costs of development. It also reinforced the tendency toward industrial localization, as most of these specialized inputs were not available elsewhere in the country."[1]

As the quote suggests, the availability of this dense network of specialized suppliers has given high-technology firms in Silicon Valley some considerable advantages over firms elsewhere. Key inputs are cheaper and more easily available because there are many firms competing to provide them, and firms can concentrate on what they do best, contracting out other aspects of their business. For example, some Silicon Valley firms that specialize in providing highly sophisticated computer chips for particular customers have chosen to become "fabless," that is, they do not have any factories in which chips can be fabricated. Instead, they concentrate on designing the chips, and then hire another firm to actually fabricate them.

A company that tried to enter the industry in another location—for example, in a country that did not have a comparable industrial cluster—would be at an immediate disadvantage because it would lack easy access to Silicon Valley's suppliers and would either have to provide them for itself or be faced with the task of trying to deal with Silicon Valley–based suppliers at long distance.

Labor Market Pooling

A second source of external economies is the way that a cluster of firms can create a pooled market for workers with highly specialized skills. Such a pooled market is to the advantage of both the producers and the workers, as the producers are less likely to suffer from labor shortages and the workers are less likely to become unemployed.

The point can best be made with a simplified example. Imagine that there are two companies that both use the same kind of specialized labor, say, two film studios that make use of experts in computer animation. Both employers are, however, uncertain about how many workers they will want to hire: If demand for their product is high, both companies will want to hire 150 workers, but if it is low, they will want to hire only 50. Suppose also that there are 200 workers with this special skill. Now compare two situations: one with both firms and all 200 workers in the same city, the other with the firms, each with 100 workers, in two different cities. It is straightforward to show that both the workers and their employers are better off if everyone is in the same place.

First, consider the situation from the point of view of the companies. If they are in different locations, whenever one of the companies is doing well, it will be confronted with a labor shortage: It will want to hire 150 workers, but only 100 will be available. If the firms are near each other, however, it is at least possible that one will be doing well when the other is doing badly, so both firms may be able to hire as many workers as they want. By locating near each other, the companies increase the likelihood that they will be able to take advantage of business opportunities.

From the workers' point of view, having the industry concentrated in one location is also an advantage. If the industry is divided between two cities, then whenever one of the firms has a low demand for workers, the result will be unemployment: The firm will be willing to hire only 50 of the 100 workers who live nearby. But if the industry is concentrated in a single city, low labor demand from one firm will at least sometimes be offset by high demand from the other. As a result, workers will have a lower risk of unemployment.

Again, these advantages have been documented for Silicon Valley, where it is common both for companies to expand rapidly and for workers to change employers. The same study of Silicon Valley that was quoted previously notes that the concentration of firms in

[1] See p. 40 of the book by Saxenian listed in Further Readings.

a single location makes it easy to switch employers. One engineer is quoted as saying that "it wasn't that big a catastrophe to quit your job on Friday and have another job on Monday. . . . You didn't even necessarily have to tell your wife. You just drove off in another direction on Monday morning."[2] This flexibility makes Silicon Valley an attractive location both for highly skilled workers and for the companies that employ them.

Knowledge Spillovers

It is by now a cliché that in the modern economy, knowledge is at least as important an input as are factors of production like labor, capital, and raw materials. This is especially true in highly innovative industries, where being even a few months behind the cutting edge in production techniques or product design can put a company at a major disadvantage.

But where does the specialized knowledge that is crucial to success in innovative industries come from? Companies can acquire technology through their own research and development efforts. They can also try to learn from competitors by studying their products and, in some cases, by taking them apart to "reverse engineer" their design and manufacture. An important source of technical know-how, however, is the informal exchange of information and ideas that takes place at a personal level. And this kind of informal diffusion of knowledge often seems to take place most effectively when an industry is concentrated in a fairly small area, so that employees of different companies mix socially and talk freely about technical issues.

