CHAPTER 24

INDUSTRY SUPPLY

We have seen how to derive a firm's supply curve from its marginal cost curve. But in a competitive market there will typically be many firms, so the supply curve the industry presents to the market will be the sum of the supplies of all the individual firms. In this chapter we will investigate the **industry supply curve**.

24.1 Short-Run Industry Supply

We begin by studying an industry with a fixed number of firms, n. We let $S_i(p)$ be the supply curve of firm i, so that the **industry supply curve**, or the **market supply curve** is

$$S(p) = \sum_{i=1}^{n} S_i(p),$$

which is the sum of the individual supply curves. Geometrically we take the sum of the quantities supplied by each firm at each price, which gives us a *horizontal* sum of supply curves, as in Figure 24.1.

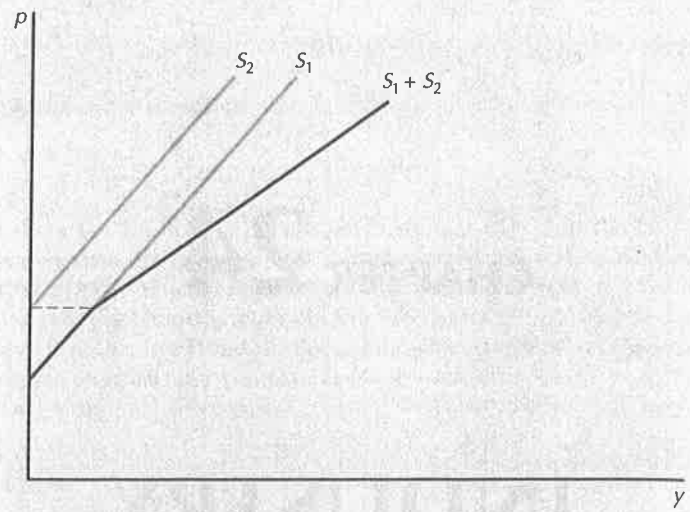


Figure 24.1 **The industry supply curve.** The industry supply curve ($S_1 + S_2$) is the sum of the individual supply curves (S_1 and S_2).

24.2 Industry Equilibrium in the Short Run

In order to find the industry equilibrium we take this market supply curve and find the intersection with the market demand curve. This gives us an equilibrium price, p^*.

Given this equilibrium price, we can go back to look at the individual firms and examine their output levels and profits. A typical configuration with three firms, A, B, and C, is illustrated in Figure 24.2. In this example, firm A is operating at a price and output combination that lies on its average cost curve. This means that

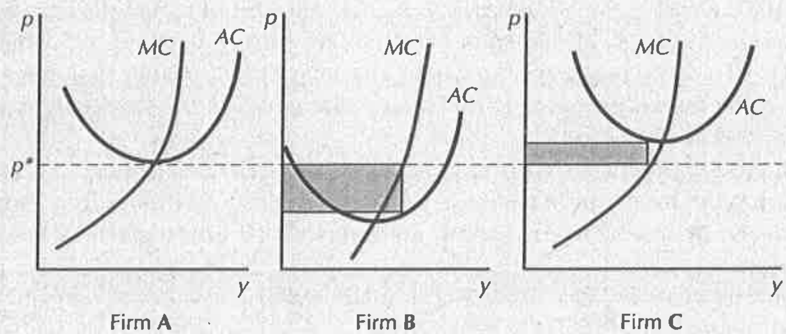
$$p = \frac{c(y)}{y}.$$

Cross multiplying and rearranging, we have

$$py - c(y) = 0.$$

Thus firm A is making zero profits.

Firm B is operating at a point where price is greater than average cost: $p > c(y)/y$, which means it is making a profit in this short-run equilibrium.



Short-run equilibrium. An example of a short-run equilibrium with three firms. Firm A is making zero profits, firm B is making positive profits, and firm C is making negative profits, that is, making a loss.

Figure 24.2

Firm C is operating where price is less than average cost, so it is making negative profits, that is, making a loss.

In general, combinations of price and output that lie above the average cost curve represent positive profits, and combinations that lie below represent negative profits. Even if a firm is making negative profits, it will still be better for it to stay in business in the short run if the price and output combination lie above the average *variable* cost curve. For in this case, it will make less of a loss by remaining in business than by producing a zero level of output.

24.3 Industry Equilibrium in the Long Run

In the long run, firms are able to adjust their fixed factors. They can choose the plant size, or the capital equipment, or whatever to maximize their long-run profits. This just means that they will move from their short-run to their long-run cost curves, and this adds no new analytical difficulties: we simply use the long-run supply curves as determined by the long-run marginal cost curve.

However, there is an additional long-run effect that may occur. If a firm is making losses in the long run, there is no reason to stay in the industry, so we would expect to see such a firm *exit* the industry, since by exiting from the industry, the firm could reduce its losses to zero. This is just another way of saying that the only relevant part of a firm's supply curve in the long run is that part that lies *on or above* the average cost curve—since these are locations that correspond to nonnegative profits.

Similarly, if a firm is making profits we would expect *entry* to occur. After all, the cost curve is supposed to include the cost of all factors necessary to produce output, measured at their market price (i.e., their opportunity cost). If a firm is making profits in the long run it means that *anybody* can go to market, acquire those factors, and produce the same amount of output at the same cost.

In most competitive industries there are no restrictions against new firms entering the industry; in this case we say the industry exhibits **free entry**. However, in some industries there are **barriers to entry**, such as licenses or legal restrictions on how many firms can be in the industry. For example, regulations on the sales of alcohol in many states prevent free entry to the retail liquor industry.

The two long-run effects—acquiring different fixed factors and the entry and exit phenomena—are closely related. An existing firm in an industry can decide to acquire a new plant or store and produce more output. Or a new firm may enter the industry by acquiring a new plant and producing output. The only difference is in who owns the new production facilities.

Of course as more firms enter the industry—and firms that are losing money exit the industry—the total amount produced will change and lead to a change in the market price. This in turn will affect profits and the incentives to exit and enter. What will the final equilibrium look like in an industry with free entry?

Let's examine a case where all firms have identical long-run cost functions, say, $c(y)$. Given the cost function we can compute the level of output where average costs are minimized, which we denote by y^*. We let $p^* = c(y^*)/y^*$ be the minimum value of average cost. This cost is significant because it is the lowest price that could be charged in the market and still allow firms to break even.

We can now graph the industry supply curves for each different number of firms that can be in the market. Figure 24.3 illustrates the industry supply curves if there are $1, \ldots, 4$ firms in the market. (We are using 4 firms only for purposes of an example; in reality, one would expect there to be many more firms in a competitive industry.) Note that since all firms have the same supply curve, the total amount supplied if 2 firms are in the market is just twice as much as when 1 firm is the market, the supply when 3 firms are in the market is just three times as much, and so on.

