Introduction

In Parts One and Two, we dealt with the "real," as opposed to the monetary, side of the economy. Money was not explicitly considered, and the discussion was in terms of relative commodity prices. We now begin our examination of the monetary aspects of international economics, or international finance. Here, money is explicitly brought into the picture, and commodity prices are expressed in terms of domestic and foreign currency units. We begin our discussion of international finance by examining the balance of payments.

The balance of payments is a summary statement in which, in principle, all the transactions of the residents of a nation with the residents of all other nations are recorded during a particular period of time, usually a calendar year. The United States and some other nations also keep such a record on a quarterly basis. The main purpose of the balance of payments is to inform the government of the international position of the nation and to help it in its formulation of monetary, fiscal, and trade policies. Governments also regularly consult the balance of payments of important trade partners in making policy decisions. The information contained in a nation's balance of payments is also indispensable to banks, firms, and individuals directly or indirectly involved in international trade and finance.

The definition of the balance of payments just given requires some clarification. First of all, it is obvious that the literally millions of transactions of the residents of a nation with the rest of the world cannot appear *individually* in the balance of payments. As a *summary statement*, the balance of payments aggregates all merchandise trade into a few major categories. Similarly, only the net balance of each type of international capital flow is included. Furthermore, the balance of payments includes some transactions in which the residents of foreign nations are not directly involved—for example, when a nation's central bank sells a portion of its foreign currency holdings to the nation's commercial banks.

An *international transaction* refers to the exchange of a good, service, or asset (for which payment is usually required) between the residents of one nation and the residents of other nations. However, gifts and certain other transfers (for which no payment is required) are also included in a nation's balance of payments. The question of who is a *resident* of a nation also requires some clarification. Diplomats, military personnel, tourists, and workers who temporarily migrate are residents of the nation in which they hold citizenship. Similarly, a corporation is the resident of the nation in which it is incorporated, but its foreign branches and subsidiaries are not. Some of these distinctions are, of course, arbitrary and may lead to difficulties. For example, a worker may start by emigrating temporarily and then decide to remain abroad permanently. International institutions such as the United Nations, the International Monetary Fund (IMF), the World Bank, and the World Trade Organization (WTO) are not residents of the nation in which they are located. Also to be remembered is that the balance of payments has a time dimension. Thus, it is the flow of goods, services, gifts, and assets between the residents of a nation and the residents of other nations *during a particular period of time*, usually a calendar year.

In this chapter, we examine the international transactions of the United States and other nations. In Section 13.2, we discuss some accounting principles used in the presentation of the balance of payments. In Section 13.3, we present and analyze the international transactions of the United States for the year 2001. Section 13.4 then examines some accounting balances and the concept and measurement of balance-of-payments disequilibrium. Section 13.5 briefly reviews the postwar balance-of-payments history of the United States. Section 13.6 then examines the international investment position of the United States. The appendix presents the method of measuring the balance of payments that all nations must use in reporting to the International Monetary Fund. This ensures consistency and permits international comparison of the balance of payments of different nations. The appendix also examines the reason that the current account balance for the world as a whole does not balance, but instead shows large and persistent deficits.

13.2 Balance-of-Payments Accounting Principles

In this section, we examine some balance-of-payments accounting principles as a necessary first step in the presentation of the international transactions of the

United States. We begin with the distinction between credits and debits, and then we examine double-entry bookkeeping.

13.2A Credits and Debits

International transactions are classified as credits or debits. **Credit transactions** are those that involve the receipt of payments *from* foreigners. **Debit transactions** are those that involve the making of payments *to* foreigners. Credit transactions are entered with a positive sign, and debit transactions are entered with a negative sign in the nation's balance of payments.

Thus, the export of goods and services, unilateral transfers (gifts) received from foreigners, and capital inflows are entered as credits (+) because they involve the receipt of payments from foreigners. On the other hand, the import of goods and services, unilateral transfers or gifts made to foreigners, and capital outflows involve payments to foreigners and are entered as debits (–) in the nation's balance of payments.

Capital inflows can take either of two forms: an increase in foreign assets in the nation or a reduction in the nation's assets abroad. For example, when a U.K. resident purchases a U.S. stock, foreign assets in the United States increase. This is a capital inflow to the United States and is recorded as a credit in the U.S. balance of payments because it involves the receipt of a payment from a foreigner. A capital inflow can also take the form of a reduction in the nation's assets abroad. For example, when a U.S. resident sells a foreign stock, U.S. assets abroad decrease. This is a capital inflow to the United States (reversing the capital outflow that occurred when the U.S. resident purchased the foreign stock) and is recorded as a credit in the U.S. balance of payments because it too involves the receipt of a payment from foreigners.

The definition of capital inflows to the United States as increases in foreign assets in the United States or reductions in U.S. assets abroad can be confusing and is somewhat unfortunate, but this is the terminology actually used in all U.S. government publications. Confusion can be avoided by remembering that when a foreigner purchases a U.S. asset (an increase in foreign assets in the United States), this involves the receipt of a payment from foreigners. Therefore, it is a capital inflow, or credit. Similarly, when a U.S. resident sells a foreign asset (a reduction in U.S. assets abroad), this also involves a payment from foreigners; therefore, it too represents a capital inflow to the United States and a credit. Both an increase in foreign assets in the United States and a reduction in U.S. assets abroad are capital inflows, or credits, because they both involve the receipt of payment from foreigners.

On the other hand, **capital outflows** can take the form of either an increase in the nation's assets abroad or a reduction in foreign assets in the nation because both involve a payment to foreigners. For example, the purchase of a U.K. treasury bill by a U.S. resident increases U.S. assets abroad and is a debit because it involves a payment to foreigners. Similarly, the sale of its U.S. subsidiary by a German firm reduces foreign assets in the United States and is also a debit because it involves a payment to foreigners. (The student should study these definitions and examples carefully, since mastery of these important concepts is crucial to understanding what follows.)

To summarize, the export of goods and services, the receipt of unilateral transfers, and capital inflows are credits (+) because they all involve the receipt of payments from foreigners. On the other hand, the import of goods and services, unilateral transfers to foreigners, and capital outflows are debits (–) because they involve payments to foreigners.

13.2B Double-Entry Bookkeeping

In recording a nation's international transactions, the accounting procedure known as **double-entry bookkeeping** is used. This means that each international transaction is recorded twice, once as a credit and once as a debit of an equal amount. The reason for this is that in general every transaction has two sides. We sell something and we receive payment for it. We buy something and we have to pay for it.

For example, suppose that a U.S. firm exports $500 of goods to be paid for in three months. The United States first credits goods exports for $500 since this goods export will lead to the receipt of a payment from foreigners. *The payment itself* is then entered as a capital debit because it represents a capital outflow from the United States. That is, by agreeing to wait three months for payment, the U.S. exporter is extending credit to, and has acquired a claim on, the foreign importer. This is an increase in U.S. assets abroad and a debit. The entire transaction is entered as follows in the U.S. balance of payments:

	Credit (+)	Debit (–)
Goods exports	$500	
Capital outflow		$500

← increase in nation's assets abroad

As another example of double-entry bookkeeping, suppose that a U.S. resident visits London and spends $200 on hotels, meals, and so on. The U.S. resident is purchasing travel services from foreigners requiring a payment. (This is similar to a U.S. import.) Thus, the U.S. debits travel services for $200. The payment itself is then entered as a credit because it represents an increase in foreign claims on the United States. Specifically, we can think of the $200 in British hands as "securities" giving the United Kingdom a claim on U.S. goods and services, equivalent to an increase in foreign assets in the United States. Therefore, it is a capital inflow to the United States and is recorded as a capital credit of $200. The entire transaction is entered as follows in the U.S. balance of payments:

	Credit (+)	Debit (–)
Travel services purchased from foreigners		$200
Capital inflow	$200	

As a third example, assume that the U.S. government gives a U.S. bank balance of $100 to the government of a developing nation as part of the U.S. aid program. The United States debits **unilateral transfers** for the $100 gift given (payment made) to foreigners. The payment itself is the U.S. bank balance given to the government of the developing nation. This represents an increase in foreign claims on, or foreign assets in, the United States and is recorded as a capital inflow, or credit, in the U.S. balance of payments. The entire transaction is thus:

