondary education GERs have expanded substantially in every region except Sub-Saharan Africa, where there has been less of an increase with a secondary GER of only 43% as of 2018. Although Sub-Saharan Africa has the lowest enrollments in secondary, other regions still do not have universal secondary access. Recently, particularly since the 1990s, tertiary education access has expanded rapidly in the developing world, with between 41-51% GERs in Latin America & the Caribbean, East Asia & the Pacific, and the Middle East and North Africa in 2018. Sub-Saharan Africa’s tertiary GER is only 9% at the same time.

\(^{10}\) World Bank, 2020.
Figure 7.2. Gross enrollment ratios by level and region (percentage)
Reductions in inequality

Expanding access to education has also reduced some forms of inequality. Figure 7.3\textsuperscript{11} shows the ratio of years of schooling for women to men, as a percentage, for adults ages 15-64. Although in the 1800s, no country had gender equity, by 2010 Latin America and the Caribbean, Eastern Europe, and the Advanced Economies had all achieved gender equity in years of schooling. Other regions had made major progress as well, with women having, on average, between 82% and 87% the amount of schooling of men.

![Figure 7.3. Gender ratio (years of schooling for women/men) as a percentage, by region, 1870-2010. Notes: For adults ages 15-64.](image)

Problems with education quality

There has been an enormous expansion in access to education throughout the world, with the potential to dramatically increase human capital and improve a wide variety of economic and social outcomes. The payoffs

\textsuperscript{11} Lee and Lee, 2016.
to education depend, crucially, on how much students learn in school. In this regard, there has been less progress than enrollment trends suggest. Figure 7.4\(^{12}\) shows the percentage of students in grade two who are unable to read even one word in a short text. The share unable to read is as high as 90% in Malawi, 85% in India, and 83% in Ghana. Only in one country—Jordan—of the countries examined is the share unable to read below 25% (11% are unable to read in Jordan). The failure to learn to read—even one word—after two years of school is a symptom of a global “learning crisis.”\(^{13}\) Enrollment—being in school—is not the same as learning. Learning, acquiring human capital, is what contributes to economic and social outcomes, so the learning crisis means that globally education is falling short of its potential.

![Figure 7.4. Percentage of grade 2 students not able to read a single word of a short text. Notes: Among countries with available Early Grade Reading Assessment data.](image)

The shortfall in education’s potential is an issue not only in developing countries but developed ones as well. Figure 7.5\(^{14}\) shows countries’ average scores on the Trends in International Mathematics and Science Study (TIMSS). The TIMSS is an international standardized test. Scores range from 0 to 1,000. The tests also have

13. Ibid.
benchmark scores for different levels of achievement, from low (400) to advanced (625). There are a number of countries, such as Saudi Arabia and Botswana, whose average scores do not even meet the low benchmark. The United States, with an average score of 518, is just two points above Slovenia, and passes the intermediate but not the high benchmark. Only a few countries, including Japan (586), Hong Kong (594), Taipei (599), South Korea (606), and Singapore (621) surpass the high benchmark. The TIMSS results show that there is substantial room for improvement in the quality of education in both developed and developing countries. How to improve quality is a thorny question we will examine in subsequent sections.
Global development goals for education

The progress in expanding enrollments, coupled with the deficit in learning, has been reflected in global development goals for education. In 1990, the international community committed to Education for All, a principle subsequently updated in 2000 to expand access, reduce inequality, and improve quality in
education. The first set of global development goals were the Millennium Development Goals (MDGs), approved in 2000, which set targets for 2015. One of the goals of the MDGs was universal primary education (UPE), ensuring every child completes primary. These goals often framed education as a human right—but primarily focused on access to education, not its quality. The MDGs were replaced in 2015 by the Sustainable Development Goals (SDGs), goals to achieve by 2030. The SDGs shifted the focus from enrollment to learning. For example, one of the indicators for the SDGs is the proportion of children in grade 2 or 3 who have achieved minimum proficiency in reading—one of the deficits shown in Figure 7.4.

How do families and societies make decisions about education?

To unlock the potential of education, we need to understand how education decisions are made, before we can think about policies that might change the current landscape. We will first examine the way families make decisions about education. Then we will examine why society might want to have different levels of education than families would choose on their own.

Supply and demand for education in a private market

The typical model for the market for a good or service is supply and demand. A supply and demand model of the “market” for education will be our starting point for understanding the economics of education. Figure 7.6 shows supply and demand for education in Peru, where the currency is the Peruvian sol. We are starting with a case of just a private market for education. In this private market, families pay tuition each year in Peruvian sol. How many years of school they will buy at each tuition rate is shown by the demand curve. You can think of this as the case for an “average” or “representative” family. Although the idea of demand is the same as for other goods and services, in this model we delve a bit deeper into what demand is. Although families may get some enjoyment out of education, a major motivation for sending kids to school is the benefit of school. One of those benefits, earning higher wages as an adult, is called the return to education. Globally, the return to education is an approximately 9% increase in wages for each additional year of school.

Since demand is based on the benefits of education—whether enjoyment or wages—we can rename the demand curve to be the marginal private benefit (MPB). The benefit is “private” in the sense that the family decides how much education to demand based on their benefits—not the benefits to society. Demand, as is

17. Psacharopoulos and Patrinos, 2018.
typical, increases as the price of schooling drops. Since we know that demand is affected by a variety of factors, such as income, different families are likely to have different demand depending on their income, preferences, expectations for the future, or alternatives for their children. For example, families may have lower demand for formal schooling when alternatives such as apprenticeships are more lucrative. Demand will also depend on the quality of schooling. Students are likely to drop out earlier when education quality is poor.

We can now rename supply to be the marginal private cost (MPC). Since we are operating (for the moment) in a world with a purely private market, you can think of for-profit schools deciding how much to charge for different amounts of schooling. The MPC rises with years of schooling because it is more expensive to deliver additional education; teacher’s wages must be raised to attract more teachers, more schools must be built in remote locations, and at higher levels teachers need more qualifications and training. Thus, supply is upward sloping. Supply here is related to the usual factors, including input costs (for example, teachers and textbooks).

In the example in Figure 7.6, the education outcome that occurs is determined by equilibrium, where $D=S$ or $MPB=MPC$. In this example, equilibrium is six years of school costing 2,000 sol per year. As the next section shows, although this is the equilibrium, it is not necessarily efficient.