Now add two more lines to the diagram: a horizontal line at p^*, the minimum price consistent with nonnegative profits, and the market demand curve. Consider the intersections of the demand curve and the supply curves for $n = 1, 2, \ldots$ firms. If firms enter the industry when positive profits are being made, then the relevant intersection is the *lowest price consistent with nonnegative profits.* This is denoted by p' in Figure 24.3, and it happens to occur when there are three firms in the market. If one

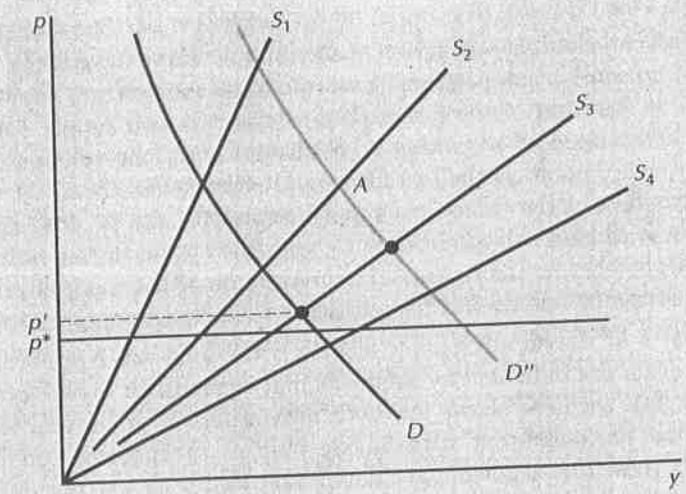


Figure 24.3

Industry supply curves with free entry. Supply curves for $1, \ldots, 4$ firms. The equilibrium price, p', occurs at the lowest possible intersection of demand and supply such that $p' \geq p^*$.

more firm enters the market, profits are pushed to be negative. In this case, the maximum number of competitive firms this industry can support is three.

24.4 The Long-Run Supply Curve

The construction given in the last section—draw the industry supply curves for each possible number of firms that could be in the market and then look for the largest number of firms consistent with nonnegative profits—is perfectly rigorous and easy to apply. However, there is a useful approximation that usually gives something very close to the right answer.

Let's see if there is some way to construct *one* industry supply curve out of the n curves we have above. The first thing to note is that we can rule out all of the points on the supply curve that are below p^*, since those can never be long-run operating positions. But we can also rule out some of the points on the supply curves *above* p^*.

We typically assume that the market demand curve is downward sloping. The steepest possible demand curve is therefore a vertical line. This implies that points like A in Figure 24.3 would never be observed—for any downward-sloping demand curve that passed through A would also have to intersect a supply curve associated with a larger number of firms, as shown by the hypothetical demand curve D'' passing through the point A in Figure 24.3.

Thus we can eliminate a portion of each supply curve from being a possible long-run equilibrium position. Every point on the one-firm supply curve that lies to the right of the intersection of the two-firm supply curve and the line determined by p^* cannot be consistent with long-run equilibrium. Similarly, every point on the two-firm supply curve that lies to the right of the intersection of the three-firm supply curve with the p^* line cannot be consistent with long-run equilibrium ... and every point on the n-firm supply curve that lies to the right of the intersection of the $n + 1$-firm supply curve with the p^* line cannot be consistent with equilibrium.

The parts of the supply curves on which the long-run equilibrium can actually occur are indicated by the black line segments in Figure 24.4. The n^{th} black line segment shows all the combinations of prices and industry output that are consistent with having n firms in long-run equilibrium. Note that these line segments get flatter and flatter as we consider larger and larger levels of industry output, involving more and more firms in the industry.

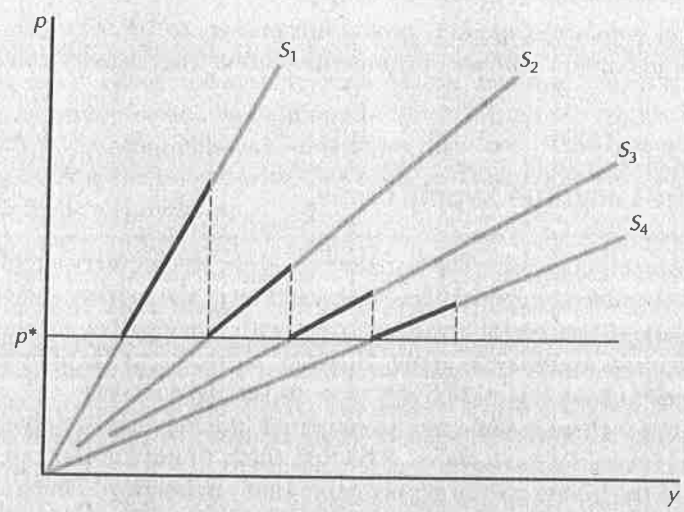
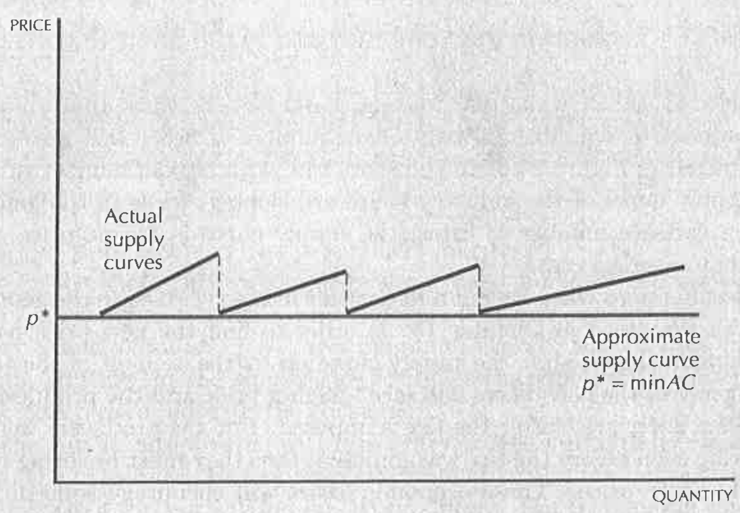


Figure 24.4

The long-run supply curve. We can eliminate portions of the supply curves that can never be intersections with a downward-sloping market demand curve in the long run, such as the points on each supply curve to the right of the dotted lines.

Why do these curves get flatter? Think about it. If there is one firm in the market and the price goes up by Δp, it will produce, say, Δy more output. If there are n firms in the market and the price goes up by Δp, *each* firm will produce Δy more output, so we will get $n\Delta y$ more output in total. This means that the supply curve will be getting flatter and flatter as there are more and more firms in the market, since the supply of output will be more and more sensitive to price.