	Credit (+)	Debit (−)
Unilateral transfers made		$100
Capital inflow	$100	

As a fourth example, suppose that a U.S. resident purchases a foreign stock for $400 and pays for it by increasing *foreign* bank balances in the United States. The purchase of the foreign stock increases U.S. assets abroad. This is a capital outflow from the United States and is recorded as a capital debit of $400 in the U.S. balance of payments. The increase in foreign bank balances in the United States is an increase in foreign assets in the United States (a capital inflow to the United States) and is entered as a credit in the U.S. balance of payments. The result would be the same if the U.S. resident paid for the foreign stock by reducing bank balances abroad. (This would be a reduction in U.S. assets abroad, which is also a capital inflow to the United States and a credit.) Note that both sides of this transaction are financial:

	Credit (+)	Debit (−)
Capital outflow (the purchase of the foreign stock by the U.S. resident)		$400
Capital inflow (the increase in foreign bank balances in the U.S.)	$400	

Finally, suppose that a foreign investor purchases $300 of U.S. treasury bills and pays by drawing down his bank balances in the United States by an equal amount. The purchase of the U.S. treasury bills increases foreign assets in the United States. This is a capital inflow to the United States and is recorded as a credit in the U.S. balance of payments. The drawing down of U.S. bank balances by the foreigner is a reduction in foreign assets in the United States. This is a capital outflow from the United States and is recorded as a capital outflow of the U.S. balance of payments:

	Credit (+)	Debit (−)
Capital inflow (the purchase of U.S. treasury bills by a foreigner)	$300	
Capital outflow (the reduction in foreign bank balances in the U.S.)		$300

If we assume that these five transactions are all the international transactions of the United States during the year, then the U.S. balance of payments is as follows:

	Credit (+)	Debit (−)
Goods	$500	
Services		$200
Unilateral transfers		100
Capital, net	___	200
Total debits and credits	$500	$500

The net capital debit balance of –$200 is obtained by adding together the seven capital entries (–$500, $200, $100, –$400, $400, $300, –$300) previously examined separately. Total debits equal total credits because of double-entry bookkeeping.

The traditional distinction between short-term capital and long-term capital (i.e., capital with maturity of more than one year, such as a bond or a stock, as opposed to three-month treasury bills) is usually no longer made because bonds and stocks are liquid (i.e., can be sold and bought almost immediately).

13.4 Accounting Balances and Disequilibrium in International Transactions

The first accounting balance in the memoranda at the bottom of Table 13.1 is the balance on goods trade. In 2001, the United States exported $719 billion and imported $1,146 billion of goods, for a net debit balance on goods of (–)$427 billion. On the other hand, the United States had a net credit balance on services of $69 billion (from the $279 billion export of services minus the $210 billion import of services). Thus, the United States had a net debit balance on goods and services of (–)$358 billion. The United States also had a net credit balance of (+)$14 billion on investment income (from the $284 billion interest and dividends earned on U.S. investment abroad minus the $269 billion income payments on foreign assets in the United States and a $1 billion rounding error). The United States therefore had a net debit balance on goods, services, and income of (–)$344 billion.

Adding the net debit balance of (–)$49 billion of unilateral transfers to the net debit balance of (–)$344 billion on goods, services, and income, we get the current account net debit balance of (–)$393 billion. Thus, the **current account** lumps together all sales and purchases of currently produced goods and services, investment incomes, and unilateral transfers and provides the link between the nation's international transactions and its national income. Specifically, a current account surplus stimulates domestic production and income, while a current account deficit dampens domestic production and income. (This link between the nation's international trade and current account and its national income will be examined in detail in Chapter 17.)

The change in U.S.-owned assets abroad and foreign-owned assets in the United States, other than official reserve assets, gives the **capital account** of the United States. This measures the change in the stock of all nonreserve financial assets. The justification for excluding financial reserve assets from the capital account is that changes in reserves reflect government policy rather than market forces. Thus, the capital account of the United States shows a net increase in U.S.-owned assets abroad (a capital outflow of the United States) of (–)$366 billion (from the total of –$371 billion minus the –$5 billion net increase in U.S. official reserve assets) and a net increase in foreign-owned assets in the United States (a capital inflow to the United States) of $748 billion (from the total of $753 billion minus the +$5 billion net *increase* in foreign official assets in the United States). Thus, the United States had a net credit balance (a net capital inflow) in its capital account of (+)$382 billion (–$366 billion +$748 billion) in 2001.

All transactions in the current and capital accounts are called **autonomous transactions** because they take place for business or profit motives (except for unilateral transfers) and independently of balance-of-payments considerations. Autonomous items are sometimes referred to as "the items above the line." On the other hand, transactions in official reserve assets are called **accommodating transactions** (or "items below the line") because they result from and are needed to balance international transactions. The accommodating or below-the-line items form the **official reserve account**, and the balance on the official reserve account is called the **official settlements balance**.

If total debits exceed total credits in the current and capital accounts, the net debit balance measures the deficit in the nation's balance of payments. This deficit must then be settled with an equal net credit in the official reserve account. Thus, a **deficit in the balance of payments** can be measured either by the excess of debits over credits in the current and capital accounts or by the excess of credits over debits in the official reserve account. On the other hand, a nation has a **surplus in the balance of payments** if its total credits exceed its total debits in the current and capital accounts. The net credit balance gives the size of the surplus and is settled by an equal debit balance in the official reserve account.

From Table 13.1, we see that during 2001 there was a net increase in U.S. official reserve assets (a capital outflow of the United States) of (–)$5 billion. At the same time, there was a net increase in foreign official assets in the United States (a capital inflow to the United States) of (+)$5 billion, resulting in a zero net balance on transactions in official reserve assets or the official reserve account. This is equal to the zero net balance on the items above the line (obtained from the net debit balance of (–)$393 billion in the current account, plus the net credit balance of +$382 billion in the capital account, and the +$11 billion credit for the statistical discrepancy—see Table 13.1).

This method of measuring the deficit (or surplus) in the balance of payments (as well as the concept itself) is strictly correct, however, only under a fixed exchange rate system (as we had from the end of World War II until 1973) and not under a flexible exchange rate system or a managed exchange rate system (that we have had since then). The reason for this is explained in Chapter 14 after we have defined the exchange rate and examined how it is determined.

1. What is meant by the balance of payments? In what way is the balance of payments a summary statement? What is meant by an international transaction? How is a resident of a nation defined? In what way is the time element involved in measuring a nation's balance of payments?

2. What is a credit transaction? a debit transaction? Which are the broad categories of international transactions classified as credits? as debits?

3. What is double-entry bookkeeping? Why does double-entry bookkeeping usually involve an entry called statistical discrepancy? How does such a statistical discrepancy arise?

Equilibrium Foreign Exchange Rates

Assume for simplicity that there are only two economies, the United States and the European Monetary Union (EMU), with the dollar ($) as the domestic currency and the euro (€) as the foreign currency. The **exchange rate** between the dollar and the euro (R) is equal to the number of dollars needed to purchase one euro. That is, $R = \$/€$. For example, if $R = \$/€ = 1$, this means that one dollar is required to purchase one pound.

Under a flexible exchange rate system of the type we have today, the dollar price of the euro (R) is determined, just like the price of any commodity, by the intersection of the market demand and supply curves for euros. This is shown in Figure 14.1, where the vertical axis measures the dollar price of the euro, or the exchange rate, $R = \$/€$, and the horizontal axis measures the quantity of euros. The market demand and supply curves for euros intersect at point E, defining the equilibrium exchange rate of $R = 1$, at which the quantity of euros demanded and the quantity supplied are equal at €200 million per day. At a higher exchange rate, the quantity of euros supplied exceeds the quantity demanded, and the exchange rate will fall toward the equilibrium rate of $R = 1$. At an exchange rate lower than $R = 1$, the quantity of euros demanded exceeds the quantity supplied, and the exchange rate will be bid up toward the equilibrium rate of $R = 1$. If the exchange rate were not allowed to rise to its equilibrium level (as under the fixed exchange rate system that prevailed until March 1973), then either restrictions would have to be imposed on the demand for euros of U.S. residents or the U.S. central bank (the Federal Reserve System) would have to fill or satisfy the excess demand for euros out of its international reserves.

