Section 2: Efficiency in the Economy as a Whole

We’ve seen how the equilibrium outcome of an individual competitive market usually maximizes the total surplus of participants in that market. Is there an equivalent result for the economy as a whole? That is, is there a corresponding concept of equilibrium for the whole economy? And if so, does this equilibrium outcome maximize the welfare of the economy’s participants?

The economy as a whole consists not of one but of many, many markets, all interrelated in two ways:

- On the consumption side, the demand for each good is affected by the prices of other goods.
- On the production side, producers of different goods compete for the same factors of production.

To think about the economy as a whole, then, we have to think of many markets, for both goods and factors. A competitive market economy is an economy in which all of these markets are perfectly competitive, with equilibrium prices determined by supply and demand. In each market both the supply and demand curves are likely to be affected by events in other markets.
An economy is in **general equilibrium** when the quantity supplied is equal to the quantity demanded in all markets.

When all markets have reached equilibrium—when the quantity of each good and factor demanded is equal to the quantity of each good and factor supplied at the going market prices—we say that the economy is in **general equilibrium**. To put it another way, general equilibrium is the economy-wide counterpart of ordinary equilibrium in a single market.

Our next task is to show that, as with an individual competitive market in equilibrium, a competitive market economy in general equilibrium is usually **efficient**—that is, it is efficient except in certain well-defined cases. What do we mean by saying that the economy as a whole is efficient? Actually, we defined efficiency way back in Chapter 1. We will start by revisiting that definition to see why it is the right approach to analyzing the economy as a whole. Next, we will describe the three criteria that an economy as a whole must satisfy in order to be efficient. Finally, we will learn how failures of individual markets can lead to inefficiency in the economy as a whole—failures we can view as cases in which prices fail to perform as economic signals.

**Efficiency, Revisited**

When economists **discuss efficiency in an individual market**, they usually use the concepts of **consumer and producer surplus**, which measure costs and benefits in monetary terms. This makes sense when you are talking about the market for just one good, because you can take the prices of other goods—and therefore the value of a dollar—as given. **When we are analyzing the economy as a whole, however, measuring costs and benefits in dollar terms no longer makes sense, because all prices are “to be determined.”**

Instead, economists focus on the **basic definition of efficiency**. Recall from Chapter 1: an economy is efficient if it does not pass up any opportunities to make some people better off without making other people worse off.

To achieve efficiency, an economy must meet three criteria, which closely parallel the three features ensuring that total surplus is maximized in an individual market. The economy must be **efficient in consumption, efficient in production**, and **efficient in output levels**. Let’s look at these criteria and see how a competitive market economy satisfies them.
Efficiency in Consumption

An economy is **efficient in consumption** if there is no way to redistribute goods among consumers that makes some consumers better off without making others worse off.

To see what efficiency in consumption involves, it helps to imagine scenarios for inefficiency. Imagine, for example, an economy that produces both cornflakes and shredded wheat but that provides those who prefer shredded wheat with cornflakes, and vice versa. Then it would be possible to make at least one person better off without making anyone else worse off by redistributing the goods, giving people the breakfast cereal they prefer.

The first piece of good news is that as long as prices perform properly as economic signals, this kind of inefficiency won’t occur in a competitive market economy. We’ve seen this already in the case of market equilibrium in one individual market: the consumers who actually receive a good at the market equilibrium are those with the greatest willingness to pay—thanks to the role prices play in helping people make the right economic decisions. Consumers who prefer an additional box of cornflakes will be willing to pay more for that box than consumers who would rather have an additional box of shredded wheat. So if the markets for cornflakes and shredded wheat are both in equilibrium, there won’t be any way to make at least one consumer better off without making others worse off by redistributing the available quantities of breakfast cereals.

In other words, prices in goods markets ensure that you can’t increase total surplus in a single market by taking a good away from one person and giving it to another. Similarly, prices also ensure that when all markets in an economy are in perfectly competitive general equilibrium, there is no way to redistribute goods that makes some consumers better off without making others worse off.

It’s important, however, to realize the limitations of that statement: even though an economy is efficient, you can always make some consumers better off if you are willing to make others worse off. We’ll come back to that point shortly.
Efficiency in Production

Economists say that an economy is **efficient in production** if it is not possible to produce more of some goods without producing less of others.