By the time we get a reasonable number of firms in the market, the slope of the supply curve will be very flat indeed. Flat enough so that it is reasonable to take it as having a slope of zero—that is, as taking the long-run industry supply curve to be a flat line at price equals minimum average cost. This will be a poor approximation if there are only a few firms in the industry in the long run. But the assumption that a small number of firms behave competitively will also probably be a poor approximation! If there are a reasonable number of firms in the long run, the equilibrium price cannot get far from minimum average cost. This is depicted in Figure 24.5.



Approximate long-run supply curve. The long-run supply curve will be approximately flat at price equals minimum average cost.

Figure 24.5

This result has the important implication that in a competitive industry with free entry, profits cannot get very far from zero. If there are significant levels of profits in an industry with free entry, it will induce other firms to enter that industry and thereby push profits toward zero.

Remember, the correct calculation of economic costs involves measuring all factors of production at their market prices. As long as *all* factors are being measured and properly priced, a firm earning positive profits can be exactly duplicated by anyone. Anyone can go to the open market and purchase the factors of production necessary to produce the same output in the same way as the firm in question.

In an industry with free entry and exit, the long-run average cost curve should be essentially flat at a price equal to the minimum average cost. This is just the kind of long-run supply curve that a single firm with constant returns to scale would have. This is no accident. We argued that constant returns to scale was a reasonable assumption since a firm could always replicate what it was doing before. But another firm could replicate it as well! Expanding output by building a duplicate plant is just like a new firm entering the market with duplicate production facilities. Thus the long-run supply curve of a competitive industry with free entry will look like the long-run supply curve of a firm with constant returns to scale: a flat line at price equals minimum average cost.

EXAMPLE: Taxation in the Long Run and in the Short Run

Consider an industry that has free entry and exit. Suppose that initially it is in a long-run equilibrium with a fixed number of firms, and zero profits, as depicted in Figure 24.6. In the short run, with a fixed number of firms, the supply curve of the industry is upward sloping, while in the long run, with a variable number of firms, the supply curve is flat at price equals minimum average cost.

What happens when we put a tax on this industry? We use the geometric analysis discussed in Chapter 16: in order to find the new price paid by the demanders, we shift the supply curve up by the amount of the tax.

In general, the consumers will face a higher price and the producers will receive a lower price after the tax is imposed. But the producers were just breaking even before the tax was imposed; thus they must be losing money at any lower price. These economic losses will encourage some firms to leave the industry. Thus the supply of output will be reduced, and the price to the consumers will rise even further.

In the long run, the industry will supply along the horizontal long-run supply curve. In order to supply along this curve, the firms will have to receive a price equal to the minimum average cost—just what they were receiving before the tax was imposed. Thus the price to the consumers will have to rise by the entire amount of the tax.

In Figure 24.6, the equilibrium is initially at $P_D = P_S$. Then the tax is imposed, shifting the short-run supply curve up by the amount of the tax, and the equilibrium price paid by the demanders increases to P'_D. The

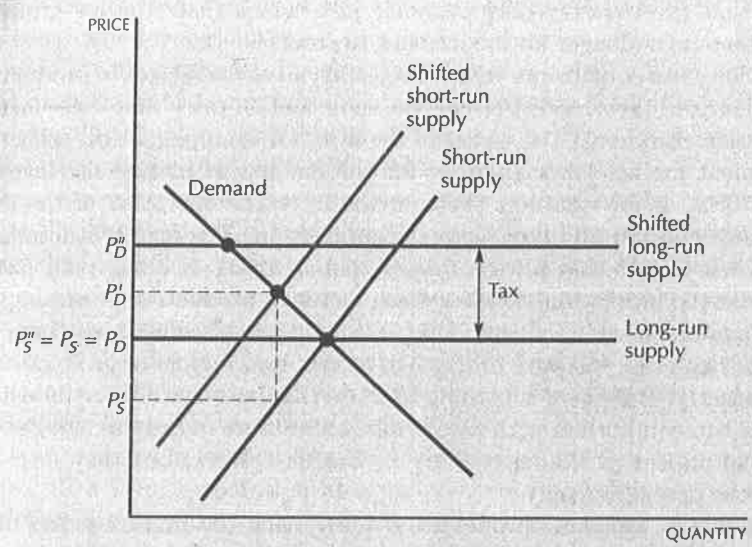


Figure 24.6

Taxation in the short run and long run. In the short run, with a fixed number of firms, the industry supply curve will have an upward slope, so that part of the tax falls on the consumers and part on the firms. In the long run, the industry supply curve will be horizontal so all of the tax falls on the consumers.

equilibrium price received by the suppliers falls to $P'_S = P'_D - t$. But this is only in the short run—when there are a fixed number of firms in the industry. Because of free entry and exit, the *long-run* supply curve in the industry is horizontal at $P_D = P_S =$ minimum average cost. Hence, in the long run, shifting up the supply curve implies that the entire amount of the tax gets passed along to the consumers.

To sum up: in an industry with free entry, a tax will initially raise the price to the consumers by less than the amount of the tax, since some of the incidence of the tax will fall on the producers. But in the long run the tax will induce firms to exit from the industry, thereby reducing supply, so that consumers will eventually end up paying the entire burden of the tax.

24.5 The Meaning of Zero Profits

In an industry with free entry, profits will be driven to zero by new entrants: whenever profits are positive, there will be an incentive for a new firm to come in to acquire some of those profits. When profits are zero it doesn't mean that the industry disappears; it just means that it stops growing, since there is no longer an inducement to enter.

In a long-run equilibrium with zero profits, all of the factors of production are being paid their market price—the same market price that these factors could earn elsewhere. The owner of the firm, for example, is still collecting a payment for her labor time, or for the amount of money she invested in the firm, or for whatever she contributes to the operation of the firm. The same goes for all other factors of production. The firm is still making money—it is just that all the money that it makes is being paid out to purchase the inputs that it uses. Each factor of production is earning the same amount in this industry that it could earn elsewhere, so there are no extra rewards—no pure profits—to attract new factors of production to this industry. But there is nothing to cause them to leave either. Industries in long-run equilibrium with zero profits are mature industries; they're not likely to appear as the cover story in *Business Week*, but they form the backbone of the economy.

Remember, economic profits are defined using the market prices of all factors of production. The market prices measure the opportunity cost of those factors—what they could earn elsewhere. Any amount of money earned in excess of the payments to the factors of production is a pure economic profit. But whenever someone finds a pure economic profit, other people will try to enter the industry and acquire some of that profit for themselves. It is this attempt to capture economic profits that eventually drives them to zero in a competitive industry with free entry.

In some quarters, the profit motive is regarded with some disdain. But when you think about it purely on economic grounds, profits are providing exactly the right signals as far as resource allocation is concerned. If a firm is making positive profits, it means that people value the output of the firm more highly than they value the inputs. Doesn't it make sense to have more firms producing that kind of output?