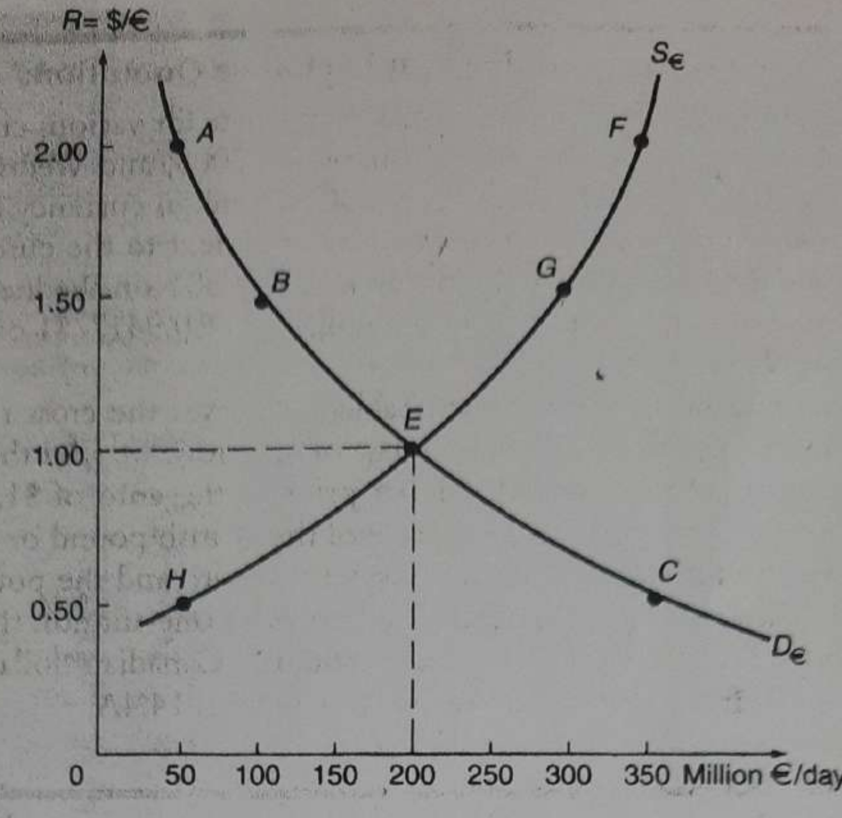

FIGURE 14.1. The Exchange Rate under a Flexible Exchange Rate System. The vertical axis measures the dollar price of the euro (R = \$/€), and the horizontal axis measures the quantity of euros. With a flexible exchange rate system, the equilibrium exchange rate is $R = 1$, at which the quantity demanded and the quantity supplied are equal at €200 million per day. This is given by the intersection at point E of the U.S. demand and supply curves for euros. At a higher exchange rate, a surplus of euros would result that would tend to lower the exchange rate toward the equilibrium rate. At an exchange rate lower than R = 1, a shortage of euros would result that would drive the exchange rate up toward the equilibrium level.

The U.S. demand for euros is negatively inclined, indicating that the lower the exchange rate *(R)*, the greater the quantity of euros demanded by the U.S. residents. The reason is that the lower the exchange rate (i.e., the fewer the number of dollars required to purchase a euro), the cheaper it is for the U.S. residents to import from and to invest in the European Monetary Union, and thus the greater the quantity of euros demanded by U.S. residents. On the other hand, the U.S. supply of euros is usually positively inclined (see Figure 14.1), indicating that the higher the exchange rate *(R)*, the greater the quantity of euros earned by U.S. residents and supplied to the United States. The reason is that at higher exchange rates, EMU residents receive more dollars for each of their euros. As a result, they find U.S. goods and investments cheaper and more attractive and spend more in the United States, thus supplying more euros to the United States.

If the U.S. demand curve for euros shifted up (for example, as a result of increased U.S. tastes for EMU goods) and intersected the U.S. supply curve for euros at point G (see Figure 14.1), the equilibrium exchange rate would be R = 1.50, and the equilibrium quantity of pounds would be €300 million per day. The dollar is then said to have depreciated since it now requires \$1.50 (instead of the previous \$1) to purchase one euro. **Depreciation** thus refers to an increase in the domestic price of the foreign currency. On the other hand, if the U.S. demand curve for euros shifted down so as to intersect the U.S. supply curve for euros at point H (see Figure 14.1), the equilibrium exchange rate would fall to $R = 0.5$ and the dollar is said to have appreciated (because fewer dollars are now required to purchase one euro). **Appreciation** thus refers to a decline in the domestic price of the foreign currency. An appreciation of the domestic currency means a depreciation of the foreign currency and vice versa. Shifts in the U.S. supply curve for euros would similarly affect the equilibrium exchange rate and equilibrium quantity of euros (these are left as end-of-chapter problems).

The exchange rate could also be defined as the foreign currency price of a unit of the domestic currency. This is the inverse, or reciprocal, of our previous definition. Since in the case we examined above, the dollar price of the euro is $R = 1$, its inverse is also 1. If the dollar price of the euro were instead $R = 2$, then the euro price of the dollar would be $1/R = 1/2$, or it would take half a euro to purchase one dollar. Although this definition of the exchange rate is sometimes used, we will use the previous one, or the dollar price of the euro (R), unless clearly stated to the contrary. In the real world, the particular definition of the exchange rate being used is generally spelled out to avoid confusion (see Case Study 14-3).

Finally, while we have dealt with only two currencies for simplicity, in reality there are numerous exchange rates, one between any pair of currencies. Thus, besides the exchange rate between the U.S. dollar and the euro, there is an exchange rate between the U.S. dollar and the British pound (£), between the U.S. dollar and the Swiss franc, the Canadian dollar and the Mexican peso, the British pound and the euro, the euro and the Swiss franc, and between each of these currencies and the Japanese yen. Once the exchange rate between each of a pair of currencies with respect to the dollar is established, however, the exchange rate between the two currencies themselves, or **cross exchange rate**, can easily be determined. For example, if the exchange rate *(R)* were 2 between the U.S. dollar and the British pound and 1.25 between the dollar and the euro, then the exchange rate between the pound and the euro would be 1.60 (i.e., it takes €1.6 to purchase 1£). Specifically,

$$R = €/£ = \frac{\$ \text{ value of } £}{\$ \text{ value of } €} = \frac{2}{1.25} = 1.60$$

The lower portion of Table 14.2 in Case Study 14-3 gives the actual cross rates for several leading currencies.

Since over time a currency can depreciate with respect to some currencies and appreciate against others, an **effective exchange rate** is calculated. This is a weighted average of the exchange rates between the domestic currency and the nation's most important trade partners, with weights given by the relative importance of the nation's trade with each of these trade partners (see Section 14.5A). Finally, we must also distinguish between the nominal exchange rate (the one we have been discussing) and the real exchange rate (to be discussed in Chapter 15).

Arbitrage

The exchange rate between any two currencies is kept the same in different monetary centers by **arbitrage**. This refers to the purchase of a currency in the monetary center where it is cheaper, for immediate resale in the monetary center where it is more expensive, in order to make a profit.

For example, if the dollar price of the euro was \$0.99 in New York and \$1.01 in Frankfurt, an arbitrageur (usually a foreign exchange dealer of a commercial bank) would purchase euros at \$0.99 in New York and immediately resell them in Frankfurt for \$1.01, thus realizing a profit of \$0.02 per euro. While the profit per euro transferred seems small, on €1 million the profit would be \$20,000 for only a few minutes' work. From this profit must be deducted the cost of the electronic transfer and the other costs associated with arbitrage. Since these costs are very small, we shall ignore them here.

As arbitrage takes place, however, the exchange rate between the two currencies tends to be equalized in the two monetary centers. Continuing our example, we see that arbitrage increases the demand for euros in New York, thereby exerting an upward pressure on the dollar price of euros in New York. At the same time, the sale of euros in Frankfurt increases the supply of euros there, thus exerting a downward pressure on the dollar price of euros in Frankfurt. This continues until the dollar price of the euro quickly becomes equal in New York and Frankfurt (say at \$1 = €1), thus eliminating the profitability of further arbitrage.

When only two currencies and two monetary centers are involved in arbitrage, as in the preceding example, we have *two-point arbitrage*. When three currencies and three monetary centers are involved, we have *triangular, or three-point, arbitrage*. While triangular arbitrage is not very common, it operates in the same manner to ensure *consistent indirect, or cross, exchange rates* between the three currencies in the three monetary centers. For example, suppose exchange rates are as follows:

\$1 = €1 in New York
€1 = £0.64 in Frankfurt
£0.64 = \$1 in London

These cross rates are consistent because

$$\$1 = €1 = £0.64$$

and there is no possibility of profitable arbitrage. However, if the dollar price of the euro were $0.96 in New York, with the other exchange rates as indicated previously, then it would pay to use $0.96 to purchase €1 in New York, use the €1 to buy £0.64 in Frankfurt, and exchange the £0.64 for $1 in London, thus realizing a $0.04 profit on each euro so transferred. On the other hand, if the dollar price of the euro was $1.04 in New York, it would pay to do just the opposite—that is, use $1 to purchase £0.64 in London, exchange the £0.64 for €1 in Frankfurt, and exchange the €1 for $1.04 in New York, thus making a profit of $0.04 on each euro so transferred.