We can use the *production possibility frontier* model from Chapter 2 to understand this. This model uses a diagram like Figure 13-2 to illustrate the economy’s trade-offs: the more wheat it produces, the less corn it can produce, and vice versa. If the economy produces the quantities at either point A or point B on the production possibility frontier, it is efficient in production: it is possible to produce more corn than the economy produces at A, but only by producing less wheat; it is possible to produce more wheat than the economy produces at B, but only by producing less corn. The economy is not efficient in production, however, if it produces at point C: it is possible to produce more wheat and more corn than the economy does at that point.

An economy will be efficient in production if it has an **efficient allocation of resources**—if there is no way to reallocate factors of production among producers to produce more of some goods without producing less of others. This is an important result: *An economy that is efficient in allocation of resources is efficient in production, and vice versa.*

Here is another way to think about Figure 13-2: at point A the economy can produce more corn only by taking resources away from wheat production. Similarly, at point B the economy can produce more wheat only by taking resources away from corn production.

Just as in the case of efficiency in consumption, it helps to imagine scenarios for inefficiency. In the United States, land in Iowa is ideally suited for growing corn, land in Minnesota is much better suited for growing wheat, and most land in Maine isn’t suitable for growing either. It would clearly be inefficient if good land in the Midwest were left idle but farmers struggled with the stony soil of Maine; it would also be inefficient if Iowa farms grew wheat and Minnesota farms grew corn.

The second piece of good news is that, just as in the case of consumption, the role of prices as economic signals ensures that a competitive market economy in general...
equilibrium achieves efficiency in production. The logic is similar, but this time it applies to prices in factor markets rather than prices in goods markets. Corn farmers are willing to pay more for Iowa land than wheat farmers; wheat farmers are willing to pay more for Minnesota land than are corn farmers. And much of New England is no longer farmed but has returned to forest, because labor and capital can be more productively employed on richer lands elsewhere. In short, when factor markets are

**Figure 13-2**

**The Production Possibility Frontier and Efficiency in Production**

An economy is efficient in production if it cannot produce more of one good without producing less of other goods. Equivalently, an economy is efficient in production if it is on its production possibility frontier. Here $A$ and $B$ are efficient production points—at each point the economy can produce more of one good only by producing less of the other. $C$ is not an efficient production point because more corn and more wheat can be produced.
competitive, resources are allocated to the producers that can make the best use of them, and the economy is indeed efficient in production.

Notice, however, that this does not say what the economy produces. Both A and B in Figure 13-2 represent efficient production. We still need to ask whether the economy produces at the “right” place on the production possibility frontier—or, rather, a right place, because there may be many efficient outcomes. But let’s hold off on that for a moment and finish our description of efficiency in the competitive market economy as a whole.

Efficiency in Output Levels

Suppose that a competitive market economy is efficient in production—it cannot produce more of some goods without producing less of others. Suppose also that it is efficient in consumption—there is no way to redistribute the goods produced that will make some consumers better off without making others worse off. There is still the question of whether the competitive market economy is producing the right mix of goods to start with. For example, suppose that point A in Figure 13-3 corresponds to producing enough wheat to let everyone have shredded wheat five times a week and cornflakes two times a week. This will still be inefficient if everyone prefers to have shredded wheat only three times a week but cornflakes four times a week—and if point B would allow them to do so. In that case, moving from A to B—that is, shifting resources into corn production—would make everyone better off. Our third criterion for efficiency, then, is that the economy be efficient in output levels: there must not be a different mix of output that would make some people better off without making others worse off.

The third and final piece of good news about the equilibrium of a competitive market economy is that it will be efficient in output levels when prices perform properly as economic signals.
How do we know this? We already saw that in an individual competitive market producers produce the quantity of output that maximizes total surplus. The reason is that consumers and producers face the same price—the market price is an economic signal telling producers what one more unit of output is worth to consumers. This signal induces producers to produce that extra unit of output if the cost of the resources they would need to produce it is less than the market price.