**Why the private market is inefficient**

There are a number of reasons the private market in education is not efficient. First, families are making decisions for their children about education. Parents pay the costs: tuition, as well as the opportunity cost of children’s time taken away from other activities, such as the family business or caring for siblings. Parents do
not receive all the benefits of education. They will not be the ones receiving additional future income, their children will. In contexts where girls are more likely to die, less likely to work, where girls earn less, or where girls leave the family or community at marriage, families will receive lower benefits for girls’ education than boys’ education. As a result, there will be lower demand for girls’ education and gender inequality.\footnote{Aslam, 2009; Jayachandran and Lleras-Muney, 2009; Jayachandran, 2015.}

Not only do benefits often accrue to children rather than their families, but also some benefits of education go to neither the children receiving the education nor their families. Benefits of education include improved health for the educated individual’s children.\footnote{Currie and Moretti, 2003; Glewwe, 1999.} Especially since women disproportionately care for the next generation, this benefit is greater for educating girls than for boys. However, it is not a private benefit. This benefit is an \textit{externality}, also called a spillover, specifically an externality in consumption. An \textit{externality in consumption} occurs when the person consuming (demanding) a good does not receive the full benefits\footnote{Or, as we will see in later cases, with pollution, the costs.} of that good. When a girl going to school improves her future children’s health, this is an externality in consumption. The benefits to society of reduced crime, increased political participation, and reduced intergenerational inequality are just a few of the externalities to education. These externalities are additional benefits on top of the direct, private benefits to families.

Figure 7.7 incorporates the externalities that occur in education into our supply and demand model. The externalities in education are benefits added to the MPB. The MPB the externality is what we call the marginal social benefit (MSB). MSB is the total benefit to society, including to the individual, his or her family, and society as a whole. For now, we are going to assume that marginal private costs and \textit{marginal social costs} (MSC) are the same. We will revisit that assumption in the next chapter, when we examine pollution. We can now define a new concept: the \textit{social optimum}, where the MSB=MSC. On Figure 7.7, this point is ten years of school at a cost of 2,285 sol. This is the point where the benefits to society are equal to the costs. The social optimum is efficient for society, because if there were more education than ten years, the costs would exceed the social benefits. If there were less than ten years of schooling, for instance the six at the equilibrium, there would be unrealized benefits that would be worth additional investment in schooling—up to ten years.
In a market without externalities, such as that for wheat, from Chapter 2, the market equilibrium is efficient and socially optimal. For education, because of the externality, the market equilibrium is inefficient. Cases where the market equilibrium is inefficient are referred to as market failures. Public goods, discussed in the crime chapter, are another example of market failures, where the private market will not provide the efficient or optimal amount of a good or service.

In Figure 7.7, the externality is shown as a constant additional 952 sol worth of benefits to society. It is, however, entirely possible that the externality varies by level of education (the MSB line may not be parallel to MPB). Private benefits may also vary by education (the MPB line may not be straight). For example, researchers
have argued the return to investing in pre-primary education is higher than for other levels. One study of the impact of a pre-primary education program on disadvantaged children in Chicago found that the benefit of investing in the preschool program was $74,981 for $7,384 of average costs, a benefit/cost ratio of $10.15. Most of the benefit was an externality, benefits to society beyond those to individuals, specifically $6.87 of the $10.15 benefit/cost ratio. Externalities can be very large—but they are difficult to measure and quantify.

How can policy help solve education market failures?

Now that we know how and why the private market fails to deliver a socially optimal level of education, we can better understand the role of the public sector in education, and in particular different potential approaches to addressing the externality. Our goal with these policies is to achieve the socially optimal level of education.

Subsidizing education

The first approach we will consider for education is one of subsidizing education. A subsidy reduces the cost of a good or service, by the government paying for part of that cost. A subsidy does not mean a good is free, just that its price is lower for consumers. The government must pay the difference between what suppliers receive and what consumers pay. Figure 7.8 shows how an education subsidy would work for education in Peru. The government must pay the amount of the externality—the gap between the social optimum (MSB) and MPB. Since the amount of the externality is 952 sol and the MPC(=MSC) is 2,285 sol per year for 10 years of school, the subsidized price will be 1,333 sol (=2,285-952). At this price, families will choose to consume ten years of school based on their demand (MPB). An education subsidy will achieve the social optimum, so long as the government knows the size of the externality and sets the subsidy accordingly.

Countries do subsidize education, but more commonly, they offer education free of charge. Figure 7.9 analyzes the economic results of free education in Peru. With our assumptions about the externality, we know that the socially optimal level of education is ten years of school. When the price of school is lowered to zero for families, they will still choose the quantity where the price equals their demand (MPB). In this example, at a price of zero, families demand 18 years of school (all the way through six years of primary, six years of secondary, four years of university, and two years of a master’s degree). Although this education is free to families, it is not free to society. Society has to pay 2,857 sol per year, for 18 years. Free education is not socially optimal in this case;
it is inefficient. For years past ten, the additional years of schooling have social benefits that are less than their costs.

If free schooling is inefficient, as Figure 7.9 suggests, why is it so common to have free education? One reason is that we do not actually know the exact, total size of all the externalities to education. It is entirely possible the externalities are so large that free education would lead to (close) to the social optimum. It may also make sense to make certain levels of education free. In this case for Peru, although society would bear the costs rather than families, free education through grade 10 (age 16) would lead to the optimal amount of education.
Society would have to fund this education through taxes, but it might be more politically feasible to have compulsory (required) free education through age 16 financed by taxes. A similar argument could be made for free education past compulsory schooling if the private benefits could be recouped through taxes on additional wages that result from education. Although directly taxing graduates of education is not common, it is another potential approach to funding education.25

Separating education financing from education provision

In the preceding discussion, we did not specify who provided education. Although there is a strong economic argument for subsidizing education, that does not (necessarily) mean that the public sector has to provide education. However, publicly provided education is the norm for many countries, including the United States. Figure 7.1026 shows public and private institutions’ spending on education as a percentage of GDP in the U.S. There was a substantial increase from the 1950s through 1970s for both public and private spending. Yet, public institutions remain dominant, at 5% of GDP, compared to 1% for private institutions.

Is public provision of education the best option? Public financing can, potentially, co-exist with private provision of education. For example, **school vouchers** are public funding that can be used in (participating) private schools to pay for education. **Charter schools** are publicly-funded but privately-run schools. They typically have more flexible regulations than standard public schools. School vouchers and charter schools are part of what is referred to as “school choice.” The arguments in favor of school choice are that more choice makes families happier, that more options foster competition that improves all schools, and that private providers are better. The second two arguments can be assessed empirically. First, although some research from the U.S. and abroad suggests that having additional school options can improve educational outcomes,
including learning, this result is hotly contested and not a consistent or consensus finding.²⁷ Policy design, such as having high standards, may be particularly important.²⁸

Second, whether private providers are better or worse, in terms of costs, quality, or equity, is uncertain.²⁹ There are major tradeoffs that occur in private provision. An experiment with school vouchers in Louisiana, which targeted low-income students in low-performing schools, allowed numerous students to enroll in private schools. Initial effects on learning were negative, although in the long run the differences were not statistically significant.³⁰ Liberia recently began an experiment with a number of private providers (and public funding) for its education system. Although learning improved, the costs were also higher than public schools, and varied substantially across different private providers.³¹ Two of my papers, examining higher education in Egypt and Jordan, showed that private higher education was no better (and no worse) in terms of either quality or labor market outcomes.³² School choice can, potentially, worsen racial segregation.³³

**Inequalities exacerbated through education financing**

In addition to debates about public versus private funding and provision of education, there are important debates around how schools are publicly funded and what that means for equality. Figure 7.11³⁴ shows how the United States funds public schools, specifically the percentage of GDP that goes towards schools from local, state, and federal sources. Historically, funding for schools was predominantly local, but since the 1970s there have been equal local and state shares in overall funding (around 2% each in 2010). Federal funding remains very low, less than 1%. Local funding in particular may contribute to inequality.³⁵ When children live in low-income communities, they will already be at a disadvantage due to community poverty (and often family poverty as well). When local property taxes fund their schools, and the tax base that votes on school matters has lower income and lower property values, schools will be underfunded. Local property taxes as the basis for school funding translate into students already at risk for poorer outcomes attending schools with

---

fewer resources and lower quality. Although state and federal policies may direct some money towards low-income schools or children, spending is not equal, much less learning. Spending equalization has the potential to narrow test score gaps that depend on family background. Box 7.1 discusses another idea, “baby bonds” targeted to low-income families.

