CHAPTER

International Factor Movements

p to this point we have concerned ourselves entirely with international *trade*. That is, we have focused on the causes and effects of international exchanges of goods and services. Movement of goods and services is not, however, the only form of international integration. This chapter is concerned with another form of integration, international movements of factors of production, or **factor movements.** Factor movements include labor migration, the transfer of capital via international borrowing and lending, and the subtle international linkages involved in the formation of multinational corporations.

The principles of international factor movement do not differ in their essentials from those underlying international trade in goods. Both international borrowing and lending and international labor migration can be thought of as analogous in their causes and effects to the movement of goods analyzed in Chapters 3 through 5. The role of the multinational corporation may be understood by extending some of the concepts developed in Chapter 6. So when we turn from trade in goods and services to factor movements, we do not make a radical shift in emphasis.

Although there is a fundamental economic similarity between trade and factor movements, however, there are major differences in the political context. A labor-abundant country may under some circumstances import capital-intensive goods; under other circumstances it may acquire capital by borrowing abroad. A capital-abundant country may import labor-intensive goods or begin employing migrant workers. A country that is too small to support firms of efficient size may import goods where large firms have an advantage or allow those goods to be produced locally by subsidiaries of foreign firms. In each case the alternative strategies may be similar in their purely economic consequences but radically different in their political acceptability.

On the whole, international factor movement tends to raise even more political difficulties than international trade. Thus factor movements are subject to more restriction than trade in goods. Immigration restrictions are nearly universal. Until the 1980s several European countries, such as France, maintained controls on capital movements even though they had virtually free trade in goods with their neighbors. Investment by foreign-based multinational corporations is regarded with suspicion and tightly regulated through much of the world. The result is that factor movements are probably less important in practice than trade in goods, which is why we took an analysis of trade in the absence of factor movements as our starting point. Nonetheless, factor movements are very important, and it is valuable to spend a chapter on their analysis.

This chapter is in three parts. We begin with a simple model of international labor mobility. We then proceed to an analysis of international borrowing and lending, in which we show that this lending can be interpreted as trade *over time:* The lending country gives up resources now to receive repayment in the future, while the borrower does the reverse. Finally, the last section of the chapter analyzes multinational corporations.

Learning Goals

After reading this chapter, you will be able to:

- Discuss the causes as well as the winners and losers from migration and labor mobility between nations.
- Describe the concept of intertemporal comparative advantage and explain how it relates to international capital flows, international lending, and foreign investment.
- Understand theories that explain the existence of multinational firms and the motivation for foreign direct investment across economies.

International Labor Mobility

We begin our discussion with an analysis of the effects of labor mobility. In the modern world, restrictions on the flow of labor are legion—just about every country imposes restrictions on immigration. Thus labor mobility is less prevalent in practice than capital mobility. It remains important, however; it is also simpler in some ways to analyze than capital movement, for reasons that will become apparent later in the chapter.

A One-Good Model Without Factor Mobility

As in the analysis of trade, the best way to understand factor mobility is to begin with a world that is not economically integrated, then examine what happens when international transactions are allowed. Let's assume that we have, as usual, a two-country world consisting of Home and Foreign, each with two factors of production, land and labor. We assume for the moment, however, that this world is even simpler than the one we examined in Chapter 4, in that the two countries produce only *one* good, which we will simply refer to as "output." Thus there is no scope for ordinary trade, the exchange of different goods, in this world. The only way for these economies to become integrated with each other is via movement of either land or labor. Land almost by definition cannot move, so this is a model of integration via international labor mobility.

Before we introduce factor movements, however, let us analyze the determinants of the level of output in each country. Land (T) and labor (L) are the only scarce resources. Thus the output of each country will depend, other things equal, on the quantity of these factors available. The relationship between the supplies of factors on one side and the output of the economy on the other is referred to as the economy's production function, which we denote by Q(T, L).



A useful way to look at the production function is to ask how output depends on the supply of one factor of production, holding fixed the supply of land. The slope of the production function measures the increase in output that would be gained by using a little more labor and is thus referred to as the *marginal product of labor*. As the curve is drawn in Figure 7-1, the marginal product of labor is assumed to fall as the ratio of labor to land rises. This is the normal case: As a country seeks to employ more labor on a given amount of land, it must move to increasingly labor-intensive techniques of production, and this will normally become increasingly difficult the further the substitution of labor for land goes.

Figure 7-2 contains the same information as Figure 7-1 but plots it in a different way. We now show directly how the marginal product of labor depends on the quantity of labor employed. We also indicate that the real wage earned by each unit of labor is equal to labor's marginal product. This will be true as long as the economy is perfectly competitive, which we assume to be the case.

Figure 7-2

The Marginal Product of Labor

The marginal product of labor declines with employment. The area under the marginal product curve equals total output. Given the level of employment, the marginal product determines the real wage; thus the total payment to labor (the real wage times the number of employees) is shown by the rectangle in the figure. The rest of output consists of land rents.





What about the income earned by land? As we show in Appendix 1 to this chapter, the total output of the economy can be measured by the area under the marginal product curve. Of that total output, wages earned by workers equal the real wage rate times the employment of labor, and hence equal the indicated area on the figure. The remainder, also shown, equals rents earned by landowners.

Assume that Home and Foreign have the same technology but different overall landlabor ratios. If Home is the labor-abundant country, workers in Home will earn less than those in Foreign, while land in Home earns more than in Foreign. This obviously creates an incentive for factors of production to move. Home workers would like to move to Foreign; Foreign landowners would also like to move their land to Home, but we are supposing that this is impossible. Our next step is to allow workers to move and see what happens.

International Labor Movement

Now suppose that workers are able to move between our two countries. Workers will move from Home to Foreign. This movement will reduce the Home labor force and thus raise the real wage in Home, while increasing the labor force and reducing the real wage in Foreign. If there are no obstacles to labor movement, this process will continue until the marginal product of labor is the same in the two countries.

Figure 7-3 illustrates the causes and effects of international labor mobility. The horizontal axis represents the total world labor force. The workers employed in Home are measured from the left, the workers employed in Foreign from the right. The left vertical axis shows the marginal product of labor in Home; the right vertical axis shows the marginal product of labor in Foreign. Initially we assume that there are OL^1 workers in Home, L^1O^* workers in Foreign. Given this allocation, the real wage rate would be lower in Home (point *C*) than in Foreign (point *B*). If workers can move freely to whichever country offers the higher real wage, they will move from Home to Foreign until the real wage rates are

equalized. The eventual distribution of the world's labor force will be one with workers in Home, L^2O^* workers in Foreign (point A).