As in the case of two-point arbitrage, triangular arbitrage increases the demand for the currency in the monetary center where the currency is cheaper, increases the supply of the currency in the monetary center where the currency is more expensive, and quickly eliminates inconsistent cross rates and the profitability of further arbitrage. As a result, arbitrage quickly equalizes exchange rates for each pair of currencies and results in consistent cross rates among all pairs of currencies, thus unifying all international monetary centers into a single market.

.3c The Exchange Rate and the Balance of Payments

We can examine the relationship between the exchange rate and the nation's balance of payments with Figure 14.2, which is identical to Figure 14.1 except for the addition of the new demand curve for euros labeled $D'_€$. We have seen in Chapter 13 that the U.S. demand for euros ($D_€$) arises from the U.S. demand for imports of goods and services from the European Union, from U.S. unilateral transfers to the European Union, and from U.S. investments in the European Monetary Union (a capital outflow from the United States). These are the autonomous debit transactions of the United States that involve payments to the European Monetary Union. On the other hand, the supply of euros ($S_€$) arises from U.S. exports of goods and services to the European Monetary Union, from unilateral transfers received from the European Monetary Union, and from the EMU investments in the United States (a capital inflow to the United States). These are the autonomous credit transactions of the United States that involve payments from the European Monetary Union. (We are assuming for simplicity that the United States and the European Monetary Union are the only two economies in the world and that all transactions between them take place in euros.)

With $D_€$ and $S_€$, the equilibrium exchange rate is $R = \$/€ = 1$ (point E in Figure 14.2), at which €200 million are demanded and supplied per day (exactly as in Figure 14.1). Now suppose that for whatever reason (such as an increase in U.S. tastes for EMU products) the U.S. autonomous demand for euros shifts up to $D'_€$. If the United States wanted to maintain the exchange rate fixed at $R = 1$, U.S. monetary authorities would have to satisfy the excess demand for euros of TE (€450 million per day in Figure 14.2) out of their official reserve holdings of euros. Alternatively, EMU monetary authorities would have to purchase dollars (thus adding to

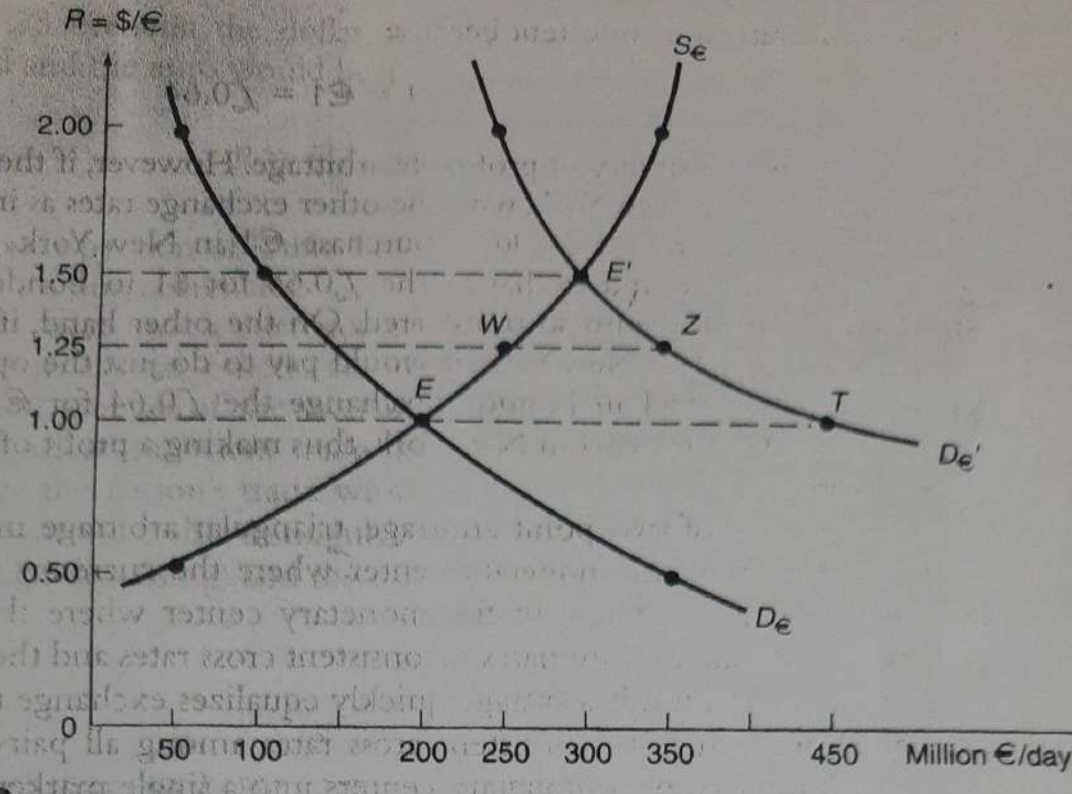

FIGURE 14.2. Disequilibrium under a Fixed and a Flexible Exchange Rate System. With $D_€$ and $S_€$, equilibrium is at point *E* at the exchange rate of R = \$/€ = 1, at which the quantities of euros demanded and supplied are equal at €200 million per day. If $D_€$ shifted up to $D'_€$, the United States could maintain the exchange rate at $R = 1$ by satisfying (out of its official euro reserves) the excess demand of €250 million per day (*TE* in the figure). With a freely flexible exchange rate system, the dollar would depreciate until $R = 1.50$ (point *E'* in the figure). If, on the other hand, the United States wanted to limit the depreciation of the dollar to $R = 1.25$ under a managed float, it would have to satisfy the excess demand of €100 million per day (*WZ* in the figure) out of its official euro reserves.

their official dollar reserves) and supply euros to the foreign exchange market to prevent an appreciation of the euro (a depreciation of the dollar). In either case, the U.S. official settlements balance would show a deficit of €450 million (\$450 million at the official exchange rate of $R = 1$) per day, or €164.25 billion (\$164.25 billion) per year.

If, on the other hand, the United States operated under a freely flexible exchange rate system, the exchange rate would rise (i.e., the dollar would depreciate) from $R = 1.00$ to $R = 1.50$, at which the quantity of euros demanded (€300 million per day) exactly equals the quantity supplied (point *E'* in Figure 14.2). In this case, the United States would not lose any of its official euro reserves. Indeed, international reserves would be entirely unnecessary under such a system. The tendency for an excess demand for euros on autonomous transactions would be completely eliminated by a sufficient depreciation of the dollar with respect to the euro.

However, under a managed floating exchange rate system of the type in operation since 1973, U.S. monetary authorities can intervene in foreign exchange markets to moderate the depreciation (or appreciation) of the dollar. In the preceding example, the United States might limit the depreciation of the dollar to $R = 1.25$

(instead of letting the dollar depreciate all the way to $R = 1.50$ as under a freely fluctuating exchange rate system). The United States could do this by supplying to the foreign exchange market the excess demand for euros of WZ, or €100 million per day, out of its official euro reserves (see the figure). Under such a system, part of the potential deficit in the U.S. balance of payments is covered by the loss of official reserve assets of the United States, and part is reflected in the form of a depreciation of the dollar. Thus, we cannot now measure the deficit in the U.S. balance of payments by simply measuring the loss of U.S. international reserves or by the amount of the net credit balance in the official reserve account of the United States. Under a managed float, the loss of official reserves only indicates the degree of official intervention in foreign exchange markets to influence the level and movement of exchange rates, and not the balance-of-payments deficit.

For this reason, since 1976 the United States has suspended the calculation of the balance-of-payments deficit or surplus. The statement of international transactions does not even show the net balance on the official reserve account (although it can be easily calculated) in order to be neutral and not to focus undue attention on such a balance, in view of the present system of floating but managed exchange rates (see Table 13.1).