Figure 7.11. United States source of funding for public schools as a percentage of GDP

Box 7.1. Economists in action: William Darity, Jr. and Darrick Hamilton propose “baby bonds”

37. Hamilton and Darity Jr., 2010; Darity Jr., Hamilton, and Stewart, 2015; The Ohio State University, 2020.
Economists play a key role in the field of public policy, as the proposal for “baby bonds” by William Darity Jr. and Darrick Hamilton illustrates. Dr. William Darity Jr. is the Samuel DuBois Cook Professor of Public Policy, African and African American Studies and Economics, as well as the director of the Samuel DuBois Cook Center on Social Equity at Duke University. Dr. Darrick Hamilton is an economist and executive director of the Kirwan Institute for the Study of Race and Ethnicity and a Professor at the John Glenn College of Public Affairs at Ohio State University. They received PhDs in economics “to study the problems of poverty and inequality and to develop policies to combat them,” and now work in the field of public policy, believing that “economic justice is a moral imperative.” The two are pioneers in developing the field of “stratification economics,” which focuses on the causes and remedies of intergroup disparities (for instance, differences in wealth and education by race). One of their proposals to address persistent racial disparities is to offer “baby bonds,” investments that become available to children from low-income families when they turn 18, which would allow them to afford college as well as address persistent racial wealth gaps. Former presidential candidate and Senator Cory Booker has even drafted this idea into legislation.

What works to improve education?

What works to improve education? This section examines two key issues in education: enrollment and learning. Distinct interventions work for these different challenges, however, interventions tend to be one of a few different types: resources (money, school buildings, more teachers, etc.), pedagogy (how teachers teach), incentives (how schools are managed, how teachers are evaluated and paid), or supporting students’ well-being (lunch programs, school nurses, etc.).

Enrollment and Attendance

Once a local school exists (a key prerequisite!), what helps families enroll their children in school and children attend school? Family resources are particularly important for enrollment. Especially in contexts (families or countries) where income is a major constraint on enrollment, cash transfers can be very effective for increasing enrollment.\(^\text{38}\) Transfers may be conditioned on attending education, but there are also benefits to unconditional transfers. In contexts where enrollment is low and food security is particularly a challenge,
school feeding programs (meals eaten at school or taken home from school) improve enrollment. Not every resource makes a difference in ensuring attendance. For example, policy makers had identified menstruation (a girl having her period) and lack of sanitary products as a barrier to girls’ schooling. However, a study in Nepal showed menstruation had little impact on attendance and a random experiment giving out sanitary products had no effect on attendance.  

Learning

What improves learning when students are in school? Cash transfers are not effective for improving learning (although they help with enrollment). Educational technology, such as giving a laptop to every student (“one laptop per child”) may sound promising, but is not effective. Narrower applications of education technology are more effective. Specifically, computer-adaptive technology for math, where questions are adjusted and explanations offered at the student’s current level, does improve learning. Reducing class size can improve learning, particularly in early grades, but is also costly.

When it comes to learning, pedagogy is particularly important. Teaching at the right level (addressing children where they are at, rather than adhering to the official curriculum) causes substantial gains in learning. Remedial tutoring, particularly in addition to regular school, rather than as a substitute, can help children who have fallen behind catch up. Teacher training, so long as it is ongoing and supportive, not one-off, can help improve teachers’ skills and students’ learning.

There is mixed evidence about the role of teachers’ incentives in learning. Some policy makers argue that it should be easier to hire and fire teachers, and that their pay should be based on performance. Contract teachers are less expensive and equally effective as regular civil-service teachers in India. Paying teachers based on student test scores sometimes improves test scores and other times does not. Performance pay may increase “teaching to the test.” Providing stronger local control of schools (school-based management) does not

41. Ponce and Bedi, 2010.
42. Cristia et al., 2012; Escueta et al., 2017.
45. Banerjee et al., 2016.
46. Banerjee et al., 2007.
47. Kerwin and Thornton, 2020; Popova, Evans, and Arancibia, 2016; Popova et al., 2018.
School-based management works better in middle-income countries when parents are more able to engage in the process.

Although the evidence suggests what interventions are likely to be more (or less) effective, an important part of addressing learning is assessing what the barriers are for a particular context or country. For example, if rigid adherence to a curriculum that leaves students behind is common, then teaching at the right level may be more effective. Therefore, as well as evaluating “what works” education economics can play a role in first evaluating “what’s the problem?”

Box 7.2. Economists in action: Esther Duflo and the Poverty Action Lab

Esther Duflo is a Professor of Poverty Alleviation and Development Economics at the Massachusetts Institute of Technology (MIT). She got her undergraduate degrees in History and Economics at Ecole Normale Superieure, Paris and her PhD in Economics from MIT. While working at MIT, she cofounded (and currently co-directs) the Abdul Latif Jameel Poverty Action Lab (J-PAL), which undertakes evaluations of programs to reduce poverty—including in the education sector.

Dr. Duflo has studied education topics such as the expansion of education in Indonesia, the impact of remedial education on learning in India, and the impact of cash transfers for education on enrollment in Morocco. She has been particularly prominent in increasing the role of impact evaluations—to figure out what works and why—in development economics. Her innovative work earned her the Nobel Prize in Economics, along with Abhijit Banerjee and Michael Kremer.

Summary and conclusions

Education has enormous potential to improve individuals’ well-being and countries’ development. Education benefits individuals, who earn more as they learn more. However, education also has externalities, spillovers that benefit society as a whole. Due to externalities in education, it is inefficient to leave education to private markets alone—education will be under-provided. The public sector must intervene to ensure there is enough

51. Massachusetts Institute of Technology, 2017; Banerjee et al., 2007; Benhassine et al., 2015; Duflo, 2000; Duflo, 2019.
education. Whether the public sector should provide, as well as fund, education is a hotly debated issue. The balance of evidence is that private provision will not be sufficient for improving education access, quality, or equality. A growing body of research indicates what policies are more (or less) likely to work in education, but much more research is needed to ensure education achieves its full potential.