Three points should be noted about this redistribution of the world's labor force.

1. It leads to a convergence of real wage rates. Real wages rise in Home, fall in Foreign.

2. It increases the world's output as a whole. Foreign's output rises by the area under its marginal product curve from L^1 to L^2 , while Home's falls by the corresponding area under its marginal product curve. We see from the figure that Foreign's gain is larger than Home's loss, by an amount equal to the colored area *ABC* in the figure.

3. Despite this gain, some people are hurt by the change. Those who would originally have worked in Home receive higher real wages, but those who would originally have worked in Foreign receive lower real wages. Landowners in Foreign benefit from the larger labor supply, but landowners in Home are made worse off. As in the case of the gains from international trade, then, international labor mobility, while allowing everyone to be made better off in principle, leaves some groups worse off in practice.

Extending the Analysis

We have just seen that a very simple model tells us quite a lot about both why international factor movements occur and what effects they have. Labor mobility in our simple model, like trade in the model of Chapter 4, is driven by international differences in resources; also like trade, it is beneficial in the sense that it increases world production yet is associated with strong income distribution effects that make those gains problematic.

Let us consider briefly how the analysis is modified when we add some of the complications we have assumed away.

We need to remove the assumption that the two countries produce only one good. Suppose, then, that the countries produce two goods, one more labor-intensive than the other. We already know from our discussion of the factor proportions model in Chapter 4 that in this case trade offers an alternative to factor mobility. Home can in a sense export labor and import land by exporting the labor-intensive good and importing the land-intensive good. It is possible in principle for such trade to lead to a complete equalization of factor prices without any need for factor mobility. If this happened, it would of course remove any incentive for labor to move from Home to Foreign.

In practice, while trade is indeed a substitute for international factor movement, it is not a perfect substitute. The reasons are those already summarized in Chapter 4. Complete factor-price equalization is not observed in the real world because countries are sometimes too different in their resources to remain unspecialized; there are barriers to trade, both natural and artificial; and there are differences in technology as well as resources between countries.

We might wonder on the other side whether factor movements do not remove the incentive for international trade. Again the answer is that while in a simple model movement of factors of production can make international trade in goods unnecessary, in practice there are substantial barriers to free movement of labor, capital, and other potentially mobile resources. And some resources cannot be brought together—Canadian forests and Caribbean sunshine cannot migrate.

Extending the simple model of factor mobility, then, does not change its fundamental message. The main point is that trade in factors is, in purely economic terms, very much like trade in goods; it occurs for much the same reasons and produces similar results.

Case Study

Wage Convergence in the Age of Mass Migration

Although there are substantial movements of people between countries in the modern world, the truly heroic age of labor mobility—when immigration was a major source of population growth in some countries, while emigration caused population in other countries to decline—was in the late 19th and early 20th centuries. In a global economy



newly integrated by railroads, steam-ships, and telegraph cables, and not yet subject to many legal restrictions on migration, tens of millions of people moved long distances in search of a better life. Chinese moved to Southeast Asia and California; Indians to Africa and the Caribbean; a substantial number of Japanese moved to Brazil. Above all, people from the periphery of Europe—from Scandinavia, Ireland, Italy, and Eastern Europe—moved to places where land was abundant and wages were high: the United States, but also Canada, Argentina, and Australia.

Did this process cause the kind of real wage convergence that our model predicts? Indeed it did. The accompanying table shows real wages in 1870, and the change in these wages up to the eve of World War I, for

four major "destination" countries and for four important "origin" countries. As the table shows, at the beginning of the period real wages were much higher in the destination than the origin countries. Over the next four decades real wages rose in all countries, but (except for a surprisingly large increase in Canada) they increased much more rapidly in the origin than the destination countries, suggesting that migration actually did move the world toward (although not by any means all the way to) wage equalization.

As documented in the case study on the U.S. economy, legal restrictions put an end to the age of mass migration after World War I. For that and other reasons (notably a decline in world trade, and the direct effects of two world wars), convergence in real wages came to a halt and even reversed itself for several decades, only to resume in the postwar years.

	Real Wage, 1870 (U.S. = 100)	Percentage Increase in Real Wage, 1870–1913
Destination Countries		
Argentina	53	51
Australia	110	1
Canada	86	121
United States	100	47
Origin Countries		
Ireland	43	84
Italy	23	112
Norway	24	193
Sweden	24	250

Source: Jeffrey G. Williamson, "The Evolution of Global Labor Markets Since 1830: Background Evidence and Hypotheses," *Explorations in Economic History* 32 (1995), pp. 141–196.

Case Study

Immigration and the U.S. Economy

As Figure 7-4 shows, the share of immigrants in the U.S. population has varied greatly over the past century. In the early 20th century, the number of foreign-born U.S. residents was swelled by vast immigration from Eastern and Southern Europe. Tight restrictions on immigration imposed in the 1920s brought an end to this era, and by the1960s immigrants were a minor factor on the American scene. A new wave of immigration began around 1970, this time with most immigrants coming from Latin America and Asia.

How has this new wave of immigration affected the U.S. economy? The most direct effect is that immigration has expanded the work force. As of 2006, foreign-born workers made up 15.3 percent of the U.S. labor force—that is, without immigrants the United States would have 15 percent fewer workers.

Other things equal, we would expect this increase in the work force to reduce wages. One widely cited estimate is that average wages in the United States are 3 percent lower than they would have been in the absence of immigration.¹ However, comparisons of average wages can be misleading. Immigrant workers are much more likely than native-born workers to have low levels of education: In 2006, 28 percent of the immigrant labor force had not completed high school or its equivalent, compared with only 6 percent of native-born workers. As a result, most estimates suggest that immigration has actually raised the wages of native-born Americans with a college education or above. Any negative effects on wages fall on less-educated Americans. There is, however, considerable dispute among economists about how large these negative wage effects are, with estimates ranging from an 8 percent decline to much smaller numbers.



¹George Borjas, "The Labor Demand Curve *Is* Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* 118(November 2003), pp. 1335–1374.