The concept and measurement of international transactions and the balance of payments are still very important and useful, however, for several reasons. First, as pointed out in Chapter 13, the flow of trade provides the link between international transactions and the national income. (This link will be examined in Chapter 17.) Second, many developing countries still operate under a fixed exchange rate system and peg their currency to a major currency, such as the U.S. dollar and the euro, or to SDRs. Third, the International Monetary Fund requires all member nations to report their balance-of-payments statement annually to it (in the specific format shown in Section A13.1). Finally, and perhaps more important, while not measuring the deficit or surplus in the balance of payments, the balance of the official reserve account gives an indication of the degree of intervention by the nation's monetary authorities in the foreign exchange market to reduce exchange rate volatility and to influence exchange rate levels.

15.3 Monetary Approach to the Balance of Payments and Exchange Rates

In this section we examine the **monetary approach to the balance of payments**. This approach was started toward the end of the 1960s by *Robert Mundell* and *Harry Johnson* and became fully developed during the 1970s. The monetary approach represents an extension of domestic monetarism (stemming from the Chicago school) to the international economy in that it views the balance of payments as an *essentially monetary phenomenon*. That is, money plays the crucial role in the long run both as a disturbance and as an adjustment in the nation's balance of payments. In Section 15.3A we examine the monetary approach under fixed exchange rates, in Section 15.3B we examine the monetary approach under flexible exchange rates, in Section 15.3C we show how exchange rates are determined according to the monetary approach, and in Section 15.3D we discuss the effect of expectations on exchange rates.

15.3A Monetary Approach under Fixed Exchange Rates

The monetary approach begins by postulating that the demand for *nominal* money balances is positively related to the level of *nominal* national income and is stable in the long run. Thus, the equation for the **demand for money** can be written as:

$$M_d = kPY \tag{15-3}$$

where M_d = quantity demanded of nominal money balances
k = desired ratio of nominal money balances to nominal national income
P = domestic price level
Y = real output

In Equation (15-3), PY is the nominal national income or output (GDP). This is assumed to be at or to tend toward full employment in the long run. The symbol k is the desired ratio of nominal money balances to nominal national income; k is also equal to $1/V$, where V is the velocity of circulation of money or the number of times a dollar turns over in the economy during a year. With V (and thus k) depending on institutional factors and assumed to be constant, M_d is a stable and positive function of the domestic price level and real national income.

For example, if GDP = PY = \$1 billion and $V = 5$ (so that $k = 1/V = 1/5$), then $M_d = (1/5)PY = (1/5)(\$1 \text{ billion}) = \$200$ million. Although not included in Equation (15-3), the demand for money is also related, but inversely, to the interest rate (i) or opportunity cost of holding inactive money balances rather than interest-bearing securities. Thus, M_d is directly related to PY and inversely related to i. (This more complete money demand function is formally presented in the appendix to this chapter.) To simplify the analysis, however, we assume for now that M_d is related only to PY, or the nation's nominal GDP, and will work with Equation (15-3).

On the other hand, the nation's **supply of money** is given by

$$M_s = m(D + F) \tag{15-4}$$

where M_s = the nation's total money supply
m = money multiplier
D = domestic component of the nation's monetary base
F = international or foreign component of the nation's monetary base

The domestic component of the nation's monetary base (D) is the domestic credit created by the nation's monetary authorities or the domestic assets backing the nation's money supply. The international or foreign component of the nation's money supply (F) refers to the international reserves of the nation, which can be increased or decreased through balance-of-payments surpluses or deficits, respectively. $D + F$ is called the **monetary base** of the nation, or *high-powered money*. Under a fractional-reserve banking system (such as we have today), each new dollar of D or F deposited in any commercial bank results in an increase in the nation's money supply by a multiple of \$1. This is the money multiplier, m, in Equation (15-4).

For example, a new deposit of \$1 in a commercial bank allows the bank to lend (i.e., to create demand deposits for borrowers) \$0.80, if the legal reserve requirement (LRR) is 20 percent. The \$0.80 lent by the first bank is usually used by the borrower to make a payment and ends up as a deposit in another bank of the system, which proceeds to lend 80 percent of it (\$0.64), while retaining 20 percent (\$0.16) as reserve. The process continues until the original \$1 deposit has become the reserve base of a total of \$1.00 + \$0.80 + \$0.64 + … = \$5 in demand deposits (which are part of the nation's total money supply). The figure of \$5 is obtained by dividing the original deposit of \$1 by the legal reserve requirement of 20 percent, or 0.2. That is, $\$1/0.2 = 5 = m$. However, due to excess reserves and leakages, the real-world multiplier is likely to be smaller. In what follows, we assume for simplicity that the money multiplier (m) is constant over time.

Starting from a condition of equilibrium where $M_d = M_s$, an increase in the demand for money (resulting, say, from a once-and-for-all increase in the nation's GDP) can be satisfied either by an increase in the nation's domestic monetary base (D) or by an inflow of international reserves, or balance-of-payments surplus (F). If the nation's monetary authorities do not increase D, the excess demand for money will be satisfied by an increase in F. On the other hand, an increase in the domestic component of the nation's monetary base (D) and money supply (M_s), in the face of unchanged money demand (M_d), flows out of the nation and leads to a fall in F (a deficit in the nation's balance of payments). Thus, a surplus in the nation's balance of payments results from an excess in the stock of money demanded that is not satisfied by an increase in the domestic component of the nation's monetary base, while a deficit in the nation's balance of payments results from an excess in the stock of the money supply of the nation that is not eliminated by the nation's monetary authorities but is corrected by an outflow of reserves.

For example, an increase in the nation's GNP from \$1 billion to \$1.1 billion increases M_d from \$200 million (1/5 of \$1 billion) to \$220 million (1/5 of \$1.1 billion). If the nation's monetary authorities keep D constant, F will ultimately have to increase (a surplus in the nation's balance of payments) by \$4 million, so that the nation's money supply also increases by \$20 million (the \$4 million increase in F times the money multiplier of $m = 5$). Such a balance-of-payments surplus could be generated from a surplus in the current account or the capital account of the nation. How this surplus arises is not important at this time, except to note that the excess demand for money will lead to a balance-of-payments surplus that increases M_s by the same amount. On the other hand, an excess in the stock of money supplied will lead to an outflow of reserves (a balance-of-payments deficit) sufficient to eliminate the excess supply of money in the nation.

The nation, therefore, has no control over its money supply under a fixed exchange rate system in the long run. That is, the size of the nation's money supply will be the one that is consistent with equilibrium in its balance of payments in the long run. Only a reserve-currency country, such as the United States, retains control over its money supply in the long run under a fixed exchange rate system because foreigners willingly hold dollars.

To summarize, a *surplus* in the nation's balance of payments *results* from an *excess in the stock of money demanded* that is not satisfied by domestic monetary authorities. On the other hand, a *deficit* in the nation's balance of payments *results* from an *excess in the stock of money supplied* that is not eliminated or corrected by the nation's monetary authorities. The nation's balance-of-payments surplus or deficit is temporary and self-correcting in the long run; that is, after the excess demand for or supply of money is eliminated through an inflow or outflow of funds, the balance-of-payments surplus or deficit is corrected and the international flow of money dries up and comes to an end. Thus, except for a currency-reserve country, such as the United States, the nation has no control over its money supply in the long run under a fixed exchange rate system.

B Monetary Approach under Flexible Exchange Rates

Under a flexible exchange rate system, balance-of-payments disequilibria are immediately corrected by automatic changes in exchange rates without any international flow of money or reserves. Thus, under a flexible exchange rate system, the nation retains dominant control over its money supply and monetary policy. Adjustment takes place as a result of the change in domestic prices that accompanies the change in the exchange rate. For example, a deficit in the balance of payments (resulting from an excess money supply) leads to an automatic depreciation of the nation's currency, which causes prices and therefore the demand for money to rise sufficiently to absorb the excess supply of money and automatically eliminate the balance-of-payments deficit.

On the other hand, a surplus in the balance of payments (resulting from an excess demand for money) automatically leads to an appreciation of the nation's currency, which tends to reduce domestic prices, thus eliminating the excess demand for money and the balance-of-payments surplus. Whereas under fixed exchange rates, a balance-of-payments disequilibrium is defined as and results from an international flow of money or reserves (so that the nation has no control over its money supply in the long run); under a flexible exchange rate system, a balance-of-payments disequilibrium is immediately corrected by an automatic change in exchange rates and without any international flow of money or reserves (so that the nation retains dominant control over its money supply and domestic monetary policy).