**List of terms**

- Literacy
- Literacy
- Enrollment
- Gross enrollment ratio
- Education for All
- Millennium Development Goals (MDGs)
- Universal Primary Education (UPE)
- Sustainable Development Goals (SDGs)
- Return to education
- Marginal Private Benefit (MPB)
- Marginal Private Cost (MPC)
- Externality
- Externality in consumption
- Marginal Social Benefit (MSB)
- Marginal Social Cost (MSC)
- Social optimum
- Market failure
- Subsidy
- School voucher
- Charter school

**References**


Ashley, Laura Day, Claire Mcloughlin, Monazza Aslam, Jakob Engel, Joseph Wales, Shenila Rawal, Richard


Carr-Hill, Roy, Caine Rolleston, Tejendra Pherali, and Rebecca Schendel. “The Effects of School-Based


Krafft, Caroline. “Is School the Best Route to Skills? Returns to Vocational School and Vocational Skills


Oster, Emily, and Rebecca Thornton. “Menstruation, Sanitary Products, and School Attendance: Evidence


WHY IS THERE POLLUTION AND WHAT CAN WE DO ABOUT IT?

Caroline Krafft

Global impact of pollution

Pollution places great costs on society. Diseases caused by pollution reduce the GDP of developing countries by as much as 2% per year.¹ Pollution contributes to almost 7% of health costs in developing countries and 2% in developed countries. According to one estimate, damages from pollution cost $4.6 trillion dollars per year, around 6% of global GDP.

Pollution takes a toll in terms of human lives and disproportionately affects children, lower income countries, minorities, and other marginalized groups. In 2015 alone, nine million people died prematurely due to pollution.² In the same year, deaths due to pollution were 16% of all deaths worldwide. The role of pollution in global mortality was triple that of malaria, AIDS, and tuberculosis combined. Pollution killed 15 times more people than war and violence.

Why is there so much pollution, doing so much harm across the globe? What can we do to reduce pollution? This chapter investigates the causes of pollution through an economic lens and evaluates a variety of policy options for reducing pollution. One of the challenges in addressing pollution is that it does not respect local, state, or national boundaries. Pollution is a global problem; this chapter also discusses the global policies and institutions to address pollution.

Pollution imposes costs on society

Pollution can take a variety of forms. It can be the release of harmful substances into the air, the water, or the soil. It can occur in food, in buildings (for instance, through mold), or in the form of energy. Pollution can take the form of light, heat, or noise. Noise pollution from ambient noise, such as traffic, affects children’s ability

¹. Landrigan et al., 2017.
². Ibid.
to learn. Pollution can have negative effects on humans directly, such as when they breathe in sooty air, or indirectly, when pollutants damage our forests. Pollution imposes a cost on society, either directly or indirectly.

A helpful framework for thinking about the impacts of pollution when they are indirect or on the environment is the idea of **environmental services** (also sometimes called ecosystem services), the contributions of ecosystems to people’s wellbeing. For example, wetlands provide services by cleaning our water and preventing flooding. Trees help dissipate rain, absorb carbon dioxide and produce oxygen. Pollution damages or destroys these environmental services, depriving society of their valuable functions.

### Pollution as an externality in production

We can model pollution as a cost to society by returning to our concept of an externality. When examining education, we focused on how, when children “consumed” education, there were positive spillovers to society. This was a positive externality in consumption. The **marginal social benefit** (MSB) was greater than the **marginal private benefit** (MPB). Now we will apply a similar concept to pollution. Pollution is typically generated through the production (supply) side of the market, and imposes a negative spillover on society, so it is a **negative externality in production**.

To model a negative externality, we can consider the case of corn. Agriculture along the Mississippi River, for instance corn farming, uses intensive quantities of fertilizers that enter the water as nitrogen and phosphorous. These harmful chemicals flow through the river’s basin, eventually flowing downstream and creating a “dead zone” in the Gulf of Mexico. This harm is a negative externality from producing corn. The farmers making the decision of how much fertilizer to apply in Minnesota do not live in the Gulf of Mexico, where the fish will die due to algae blooms brought on by fertilizer. The farmers are imposing costs on others, costs that they do not consider in their decision to use fertilizer and grow corn.

To model how this externality affects the economy, Figure 8.1 shows corn production in terms of thousands of square miles of corn and the price of corn ($/bushel). The supply curve is based on the **marginal private cost** (MPC). Corn farmers decide how much corn to produce based on their private costs. They have to pay for the bags of fertilizer, the seeds, and the tractor. However, they do not have to pay for the harm caused by the bag of fertilizer to the waterways, environment, and ultimately society. This negative externality means that the **marginal social cost** (MSC) differs from the MPC. Akin to MSB, MSC=MPC externality. Society bears the private costs plus the spillovers. However, the equilibrium outcome in the corn market is determined by supply (MPC) meeting demand. In the figure, this is a price of $4/bushel and 8,000 square miles of corn being farmed.

When there is an externality, the market equilibrium is different than the social optimum, a case of **market**

---

failure. Here, there is an externality of $2/bushel in environmental harms. In this case the social optimum is where the demand (=MSB, assuming no externalities in consumption) meets the MSC. A price of $5/bushel and 6,000 square miles farmed would be socially optimal. With a negative externality in production, too much is produced (8,000 square miles), and the price is too low ($4/bushel). We can actually estimate the loss to society from producing too much corn, after taking into account the harms to society from agricultural pollution. The area in red on the figure is what is referred to as the deadweight loss to society; the amount of money we, as a society, lose when pollution occurs because of production at the equilibrium level rather than the social optimum. This loss is the area of a triangle determined by the size of the externality. It is the area where the social costs (MSC) exceed the social benefits (MSB=demand), generating a loss to society.

Figure 8.1. Corn market with a (negative) externality in production
In the spring of 2014, Flint, Michigan was under emergency financial management. In order to save the city money, the management switched the water supply from Detroit municipal water to the Flint River. The river water, unlike the Detroit municipal water, was corrosive. Flint had an aging water system with a high percentage of lead pipes. Due to corrosion, water lead and then blood lead levels rose. Lead is extremely toxic to children, and lead poisoning disproportionately affects low-income communities and communities of color, as in Flint. The savings to the city from switching water sources were $5 million; the switch created externalities to the extent of at least $395 million in injury to Flint (and counting). Systemic racism and externalities created a toxic situation for the families of Flint.

Climate change: a pollution case study

Climate change is an effect of certain types of pollution. To understand what climate change is, we can start with the idea of weather. *Weather* changes day to day (if you don’t like the weather, wait five minutes!). *Climate* is the long-run trend or pattern in weather. *Climate change* is the shift in climate *over time,* including higher temperatures and more extreme weather events, that human behavior has caused. Figure 8.2 shows a number of measures of climate change and humans’ role in that change. The first panel shows the human-created emissions of greenhouse gases. The greatest emissions are in the form of carbon dioxide. The other gases are shown in their carbon dioxide equivalents. Emissions have been rising steadily for some time, particularly as we burn more gas and fuel. As a result of our emissions, the concentration of carbon dioxide in the atmosphere is rising (panel b). This carbon dioxide in turn contributes to rising temperatures (panel c). Temperatures since the 1970s have been rising above the long-term average, to the point that they were, by 2021, nearly one degree Celsius above the historic trend.