What about the overall effects on America's income? America's gross domestic product—the total value of all goods and services produced here—is clearly larger because of immigrant workers. However, much of this increase in the value of production is used to pay wages to the immigrants themselves. Estimates of the "immigration surplus"—the difference between the gain in GDP and the cost in wages paid to immigrants—are generally small, on the order of 0.1% of GDP.²

There's one more complication in assessing the economic effects of immigration: the effects on tax revenue and government spending. On one side, immigrants pay taxes, helping cover the cost of government. On the other side, they impose costs on the government, because their cars need roads to drive on, their children need schools to study in, and so on. Because many immigrants earn low wages and hence pay low taxes, some estimates suggest that immigrants cost more in additional spending than they pay in. However, estimates of the net fiscal cost, like estimates of the net economic effects, are small, again on the order of 0.1% of GDP.

Immigration is, of course, an extremely contentious political issue. The economics of immigration, however, probably don't explain this contentiousness. Instead, it may be helpful to recall what the Swiss author Max Frisch once said about the effects of immigration into his own country, which at one point relied heavily on workers from other countries: "We asked for labor, but people came." And it's the fact that immigrants are people that makes the immigration issue so difficult.

International Borrowing and Lending

International movements of capital are a prominent feature of the international economic landscape. It is tempting to analyze these movements in a way parallel to our analysis of labor mobility and this is sometimes a useful exercise. There are some important differences, however. When we speak of international labor mobility, it is clear that workers are physically moving from one country to another. International capital movements are not so simple. When we speak of capital flows from the United States to Mexico, we do not mean that U.S. machines are literally being unbolted and shipped south. We are instead talking of a *financial* transaction. A U.S. bank lends to a Mexican firm, or U.S. residents buy stock in Mexico, or a U.S. firm invests through its Mexican subsidiary. We focus for now on the first type of transaction, in which U.S. residents make loans to Mexicans—that is, the U.S. residents grant Mexicans the right to spend more than they earn today in return for a promise to repay in the future.

The analysis of financial aspects of the international economy is the subject of the second half of this book. It is important to realize, however, that financial transactions do not exist simply on paper. They have real consequences. International borrowing and lending, in particular, can be interpreted as a kind of international trade. The trade is not of one good for another at a point in time but of goods today for goods in the future. This kind of

² See Gordon Hanson, "Challenges for Immigration Policy," in C. Fred Bergsten, ed., *The United States and the World Economy: Foreign Economic Policy for the Next Decade*, Washington, D.C.: Institute for International Economics, 2005, 343–372.



trade is known as **intertemporal trade**; we will have much more to say about it later in this text, but for present purposes a simple model will be sufficient to make our point.³

Intertemporal Production Possibilities and Trade

Even in the absence of international capital movements, any economy faces a trade-off between consumption now and consumption in the future. Economies usually do not consume all of their current output; some of their output takes the form of investment in machines, buildings, and other forms of productive capital. The more investment an economy undertakes now, the more it will be able to produce and consume in the future. To invest more, however, an economy must release resources by consuming less (unless there are unemployed resources, a possibility we temporarily disregard). Thus there is a trade-off between current and future consumption.

Let's imagine an economy that consumes only one good and will exist for only two periods, which we will call present and future. Then there will be a trade-off between present and future production of the consumption good, which we can summarize by drawing an **intertemporal production possibility frontier.** Such a frontier is illustrated in Figure 7-5. It looks just like the production possibility frontiers we have been drawing between two goods at a point in time.

The shape of the intertemporal production possibility frontier will differ among countries. Some countries will have production possibilities that are biased toward present output, while others are biased toward future output. We will ask what real differences these biases correspond to in a moment, but first let's simply suppose that there are two countries, Home and Foreign, with different intertemporal production possibilities. Home's possibilities are biased toward current consumption, while Foreign's are biased toward future consumption.

 $^{^{3}}$ This chapter's Appendix 2 contains a more detailed examination of the model developed in this section.

Reasoning by analogy, we already know what to expect. In the absence of international borrowing and lending, we would expect the relative price of future consumption to be higher in Home than in Foreign, and thus if we open the possibility of trade over time, we would expect Home to export present consumption and import future consumption.

This may, however, seem a little puzzling. What is the relative price of future consumption, and how does one trade over time?

The Real Interest Rate

How does a country trade over time? Like an individual, a country can trade over time by borrowing or lending. Consider what happens when an individual borrows: She is initially able to spend more than her income or, in other words, to consume more than her production. Later, however, she must repay the loan with interest, and therefore in the future she consumes *less* than she produces. By borrowing, then, she has in effect traded future consumption for current consumption. The same is true of a borrowing country.

Clearly the price of future consumption in terms of present consumption has something to do with the interest rate. As we will see in the second half of this book, in the real world the interpretation of interest rates is complicated by the possibility of changes in the overall price level. For now, we bypass that problem by supposing that loan contracts are specified in "real" terms: When a country borrows, it gets the right to purchase some quantity of consumption at present in return for repayment of some larger quantity in the future. Specifically, the quantity of repayment in future will be (1 + r) times the quantity borrowed in present, where *r* is the **real interest rate** on borrowing. Since the trade-off is one unit of consumption in present for (1 + r) units in future, the relative price of future consumption is 1/(1 + r).

The parallel with our standard trade model is now complete. If borrowing and lending are allowed, the relative price of future consumption, and thus the world real interest rate, will be determined by the world relative supply and demand for future consumption. Home, whose intertemporal production possibilities are biased toward present consumption, will export present consumption and import future consumption. That is, Home will lend to Foreign in the first period and receive repayment in the second.

Intertemporal Comparative Advantage

We have assumed that Home's intertemporal production possibilities are biased toward present production. But what does this mean? The sources of intertemporal comparative advantage are somewhat different from those that give rise to ordinary trade.

A country that has a comparative advantage in future production of consumption goods is one that in the absence of international borrowing and lending would have a low relative price of future consumption, that is, a high real interest rate. This high real interest rate corresponds to a high return on investment, that is, a high return to diverting resources from current production of consumption goods to production of capital goods, construction, and other activities that enhance the economy's future ability to produce. So countries that borrow in the international market will be those where highly productive investment opportunities are available relative to current productive capacity, while countries that lend will be those where such opportunities are not available domestically.