The actual exchange value of a nation's currency in terms of the currencies of other nations is determined by the rate of growth of the money supply and real income in the nation relative to the growth of the money supply and real income in the other nations. For example, assuming zero growth in real income and the demand for money, as well as in the supply of money, in the rest of the world, the growth in the nation's money supply in excess of the growth in its real income and demand for money leads to an increase in prices and in the exchange rate (a depreciation of the currency) of the nation. On the other hand, an increase in the nation's money supply that falls short of the increase in its real income and demand for money tends to reduce prices and the exchange rate (an appreciation of the currency) of the nation. (The actual process by which exchange rates are determined under the monetary approach is examined in the next section.)

Thus, according to the monetary approach, a currency depreciation results from excessive money growth in the nation over time, while a currency appreciation results from inadequate money growth in the nation. Put differently, a nation facing greater inflationary pressure than other nations (resulting from more rapid growth of its money supply in relation to the growth in its real income and demand for money) will find its exchange rate rising (its currency depreciating—see Figure 15.2). On the other hand, a nation facing lower inflationary pressure than the rest of the world will find its exchange rate falling (its currency appreciating). According to *global monetarists,* the depreciation of the U.S. dollar and the appreciation of the German mark during the 1970s were due to excessive money growth and inflationary pressure in the United States, and to the much smaller rate of money growth and inflationary pressure in Germany than in the rest of the world.

With flexible exchange rates, the rest of the world is to some extent shielded from the monetary excesses of some nations. The nations with excessive money growth and depreciating currencies will now transmit inflationary pressures to the rest of the world primarily through their increased imports rather than directly through the export of money or reserves. This will take some time to occur and will depend on how much slack exists in the world economy and on structural conditions abroad.

Under a managed floating exchange rate system of the type in operation today, the nation's monetary authorities intervene in foreign exchange markets and either lose or accumulate international reserves to prevent an "excessive" depreciation or appreciation of the nation's currency, respectively. Under such a system, part of a balance-of-payments deficit is automatically corrected by a depreciation of the nation's currency, and part is corrected by a loss of international reserves (refer back to Figure 14.2). As a result, the nation's money supply is affected by the balance-of-payments deficit, and domestic monetary policy loses some of its effectiveness. Under a managed float, the nation's money supply is similarly affected by excessive or inadequate growth of the money supply in other nations, although to a smaller extent than under a fixed exchange rate system. The operation of the present floating exchange rate system is discussed in detail in Chapters 20 and 21.

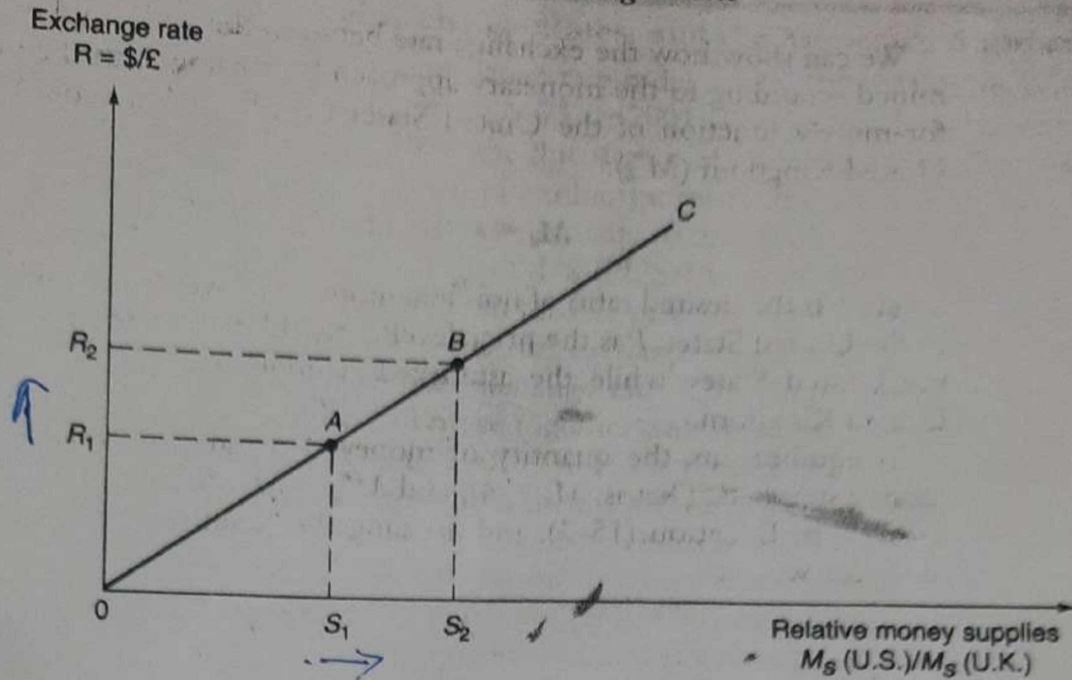

FIGURE 15.2. Relative Money Supplies and Exchange Rates. Line *OC* shows the relationship between the money supply in the United States relative to the money supply in the United Kingdom [$S = M_s(\text{U.S.})/M_s(\text{U.K.})$] and the dollar–pound exchange rate (R = \$/£). Line *OC* thus shows that a change from S_1 to S_2 causes a proportional change in R from R_1 to R_2.

3C Monetary Approach to Exchange Rate Determination

In Section 14.3A, we defined the exchange rate as the domestic currency price of a unit of the foreign currency. With the dollar (\$) as the domestic currency and the pound sterling (£) as the foreign currency, the exchange rate (R) was defined as the number of dollars per pound, or R = \$/£. For example, if R = \$2/£1, this means that two dollars are required to purchase one pound.

If markets are competitive and if there are no tariffs, transportation costs, or other obstructions to international trade, then according to the law of one price postulated by the purchasing-power parity (PPP) theory, the price of a commodity must be the same in the United States as in the United Kingdom. That is, $P_X(\$) = RP_X(£)$. For example, if the price of a unit of commodity X is P_X = £1 in the United Kingdom, then P_X = \$2 in the United States. The same is true for every other traded commodity and for all commodities together (price indices). That is,

$$P = RP^*$$

and

$$R = \frac{P}{P^*} \tag{15-1}$$

where R is the exchange rate of the dollar, P is the index of dollar prices in the United States, and P^* is the index of pound prices in the United Kingdom.

We can show how the exchange rate between the dollar and the pound is determined according to the monetary approach by starting with the nominal demand-for-money function of the United States (M_d, from Equation 15-3) and for the United Kingdom (M^*_d):

$$M_d = kPY \text{ and } M^*_d = k^*P^*Y^*$$

where k is the desired ratio of nominal money balances to nominal national income in the United States, P is the price level in the United States, and Y is real output in the United States, while the asterisked symbols have the same meaning for the United Kingdom.

In equilibrium, the quantity of money demanded is equal to the quantity of money supplied. That is, $M_d = M_s$ and $M^*_d = M^*_s$. Substituting M_s for M_d and M^*_s for M^*_d in Equation (15-3), and dividing the resulting U.K. function by the U.S. function, we get

$$\frac{M^*_s}{M_s} = \frac{k^*P^*Y^*}{kPY} \tag{15-5}$$

By then dividing both sides of Equation (15-5) by P^*/P and M^*_s/M_s we get

$$\frac{P}{P^*} = \frac{M_s k^* Y^*}{M^*_s kY} \tag{15-6}$$

But since $R = P/P^*$ (from Equation 15-1), we have

$$R = \frac{M_s k^* Y^*}{M^*_s kY} \tag{15-7}$$

Since k^* and Y^* in the United Kingdom and k and Y in the United States are assumed to be constant, R is constant as long as M_s and M^*_s remain unchanged. For example, if $k^*Y^*/kY = 1/2$ and $M_s/M^*_s = 4$, then $R = \$2/£1$. In addition, changes in R are proportional to changes in M_s and inversely proportional to changes in M^*_s. For example, if M_s increases by 10 percent in relation to M^*_s, R will increase (i.e., the dollar will depreciate) by 10 percent, and so on.

Several important things need to be noted with respect to Equation (15-7). First, it depends on the purchasing-power parity (PPP) theory and the law of one price (Equation 15-1). Second, Equation (15-7) was derived from the demand for nominal money balances in the form of Equation (15-3), which does not include the interest rate. The relationship between interest rates and the exchange rate is examined in Section 15.3D, which deals with expectations. Third, the exchange rate adjusts to clear money markets in each country without any flow or change in reserves. Thus, for a small country (one that does not affect world prices by its trading), the PPP theory determines the price level under fixed exchange rates and the exchange rate under flexible rates. Case Study 15-4 shows the relationship between increases in the money supply and inflation rates (Equation 15-6), while Case Study 15-5 shows the relationship between the nominal and the real exchange rate and provides a further test of the monetary approach under flexible exchange rates.