---

(a) Emissions in millions of tons and carbon dioxide (CO₂) equivalent

(b) Carbon dioxide (CO₂) parts per million (ppm) in the atmosphere

(c) Global temperature anomaly relative to 1961-1990 average in degrees Celsius
This pollution and global warming have important consequences for our environment and communities. The people and countries that climate change will impact the most are the most vulnerable. Communities that are highly reliant on agriculture are being exposed to the greatest increase in risk and environmental degradation. Not only are these communities the most at risk from the harms of climate change, but they are also the ones who have contributed the least to climate change.

**How can we address pollution?**

Pollution is an enormous problem, but how can we address it? **Abatement** is the process or preventing or cleaning up pollution. There are a number of different ways to prevent pollution or to incentivize individuals and firms to create less pollution. One approach is direct government action, through preserving wetlands or cleaning waterways. Cost-benefit analysis can be used to compare the effectiveness of these different approaches. However, these approaches do not necessarily prevent pollution from happening in the first place or prevent the loss to society. In this section, we will examine several different policy options for addressing pollution through regulations or incentives.

**Limiting pollution through command-and-control**

**Command-and-control** approaches to pollution specify how much pollution a particular polluter can emit, or particular pollution-abatement technologies firms must adopt. For example, numerous countries have regulations around Nitrous Oxide (NO\textsubscript{x}) emissions, which are a form of air pollution primarily emitted from diesel vehicles. Countries have regulations on how much NO\textsubscript{x} vehicles can emit. This is a command-and-control approach applied to the emissions from each tailpipe.

Evasion of such regulations may be a problem; the recent Volkswagen scandal, in which VW vehicles were designed to “cheat” pollution tests, illustrates the harms of evasion—up to a million extra tons of NO\textsubscript{x} per year. Command-and-control solutions require vigorous monitoring and enforcement. For example, the Chinese government, in the fall of 2017, started cracking down on pollution after public outcry around air quality. They took a command-and-control approach, including banning coal in certain areas, closing down thousands of coal-fired furnaces, and coal-fired generators.

---

An underlying assumption of command-and-control regulations is that firms should all take the same approach to addressing pollution. In reality, firms may have different technologies of production, which means they face different costs, so command-and-control is inefficient. Additionally, once they are within the command-and-control standards, firms face no incentives to keep abating. For example, car manufacturers face no incentive to have lower pollution than is mandated.

**Taxing pollution**

A more flexible and efficient approach to addressing pollution is to **tax polluting goods**. Assuming we know (at least approximately) the size of the externality from a good, we can create a tax that is equivalent to the externality. With this tax, firms face an additional input cost. Ideally, they will now face essentially the MSC as their costs, and choose the social optimum, as shown in Figure 8.3 for corn. Implementing this tax policy requires substantial information on what the size of the tax should be, as well as how much corn contributes to pollution. When the tax makes corn production too expensive for certain farmers, they will stop producing corn, reducing the pollution in the environment and achieving the social optimum.
Particularly when the government struggles to figure out how much pollution a particular good or service creates or how large the externality from that pollution is, taxing pollutants directly may be a more effective approach. Figure 8.4 models this possibility. The good we are considering is the quantity of abatement in thousands of tons of NO$_x$ abated (not polluted). The price is dollars per ton to abate. For some firms it will be very easy to abate, but the costs of abatement will rise as “easy” strategies are implemented and only the
more difficult forms of abatement remain. This leads to a supply curve sloping upward with increasing MPC of abatement. The social benefits of abatement (the demand curve for abatement) will also decline as we move from acute health hazards to less dangerous levels. Where the MSB of abatement equal the MPC of abatement, this is the price where we should set our tax. In the example, this is a tax of $3/ton, with 6,000 tons of nitrous oxide abated.

The advantage of these pollution taxes is that they encourage firms that have lower costs of abatement to undertake abatement (for instance switching to cleaner energy), while those that face high costs of abatement (for example, who have poor access to solar or wind energy) can continue to pollute and pay the tax. A global example of a pollution tax is carbon taxes, which directly tax carbon dioxide. One challenge with implementing a tax on pollution is that it requires the government to have information on abatement costs, which may be difficult to obtain accurately, as firms have incentives to lie.
Pollution permits: Cap and trade

An alternative to placing a price (tax) on pollution is to set the quantity of pollution. Permits for the optimal quantity can be auctioned off and traded. This is a **cap-and-trade** (pollution permit) strategy. Figure 8.5 shows this abatement strategy, where again we are looking at the price and quantity of abatement for NOx.
Again, we pick the optimal point, where MSB=MPC for abatement. Here we set that number of permits rather than the tax. However, this will lead to exactly the price of the tax. Economically, taxes and permits are equivalent. In both cases (from the tax or the permit) the government gets revenue. In contrast to tradable permits, command-and-control approaches can be considered a non-tradable permit to pollute a certain amount. The advantage of tradable permits, as with taxes, is that they allow firms with lower costs of abatement to abate more, while firms with higher costs of abatement can buy permits and abate less. The total amount of abatement is the same, but the costs to firms to achieve that abatement are reduced relative to command and control. Additionally, unlike with taxes, as the costs of abatement change over time, the permits can change price.

Should we aim for no pollution?

The theme of the 2017 annual UN Environment Assembly meeting in Nairobi, Kenya was “Towards a Pollution-Free Planet.” Should we aspire to this goal of zero pollution? Certainly, our economic analysis showed that we must account for externalities and the social costs of pollution. However, we know that in
general—as for our other economic analyses in areas such as crime and education—we want to undertake action so long as the economic benefits of doing so outweigh the economic costs. The analysis above did not lead us to a state where we had no corn farming at all, but it did reduce the quantity of corn. Likewise, we did not set the number of permits at zero. We would only want to reduce pollution to the point that the benefits of reducing pollution outweighed the costs. This might be no pollution at all, or instead, it might be a lower level of pollution, due to the rising costs of abatement as the quantity of abatement increases.

Global agreements to address pollution

One of the challenges in tackling the global issue of pollution is that its effects are felt across national borders. Air pollution in one city affects air quality for its neighbors. Wildfires in Washington reduce air quality in Minnesota. Pollution challenges also cross a number of sectors. There are effects on health, the environment, and the economy. One of the organizations that focuses on the health effects of pollution is the World Health Organization (WHO). The “health impacts of climate and environmental change” are one of the WHO’s current top priorities in health. Achieving the global Sustainable Development Goals (SDGs) depends on addressing pollution. For example, reducing pollution is critical to improving health, reducing poverty, improving access to water and sanitation, and other goals.

The current global framework for addressing climate change is the Paris Agreement. Every country in the world except the United States has signed the Paris Agreement. The agreement aims to keep global temperature increases below 1.5-2.0 degrees Celsius, the critical tipping point for irreversible environmental change. However, the specific targets and contributions set forth in agreement are estimated to lead to an increase of 2.4-2.7 degrees Celsius by 2100. Further commitments and more progress will be needed. The agreement includes commitments by developed countries to help support developing countries, an attempt to address the disproportionately high impact of climate change on developing countries, despite their disproportionately low contributions to climate change.