The pattern of international borrowing and lending in the 1970s illustrates the point. Table 22-3 compares the international lending of three groups of countries: industrial countries, non-oil developing countries, and major oil exporters. From 1974 to 1981, the oil exporters lent \$395 billion, the less-developed countries borrowed \$315 billion, and the (much larger) industrial countries borrowed a smaller amount, \$265 billion. In the light of

our model, this is not surprising. During the 1970s, as a result of a spectacular increase in oil prices, oil exporters like Saudi Arabia found themselves with very high current income. They did not, however, find any comparable increase in their domestic investment opportunities. That is, they had a comparative advantage in current consumption. With small populations, limited resources other than oil, and little expertise in industrial or other production, their natural reaction was to invest much of their increased earnings abroad. By contrast, rapidly developing countries such as Brazil and South Korea expected to have much higher incomes in the future and saw highly productive investment opportunities in their growing industrial sectors; they had a comparative advantage in future income. Thus in this time frame (1974 to 1981) the oil exporters also exported current consumption by lending their money, in part, to less-developed countries.

Direct Foreign Investment and Multinational Firms

In the last section we focused on international borrowing and lending. This is a relatively simple transaction, in that the borrower makes no demands on the lender other than that of repayment. An important part of international capital movement, however, takes a different form, that of **direct foreign investment**. By direct foreign investment we mean international capital flows in which a firm in one country creates or expands a subsidiary in another. The distinctive feature of direct foreign investment is that it involves not only a transfer of resources but also the acquisition of *control*. That is, the subsidiary does not simply have a financial obligation to the parent company; it is part of the same organizational structure.

When is a corporation multinational? In U.S. statistics, a U.S. company is considered foreign-controlled, and therefore a subsidiary of a foreign-based multinational, if 10 percent or more of the stock is held by a foreign company; the idea is that 10 percent is enough to convey effective control. A U.S.-based company is considered multinational if it has a controlling share of companies abroad.

Alert readers will notice that these definitions make it possible for a company to be considered both a U.S. subsidiary of a foreign company and a U.S. multinational. And this sometimes happens: From 1981 until 1995 the chemical company DuPont was officially foreign-controlled (because the Canadian company Seagram owned a large block of its stock) but was also considered an American multinational. In practice, such strange cases are rare: Usually multinational companies have a clear national home base.

Multinational firms are often a vehicle for international borrowing and lending. Parent companies often provide their foreign subsidiaries with capital, in the expectation of eventual repayment. To the extent that multinational firms provide financing to their foreign subsidiaries, direct foreign investment is an alternative way of accomplishing the same things as international lending. This still leaves open the question, however, of why direct investment rather than some other way of transferring funds is chosen. In any case, the existence of multinational firms does not necessarily reflect a net capital flow from one country to another. Multinationals sometimes raise money for the expansion of their subsidiaries in the country where the subsidiary operates rather than in their home country. Furthermore, there is a good deal of two-way foreign direct investment among industrial countries, U.S. firms expanding their European subsidiaries at the same time that European firms expand their U.S. subsidiaries, for example.

The point is that while multinational firms sometimes act as a vehicle for international capital flows, it is probably a mistake to view direct foreign investment as primarily an alternative way for countries to borrow and lend. Instead, the main point of direct foreign investment is to allow the formation of multinational organizations. That is, the extension of control is the essential purpose.

Does Capital Movement to Developing Countries Hurt Workers in High-Wage Countries?

We have turned repeatedly in this textbook to concerns created by the rapid economic growth of newly industrializing economies (NIEs), mainly in Asia. In Chapter 4 we discussed the concern that trade with the NIEs might, via the Stolper-Samuelson effect, reduce the real wages of less-skilled workers in advanced nations and saw that it had some justification. In Chapter 5 we turned to the possibility that growth in the NIEs might, by worsening the terms of trade of advanced nations, lower their overall real income but saw that this was unlikely. In the 1990s there was growing worry among some commentators that the export of capital to the NIEs would have a severe impact on the wages of workers in advanced countries.

The logic of this view is as follows: If high-wage countries finance investment in low-wage countries, this will mean less savings available to build up the capital stock at home. Because each worker at home will have less capital to work with than she otherwise would, her marginal product—and hence her wage rate—will be lower than it would have been in the absence of the capital movement. Overall real income, including the returns from capital invested abroad, may be higher for the home country than it would otherwise have been, but more than all the gains will go to capital, with labor actually worse off.

The actual facts of capital flows to low-wage countries, however, haven't supported such fears. Figure 7-6 shows total capital flows to developing economies since 1990, expressed as a share of advanced (i.e., high-wage) countries' GDP. During the 1990s, there was considerable movement of capital from high-wage to low-wage countries, but it was never more than a fraction of the GDP of the high-wage economies, and therefore can't have been responsible for more than a small reduction in wages.

In the late 1990s, a financial crisis in Asia caused investors to rethink the idea of investing in low-wage economies, and capital flows dried up. And then a strange thing happened: Capital began flowing *out* of low-wage countries and into high-wage countries, mainly the United States. The principal cause of this "uphill" capital flow was the actions of Asian governments, especially the government of China, which bought up large quantities of dollars in an attempt to keep the value of China's currency, the yuan,

But why do firms seek to extend control? Economists do not have as fully developed a theory of multinational enterprise as they do of many other issues in international economics. There is some theory on the subject, however, which we now review.

The Theory of Multinational Enterprise

The basic necessary elements of a theory of multinational firms can best be seen by looking at an example. Consider the European operations of American auto manufacturers. Ford and General Motors, for example, sell many cars in Europe, but nearly all those cars are manufactured in plants in Germany, Britain, and Spain. This arrangement is familiar, but we should realize that there are two obvious alternatives. On one side, instead of producing in Europe the U.S. firms could produce in the United States and export to the European market. On the other side, the whole market could be served by European producers such as Volkswagen and Renault. Why, then, do we see this particular arrangement, in which the *same* firms produce in *different* countries?