24.6 Fixed Factors and Economic Rent

If there is free entry, profits are driven to zero in the long run. But not every industry has free entry. In some industries the number of firms in the industry is fixed.

A common reason for this is that there are some factors of production that are available in fixed supply. We said that in the long run the fixed factors could be bought or sold by an individual firm. But there are some factors that are fixed for the *economy as a whole* even in the long run.

The most obvious example of this is in resource-extraction industries: oil in the ground is a necessary input to the oil-extraction industry, and there is only so much oil around to be extracted. A similar statement could be made for coal, gas, precious metals, or any other such resource.

Agriculture gives another example. There is only a certain amount of land that is suitable for agriculture.

A more exotic example of such a fixed factor is talent. There are only a certain number of people who possess the necessary level of talent to be professional athletes or entertainers. There may be "free entry" into such fields—but only for those who are good enough to get in!

There are other cases where the fixed factor is fixed not by nature, but by law. In many industries it is necessary to have a license or permit, and the number of these permits may be fixed by law. The taxicab industry in many cities is regulated in this way. Liquor licenses are another example.

If there are restrictions such as the above on the number of firms in the industry, so that firms cannot enter the industry freely, it may appear that it is possible to have an industry with positive profits in the long run, with no economic forces to drive those profits to zero.

This appearance is wrong. There is an economic force that pushes profits to zero. If a firm is operating at a point where its profits appear to be positive in the long run, it is probably because we are not appropriately measuring the market value of whatever it is that is preventing entry.

Here it is important to remember the economic definition of costs: we should value each factor of production at its *market price*—its opportunity cost. If it appears that a farmer is making positive profits after we have subtracted his costs of production, it is probably because we have forgotten to subtract the cost of his land.

Suppose that we manage to value all of the inputs to farming except for the land cost, and we end up with π dollars per year for profits. How much would the land be worth on a free market? How much would someone pay to rent that land for a year?

The answer is: they would be willing to rent it for π dollars per year, the "profits" that it brings in. You wouldn't even have to know anything about farming to rent this land and earn π dollars—after all, we valued the farmer's labor at its market price as well, and that means that you can hire a farmer and still make π dollars of profit. So the market value of that land—its competitive rent—is just π. The economic profits to farming are zero.

Note that the rental rate determined by this procedure may have nothing whatsoever to do with the historical cost of the farm. What matters is not what you bought it for, but what you can sell it for—that's what determines opportunity cost.

Whenever there is some fixed factor that is preventing entry into an industry, there will be an equilibrium rental rate for that factor. Even with fixed factors, you can always enter an industry by buying out the position of a firm that is currently in the industry. Every firm in the industry has the option of selling out—and the opportunity cost of not doing so is a cost of production that it has to consider.

Thus in one sense it is always the *possibility* of entry that drives profits to zero. After all, there are two ways to enter an industry: you can form a new firm, or you can buy out an existing firm that is currently in the industry. If a new firm can buy everything necessary to produce in an industry and still make a profit, it will do so. But if there are some factors that are in fixed supply, then competition for those factors among potential entrants will bid the prices of these factors up to a point where the profit disappears.

EXAMPLE: Taxi Licenses in New York City

Earlier we said that licenses to operate New York City taxicabs sell for about $100,000. Yet in 1986 taxicab drivers made only about $400 for a 50-hour week; this translated into less than an $8 hourly wage. The New York Taxi and Limosine Commission argued that this wage was too low to attract skilled drivers and that taxi fares should be raised in order to attract better drivers.

An economist would argue that allowing the fares to increase would have virtually no effect on the take-home pay of the drivers; all that would happen is that the value of the taxicab license would increase. We can see why by examining the commission's figures for the costs of operating a taxi. In 1986, the lease rate was $55 for a day shift and $65 for a night shift. The driver who leased the taxi paid for the gasoline and netted about $80 a day in income.

But note how much the owner of the taxicab license made. Assuming that the cab could be rented for two shifts for 320 days a year, the lease income comes to $38,400. Insurance, depreciation, maintenance, and so on amounted to about $21,100 a year; this leaves a net profit of $17,300 per year. Since the license cost about $100,000, this indicates a total return of about 17 percent.

An increase in the rate that taxis were allowed to charge would be reflected directly in the value of the license. A fare increase that brought in an extra $10,000 a year would result in a license's value increasing by about $60,000. The wage rate for the cab drivers—which is set in the labor market—would not be affected by such a change.[1]

24.7 Economic Rent

The examples in the last section are instances of **economic rent**. Economic rent is defined as those payments to a factor of production that are in excess of the minimum payment necessary to have that factor supplied.

[1] Figures are taken from an unsigned editorial in the *New York Times,* August 17, 1986.

Consider, for example, the case of oil discussed earlier. In order to produce oil you need some labor, some machinery, and, most importantly, some oil in the ground! Suppose that it costs \$1 a barrel to pump oil out of the ground from an existing well. Then any price in excess of \$1 a barrel will induce firms to supply oil from existing wells. But the actual price of oil is much higher than \$1 a barrel. People want oil for various reasons, and they are willing to pay more than its cost of production to get it. The excess of the price of oil over its cost of production is economic rent.

Why don't firms enter this industry? Well, they try. But there is only a certain amount of oil available. Oil will sell for more than its cost of production because of the limited supply.

Now consider taxicab licenses. Viewed as pieces of paper, these cost almost nothing to produce. But in New York City a taxicab license can sell for \$100,000! Why don't people enter this industry and produce more taxicab licenses? The reason is that entry is illegal—the supply of taxicab licenses is controlled by the city.

Farmland is yet another example of economic rent. In the aggregate, the total amount of land is fixed. There would be just as much land supplied at zero dollars an acre as at \$1000 an acre. Thus in the aggregate, the payments to land constitute economic rent.

From the viewpoint of the economy as a whole, it is the price of agricultural products that determines the value of agricultural land. But from the viewpoint of the individual farmer, the value of his land is a cost of production that enters into the pricing of his product.

This is depicted in Figure 24.7. Here AVC represents the average cost curve for all factors of production *excluding* land costs. (We are assuming that land is the only fixed factor.) If the price of the crop grown on this land is p^*, then the "profits" attributable to the land are measured by the area of the box: these are the economic rents. This is how much the land would rent for in a competitive market—whatever it took to drive the profits to zero.

The average cost curve *including* the value of the land is labeled AC. If we measure the value of the land correctly, the economic profits to operating the farm will be exactly zero. Since the equilibrium rent for the land will be whatever it takes to drive profits to zero, we have

$$p^*y^* - c_v(y^*) - \text{rent} = 0$$

or

$$\text{rent} = p^*y^* - c_v(y^*). \tag{24.1}$$

This is precisely what we referred to as producer's surplus earlier. Indeed, it is the same concept, simply viewed in a different light. Thus we can also measure rent by taking the area to the left of the marginal cost curve, as we saw earlier.