Figure 13-3

Efficiency in Output Levels

The output levels $A$ and $B$ are both efficient in production. However, suppose that $B$ represents a mix of quantities of corn and wheat that everyone prefers to the mix at $A$—at $A$ everyone prefers more corn and less wheat. Then the economy that produces output mix $A$ is not efficient in output levels.
In the economy as a whole, producers learn how much consumers are willing to pay for a bit more of one good versus a bit more of another when market prices operate as economic signals. This process ensures that a competitive market economy in general equilibrium produces the right mix of goods.

To see how this happens, imagine an economy in which the only resource that can be shifted between industries is labor, and all producers are hiring from the same labor market. (We’ll also assume there are no complications like compensating differentials that make wages differ.) Imagine that right now consumers would prefer more corn and less wheat than the economy is currently producing. The economy can provide what consumers want by transferring labor from wheat production to corn production—by forgoing some wheat output in order to produce more corn. But will this adjustment take place?

Yes, it will, because consumers are willing to pay more for the additional corn that one more worker employed in corn production can produce than they are willing to pay for the wheat forgone by employing one fewer worker in wheat production. We can express this algebraically. The extra corn that a unit of labor can produce is $MPL_{corn}$, the marginal product of labor in corn. The wheat that unit of labor would have produced is $MPL_{wheat}$, the marginal product of labor in wheat. When we say that consumers are willing to pay more for the extra corn than for the wheat, we are saying that at the current employment and output levels in the corn and wheat sectors

\[(13-1) \quad P_{corn} \times MPL_{corn} > P_{wheat} \times MPL_{wheat}\]

where $P_{corn}$ and $P_{wheat}$ are the prices of corn and wheat respectively.

We’ve already seen the expressions in Equation 13-1 in Chapter 12. $P_{corn} \times MPL_{corn}$ is the value of the marginal product of labor in corn production, and $P_{wheat} \times MPL_{wheat}$ is the value of the marginal product of labor in wheat production. So we can restate Equation 13-1 as

\[(13-2) \quad VMPL_{corn} > VMPL_{wheat}\]
This expression says that the value produced by an additional unit of labor employed in corn production is greater than that of an additional unit of labor employed in wheat production when consumers prefer more corn and less wheat than is being produced.

Can this be an equilibrium? No; we learned in Chapter 12 that producers maximize profits by hiring labor up until the point that \( VMPL = W \), where \( W \) is the wage rate. That is, a producer hires labor until the value of the output produced by the last worker employed is equal to the current market wage rate. In this example, corn producers and wheat producers hire workers from the same labor market. So the direct implication of \( VMPL_{\text{corn}} > VMPL_{\text{wheat}} \) is that, at current employment levels, corn producers are willing to pay a higher wage rate than wheat producers. Corn producers will hire workers away from wheat producers.

When will this process stop? When the wage rate that corn producers are willing to pay is equal to the wage rate that wheat producers are willing to pay; that is, when \( VMPL_{\text{corn}} = VMPL_{\text{wheat}} \). The evolution of this process is illustrated in Figure 13-4. In panel (a), a corn producer starts at a \( VMPL_{\text{corn}} \) greater than the current market wage rate. She increases her profits by hiring more workers and moves down her \( VMPL \) curve until she reaches her optimal employment level, at which \( VMPL_{\text{corn}} = W \).

Where are these new workers in the corn sector coming from? They are coming from the wheat sector. This is illustrated in panel (b), where a wheat producer is losing workers and in the process is moving up his \( VMPL \) curve. He increases his profits by laying off workers, letting them go until he reaches his optimal employment level, at which \( VMPL_{\text{wheat}} = W \).