Marshall described this process memorably when he wrote that in a district with many firms in the same industry, "The mysteries of the trade become no mystery, but are as it were in the air. . . . Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organization of the business have their merits promptly discussed: If one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas."[3]

A journalist described how these knowledge spillovers worked during the rise of Silicon Valley (and also gave an excellent sense of the amount of specialized knowledge involved in the industry) as follows: "Every year there was some place, the Wagon Wheel, Chez Yvonne, Rickey's, the Roundhouse, where members of this esoteric fraternity, the young men and women of the semiconductor industry, would head after work to have a drink and gossip and trade war stories about phase jitters, phantom circuits, bubble memories, pulse trains, bounceless contacts, burst modes, leapfrog tests, p-n junctions, sleeping sickness modes, slow-death episodes, RAMs, NAKs, MOSes, PCMs, PROMs, PROM blowers, PROM blasters, and teramagnitudes. . . ."[4] This kind of informal information flow means that it is easier for companies in the Silicon Valley area to stay near the technological frontier than it is for companies elsewhere; indeed, many multinational firms have established research centers and even factories in Silicon Valley simply in order to keep up with the latest technology.

External Economies and Market Equilibrium

As we've just seen, a geographically concentrated industry is able to support specialized suppliers, provide a pooled labor market, and facilitate knowledge spillovers in a way that a geographically dispersed industry cannot. But the strength of these economies presumably depends on the industry's size: Other things equal, a bigger industry will generate stronger external economies. What does this say about the determination of output and prices?

---

[2] Saxenian, p. 35.

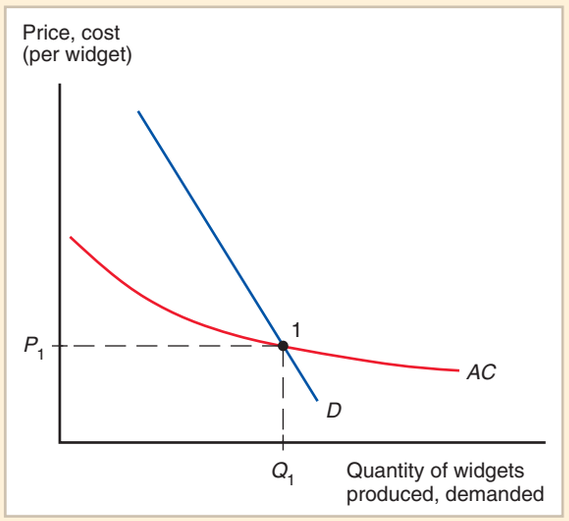
[3] Alfred Marshall, *Principles of Economics* (London: MacMillan, 1920).

[4] Tom Wolfe, quoted in Saxenian, p. 33.

Figure 7-1

External Economies and Market Equilibrium

When there are external economies of scale, the average cost of producing a good falls as the quantity produced rises. Given competition among many producers, the downward-sloping average cost curve *AC* can be interpreted as a *forward-falling supply curve*. As in ordinary supply-and-demand analysis, market equilibrium is at point 1, where the supply curve intersects the demand curve, *D*. The equilibrium level of output is Q_1, the equilibrium price P_1.



While the details of external economies in practice are often quite subtle and complex (as the example of Silicon Valley shows), it can be useful to abstract from the details and represent external economies simply by assuming that the larger the industry, the lower the industry's costs. If we ignore international trade for the moment, then market equilibrium can be represented with a supply-and-demand diagram like Figure 7-1, which illustrates the market for widgets. In an ordinary picture of market equilibrium, the demand curve is downward sloping, while the supply curve is upward sloping. In the presence of external economies of scale, however, there is a **forward-falling supply curve:** the larger the industry's output, the lower the price at which firms are willing to sell, because their **average cost of production** falls as industry output rises.