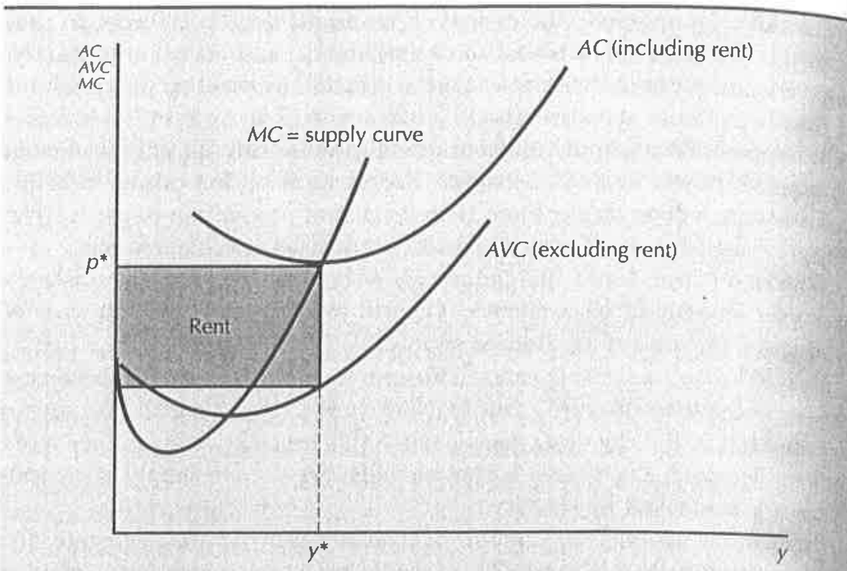


Figure 24.7 **Economic rent for land.** The area of the box represents the economic rent on the land.

Given the definition of rent in equation (24.1), it is now easy to see the truth of what we said earlier: it is the equilibrium price that determines rent, not the reverse. The firm supplies along its marginal cost curve—which is independent of the expenditures on the fixed factors. The rent will adjust to drive profits to zero.

24.8 Rental Rates and Prices

Since we are measuring output in flow units—so much output per unit of time, we should be careful to measure profits and rents in dollars per unit of time. Thus in the above discussion we talked about the rent per year for land or for a taxicab license.

If the land or the license is to be sold outright rather than rented, the equilibrium price would be the present value of the stream of rental payments. This is a simple consequence of the usual argument that assets generating a stream of payments should sell for their present values in a competitive market.

EXAMPLE: Liquor Licenses

In the United States, each state sets its own policy with respect to sales of alcohol. Some states have a liquor monopoly; other states issue licenses to those who wish to sell alcohol. In some cases, licenses are issued on payment of a fee; in other cases, the number of licenses is fixed. In Michigan, for example, the number of licenses for sales of beer and wine for consumption on premises is limited to one for every 1,500 residents.

After each Federal census, a state liquor control board allocates licenses to communities whose populations have grown. (Licenses are not taken away from communities whose populations have fallen, however.) This artificial scarcity of licenses has created a vibrant market for licenses to serve liquor in many fast-growing communities. For example, in 1983 Ann Arbor, Michigan, had sixty-six existing liquor licenses. Six new licenses were allowed to be issued as a result of the 1980 census, and 33 applicants lined up to lobby for these licenses. At the time, the market value of a liquor license was about $80,000. The local newspaper ran a story asserting that "demand exceeds supply for liquor licenses." It was hardly surprising to the local economists that giving away an $80,000 asset for a zero price resulted in excess demand!

There have been many proposals to relax the liquor control laws in Michigan by allowing the state to issue new licenses. However, these proposals have never been enacted into law due to the opposition of various political groups. Some of these groups are opposed to the consumption of alcohol on grounds of public health or religion. Others have somewhat different motives. For example, one of the most vociferous opponents of relaxed liquor laws is the Michigan Licensed Beverage Association, a group that represents the sellers of alcoholic beverages in Michigan. Though at first glance it appears paradoxical that this group would oppose liberalization of the liquor laws, a little reflection indicates a possible reason: issuing more liquor licenses would undoubtably lower the resale value of *existing* licenses—imposing significant capital losses on current holders of such licenses.

24.9 The Politics of Rent

Often economic rent exists because of legal restrictions on entry into the industry. We mentioned two examples above: taxicab licenses and liquor licenses. In each of these cases the number of licenses is fixed by law, thus restricting entry to the industry and creating economic rents.

Suppose that the New York City government wants to increase the number of operating taxicabs. What will happen to the market value of the existing taxicab licenses? Obviously they will fall in value. This reduction in value hits the industry right in the pocketbook, and it is sure to create a lobbying force to oppose any such move.

The federal government also artificially restricts output of some products in such a way as to create a rent. For example, the federal government has declared that tobacco can only be grown on certain lands. The value of this land is then determined by the demand for tobacco products. Any attempt to eliminate this licensing system has to contend with a serious lobby. Once the government creates artificial scarcity, it is very hard to eliminate it. The beneficiaries of the artificial scarcity—the people who have acquired the right to operate in the industry—will vigorously oppose any attempts to enlarge the industry.

The incumbents in an industry in which entry is legally restricted may well devote considerable resources to maintaining their favored position. Lobbying expenses, lawyers' fees, public relations costs, and so on can be substantial. From the viewpoint of society these kinds of expenses represent pure social waste. They aren't true costs of production; they don't lead to any *more* output being produced. Lobbying and public relations efforts just determine who gets the money associated with existing output.

Efforts directed at keeping or acquiring claims to factors in fixed supplies are sometimes referred to as **rent seeking**. From the viewpoint of society they represent a pure deadweight loss since they don't create any more output, they just change the market value of existing factors of production.

EXAMPLE: Farming the Government

There is only one good thing to say about the U.S. program of farm subsidies: it produces a never-ending source of examples for economics textbooks. Every new reform of the farm program brings new problems. "If you want to find the holes in a program, just toss them out to farmers. No one is more innovative in finding ways to use them," says Terry Bar, the vice president of the National Council of Farm Cooperatives.[2]

Up until 1996 the basic structure of farm subsidies in the U.S. involved price supports: the Federal government guaranteed a support price for a crop and would make up the difference if the price fell below the support price. In order to qualify for this program, a farmer had to agree not to farm a certain fraction of his land.

By the very nature of this plan, most of the benefits accrued to the large farmers. According to one calculation, 13 percent of the direct Federal subsidies were going to the 1 percent of the farmers who had sales over $500,000 a year. The Food Security Act of 1985 significantly restricted the payments to large farmers. As a result, the farmers broke up their holdings by leasing the land to local investors. The investors would acquire parcels large enough to take advantage of the subsidies, but too small to run into the restrictions aimed at large farmers. Once the land was acquired the investor would register it with a government program that would pay the

[2] Quoted in William Robbins, "Limits on Subsidies to Big Farms Go Awry, Sending Costs Climbing," *New York Times*, June 15, 1987, A1.

investor *not* to plant the land. This practice became known as "farming the government."