Box 8.2. Climate treaties can work: The 1987 Montreal Protocol and the Ozone Layer

The ozone layer of the atmosphere helps absorb most of the sun’s ultraviolet radiation, which can cause skin cancer as well as environmental harms. In the 1970s, certain chemicals were identified as pollutants that had a harmful externality: they were damaging the ozone layer, creating an ozone hole over the Antarctic. In 1987, the Montreal Protocol was an international treaty signed to limit ozone-depleting chemicals. The prevalence of the chemicals peaked in 1993 and began to decline thereafter. Recent research suggests that the protocol and averted between 8% and 15% of certain skin cancers in affected areas.

Summary and conclusions

There has been growing recognition of the global threat to our lives, ecosystems, and economies posed by pollution. Pollution is generated by firms’ production and imposes negative externalities on society. These negative externalities are a form of market failure; firms will produce too much of these polluting goods. Government intervention is required to overcome this market failure. There are multiple forms this intervention can take, including command-and-control approaches, taxes on goods that pollute or directly on pollution, or tradable permits for pollution. Command-and-control approaches regulating pollution are not efficient, but in certain situations may be easier to implement than taxes or permits. All of these methods require good information on the effect of externalities, such as the health costs of pollution, as well as the costs of abatement. Gathering this information is an important area of work in the field of environmental economics.

The externalities of pollution know no borders; pollution is a global challenge. When local, state, or national governments take action on pollution, their decisions will not be optimal if they only account for the harms and externalities in their own communities. A global analysis of the consequences of pollution is required. The current global framework for tackling climate change—the Paris Agreement—is insufficient to prevent long-term, irreversible damage. More needs to be done to address climate change and other forms of pollution.

List of terms

- Negative externality in production
- Marginal Private Benefit (MPB)
- Marginal Social Benefit (MSB)
- Marginal Private Cost (MPC)
- Marginal Social Cost (MSC)
- Social Optimum
References


What are Social Safety Nets?

What happens if you get sick? Will you have enough money to retire one day? How do governments support our health and old age? This chapter answers these important questions about the social safety net, the supports that people can rely on should they get sick, become disabled, or lose their job. The social safety net goes by different names. Sometimes it is referred to as social protection or social insurance. There are a variety of different programs that are part of the social safety net in different countries, often with confusing names. For example, in the United States we have Medicare, which provides health insurance to seniors, and Medicaid, which provides health insurance to low-income families. All the policies we discussed in the poverty chapter are part of the social safety net as well.

In this chapter we will focus on the insurance aspects of the social safety net. Private insurance programs are when you make a regular payment, called a premium, to a private company. In return, you receive a benefit if some (usually bad) event occurs in the future. For example, you may pay renter’s insurance. If the pipes leak on to your computer in your apartment, renter’s insurance will pay out a benefit to repair or replace your computer. Most private insurance is voluntary—you can choose whether or not to get it—although there are exceptions. You are required by the government to have car insurance in the United States. Banks will only lend for a home mortgage if you have homeowner’s insurance. Although these forms of insurance may be mandatory, they are still private insurance because you are paying a private firm the premium.

Social insurance programs require paying a premium to the government rather than a private company. The premium is often a contribution from your paycheck if you are working. For example, in the United States, 6.2% of each paycheck goes towards Social Security, the program that supports seniors in their old age.¹ Employers pay an equivalent amount as well. In Egypt, workers contribute 14% of their base pay towards social insurance for old age,² similar to social security in the United States.

2. Roushdy and Selwaness, 2015.
Social insurance programs may overlap with private insurance programs. For example, there are both private ways to save for retirement, through retirement savings accounts, and social insurance programs such as Social Security. There are also social safety net programs that pay out based on need rather than contribution. For example, while Medicare is funded by contributions of workers before they reach old age, Medicaid helps low-income families based on their income, and is funded primarily by tax dollars from society overall.

Countries also vary in how extensive the social safety net is. Some countries have broad and generous social safety net programs while others rely primarily on the private sector or individuals to take care of old age, insurance, or health care. An important purpose of this chapter is understanding why there may be government involvement in these sectors—why is the safety net a social undertaking? The chapter specifically examines two major social insurance topics. First, we examine the case of Social Security and saving for old age in the United States. Then we study the issue of health insurance, which is a mix of public and private programs in the United States. Lastly, we examine funding for the social safety net in the context of government spending and debt.

### Saving for old age

#### Private market options

How can people make sure they have enough income for their old age? Individuals may want to retire. For health reasons, it may not even be possible to keep working at a certain point. However, retiring requires having sufficient income to meet your needs in old age. Individuals can save for retirement, for example putting money into a savings account or investing in property. Retirement savings are often in stocks, shares in a company, that pay dividends (regular payouts from the company’s profits). People can also buy bonds, essentially IOUs (for example, a bond to fund Amazon’s takeover of Whole Foods) that generate interest (pay a percentage out every year). Savings accounts may also generate interest, as banks will pay you to keep your money there (so they can loan it out to others and receive even more interest). Another option is an annuity, a fixed sum of money paid every year, bought in either in a lump sum or over time. These options have varying degrees of risk. Different types of bonds may be more or less risky, while stock prices, dividends, and savings interest are highly variable. Annuities are a safer investment, since payouts are fixed, but they may lead to less total income than riskier investments. There are tradeoffs between the degree of risk and the rate of return (how much you receive).

Another way individuals may save for old age is through their workplace. People can often invest in stocks, bonds, and annuities through private firms at their workplace in a way that has special tax status, deferring taxes until funds they are withdrawn. These accounts referred to as 401(k)s have this special tax status.

---

and are available from many employers. Alternatively, some firms may provide pensions. **Pensions** are like annuities—fixed annual payments—that are funded by the firm. Firms build up savings in a pension fund to pay out to retired workers.

### Why is the government involved in savings?

All of the options discussed so far are based on private firms and markets providing financial products. Why would the government be involved in these markets at all? Shouldn’t people be allowed to choose how to save for old age? The main reason for government involvement is that people, when left to their own devices, are terrible at saving for old age (or saving for anything else). For example, in the United States only 47% of people report that they have savings or a rainy day fund that could cover three months of expenses.⁴ A variety of reasons could account for individuals’ difficulty saving, including information issues and uncertainty. It is hard to know how much you need for retirement or to understand complex financial products that can help with retirement.⁵ People are also bad at self-control, tend to procrastinate, and prefer money today to money tomorrow.

---

**Box 9.1. Behavioral economics: Understanding barriers to savings**⁶

An important field for understanding savings (as well as health care, discussed below) is behavioral economics. **Behavioral economics** brings together economics and psychology. The field challenges the assumption of traditional economic theory that individuals have all the information they need and are perfectly rational decision-makers.