In the absence of international trade, the unusual slope of the supply curve in Figure 7-1 doesn't seem to matter much. As in a conventional supply-and-demand analysis, the equilibrium price, P_1, and output, Q_1, are determined by the intersection of the demand curve and the supply curve. As we'll see next, however, external economies of scale make a huge difference to our view of the causes and effects of international trade.

External Economies and International Trade

External economies drive a lot of trade both within and between countries. For example, New York exports financial services to the rest of the United States, largely because external economies in the investment industry have led to a concentration of financial firms in Manhattan. Similarly, Britain exports financial services to the rest of Europe, largely because those same external economies have led to a concentration of financial firms in London. But what are the implications of this kind of trade? We'll look first at the effects of trade on output and prices; then at the determinants of the pattern of trade; and finally at the effects of trade on welfare.

External Economies, Output, and Prices

Imagine, for a moment, that we live in a world in which it is impossible to trade buttons across national borders. Assume, also, that there are just two countries in this world,

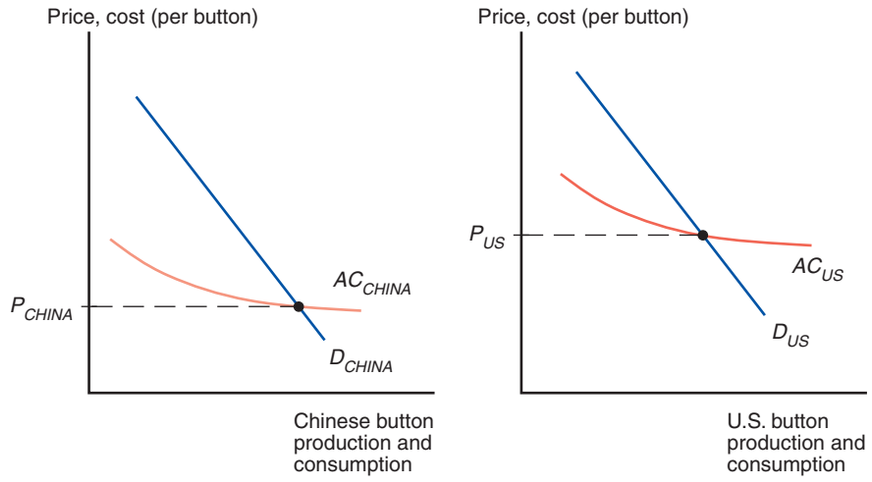


Figure 7-2

External Economies Before Trade

In the absence of trade, the price of buttons in China, P_{CHINA}, is lower than the price of buttons in the United States, P_{US}.

China and the United States. Finally, assume that production of buttons is subject to external economies of scale, which lead to a forward-falling supply curve for buttons in each country. (As the Case Study on page 147 shows, this is actually true of the button industry.)

In that case, equilibrium in the world button industry would look like the situation shown in Figure 7-2.[5] In both China and the United States, equilibrium prices and output would be at the point where the domestic supply curve intersects the domestic demand curve. In the case shown in Figure 7-2, Chinese button prices in the absence of trade would be lower than U.S. button prices.

Now suppose that we open up the potential for trade in buttons. What will happen?

It seems clear that the Chinese button industry will expand, while the U.S. button industry will contract. And this process will feed on itself: As the Chinese industry's output rises, its costs will fall further; as the U.S. industry's output falls, its costs will rise. In the end, we can expect all button production to be concentrated in China.

The effects of this concentration are illustrated in Figure 7-3. Before the opening of trade, China supplied only its own domestic button market. After trade, it supplies the world market, producing buttons for both Chinese and U.S. consumers.