So as labor in the economy is reallocated from wheat production to corn production, the output of corn rises and the output of wheat falls. Eventually, workers cease moving from the wheat sector to the corn sector when \( VMPL_{\text{corn}} = VMPL_{\text{wheat}} \). At this point the value of the additional output produced by a worker in the corn sector no longer exceeds the value of the additional output produced by a worker in the wheat sector. That is, we have finally reached an equilibrium.
If at current employment levels $\text{VMPL}_{\text{corn}} > \text{VMPL}_{\text{wheat}}$, then corn producers will increase their profits by hiring workers away from wheat producers, who will, in turn, increase their profits by laying off workers. This process is illustrated for a corn producer in panel (a). As a corn producer hires workers, she increases her corn production and moves down her $\text{VMPL}$ curve until she reaches her optimal employment level, the number of workers at which $\text{VMPL}_{\text{corn}} = W$, the market wage rate. Panel (b) shows the corresponding changes for a wheat producer: he decreases his wheat production and moves up his $\text{VMPL}$ curve as he lays off workers. He also reaches his optimal employment level at $\text{VMPL}_{\text{wheat}} = W$. Workers cease moving between sectors once $\text{VMPL}_{\text{corn}} = \text{VMPL}_{\text{wheat}}$. It is an equilibrium because at that point the value of the additional output produced by a worker in the corn sector no longer exceeds the value of the additional output produced by a worker in the wheat sector and corn producers are no longer willing to pay a higher wage than wheat producers.
This example helps us understand that, in a market economy, markets for goods and services are linked via the factor markets. Or to put it a slightly different way, any change in the amount of one good or service produced will ultimately affect the amounts of other goods and services as factors of production shift from one sector to another. Figure 13-5 helps us make sense of the interconnectedness of markets for goods and services and factor markets in a market economy. To appreciate its significance, it may help recalling the circular-flow diagram of Chapter 2. There we saw how firms and households are linked via the factor markets and the markets for goods and services through flows of money. That is, in the factor markets, firms pay households for factors of production; in the markets for goods and services, households pay firms for goods and services.

Figure 13-5 presents an alternative and richer version of the same phenomenon, this time expressed in terms of the forces of supply and demand for resources that underlie the money flows of the circular-flow diagram. Here, we imagine that the only goods in the economy are corn and wheat and that labor is the only factor of production. The right portion of the figure represents how the supply of labor from households and the demand for labor from firms meet and are brought into equilibrium in the factor markets. But supply and demand in the labor market are themselves affected by supply and demand in the markets for goods and services: firms’ demand for labor is derived from the demand for their goods and services, and households’ supply of labor is determined by the earnings required for purchases in the markets for goods and services.

Similarly, the left portion of Figure 13-5 represents how the supply of goods and services from firms and the demand for goods and services from households meet and are brought into equilibrium in the markets for goods and services. Again, there is an interrelationship between the market for goods and services and factor markets: households’ demand for goods and services is determined by the wages they earn in
Figure 13-5  Efficiency in Output Levels in a Circular-Flow Framework

Markets for goods and services

- Wheat demand
- Corn demand

Factor markets

- Wheat supply
- Labor supply

Households

- Corn supply

Firms

- Labor demand

E Q_{corn} P_{corn} D_{corn} S_{corn} E Q_{wheat} P_{wheat} D_{wheat} S_{wheat} E W S_{labor} VMPL_{corn} W S_{labor} VMPL_{wheat}

Efficiency in Output Levels in a Circular-Flow Framework
the labor market, and firms’ supply of goods and services is determined by the returns they generate from hiring labor in the labor market.

Moreover, any change in the demand for a good will initiate a cascade of events that will ripple through the entire economy as resources shift among various sectors in response. So, for example, an increase in demand for corn relative to wheat causes a rightward shift of the demand for corn and a leftward shift in the demand for wheat. This results in an increase in the equilibrium quantity of corn supplied and a decrease in the equilibrium quantity of wheat supplied, which in turn causes an increase in the equilibrium employment level in the corn sector and a decrease in the equilibrium employment in the wheat sector. So the incentives in a competitive market economy in which prices perform properly as economic signals will lead the economy to produce the mix of goods that consumers prefer—that is, the economy will indeed be efficient in output levels. When each market for goods and services and each factor market is in equilibrium, the economy as a whole is in general equilibrium.
How does a competitive market economy achieve this amazing result? It comes down to the point we have emphasized throughout this discussion: the role of prices as *economic signals*. The fact that everyone faces the same prices ensures that goods and services are efficiently allocated among consumers, that factors of production are efficiently allocated among producers, and that the mix of goods and services produced reflects what people want.

**Inefficiency in the Economy: When Prices Go Astray**

The efficiency of a competitive market economy, however, is not something we should take for granted. Just as individual markets can fail, the general equilibrium of a competitive market economy may be inefficient. When this happens, some mutually beneficial trades will go unexploited. That is, there is some reallocation of consumption or production, or some change in output levels, that would make someone better off without making others worse off.