Behavioral economics provides insights into why people are bad at savings, and what we can do about it. For example, when employees have to opt-in to their retirement savings (making the default zero) they are less likely to save for retirement than if they had to opt out of savings instead. Although the default should not matter for savers’ decisions, it does, and it matters a lot. In an experiment based on switching workers to automatic enrollment, 86% of employees contributed to a 401(k) plan when enrollment was automatic, but just 49% participated when they

---

⁴ Board of Governors of the Federal Reserve, 2016.
⁵ Knoll, 2010.
⁶ Bertrand, Mullainathan, and Shafir, 2004; Knoll, 2010; Madrian and Shea, 2004; Goda et al., 2019.
actively had to enroll. This insight from behavioral economics has led to more automatic-enrollment plans for savings.

**Social Security**

Social Security is the United States social insurance program designed to act as a social safety net as people age. The program also helps address the challenges workers face in saving. The United States’ current Social Security program has its roots in the Great Depression. The Social Security Act became law in 1935 as a form of old-age insurance. It provides monthly retirement benefits to individuals 65 and older. It originally covered few workers but expanded substantially over time. Social security is credited with bringing the rate of elderly poverty from 35.2% in 1959 to 10.2% today. The reduction in elderly poverty brought about by social security illustrates some of the arguments for government involvement in saving for old age.

Benefits for social security are based on wages while working. The program was never intended to fully replace other forms of income, as it provides benefits equivalent to 42% of individuals’ previous wages. Dependents (children) and survivors (spouses) of workers who die are eligible for benefits as well. Lower-income earners receive benefits that are a greater share, around 56%, of their earnings, compared to 35% for higher-income earners.

The program uses payroll taxes (taken directly out of workers’ paychecks) to collect workers’ contributions. Workers must contribute 6.2% of their wages. The program has an equal employer share of 6.2%. Wages are taxed up to the first $128,700.

The official rule is that employees pay 6.2% and employers pay 6.2%. However, it could

---

7. Martin and Weaver, 2005.
8. Ibid.
9. Ibid.
10. Ibid.
be the case that employers increase wages to make up for the loss to Social Security, or that they reduce wages to cover not just the employee but also their payment. This is a question of tax incidence—who, economically, pays the tax.

Figure 9.1 shows a labor market where labor supply is relatively inelastic and labor demand relatively elastic. As with Social Security, each actor must pay an equal tax. Here we have simplified to make the tax a fixed $3 per hour for both the workers and employers. This tax decreases labor supply (you take home less of every dollar) and decreases labor demand (firms have to pay more for their workers). However, although each side on paper has to pay equally, the tax incidence is not equal. Moving from equilibrium 1, without the tax, to equilibrium 2, with the tax, because labor demand is more elastic, the price that firms pay with the tax has increased only a little. However, because labor supply is relatively inelastic, the wage that workers take home has dropped substantially. Relative to where they started, Peq, firms are paying a little more and workers are making a lot less. Now, if workers receive the reduction back in the form of Social Security, this is not necessarily a problem, especially if it overcomes market failures in savings. However, it does illustrate that who theoretically pays the tax is not the true tax incidence—who is affected the most by the tax.
Taking care of our health

Another key aspect of wellbeing, insurance, and the social safety net is health care. Health care is expensive. Figure 9.2 shows health care expenditure in the United States as a percentage of GDP. It also examines the sources of spending, comparing spending from the public sector (the government) and from private sources. Health care is a growing segment of the economy. As of 2020, a total of 20% of GDP went to health care, making it one-fifth of the U.S. economy. Around 10% of our economy is private spending on health care and 10% is public spending on health care. This share has been rising substantially over time. A number of factors,

including improvements in medicine and income, have contributed to this rise. In 2020, the COVID-19 pandemic contributed to the substantial rise in public health care spending particularly.

![Health care expenditure in the United States as a percentage of GDP, 1987-2020](image)

**Figure 9.2 Health care expenditure in the United States as a percentage of GDP, 1987-2020**

However, the United States is also an outlier in global health spending. We spend far more, yet have worse outcomes, in comparison to other countries. While we spent 17% of our GDP on health care as of 2019, comparable countries spent only 11% of GDP.\(^\text{14}\) When examining 13 developed countries, despite high spending, the United States had the highest infant mortality (deaths) and shortest life expectancy.\(^\text{15}\) Why is health care so expensive while outcomes are so poor?

---

\(^{14}\) Wager, Ortaliza, and Cox, 2022.

\(^{15}\) Squires and Anderson, 2015.
The role of health insurance

Often, researchers point to health insurance as contributing to high costs. With health insurance (whether provided by the government or a private insurer), individuals have a complex set of potential costs. Typically, plans have a **deductible**, a fixed amount that insurance does not cover before paying benefits. For example, you might have to pay the first $1,000 of your health expenses. Insurance often requires cost-sharing as well. People with health insurance who seek medical care may have **co-insurance** payments, where they have to pay a percentage of costs. For example, individuals may have a co-insurance rate of 20%, where they pay 20% of costs and the insurer pays 80%. Alternatively, individuals may have a **co-payment** (co-pay) for their visit, a fixed amount, like $20 for an office visit or $250 for an emergency room visit.

Although individuals have to pay for insurance and also have various forms of potential cost-sharing, effectively insurance decreases the price of health care to individuals. It does not, however, decrease the price of health care for insurance companies or society. This disconnect leads to the situation modeled in Figure 9.3, for the case of a co-pay for doctors’ visits. If there were only a private market for doctors’ visits, then supply and demand would determine an equilibrium of five million visits with a price of $100 per visit. However, once there is insurance, the price to the patient drops. The patient now pays only $50 per visit as the co-pay. Now people would choose to have more visits, 6 million total when the price is $50. The additional visits cost insurance and society. Additionally, the visits become more expensive to society (or the insurance company), at $120 each. Ultimately, patients pay for higher costs through their premiums (if they have private insurance) or through their taxes (if there is government funded insurance).

---

Figure 9.3. Supply and demand for doctors’ visits with a co-pay

When individuals use more health care because they have insurance, this is a form of moral hazard. Moral hazard occurs when individuals change their behaviors as a result of having health insurance. This may mean that they go to the doctor more. It may also mean that they have less healthy or more risky behaviors because they know if they get sick or injured, they have insurance. Research shows that having insurance reduces healthy behaviors, for example, it increases smoking and drinking and reduces exercise.

If health insurance has all these problematic effects, why do people buy health insurance? Primarily, they want to insure themselves against risk of illness or injury. They also want financial predictability. In the United

17. Ibid.
States, 46% of bankruptcies included medical problems as a factor.\textsuperscript{19} That same motivation for insurance, to insure against risk, remains. However, insurance also creates all sorts of distortions in markets. Another substantial issue in insurance markets is referred to as adverse selection. \textbf{Adverse selection} occurs in insurance markets when riskier individuals buy insurance, but less risky individuals do not. For example, someone with an underlying medical condition may buy insurance but someone who is healthy will choose not to buy insurance. If only sick(er) people buy insurance, it becomes extremely expensive.