The modern theory of multinational enterprise starts by distinguishing between the two questions of which this larger question is composed. First, why is a good produced in two (or more) different countries rather than one? This is known as the question of **location**. Second, why is production in different locations done by the same firm rather than by



Capital flows to low-wage countries have never been large compared with the economies of rich countries—and in recent years capital has actually flowed the other way, from low-wage to high-wage nations.

from rising. The motivations for this behavior will be discussed in the second half of this book. The important point for now is that in recent years capital, far from flowing away from high-wage countries, has actually flowed away from low-wage countries, largely to the United States.

separate firms? This is known, for reasons that will become apparent in a moment, as the question of **internalization**. We need a theory of location to explain why Europe does not import its automobiles from the United States; we need a theory of internalization to explain why Europe's auto industry is not independently controlled.

The theory of location is not a difficult one in principle. It is, in fact, just the theory of trade that we developed in Chapters 3 through 6. The location of production is often determined by resources. Aluminum mining must be located where the bauxite is, aluminum smelting near cheap electricity. Minicomputer manufacturers locate their skill-intensive design facilities in Massachusetts or northern California and their labor-intensive assembly plants in Ireland or Singapore. Alternatively, transport costs and other barriers to trade may determine location. American firms produce locally for the European market partly to reduce transport costs; since the models that sell well in Europe are often quite different from those that sell well in the United States, it makes sense to have separate production facilities and to put them on different continents. As these examples reveal, the factors that determine a multinational corporation's decisions about where to produce are probably not much different from those that determine the pattern of trade in general.

The theory of internalization is another matter. Why not have independent auto companies in Europe? We may note first that there are always important transactions between a multinational's operations in different countries. The output of one subsidiary is often an input into the production of another. Or technology developed in one country may be used in others. Or management may usefully coordinate the activities of plants in several countries. These transactions are what tie the multinational firm together, and the firm presumably exists to facilitate these transactions. But international transactions need not be carried out inside a firm. Components can be sold in an open market, and technology can be licensed to other firms. Multinationals exist because it turns out to be more profitable to carry out these transactions within a firm rather than between firms. This is why the motive for multinationals is referred to as "internalization."

We have defined a concept, but we have not yet explained what gives rise to internalization. Why are some transactions more profitably conducted within a firm rather than between firms? Here there are a variety of theories, none as well-grounded either in theory or in evidence as our theories of location. We may note two influential views, however, about why activities in different countries may usefully be integrated in a single firm.

The first view stresses the advantages of internalization for **technology transfer.** Technology, broadly defined as any kind of economically useful knowledge, can sometimes be sold or licensed. There are important difficulties in doing this, however. Often the technology involved in, say, running a factory has never been written down; it is embodied in the knowledge of a group of individuals and cannot be packaged and sold. Also, it is difficult for a prospective buyer to know how much knowledge is worth—if the buyer knew as much as the seller, there would be no need to buy! Finally, property rights in knowledge are often hard to establish. If a European firm licenses technology to a U.S. firm, other U.S. firms may legally imitate that technology. All these problems may be reduced if a firm, instead of selling technology, sets about capturing the returns from the technology in other countries by setting up foreign subsidiaries.

The second view stresses the advantages of internalization for **vertical integration**. If one firm (the "upstream" firm) produces a good that is used as an input for another firm (the "downstream" firm), a number of problems can result. For one thing, if each has a monopoly position, they may get into a conflict as the downstream firm tries to hold the price down while the upstream firm tries to raise it. There may be problems of coordination if demand or supply is uncertain. Finally, a fluctuating price may impose excessive risk on one or the other party. If the upstream and downstream firms are combined into a single "vertically integrated" firm, these problems may be avoided or at least reduced.

It should be clear that these views are by no means as rigorously worked out as the analysis of trade carried out elsewhere in this book. The economic theory of organizations— which is what we are talking about when we try to develop a theory of multinational corporations—is still in its infancy. This is particularly unfortunate because in practice multinationals are a subject of heated controversy—praised by some for generating economic growth, accused by others of creating poverty.

Multinational Firms in Practice

Multinational firms play an important part in world trade and investment. For example, about half of U.S. imports are transactions between "related parties." By this we mean that the buyer and the seller are to a significant extent owned and presumably controlled by the same firm. Thus half of U.S. imports can be regarded as transactions between branches of multinational firms. At the same time, 24 percent of U.S. assets abroad consists of the value of foreign subsidiaries of U.S. firms. So U.S. international trade and investment, while not dominated by multinational firms, are to an important extent conducted by such firms.

Multinational firms may, of course, be either domestic or foreign-owned. Foreign-owned multinational firms play an important role in most economies and an increasingly important

TABLE 7-1	Employment by Foreign-Owned Firms in the United States	
	As Percent of Total Nonfarm Employment	As Percent of Manufacturing Employment
1977	1.5	3.8
2005	3.8	14.0
Source: U.S. Commerce Department.		

role in the United States. Table 7-1 shows how the percentage of U.S. workers employed by foreign-owned firms has increased over the past quarter-century, both in the economy as a whole and in manufacturing especially.

The important question, however, is what difference multinationals make. With only a limited understanding of why multinationals exist, this is a hard question to answer. Nonetheless, the existing theory suggests some preliminary answers.

Notice first that much of what multinationals do could be done without multinationals, although perhaps not as easily. Two examples are the shift of labor-intensive production from industrial countries to labor-abundant nations and capital flows from capital-abundant countries to capital-scarce countries. Multinational firms are sometimes the agents of these changes and are therefore either praised or condemned for their actions (depending on the commentator's point of view). But these shifts reflect the "location" aspect of our theory of multinationals, which is really no different from ordinary trade theory. If multinationals were not there, the same things would still happen, though perhaps not to the same extent. This observation leads international economists to attribute less significance to multinational enterprise than most lay observers.

Case Study

Foreign Direct Investment in the United States

Until the 1980s, the United States was almost always regarded as a "home" country for multinational companies rather than as a "host" for foreign-based multinationals. Indeed, when the French author Jean-Jacques Servan-Schreiber wrote a best-seller warning of the growing power of multinationals, his book—published in 1968—was titled *The American Challenge*.

This perspective changed in the middle of the 1980s. Figure 7-7 shows U.S. inflows of foreign direct investment—that is, capital either used to acquire control of a U.S. company or invested in a company that foreigners already controlled—as a percentage of GDP. In the second half of the 1980s these flows, which had previously averaged less than 0.5 percent of GDP, surged. Japanese companies began building automobile plants in the United States, and European companies began buying U.S. banks and insurance companies. Foreign direct investment then slumped in the early 1990s, before beginning an astonishing rise in the late 1990s.