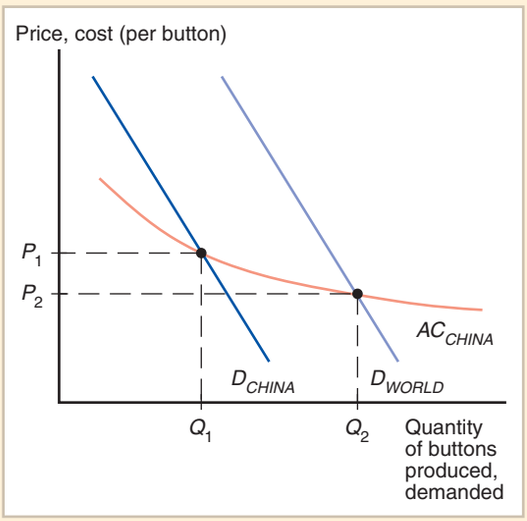
Notice the effects of this concentration of production on prices. Because China's supply curve is forward-falling, increased production as a result of trade leads to a button price that is lower than the price before trade. And bear in mind that Chinese button prices were lower than American button prices before trade. What this tells us is that trade leads to button prices that are lower than the prices in *either* country before trade.

[5] In this exposition, we focus for simplicity on *partial equilibrium* in the market for buttons, rather than on general equilibrium in the economy as a whole. It is possible, but much more complicated, to carry out the same analysis in terms of general equilibrium.

Figure 7-3

Trade and Prices

When trade is opened, China ends up producing buttons for the world market, which consists both of its own domestic market and of the U.S. market. Output rises from Q_1 to Q_2, leading to a fall in the price of buttons from P_1 to P_2, which is lower than the price of buttons in either country before trade.



This is very different from the implications of models without increasing returns. In the standard trade model, as developed in Chapter 6, relative prices converge as a result of trade. If cloth is relatively cheap in Home and relatively expensive in Foreign before trade opens, the effect of trade will be to raise cloth prices in Home and reduce them in Foreign. In our button example, by contrast, the effect of trade is to reduce prices everywhere. The reason for this difference is that when there are external economies of scale, international trade makes it possible to concentrate world production in a single location, and therefore to reduce costs by reaping the benefits of even stronger external economies.

External Economies and the Pattern of Trade

In our example of world trade in buttons, we simply assumed that the Chinese industry started out with lower production costs than the American industry. What might lead to such an initial advantage?

One possibility is comparative advantage—underlying differences in technology and resources. For example, there's a good reason that Silicon Valley is in California, rather than in Mexico. High-technology industries require a highly skilled work force, and such a work force is much easier to find in the United States, where 40 percent of the working-age population is college-educated, than in Mexico, where the number is below 16 percent. Similarly, there's a good reason that world button production is concentrated in China, rather than in Germany. Button production is a labor-intensive industry, which is best conducted in a country where the average manufacturing worker earns less than a dollar an hour rather than in a country where hourly compensation is among the highest in the world.

However, in industries characterized by external economies of scale, comparative advantage usually provides only a partial explanation of the pattern of trade. It was probably inevitable that most of the world's buttons would be made in a relatively low-wage country, but it's not clear that this country necessarily had to be China, and it certainly wasn't necessary that production be concentrated in any particular location within China.

So what does determine the pattern of specialization and trade in industries with external economies of scale? The answer, often, is historical contingency: Something gives a particular location an initial advantage in a particular industry, and this advantage gets

"locked in" by external economies of scale even after the circumstances that created the initial advantage are no longer relevant. The financial centers in London and New York are clear examples. London became Europe's dominant financial center in the 19th century, when Britain was the world's leading economy and the center of a world-spanning empire. It has retained that role even though the empire is long gone and modern Britain is only a middle-sized economic power. New York became America's financial center thanks to the Erie Canal, which made it the nation's leading port. It has retained that role even though the canal currently is used mainly by recreational boats.

Often sheer accident plays a key role in creating an industrial concentration. Geographers like to tell the tale of how a tufted bedspread, crafted as a wedding gift by a 19th-century teenager, gave rise to the cluster of carpet manufacturers around Dalton, Georgia. Silicon Valley's existence may owe a lot to the fact that a couple of Stanford graduates named Hewlett and Packard decided to start a business in a garage in that area. Bangalore might not be what it is today if vagaries of local politics had not led Texas Instruments to choose, back in 1984, to locate an investment project there rather than in another Indian city.