As we have seen in Chapter 15, international private capital flows are much larger than trade flows today, and so exchange rates reflect mostly financial rather than trade flows, especially in the short run. Trade flows, however, do have a strong effect on exchange rates in the long run. It is to isolate and identify the effect of trade flows on exchange rates and the effect of exchange rate changes on trade flows that we make the simplifying assumption of no autonomous international private capital flows in this chapter. Of course, in the real world both international financial and trade flows jointly determine exchange rates, but a fully acceptable theory of exchange rate determination that incorporates both financial and trade flows has not yet been developed. The closest we come to such a general theory is the asset market or portfolio balance model examined in Section 15.4.

In this chapter, Section 16.2 examines how the nation's current account is affected by exchange rate changes. Section 16.3 examines the effect of exchange rate changes on domestic prices (the rate of inflation) in the country. Section 16.4 deals with the closely related topic of the stability of foreign exchange markets. Section 16.5 presents estimates of trade elasticities and explains why the current account usually responds with a time lag and only partially to a change in the nation's exchange rate. Finally, Section 16.6 describes the adjustment mechanism under the gold standard (the so-called price-specie-flow mechanism). In the appendix, we illustrate graphically the effect of a change in the exchange rate on domestic prices, derive mathematically the Marshall–Lerner condition for stability in foreign exchange markets, and demonstrate that unstable foreign exchange markets will eventually become stable for large exchange rate changes. The last section shows graphically how the gold points and international gold flows were determined under the gold standard.

2 Adjustment with Flexible Exchange Rates

In this section, we examine the method of correcting a deficit in a nation's current account or balance of payments by a depreciation or a devaluation of the nation's currency. A depreciation implies a flexible exchange rate system. A **devaluation**, on the other hand, refers to the deliberate (policy) increase in the exchange rate by the nation's monetary authorities from one fixed or pegged level to another. However, since both a depreciation and a devaluation operate on prices to bring about adjustment in the nation's current account and the balance of payments, they are both referred to as the *price adjustment mechanism* and are discussed together here. This is to be distinguished from the *income adjustment mechanism*, which relies on income changes in the nation and abroad and will be examined in the next chapter. We begin by examining the process of adjustment itself, and then show how the demand and supply schedules of foreign exchange are derived.

Balance-of-Payments Adjustments with Exchange Rate Changes

The process of correcting a deficit in a nation's balance of payments by a depreciation or devaluation of its currency is shown in Figure 16.1. In the figure, it is assumed that the United States and the European Monetary Union are the only two economies in

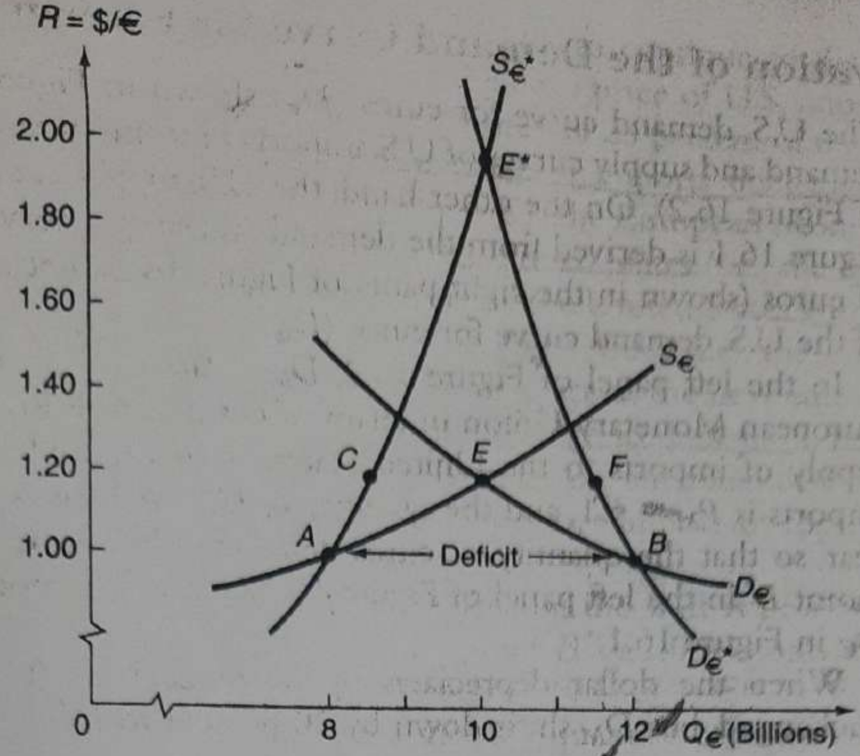

FIGURE 16.1. Balance-of-Payments Adjustments with Exchange Rate Changes. At $R = \$1/€1$, the quantity of euros demanded by the United States is €12 billion per year, while the quantity supplied is €8 billion, so that the United States has a deficit of €4 billion (*AB*) in its balance of payments. With $D_€$ and $S_€$, a 20 percent depreciation or devaluation of the dollar would completely eliminate the deficit (point *E*). With $D^*_€$ and $S^*_€$, a 100 percent depreciation or devaluation would be required to eliminate the deficit (point E^*).

the world and that there are no international capital flows, so that the U.S. demand and supply curves for euros reflect only trade in goods and services. The figure shows that at the exchange rate of $R = \$1/€1$, the quantity of euros demanded by the United States is €12 billion per year, while the quantity supplied is €8 billion. As a result, the United States has a deficit of €4 billion (*AB*) in its balance of payments.

If the U.S. demand and supply curves for euros are given by $D_€$ and $S_€$, a 20 percent devaluation or depreciation of the dollar, from $R = \$1/€1$ to $R = \$1.20/€1$, would completely eliminate the U.S. deficit. That is, at $R = \$1.20/€1$, the quantity of euros demanded and the quantity supplied would be equal at €10 billion per year (point *E* in the figure), and the U.S. balance of payments would be in equilibrium. If, however, the U.S. demand and supply curves for euros were less elastic (steeper), as indicated by $D^*_€$ and $S^*_€$, the same 20 percent devaluation would only reduce the U.S. deficit to €3 billion (*CF* in the figure), and a 100 percent devaluation or depreciation of the dollar, from $R = \$1/€1$ to $R = \$2/€1$, would be required to completely eliminate the deficit (point E^* in the figure). Such a huge devaluation or depreciation of the dollar might not be feasible (for reasons examined later).

Thus, it is very important to know how elastic the U.S. demand and supply curves for euros are. In some cases, the shape of the deficit nation's demand and supply curves for foreign exchange may be such that a devaluation or depreciation would actually increase, rather than reduce or eliminate, the deficit in its balance of payments. These crucial questions are examined next by showing how a nation's demand and supply schedules for foreign exchange are derived.

Stability of Foreign Exchange Markets

In this section, we examine the meanings of and the conditions for stability of the foreign exchange market. We have a **stable foreign exchange market** when a disturbance from the equilibrium exchange rate gives rise to automatic forces that push the exchange rate back toward the equilibrium level. We have an **unstable foreign exchange market** when a disturbance from equilibrium pushes the exchange rate further away from equilibrium.

Stable and Unstable Foreign Exchange Markets

A foreign exchange market is stable when the supply curve of foreign exchange is positively sloped or, if negatively sloped, is less elastic (steeper) than the demand curve of foreign exchange. A foreign exchange market is unstable if the supply curve is negatively sloped *and* more elastic (flatter) than the demand curve of foreign exchange. These conditions are illustrated in Figure 16.3.