What was behind these fluctuations? Rather paradoxically, the boom in direct investment in the late 1980s and the even bigger boom in the late 1990s happened for nearly opposite reasons.



Foreign direct investment flows into the United States surged in 1986–1989 and again after 1992, rapidly raising the share of U.S. production controlled by foreign firms.

Source: U.S. Commerce Department.

Much foreign direct investment in the 1980s was driven by a perception of U.S. weakness. At the time, Japanese manufacturing companies, especially in the auto industry, had pulled ahead of their U.S. competitors in productivity and technology. The lower prices and superior quality of Japanese products allowed them to take a rapidly growing share of the U.S. market; in order to serve that market better, the Japanese began to open plants in the United States.

Also, in the late 1980s the U.S. dollar was quite weak against both the Japanese yen and then-European currencies such as the German mark. This made assets in the United States appear cheap and encouraged foreign companies to move in.

Perhaps because of the perception that foreigners were taking advantage of U.S. weakness, the surge in foreign direct investment in the 1980s provoked a political backlash. The height of this backlash probably came in 1992, when Michael Crichton published the best-seller *Rising Sun*, a novel about the evil machinations of a Japanese company operating in the United States. The novel, which was made into a movie starring Sean Connery the next year, came with a long postscript warning about the dangers that Japanese companies posed to the United States.

As you can see from Figure 7-7, however, foreign direct investment in the United States was slumping even as *Rising Sun* hit the bookstores. And public concern faded along with the investment itself.

When foreign direct investment surged again, in the late 1990s, the situation was very different: Then the wave of investment was driven by perceptions of U.S. strength rather than weakness. The United States was experiencing a remarkable economic boom;

meanwhile, European growth was modest, and Japan languished in the middle of a decade of economic stagnation. Given the revived economic dominance of the United States, nearly every large company on the planet felt that it had to have a stake in the U.S. economy. And so companies flocked to the United States, mainly by acquiring control of existing U.S. companies. Whether this was a good idea is another question: The troubled acquisition of Chrysler by the German company Daimler-Benz, discussed below became a celebrated example of how investing in the United States could go wrong.

The political reception for foreign investors in the 1990s was utterly different from that given to the previous wave. It's not clear to what extent Americans were even aware of the wave of money pouring in; Michael Crichton gave up on economics and went back to writing about dinosaurs. To the extent that the inflow of direct investment was noticed, it was perceived as a tribute to U.S. strength, not as a threat.

The great foreign direct investment boom of the late 1990s abruptly ended at the beginning of the next decade, as the U.S. stock market slumped and the U.S. economy went into a recession.

Taken for a Ride

In November 1998 Germany's Daimler-Benz corporation, the makers of the Mercedes-Benz, acquired control of America's Chrysler corporation for \$40 billion about \$13 billion more than the market value of Chrysler's stock at the time. The new, merged company was named DaimlerChrysler.

For the deal to make business sense, the combined company had to be worth more than the two companies were worth separately. In fact, given the premium that Daimler-Benz paid to acquire Chrysler, the merger in effect had to create at least \$13 billion in value. Where would this gain come from?

The answer, according to executives in both companies, was that there would be "synergy" between the two companies—that the whole would be more than the sum of the parts because each company would supply something the other lacked. Skeptical analysts were not convinced. They pointed out that although both companies were in the automobile business, they occupied almost completely different market niches: Daimler-Benz had built its reputation on classy luxury sedans, while Chrysler was much more down-market: Its signature vehicles were minivans and SUVs. So it was unclear whether there would be much gain in terms of either marketing or production efficiencies. In that case, where would the extra value come from?

It soon became clear that far from generating synergies, the deal had at least initially created new problems, particularly at Chrysler. Put simply, the cultural differences between the two companiespartly a matter of national style, partly a matter of the personalities involved-created a great deal of misunderstanding and bad feelings. The initial deal was supposedly a merger of equals, but it soon became clear that the German company was the senior partner; many Chrysler executives left within a year after the merger. Partly as a result of these departures, Chrysler's product development and marketing lagged; within two years after the deal, Chrysler had gone from large profits to large losses. In 2007, Daimler handed Chrysler off to Cerberus Capital Management, a firm specializing in turnarounds of troubled companies. Remarkably, Daimler actually ended up paying Cerberus to take Chrysler, which owes billions in benefits to its current and former employees, off its hands. These developments were reflected in a plunge in the new company's stock price: Two years after the merger, far from being worth more than the sum of the two companies before the deal, Daimler-Chrysler was worth less than either company alone.

Notice, too, that in a broad sense what multinational corporations do by creating organizations that extend across national boundaries is similar to the effects of trade and simple factor mobility; that is, it is a form of international economic integration. By analogy with the other forms of international integration we have studied, we would expect multinational enterprise to produce overall gains but to produce income distribution effects that leave some people worse off. These income distribution effects are probably mostly effects *within* rather than *between* countries.

To sum up, multinational corporations probably are not as important a factor in the world economy as their visibility would suggest; their role is neither more nor less likely to be beneficial than other international linkages. This does not, however, prevent them from being cast in the role of villains or (more rarely) heroes, as we will see in our discussion of trade and development in Chapter 10.

SUMMARY

- 1. International factor movements can sometimes substitute for trade, so it is not surprising that international migration of labor is similar in its causes and effects to international trade based on differences in resources. Labor moves from countries where it is abundant to countries where it is scarce. This movement raises total world output, but it also generates strong income distribution effects, so that some groups are hurt.
- 2. International borrowing and lending can be viewed as a kind of international trade, but one that involves trade of present consumption for future consumption rather than trade of one good for another. The relative price at which this intertemporal trade takes place is one plus the real rate of interest.
- **3.** Multinational firms, while they often serve as vehicles for international borrowing and lending, primarily exist as ways of extending control over activities taking place in two or more different countries. The theory of multinational firms is not as well developed as other parts of international economics. A basic framework can be presented that stresses two crucial elements that explain the existence of a multinational: a location motive that leads the activities of the firm to be in different countries, and an internalization motive that leads these activities to be integrated in a single firm.
- **4.** The location motives of multinationals are the same as those behind all international trade. The internalization motives are less well understood; current theory points to two main motives: the need for a way to transfer technology and the advantages in some cases of vertical integration.