According to one study, the restriction on payments to the large farmers in the 1985 farm act resulted in the creation of 31,000 new applicants for farm subsidies. The cost of these subsidies was in the neighborhood of $2.3 billion.

Note that the ostensible goal of the program—restricting the amount of government subsidies paid to large farmers—has not been achieved. When the large farmers rent their land to small farmers, the market price of the rents depends on the generosity of the Federal subsidies. The higher the subsidies, the higher the equilibrium rent the large farmers receive. The benefits from the subsidy program still falls on those who initially own the land, since it is ultimately the value of what the land can earn—either from growing crops or farming the government—that determines its market value.

The Farm Act of 1996 promised a phaseout of most agricultural subsidies by 2002. However, the 1998 federal budget restored over 6 billion dollars of federal farm subsidies, illustrating once again how hard it is to reconcile politics and economics.

24.10 Energy Policy

We end this chapter with an extended example that uses some of the concepts we have developed.

In 1974 the Organization of Petroleum Exporting Countries (OPEC) levied a significant increase in the price of oil. Countries that had no domestically produced petroleum had little choice about energy policy—the price of oil and goods produced using oil had to rise.

At that time the United States produced about half of its domestic oil consumption, and Congress felt that it was unfair that the domestic producers should receive "windfall profits" from an uncontrolled increase in price. (The term windfall profits refers to an increase in profits due to some outside event, as opposed to an increase in profits due to production decisions.) Consequently, Congress devised a bizarre plan to attempt to hold down the price of products that used oil. The most prominent of these products is gasoline, so we will analyze the effect of the program for that market.

Two-Tiered Oil Pricing

The policy adopted by Congress was known as "two-tiered" oil pricing, and it went something like this. Imported oil would sell for whatever its market price was, but domestic oil—oil produced from wells that were in place before 1974—would sell for its old price: the price that it sold for before OPEC. Roughly speaking, we'll say that imported oil sold for about \$15 a barrel, while domestic oil sold for around \$5. The idea was that the average price of oil would then be about \$10 a barrel and this would help hold down the price of gasoline.

Could such a scheme work? Let's think about it from the viewpoint of the gasoline producers. What would the supply curve of gasoline look like? In order to answer this question we have to ask what the marginal cost curve for gasoline looked like.

What would you do if you were a gasoline refiner? Obviously you would try to use the cheap domestic oil first. Only after you had exhausted your supplies of domestic oil would you turn to the more expensive imported oil. Thus the aggregate marginal cost curve—the industry supply curve—for gasoline would have to look something like that depicted in Figure 24.8. The curve takes a jump at the point where the U.S. production of domestic oil is exhausted and the imported oil begins to be used. Before that point, the domestic price of oil measures the relevant factor price for producing gasoline. After that point, it is the price of foreign oil that is the relevant factor price.

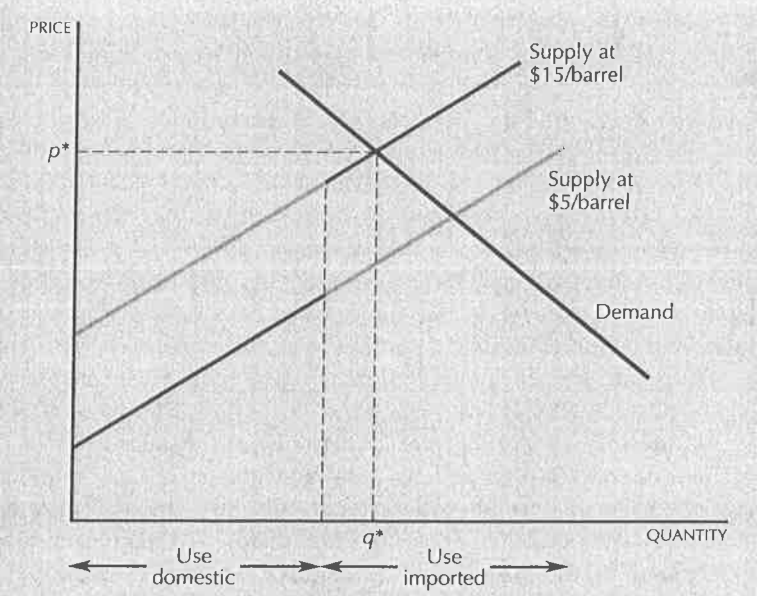
Figure 24.8 depicts the supply curve for gasoline if all oil were to sell for the world price of \$15 a barrel, and if all oil were to sell for the domestic price of \$5 a barrel. If domestic oil actually sells for \$5 a barrel and foreign oil sells for \$15 a barrel, then the supply curve for gasoline will coincide with the \$5-a-barrel supply curve until the cheaper domestic oil is used up, and then coincide with the \$15-a-barrel supply curve.

Now let's find the intersection of this supply curve with the market demand curve to find the equilibrium price in Figure 24.8. The diagram reveals an interesting fact: the price of gasoline is exactly the same in the two-tiered system as it would be if all oil sold at the price of foreign oil! The price of gasoline is determined by the *marginal* cost of production, and the *marginal* cost is determined by the cost of the imported oil.

If you think about it a minute, this makes perfectly good sense. The gasoline companies will sell their product at the price the market will bear. Just because you were lucky enough to get some cheap oil doesn't mean you won't sell your gasoline for the same price that other firms are selling theirs for.

Suppose for the moment that all oil did sell for one price, and that equilibrium was reached at the price p^*. Then the government comes along and lowers the price of the first 100 barrels of oil that each refiner used. Will this affect their supply decision? No way—in order to affect supply you have to change the incentives at the margin. The only way to get a lower price of gasoline is to increase the supply, which means that you have to make the marginal cost of oil cheaper.

The two-tiered oil pricing policy was simply a transfer from the domestic oil producers to the domestic oil refiners. The domestic producers got \$10



The supply curve for gasoline. Under the two-tiered oil pricing policy, the supply curve of gasoline would be discontinuous, jumping from the lower supply curve to the upper supply curve when the cheaper oil was exhausted.

Figure 24.8

less for their oil than they would have otherwise, and the profits they would have gotten went to the gasoline refiners. It had no effect on the supply of gasoline, and thus it could have no effect on the price of gasoline.

Price Controls

The economic forces inherent in this argument didn't take long to make themselves felt. The Department of Energy soon realized that it couldn't allow market forces to determine the price of gasoline under the two-tiered system—since market forces alone would imply one price of gasoline, which would be the same price that would prevail in the absence of the two-tiered system.