Why is a competitive market economy sometimes inefficient? For the same three reasons that individual markets sometimes fail. And when markets fail, prices fail to serve as economic signals.

First, recall that a market fails when attempts by one party to capture more resources prevents mutually beneficial transactions from occurring. We can illustrate this phenomenon by returning to our Chapter 4 example of quotas and New York taxi licenses. Recall that the current quota on taxi licenses in New York City results in a shortage of taxi rides compared to the unrestricted competitive market equilibrium. As a result, existing owners of taxi licenses gain (earning higher returns than they would have otherwise), but mutually beneficial transactions in taxi rides go unexploited. If the quota didn’t exist, the fact that the market price of a ride exceeds the unrestricted competitive market price would draw new taxi drivers into the market. So supply would increase until the willingness to pay of the last rider in the market is equal to the cost.
of the last driver in the market, and the market would achieve efficiency. Or, to put it slightly differently, the quota prevents the market price from operating as an economic signal that would draw new taxi drivers into the market to eliminate the shortage. As a result, the larger economy (of which the New York City taxi ride market is a part) is inefficient. It’s not efficient in consumption because some riders with a higher willingness to pay can’t get a taxi but some with a lower willingness to pay are able to. It is not efficient in production because some drivers with higher costs have licenses but some with lower costs do not, implying that output could be increased without using more resources by reallocating production to those drivers with the lowest cost. And it is not efficient in output levels because there are too few taxi rides provided relative to other goods. We will study a similar kind of inefficiency in detail in Chapter 14, where we investigate monopoly, an industry that has only a single seller.

The second source of market failure are side effects of actions that aren’t properly accounted for in the existing markets. For example, consider the pollution caused by a large poultry farm whose wastes foul nearby streams and reduce the welfare of nearby residents. The cost savings that the farm gains by dumping its waste rather than having it properly disposed of are smaller in value than the discomfort experienced by its neighbors—that is, what the neighbors would be willing to pay to get rid of the pollution. But a mutually beneficial trade between the farm and the neighbors is going unexploited. Efficiency could be achieved if there were a “market for pollution”—if either the farm could purchase from the residents the right to dump so many barrels of waste or the residents could purchase from the farm the right to water of a certain purity. Either way, efficiency would be achieved: consumers would buy and the seller would sell until the willingness to pay for the last unit bought and sold equals the seller’s cost of producing it. The problem here arises because the relevant market doesn’t exist—there is no market for either pollution rights or water-purity rights. As a result, there is no relevant market price that can function as an economic signal to direct individuals to buy and sell. This economy is inefficient: the total welfare of individuals...
in the economy would be higher if less pollution and, correspondingly, less poultry were produced. We will analyze these side effects, called externalities, in detail in Chapter 19. There we will see how governments can intervene in ways that assure that market prices do indeed provide the proper economic signals.

The third and final reason that markets fail is that some goods, by their nature, are unsuited for efficient management by markets. Consider a song that can be downloaded from the Internet. Once the song is recorded, the marginal cost to the record company of allowing someone to download it is virtually zero. So efficiency implies that people should be allowed to download the song until the last downloader’s willingness to pay is equal to zero. That is, efficiency is achieved only if there is free and unrestricted downloading of music. But an inconvenient fact arises at this point: if the record company can’t charge for downloads, then it won’t have revenues to pay its artists. And if it can’t pay its artists, then artists won’t record in the first place. And that, too, is inefficient. The problem here lies in the nature of the good itself, called an information good, which is the subject of Chapter 22. For information goods, efficiency in consumption means that the price to consumers must be zero, but efficiency in output levels means that the price received by producers must be greater than zero. For these goods, inefficiency is simply unavoidable. A similar inefficiency arises for what are called public goods and common resources—goods for which, by their nature, the producer cannot charge consumers. We will study public goods and common resources in Chapter 20.

In most cases, however, market failures can be corrected. Indeed, although it is important to understand how they occur, this analysis isn’t a reason to reject markets. Instead, the analysis of market failures tells policy makers how to correct them and take advantage of the remarkable ability of a competitive market economy to allocate resources efficiently.