KEY TERMS

direct foreign investment, p. 163 factor movements, p. 153 intertemporal production possibility frontier, p. 161 intertemporal trade, p. 161

location and internalization motives of multinationals, p. 164–165 real interest rate, p. 162 technology transfer, p. 166 vertical integration, p. 166

PROBLEMS



1. In Home and Foreign there are two factors of production, land and labor, used to produce only one good. The land supply in each country and the technology of production are exactly the same. The marginal product of labor in each country depends on employment as follows:

Number of Workers Employed	Marginal Product of Last Worker	
1	20	
2	19	
3	18	
4	17	
5	16	
6	15	
7	14	
8	13	
9	12	
10	11	
11	10	

Initially, there are 11 workers employed in Home, but only 3 workers in Foreign.

Find the effect of free movement of labor from Home to Foreign on employment, production, real wages, and the income of landowners in each country.

- **2.** Using the numerical example in Problem 1, assume now that Foreign limits immigration, so that only 2 workers can move there from Home. Calculate how the movement of these two workers affects the income of five different groups:
 - **a.** Workers who were originally in Foreign
 - b. Foreign landowners
 - c. Workers who stay in Home
 - d. Home landowners
 - e. The workers who do move

Studies of the effects of immigration into the United States from Mexico tend to find that the big winners are the immigrants themselves. Explain this result in terms of the example above. How might things change if the border were open, with no restrictions on immigration?

- **3.** The quantity of direct foreign investment by the United States into Mexico has increased dramatically during the past decade. How would you expect this increased quantity of direct foreign investment to affect migration flows from Mexico to the United States, all else being equal?
- **4.** Suppose that a labor-abundant country and a land-abundant country both produce labor- and land-intensive goods with the same technology. Drawing on the analysis in Chapter 4, first analyze the conditions under which trade between the two countries eliminates the incentive for labor to migrate. Then, using the analysis in Chapter 5, show that a tariff by one country will create an incentive for labor migration.
- **5.** Consider a world in which there are two countries, Guatrarica and Costamala, which share an open border such that labor flows freely between the two countries. Total income (GDP) in each country is equal to the sum of wages and rents to capital owners

that accrue from production as in Figure 7-2. Explain the impact on the two countries from a technology shock that increases the marginal product of labor in Costamala:

- **a.** The number of workers in each country.
- **b.** Wages in each country.
- **c.** GDP in each country.
- **d.** Capital rents in each country.
- **6.** Explain the analogy between international borrowing and lending and ordinary international trade.
- 7. Which of the following countries would you expect to have intertemporal production possibilities biased toward current consumption goods, and which biased toward future consumption goods?
 - **a.** A country, like Argentina or Canada in the last century, that has only recently been opened for large-scale settlement and is receiving large inflows of immigrants.
 - **b.** A country, like the United Kingdom in the late 19th century or the United States today, that leads the world technologically but is seeing that lead eroded as other countries catch up.
 - **c.** A country that has discovered large oil reserves that can be exploited with little new investment (like Saudi Arabia).
 - **d.** A country that has discovered large oil reserves that can be exploited only with massive investment (like Norway, whose oil lies under the North Sea).
 - **e.** A country like South Korea that has discovered the knack of producing industrial goods and is rapidly gaining on advanced countries.
- 8. Which of the following are direct foreign investments, and which are not?
 - a. A Saudi businessman buys \$10 million of IBM stock.
 - **b.** The same businessman buys a New York apartment building.
 - **c.** A French company merges with an American company; stockholders in the U.S. company exchange their stock for shares in the French firm.
 - **d.** An Italian firm builds a plant in Russia and manages the plant as a contractor to the Russian government.
- **9.** What are some of the reasons that a country would prefer to open its own production plant overseas rather than to outsource manufacturing to an overseas firm?
- **10.** The Karma Computer Company has decided to open a Brazilian subsidiary. Brazilian import restrictions have prevented the firm from selling into that market, while the firm has been unwilling to sell or lease its patents to Brazilian firms because it fears this will eventually hurt its technological advantage in the U.S. market. Analyze Karma's decision in terms of the theory of multinational enterprise.

FURTHER READING

- Richard A. Brecher and Robert C. Feenstra. "International Trade and Capital Mobility Between Diversified Economies." *Journal of International Economics* 14 (May 1983), pp. 321–339. A synthesis of the theories of trade and international factor movements.
- Richard E. Caves. *Multinational Enterprises and Economic Analysis*. Cambridge: Harvard University Press, 1982. A view of multinational firms' activities.
- Wilfred J. Ethier. "The Multinational Firm." *Quarterly Journal of Economics* 101 (November 1986), pp. 805–833. Models the internalization motive of multinationals.
- Irving Fisher. *The Theory of Interest*. New York: Macmillan, 1930. The "intertemporal" approach described in this chapter owes its origin to Fisher.
- Edward M. Graham and Paul R. Krugman. *Foreign Direct Investment in the United States*. Washington, D.C.: Institute for International Economics, 1989. A survey of the surge of foreign investment in the United States, with an emphasis on policy issues.

- Charles P. Kindleberger. *American Business Abroad*. New Haven: Yale University Press, 1969. A good discussion of the nature and effects of multinational firms, written at a time when such firms were primarily United States-based.
- Charles P. Kindleberger. *Europe's Postwar Growth: The Role of Labor Supply*. Cambridge: Harvard University Press, 1967. A good account of the role of labor migration during its height in Europe.
- G. D. A. MacDougall. "The Benefits and Costs of Private Investment from Abroad: A Theoretical Approach." *Economic Record* 36 (1960), pp. 13–35. A clear analysis of the costs and benefits of factor movement.
- Robert A. Mundell. "International Trade and Factor Mobility." *American Economic Review* 47 (1957), pp. 321–335. The paper that first laid out the argument that trade and factor movement can substitute for each other.
- Jeffrey Sachs. "The Current Account and Macroeconomic Adjustment in the 1970s." *Brookings Papers on Economic Activity*, 1981. A study of international capital flows that takes the approach of viewing such flows as intertemporal trade.

MYECONLAB CAN HELP YOU GET A BETTER GRADE

If your exam were tomorrow, would you be ready? For each chapter, MyEconLab Practice Tests and Study Plans pinpoint which sections you have mastered and which ones you need to study. That way, you are more efficient with your study time, and you are better prepared for your exams.