So they instituted price controls on gasoline. Each refiner was required to charge a price for gasoline that was based on the costs of producing the gasoline—which in turn was primarily determined by the cost of the oil that the refiner was able to purchase.

The availability of cheap domestic oil varied with location. In Texas the refiners were close to the major source of production and thus were able to purchase large supplies of cheap oil. Due to the price controls, the price of Texas gasoline was relatively cheap. In New England, virtually all oil had to be imported, and thus the price of gasoline in New England was quite high.

When you have different prices for the same product, it is natural for firms to try to sell at the higher price. Again, the Department of Energy had to intervene to prevent the uncontrolled shipping of gasoline from low-price regions to high-price regions. The result of this intervention was the famous gasoline shortages of the mid-seventies. Periodically, the supply of gasoline in a region of the country would dry up, and there would be little available at any price. The free market system of supplying petroleum products had never exhibited such behavior; the shortages were entirely due to the two-tiered oil pricing system coupled with price controls.

Economists pointed this out at the time, but it didn't have much effect on policy. What did have an effect was lobbying by the gasoline refiners. Much of the domestic oil was sold on long-term contracts, and some refiners were able to buy a lot of it, while others could only buy the expensive foreign oil. Naturally they objected that this was unfair, so Congress figured out another scheme to allocate the cheap domestic oil more equitably.

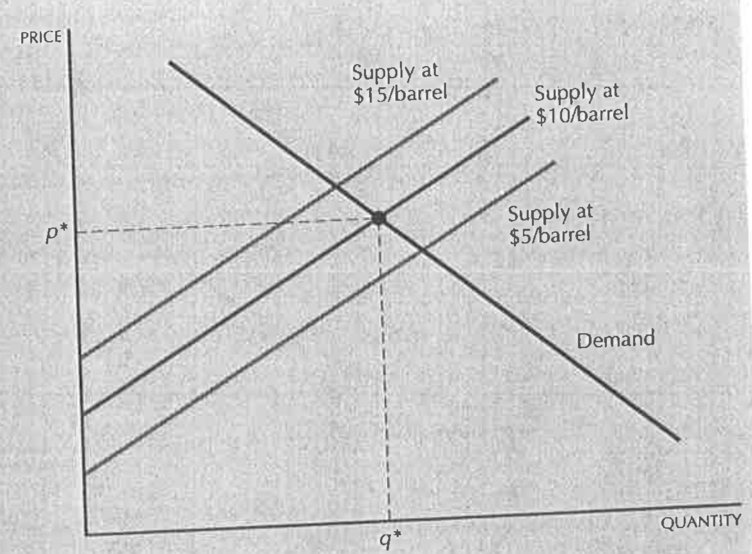
The Entitlement Program

This program was known as the "entitlement program," and it went something like this. Each time a refiner bought a barrel of expensive foreign oil he got a coupon that allowed him to buy a certain amount of cheap domestic oil. The amount that the refiner was allowed to buy depended on supply conditions, but let's say that it was one for one: each barrel of foreign oil that he bought for $15 allowed him to buy one barrel of domestic oil for $5.

What did this do to the marginal price of oil? Now the marginal price of oil was just a weighted average of the domestic price and the foreign price of oil; in the one-for-one case described above, the price would be $10. The effect on the supply curve of gasoline is depicted in Figure 24.9.

The marginal cost of oil was reduced all right, and that meant that the price of gasoline was reduced as well. But look who is paying for it: the domestic oil producers! The United States was buying foreign oil that cost $15 a barrel in real dollars and pretending that it only cost $10. The domestic oil producers were required to sell their oil for less than the market price on the world oil market. We were subsidizing the importation of foreign oil and forcing the domestic oil producers to pay the subsidy!

Eventually this program was abandoned as well, and the U.S. imposed a tax on the domestic production of oil so that the U.S. oil producers wouldn't reap windfall profits due to OPEC's action. Of course, such a tax



The entitlement program. Under the entitlement program the supply curve of gasoline would lie between the supply curve if all oil were provided at the imported price and the supply curve if all oil were provided at the domestic price.

Figure 24.9

discouraged production of domestic oil, and thereby increases the price of gasoline, but this was apparently acceptable to Congress at the time.

24.11 Carbon Tax Versus Cap and Trade

Motivated by concerns about global warming, several climatologists have urged governments to institute policies to reduce carbon emissions. Two of these reduction policies are particularly interesting from an economic point of view: **carbon taxes** and **cap and trade**.

A carbon tax imposes a tax on carbon emissions, while a cap and trade system grants licenses to emit carbon that can be traded on an organized market. To see how these systems compare, let us examine a simple model.

Optimal Production of Emissions

We begin by examining the problem of producing a target amount of emissions in the least costly way. Suppose that there are two firms that have

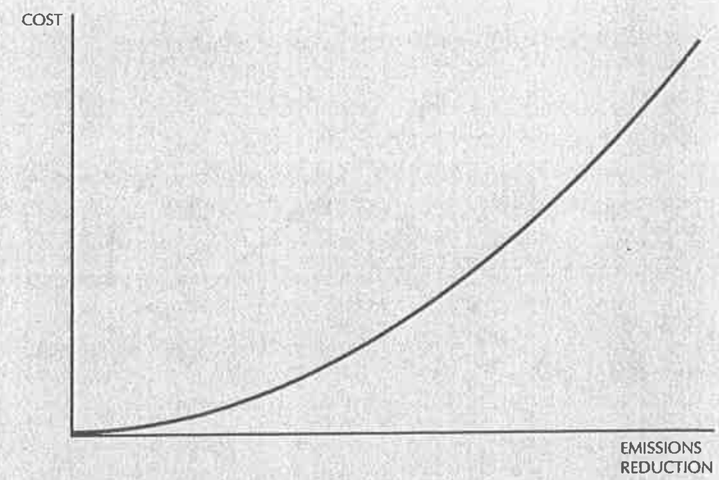


Figure 24.10

Cost function for emissions. The curve shows the cost associated with emission reductions.

current levels of carbon emissions denoted by $(\overline{x}_1, \overline{x}_2)$. Firm i can reduce its level of emissions by x_i at a cost of $c_i(x_i)$. Figure 24.10 shows a possible shape for this cost function.

The goal is to reduce emissions by some target amount, T, in the least costly way. This minimization problem can be written as

$$\min_{x_1, x_2} \; c_1(x_1) + c_2(x_2)$$
$$\text{such that } x_1 + x_2 = T.$$

If it knew the cost functions, the government could, in principle, solve this optimization problem and assign a specific amount of emission reductions to each firm. However, this is impractical if there are thousands of carbon emitters. The challenge is to find a decentralized, market-based way of achieving the optimal solution.

Let us examine the structure of the optimization problem. It is clear that at the optimal solution the marginal cost of reducing emissions must be the same for each firm. Otherwise it would pay to increase emissions in the firm with the lower marginal cost and decrease emissions in the firm with the higher marginal cost. This would keep the total output at the target level while reducing costs.