One consequence of the role of history in determining industrial location is that industries aren't always located in the "right" place: Once a country has established an advantage in an industry, it may retain that advantage even if some other country could potentially produce the goods more cheaply.

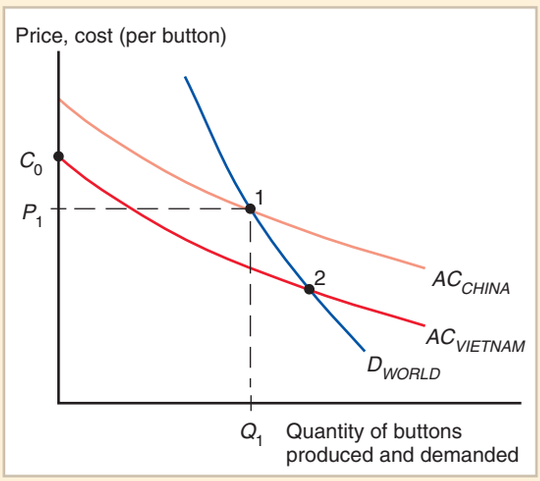
Figure 7-4, which shows the cost of producing buttons as a function of the number of buttons produced annually, illustrates this point. Two countries are shown: China and Vietnam. The Chinese cost of producing a button is shown as AC_{CHINA}, the Vietnamese cost as AC_{VIETNAM}. D_{WORLD} represents the world demand for buttons, which we assume can be satisfied either by China or by Vietnam.

Suppose that the economies of scale in button production are entirely external to firms, and that since there are no economies of scale at the level of the firm, the button industry in each country consists of many small, perfectly competitive firms. Competition therefore drives the price of buttons down to its average cost.

Figure 7-4

The Importance of Established Advantage

The average cost curve for Vietnam, AC_{VIETNAM}, lies below the average cost curve for China, AC_{CHINA}. Thus Vietnam could potentially supply the world market more cheaply than China. If the Chinese industry gets established first, however, it may be able to sell buttons at the price P_1, which is below the cost C_0 that an individual Vietnamese firm would face if it began production on its own. So a pattern of specialization established by historical accident may persist even when new producers could potentially have lower costs.



Holding the World Together

If you are reading this while fully clothed, the odds are that crucial parts of your outfit—specifically, the parts that protect you from a wardrobe malfunction—came from the Chinese town of Qiaotou, which produces 60 percent of the world's buttons and a large proportion of its zippers, too.

The Qiaotou fastener industry fits the classic pattern of geographical concentration driven by external economies of scale. The industry's origins lie in historical accident: In 1980 three brothers spotted some discarded buttons in the street, retrieved and sold them, then realized there was money to be made in the button business. There clearly aren't strong internal economies of scale: The town's button and zipper production is carried out by hundreds of small, family-owned firms. Yet there are clearly advantages to each of these small producers in operating in close proximity to the others.

Qiaotou isn't unique. As a fascinating article on the town's industry* put it, in China, "many small towns, not even worthy of a speck on most maps, have also become world-beaters by focusing on labour-intensive niches.... Start at the toothbrush town of Hang Ji, pass the tie mecca of Sheng Zhou, head east to the home of cheap cigarette lighters in Zhang Qi, slip down the coast to the giant shoe factories of Wen Ling, then move back inland to Yiwu, which not only makes more socks than anywhere else on earth, but also sells almost everything under the sun."

At a broad level, China's role as a huge exporter of labor-intensive products reflects comparative advantage: China is clearly labor-abundant compared with advanced economies. Many of those labor-intensive goods, however, are produced by highly localized industries, which benefit strongly from external economies of scale.

*"The Tiger's Teeth," *The Guardian*, May 25, 2005.