To see how it works, turn to page 9 and then go to www.myeconlab.com/krugman

Finding Total Output from the Marginal Product Curve

In the text we illustrated the production function two different ways. In Figure 7-1 we showed total output as a function of labor input, holding capital constant. We then observed that the slope of that curve is the marginal product of labor and illustrated that marginal product in Figure 7-2. We now want to demonstrate that the total output is measured by the area under the marginal product curve. (Students who are familiar with calculus will find this obvious: Marginal product is the derivative of total, so total is the integral of marginal. Even for these students, however, an intuitive approach can be helpful.)

In Figure 7A1-1 we show once again the marginal product curve. Suppose that we employ L person-hours. How can we show the total output? Let's approximate this using the marginal product curve. First, let's ask what would happen if we used slightly fewer person-hours, say dL fewer. Then output would be less. The fall in output would be approximately

$dL \times MPL$

that is, the reduction in the work force times the marginal product of labor at the initial level of employment. This reduction in output is represented by the area of the colored rectangle in Figure 7A1-1. Now subtract another few person-hours; the output loss will be another rectangle. This time the rectangle will be taller, because the marginal product of labor rises as the quantity of labor falls. If we continue this process until all the labor is gone, our approximation of the total output loss will be the sum of all the rectangles shown in the figure. When no labor is employed, however, output will fall to zero. So we can





approximate the total output by the sum of the areas of all the rectangles under the marginal product curve.

This is, however, only an approximation, because we used the marginal product of only the first person-hour in each batch of labor removed. We can get a better approximation if we take smaller groups—the smaller the better. As the groups of labor removed get infinitesimally small, however, the rectangles get thinner and thinner, and we approximate ever more closely the total area under the marginal product curve. In the end, then, we find that the total output produced with labor L is equal to the area under the marginal product of labor curve MPL up to L.

Figure 7A1-2 uses the result we just found to show the distribution of income for a given real wage. We know that employers will hire labor up to the point where the real wage, w/P, equals the marginal product. We can immediately read off the graph the total output as the area under the marginal product curve. We can also read off the graph the part of output that is paid out as wages, which is equal to the real wage times employment, and thus to the area of the rectangle shown. The part of the output that is kept by owners of capital, then, is the remainder.

More on Intertemporal Trade

This appendix contains a more detailed examination of the two-period intertemporal trade model described in the chapter. The concepts used are the same as those used in Chapter 5 to analyze international exchanges of different consumption goods at a *single* point in time. In the present setting, however, the trade model explains international patterns of investment and borrowing and the determination of the *intertemporal* terms of trade (that is, the real interest rate).

First consider Home, whose intertemporal production possibility frontier is shown in Figure 7A2-1. Recall that the quantities of present and future consumption goods produced at Home depend on the amount of present consumption goods invested to produce future goods. As currently available resources are diverted from present consumption to investment, production of present consumption, Q_P , falls and production of future consumption, Q_F , rises. Increased investment therefore shifts the economy up and to the left along the intertemporal production possibility frontier.

The chapter showed that the price of future consumption in terms of present consumption is 1/(1 + r), where r is the real interest rate. Measured in terms of present consumption, the value of the economy's total production over the two periods of its existence is therefore

$$V = Q_P + Q_F / (1 + r).$$

Figure 7A2-1 shows the isovalue lines corresponding to the relative price 1/(1 + r) for different values of V. These are straight lines with slope -(1 + r) (because future consumption is on the vertical axis). As in the standard trade model, firms' decisions lead to





a production pattern that maximizes the value of production at market prices, $Q_P + Q_F/(1 + r)$. Production therefore occurs at point Q. The economy invests the amount shown, leaving Q_P available for present consumption and producing an amount Q_F of future consumption when the first-period investment pays off.

Notice that at point Q, the extra future consumption that would result from investing an additional unit of present consumption just equals (1 + r). It would be inefficient to push investment beyond point Q because the economy could do better by lending additional present consumption to foreigners instead. Figure 7A2-1 implies that a rise in the world real interest rate r, which steepens the isovalue lines, causes investment to fall.

Figure 7A2-2 shows how Home's consumption pattern is determined for a given world interest rate. Let D_P and D_F represent the demands for present and future consumption goods, respectively. Since production is at point Q, the economy's consumption possibilities over the two periods are limited by the *intertemporal budget constraint*:

$$D_P + D_F/(1 + r) = Q_P + Q_F/(1 + r)$$

This constraint states that the value of Home's consumption over the two periods (measured in terms of present consumption) equals the value of consumption goods produced in the two periods (also measured in present consumption units). Put another way, production and consumption must lie on the same isovalue line.

Point *D*, where Home's budget constraint touches the highest attainable indifference curve, shows the present and future consumption levels chosen by the economy. Home's demand for present consumption, D_P , is smaller than its production of present consumption, Q_P , so it exports (that is, lends) $Q_P - D_P$ units of present consumption to Foreigners. Correspondingly, Home imports $D_F - Q_F$ units of future consumption from abroad when its first-period loans are repaid to it with interest. The intertemporal budget constraint implies that $D_F - Q_F = (1 + r) \times (Q_P - D_P)$, so that trade is *intertemporally* balanced.



Figure 7A2-3 shows how investment and consumption are determined in Foreign. Foreign is assumed to have a comparative advantage in producing *future* consumption goods. The diagram shows that at a real interest rate of r, Foreign borrows consumption goods in the first period and repays this loan using consumption goods produced in the second period. Because of its relatively rich domestic investment opportunities and its relative preference for present consumption, Foreign is an importer of present consumption and an exporter of future consumption.

As in the Appendix to Chapter 5, international equilibrium can be portrayed by an offer curve diagram. Recall that a country's offer curve is the result of plotting its desired exports against its desired imports. Now, however, the exchanges plotted involve present and future consumption. Figure 7A2-4 shows that the equilibrium real interest rate is determined by the intersection of the Home and Foreign offer curves *OP* and *OF* at point *E*. The ray *OE* has slope $(1 + r^1)$, where r^1 is the equilibrium world interest rate. At point *E*, Home's desired export of present consumption equals Foreign's desired import of present consumption. Put another way, at point *E*, Home's desired first-period lending equals Foreign's desired first-period borrowing. Supply and demand are therefore equal in both periods.