\section*{Health policy in the United States}

Health care and health insurance are critical issues in the United States. There are a number of reasons we have identified that these markets may not work well on their own. Market failures tend to lead to government intervention. The United States has a substantial government role in health care, a regulatory role as well as one in financing and providing healthcare. First, for individuals 65 and older, we have \textbf{Medicare}, a health insurance program for seniors. Medicare, like Social Security, is funded by payroll taxes. Individuals pay in while working and receive benefits at 65. For low-income families, we have \textbf{Medicaid}, health insurance funded by government tax dollars.

Other people may receive their health insurance from private or non-profit insurers. However, there are a number of regulations and government supports for health care. Specifically, in 2010, the U.S. Congress passed the \textbf{Patient Protection and Affordable Care Act (ACA)}, often referred to as Obamacare.\textsuperscript{20} The ACA mandated that every individual buy insurance (or face a penalty). Frequently people could buy insurance through their employer or received insurance as an employment benefit. However, individuals without employer-provided insurance could, under the new law, shop in newly created state marketplaces for for-profit or non-profit insurance plans. Individuals also received subsidies, if they qualified due to their income, to make insurance more affordable. The law includes a number of other provisions about essential health benefits that must be included in insurance, and addressed other issues as well, such as the share of insurance that may go towards profits versus medical care. Why is health insurance mandated under the ACA? Recall our earlier discussion of adverse selection. Typically, it is sicker people who buy insurance. Requiring everyone to buy insurance helps make sure that there are (currently) healthy individuals with insurance to help spread out the costs of insurance.

Although the United States has substantially increased regulations in health care, costs remain high. Empirically, the United States actually has fewer doctors and hospital visits than other countries, suggesting that higher prices may be a driving factor in our high health care spending, not higher use of health care.\textsuperscript{21}

\textsuperscript{19} Himmelstein et al., 2009.
\textsuperscript{20} Office of the Legislative Counsel, 2010.
\textsuperscript{21} Squires and Anderson, 2015.
How we pay doctors may be a central issue to high costs. Commonly, doctors in the United States are paid on a fee-for-service basis. Doctors receive fees for each service they perform. This creates an incentive to do more services—perhaps an extra test or scan. Patients lack the information or expertise to know that these additional services may be unnecessary, leading to unnecessary medical spending. Alternative models than fee-for-spending are also challenging to design but may be more effective.\(^{22}\) If doctors are paid per patient rather than procedure, they may pick healthier patients. Therefore, building in additional payments for working with patients with chronic health conditions is necessary. This is just one example of the complexities of addressing incentives within the medical system.

**Government Spending**

Social safety nets and health care have an enormous role in government spending in the United States. Figure 9.4\(^{23}\) shows the United States federal government budget for the year 2019. During 2020 and 2021, the budget was substantially different due to the pandemic, but 2019 represents a more typical year. States may also collect taxes and have substantial budgets, as do counties and cities. However, when we are thinking about social safety nets and especially social insurance, the federal government is the main player. The figure distinguishes between revenues (money coming in) and spending (money going out). Some of the money being spent is mandatory—it has to be spent based on law. Mandatory spending is also often referred to as an entitlement. Social insurance programs are common entitlements, where an individual pays in during their working years and then is entitled to a benefit.

Where does most of our federal money come from? First, 1.7 trillion dollars comes from individual income taxes. Then 1.2 trillion are from payroll taxes, which are designed to fund social insurance. Corporate income taxes provide 230 billion in revenues, and the remaining 271 billion in revenues is from other sources.

\(^{22}\) Kantarevic, Kralj, and Weinkauf, 2011.

\(^{23}\) Congressional Budget Office, 2020.
On the spending side, Social Security is the single biggest government expenditure, at 1.0 trillion dollars. Medicare is 644 billion, followed by Medicaid at 409 billion. The remaining mandatory programs, such as unemployment insurance, veterans’ benefits, the earned income tax credit, Supplemental Nutrition Assistance Program (SNAP or food stamps), and many others, together total 642 billion. These are the mandatory government expenses. Mandatory spending totals 2.9 trillion dollars compared to 1.3 trillion of discretionary spending. Discretionary spending is what lawmakers control through specific bills and acts. More than half of discretionary spending is defense (676 billion). Another important expense is interest on our national debt (375 billion).

Why do we have to pay interest on our debt? Each year when the budget is set, there may be a **budget surplus** (revenue > spending) or a **budget deficit** (revenue < spending). The deficit in 2019 was projected at 1.0 trillion. When there is a deficit, the government will pay for the extra spending by borrowing money from the public through bonds or securities. As deficits and surpluses add up over time, if we end up owing money,
this is our debt. As of 2019, public debt was equal to 79% of our GDP—16.8 trillion dollars. In one sense, this is alarming—it is a huge amount of money. However, from another angle, we have less debt than what our economy produces in a single year.

The real issue is the long-term outlook of our debt, which is poor under current law. Increases in the cost of social insurance programs—especially Medicare and Social Security—are a major problem. If current spending and revenue patterns continue and projected Medicare and Social Security trends play out, debt would reach 144% of GDP by 2049. That could be more problematic than the current situation, as the public holding government debt might start to worry about being paid back, leading to higher interest rates (another expense for the government).

Linked to our debt challenges are challenges paying for Social Security and Medicare. These programs rely on the payments of current workers—which pile up in trust funds—to help fund current benefits. Such a system is referred to as pay-as-you-go. In contrast, a system where individuals’ benefits are funded by their initial contributions would be fully funded. Demographic factors play an important role in pay-as-you-go systems. When populations are young, there are a lot of workers paying in and few people receiving benefits. As populations age—as the baby boomers are doing in the United States—there are many more people with benefits, yet relatively fewer workers paying in. Currently, the structure of Social Security is unsustainable. The shortfall in the long run (through 2093) amounts to 4.6% of taxable payroll. Although concerning, this is a solvable problem. Benefits can be reduced, the age of receiving benefits adjusted, or payroll taxes changed.

Summary and Conclusions

Safety nets are critical to ensuring that citizens are secure in their health or retirement. Although individuals can save for old age or buy health insurance on the private market, a number of challenges and market failures make the performance of the private market on its own problematic. Therefore, governments build social safety net programs. Social insurance programs, where individuals contribute from payroll taxes, were the particular focus of this chapter. Social Security (for old age income) and Medicare (for old age health care) are the largest social insurance programs in the United States. They are also our largest budget expenditures. Since these programs are not fully funded, reforms are needed to ensure their long-term sustainability. Think about how you would like to see these programs reformed—and tell your government representatives!

---

26. We also have debt held by government trust funds rather than the public, essentially different parts of the government owing each other money.
28. Ibid.
### List of terms

- Social safety net
- Medicare
- Medicaid
- Private insurance
- Premium
- Social insurance
- Social Security
- Stocks
- Dividends
- Bonds
- Interest
- Annuities
- 401(k)s
- Behavioral economics
- Tax incidence
- Deductible
- Co-insurance
- Co-payment (co-pay)
- Moral hazard
- Adverse selection
- Medicare
- Medicaid
- Patient Protection and Affordable Care Act (ACA or Obamacare)
- Fee-for-service
- Budget surplus
- Budget deficit
- Debt
- Entitlements
- Pay-as-you-go
- Fully funded

### References