We assume that the Vietnamese cost curve lies below the Chinese curve because, say, Vietnamese wages are lower than Chinese wages. This means that at any given level of production, Vietnam could manufacture buttons more cheaply than China. One might hope that this would always imply that Vietnam will in fact supply the world market. Unfortunately, this need not be the case. Suppose that China, for historical reasons, establishes its button industry first. Then, initially, world button equilibrium will be established at point 1 in Figure 7-4, with Chinese production of Q_1 units per year and a price of P_1. Now introduce the possibility of Vietnamese production. If Vietnam could take over the world market, the equilibrium would move to point 2. However, if there is no initial Vietnamese production $(Q = 0)$, any individual Vietnamese firm considering manufacture of buttons will face a cost of production of C_0. As we have drawn it, this cost is above the price at which the established Chinese industry can produce buttons. So although the Vietnamese industry could potentially make buttons more cheaply than China's industry, China's head start enables it to hold on to the industry.

As this example shows, external economies potentially give a strong role to historical accident in determining who produces what, and may allow established patterns of specialization to persist even when they run counter to comparative advantage.

Trade and Welfare with External Economies

In general, we can presume that external economies of scale lead to gains from trade over and above those from comparative advantage. The world is more efficient and thus richer because international trade allows nations to specialize in different industries and thus reap the gains from external economies as well as from comparative advantage.

temporary protection of industries to enable them to gain experience is known as the **infant industry argument;** this argument has played an important role in debates over the role of trade policy in economic development. We will discuss the infant industry argument at greater length in Chapter 10, but for now we simply note that situations like that illustrated in Figure 7-6 are just as hard to identify in practice as those involving nondynamic increasing returns.

Interregional Trade and Economic Geography

External economies play an important role in shaping the pattern of international trade, but they are even more decisive in shaping the pattern of **interregional trade**—trade that takes place between regions *within* countries.

To understand the role of external economies in interregional trade, we first need to discuss the nature of regional economics—that is, how the economies of regions within a nation fit into the national economy. Studies of the location of U.S. industries suggest that more than 60 percent of U.S. workers are employed by industries whose output is nontradable even within the United States—that is, that must be supplied locally. Table 7-2 shows some examples of tradable and nontradable industries. Thus, motion pictures made in Hollywood are shown across the country, and indeed around the world, but newspapers are mainly read in their home cities. Wall Street trades stocks and makes deals for clients across the United States, but savings banks mainly serve local depositors. Scientists at the National Institutes of Health develop medical knowledge that is applied across the whole country, but the veterinarian who figures out why your pet is sick has to be near your home.

As you might expect, the share of nontradable industries in employment is pretty much the same across the United States. For example, restaurants employ about 5 percent of the work force in every major U.S. city. On the other hand, tradable industries vary greatly in importance across regions. Manhattan accounts for only about 2 percent of America's total employment, but it accounts for a quarter of those employed in trading stocks and bonds and about one-seventh of employment in the advertising industry.

But what determines the location of tradable industries? In some cases, natural resources play a key role—for example, Houston is a center for the oil industry because east Texas is where the oil is. However, factors of production such as labor and capital play a less decisive role in interregional trade than in international trade, for the simple reason that such factors are highly mobile within countries. As a result, factors tend to move to where the industries are rather than the other way around. For example, California's Silicon Valley, near San Francisco, has a very highly educated labor force, with a high concentration of engineers and computer experts. That's not

TABLE 7-2 Some Examples of Tradable and Nontradable Industries

Tradable Industries	Nontradable Industries
Motion pictures	Newspaper publishers
Securities, commodities, etc.	Savings institutions
Scientific research	Veterinary services

Source: J. Bradford Jensen and Lori. G. Kletzer, "Tradable Services: Understanding the Scope and Impact of Services Outsourcing," in Lael Brainard and Susan M. Collins, eds., *Brookings Trade Forum 2005: Offshoring White Collar Work* (Washington, D.C.: Brookings Institution, 2005), pp. 75–116.