Hence we have a simple principle: at the optimal solution, the marginal cost of emissions reduction should be the same for every firm. In the two-firm case we are examining, we can find this optimal point using a simple diagram. Let $MC_1(x_1)$ be the marginal cost of reducing emissions by x_1

for firm 1 and write the marginal cost of emission-reduction for firm 2 as a function of firm 1's output: $MC_2(T - x_1)$, assuming the target is met. We plot these two curves in Figure 24.11. The point where they intersect determines the optimal division of emission reductions between the two firms given that T emission reductions are to be produced in total.

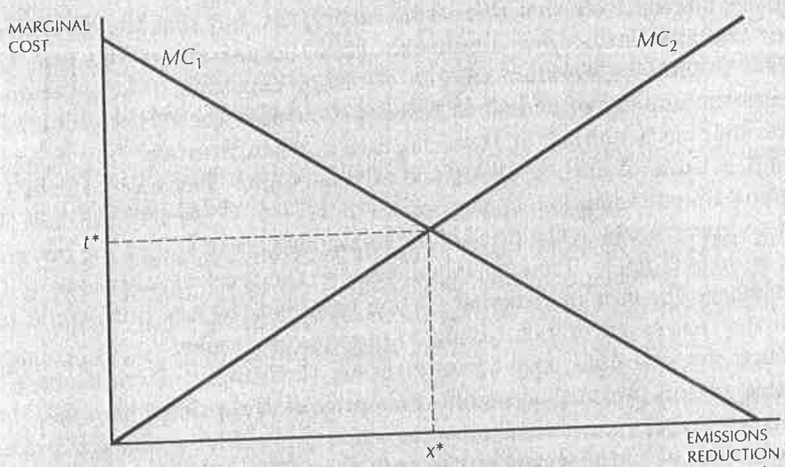


Figure 24.11

Equilibrium in the cap and trade market. The point t^* gives the optimal carbon tax and emissions license price.

A Carbon Tax

Instead of solving for the cost-minimizing solution directly, let us instead consider a decentralized solution using a carbon tax. In this framework, the government sets a tax rate t that it charges for carbon emissions.

If firm 1 starts with $\overline{x}_1$ and reduces its emissions by x_1, then it ends up with $\overline{x}_1 - x_1$ emissions. If it pays t per unit emitted, its carbon tax bill would be $t(\overline{x}_1 - x_1)$.

Faced with this tax, firm 1 would want to choose that level of emission reductions that minimized its total cost of operation: the cost of reducing emissions plus the cost of paying the carbon tax on the emissions that remain. This leads to the cost minimization problem

$$\min_{x_1} c_1(x_1) + t(\overline{x}_1 - x_1).$$

Clearly the firm will want to reduce emissions up to the point where the marginal cost of further reductions just equals the carbon tax, i.e., where $t = MC_1(x_1)$.

If the carbon tax is set to be the rate t^*, as determined in Figure 24.11, then the total amount of carbon emissions will be the targeted amount, T. Thus the carbon tax gives a decentralized way to achieve the optimal outcome.

Cap and Trade

Suppose, alternatively that there is no carbon tax, but that the government issues tradable **emissions licenses**. Each license allows the firm that holds it to produce a certain amount of carbon emissions. The government chooses the number of emissions licenses to achieve the target reduction.

We imagine a market in these licenses so each firm can buy a license to emit x units of carbon at a price of p per unit. The cost to firm 1 of reducing its emissions by x_1 is $c_1(x_1) + p(\overline{x}_1 - x_1)$. Clearly the firm will want to operate where the price of an emissions license equals the marginal cost, $p = MC_1(x_1)$. That is, it will choose the level of emissions at the point where the cost of reducing carbon emissions by one unit would just equal the cost saved by not having to purchase a license.

Hence the marginal cost curve gives us the supply of emissions as a function of the price. The equilibrium price is the price where the total supply of emissions equals the target amount T. The associated price is the same as the optimal carbon tax rate t^* in Figure 24.11.

The question that remains is how to distribute the licenses. One way would be to have the government sell the licenses to firms. This is essentially the same as the carbon tax system. The government could pick a price and sell however many licenses are demanded at that price. Alternatively, it could pick a target level of emissions and auction off permits, letting the firms themselves determine a price. This is one type of "cap and trade" system. Both of these policies should lead to essentially the same market-clearing price.

Another possibility would be for the government to hand out the licenses to the firms according to some formula. This formula could be based on a variety of criteria, but presumably an important reason to award these valuable permits would be building political support for the program. Permits might be handed out based on objective criteria, such as which firms have the most employees, or they might be handed out based on which firms have donated the most to some political causes.

From the economic point of view, it doesn't matter whether the government owns the licenses and sells them to the firms (which is basically a carbon tax system) or whether the firms are given the licenses and sell them to each other (which is basically cap and trade).

If a cap and trade system is created, firms will find it attractive to invest in ways to acquire the emission permits. For example, they would want to lobby Congress for such licenses. These lobbying expenditures should be counted as part of the cost of the system, as described in our earlier discussion of **rent seeking**. Of course, the carbon tax system would also be subject to similar lobbying. Firms would undoubtedly seek special carbon tax exemptions for one reason or another, but it has been argued that the carbon tax system is less susceptible to political manipulation than a cap and trade system.

Summary

1. The short-run supply curve of an industry is just the horizontal sum of the supply curves of the individual firms in that industry.

2. The long-run supply curve of an industry must take into account the exit and entry of firms in the industry.

3. If there is free entry and exit, then the long-run equilibrium will involve the maximum number of firms consistent with nonnegative profits. This means that the long-run supply curve will be essentially horizontal at a price equal to the minimum average cost.

4. If there are forces preventing the entry of firms into a profitable industry, the factors that prevent entry will earn economic rents. The rent earned is determined by the price of the output of the industry.

REVIEW QUESTIONS

1. If $S_1(p) = p - 10$ and $S_2(p) = p - 15$, then at what price does the industry supply curve have a kink in it?

2. In the short run the demand for cigarettes is totally inelastic. In the long run, suppose that it is perfectly elastic. What is the impact of a cigarette tax on the price that consumers pay in the short run and in the long run?

3. True or false? Convenience stores near the campus have high prices because they have to pay high rents.

4. True or false? In long-run industry equilibrium no firm will be losing money.

5. According to the model presented in this chapter, what determines the amount of entry or exit a given industry experiences?

6. The model of entry presented in this chapter implies that the more firms in a given industry, the (steeper, flatter) is the long-run industry supply curve.

7. A New York City cab operator appears to be making positive profits in the long run after carefully accounting for the operating and labor costs. Does this violate the competitive model? Why or why not?