The left panel of Figure 16.3 repeats $D_€$ and $S_€$ from Figure 16.1. With $D_€$ and $S_€$, the equilibrium exchange rate is $R = \$1.20/€1$, at which the quantity of euros demanded and the quantity supplied are equal at €10 billion per year (point E in the

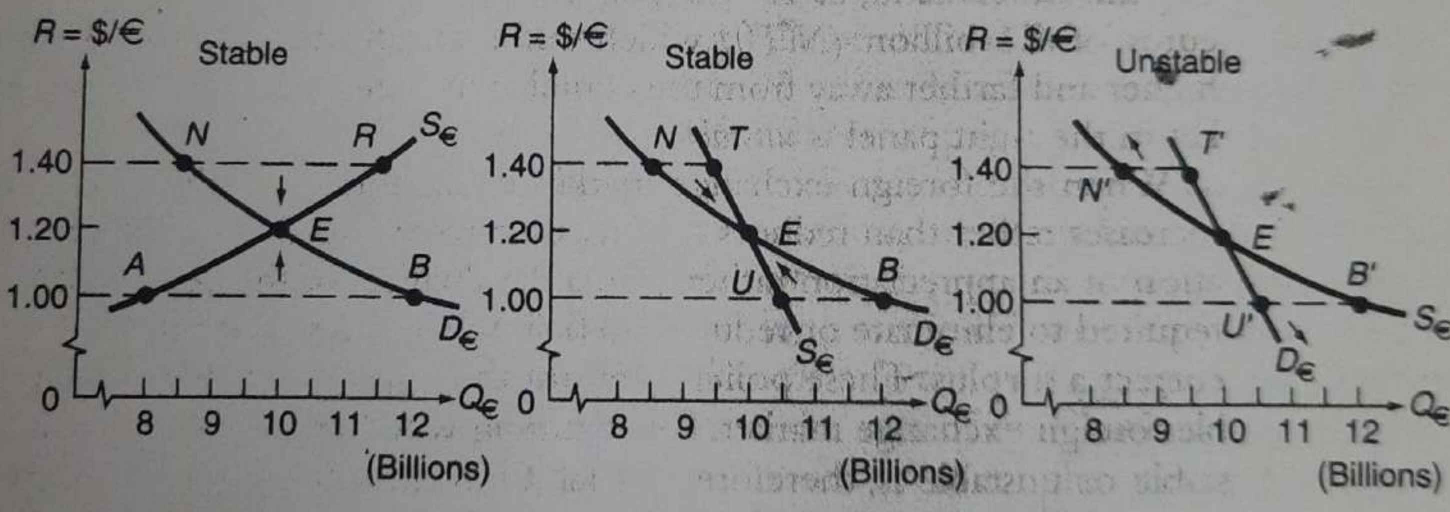

FIGURE 16.3. Stable and Unstable Foreign Exchange Markets. In all three panels, the equilibrium exchange rate is $R = \$1.20/€1$, at which €10 billion are demanded and supplied per year. If, for whatever reason, the equilibrium is disturbed and the exchange rate falls, say to $R = \$1/€1$, the excess demand for foreign exchange in the left and center panels will push the exchange rate back up toward the equilibrium rate, but the excess supply of foreign exchange in the right panel will cause the exchange rate to fall even lower. Similarly, at $R = \$1.40/€1$, the excess supply in the left and center panels will drive the exchange rate down toward $R = \$1.20/€1$, but the excess demand in the right panel will push the exchange rate even higher. Thus, the left and center panels depict stable markets, while the right panel depicts an unstable market.

left panel of Figure 16.3). If, for whatever reason, the exchange rate fell to $R = \$1/€1$, there would be an excess demand for euros (a deficit in the U.S. balance of payments) of €4 billion *(AB)*, which would automatically push the exchange rate back up toward the equilibrium rate of $R = \$1.20/€1$. On the other hand, if the exchange rate rose to $R = \$1.40/€1$, there would be an excess quantity supplied of euros (a surplus in the U.S. balance of payments) of €3 billion *(NR)*, which would automatically drive the exchange rate back down toward the equilibrium rate of $R = \$1.20/€1$. Thus, the foreign exchange market shown in the left panel of Figure 16.3 is *stable*.

The center panel of Figure 16.3 shows the same $D_€$ as in the left panel, but $S_€$ is now negatively sloped but steeper (less elastic) than $D_€$. Once again, the equilibrium exchange rate is $R = \$1.20/€1$ (point *E*). At the lower than equilibrium exchange rate $R = \$1/€1$, there is an excess demand for euros (a deficit in the U.S. balance of payments) equal to €1.5 billion *(UB)*, which automatically pushes the exchange rate back up toward the equilibrium rate of $R = \$1.20/€1$. At the higher than equilibrium exchange rate of $R = \$1.40/€1$, there is an excess supply of euros (a surplus in the U.S. balance of payments) of €1 billion (NT), which automatically pushes the exchange rate back down toward the equilibrium rate of $R = \$1.20/€1$. In this case also, the foreign exchange market is *stable*.

The right panel of Figure 16.3 looks the same as the center panel, but the labels of the demand and supply curves are reversed, so that now $S_€$ is negatively sloped and flatter (more elastic) than $D_€$. The equilibrium exchange rate is still $R = \$1.20/€1$ (point *E*). Now, however, at any exchange rate lower than equilibrium, there is an excess quantity supplied of euros, which automatically drives the exchange rate even lower and farther away from the equilibrium rate. For example, at $R = \$1/€1$, there is an excess quantity supplied of euros of €1.5 billion *(U′B′)*, which pushes the exchange rate even lower and farther away from $R = \$1.20/€1$. On the other hand, at $R = \$1.40/€1$, there is an excess quantity demanded for euros of €1 billion *(N′T′)*, which automatically pushes the exchange rate even higher and farther away from the equilibrium rate. Thus, the foreign exchange market in the right panel is *unstable*.

When the foreign exchange market is unstable, a flexible exchange rate system increases rather than reduces a balance-of-payments disequilibrium. Then a revaluation or an appreciation rather than a devaluation of the deficit nation's currency is required to eliminate or reduce a deficit, while a devaluation would be necessary to correct a surplus. These policies are just the opposite of those required under a stable foreign exchange market. Determining whether the foreign exchange market is stable or unstable is, therefore, crucial. Only after the foreign exchange market has been determined to be stable will the elasticity of $D_€$ and $S_€$ (and thus the feasibility of correcting a balance-of-payments disequilibrium with a depreciation or devaluation of the deficit nation's currency) become important.

3 The Marshall–Lerner Condition

If we knew the exact shape of the demand and supply curves of foreign exchange in the real world, it would be rather easy (as indicated above) to determine whether the foreign exchange market in a particular case was stable or unstable and, if stable, th size of the depreciation or devaluation required to correct a deficit in the balance of payments. Unfortunately, this is not the case. As a result, we can only infer whether the foreign exchange market is stable or unstable and the elasticity of the demand and supply of foreign exchange from the demand for and supply of the nation's imports and exports.

The condition that tells us whether the foreign exchange market is stable or unstable is the Marshall–Lerner condition. The general formulation of the Marshall–Lerner condition is very complex and is presented in Section A16.2 in the appendix. Here we present and discuss the simplified version that is generally used. This is valid when the supply curves of imports and exports (i.e., S_M and S_X) are both infinitely elastic, or horizontal. Then the **Marshall–Lerner condition** indicates a stable foreign exchange market if the sum of the price elasticities of the demand for imports (D_M) and the demand for exports (D_X), in absolute terms, is greater than 1. If the sum of the price elasticities of D_M and D_X is less than 1, the foreign exchange market is unstable, and if the sum of these two demand elasticities is equal to 1, a change in the exchange rate will leave the balance of payments unchanged.

For example, from the left panel of Figure 16.2 we can visualize that if D_M were vertical and S_M horizontal, a depreciation or devaluation of the dollar would leave the U.S. demand for imports and thus the quantity of euros demanded by the United States completely unchanged. By itself, this would leave the U.S. balance of payments unchanged. From the right panel of Figure 16.2, we can visualize that given a horizontal S_X that shifts down by the percentage depreciation or devaluation of the dollar, the quantity of euros supplied to the United States rises, remains unchanged, or falls, depending on whether D_X is price elastic, unitary elastic, or inelastic, respectively. Thus, the sum of the price elasticities of D_M and D_X is equal to the price elasticity of D_X (because we have here assumed D_M to have zero price elasticity), and the U.S. balance of payments improves if the elasticity of D_X is greater than 1.

If D_M is negatively sloped so that it falls or shifts down by the amount of the depreciation of the dollar, the quantity of euros demanded by the United States falls, and this, by itself, improves the U.S. balance of payments. The reduction in the quantity of euros demanded by the United States is greater the larger is the price elasticity of D_M. Now, even if the price elasticity of D_X is less than 1 so that the quantity of euros supplied falls as a result of the depreciation of the dollar, the U.S. balance of payments will still improve as long as the *reduction in the quantity of euros demanded by the United States is greater than the reduction in the quantity of euros supplied to the United States.* For this to be the case, the sum of the elasticities of D_M and D_X must be greater than 1. The greater the amount by which the sum of these two elasticities exceeds 1, the greater is the improvement in the U.S. balance of payments for a given depreciation or devaluation of the dollar.

5 Elasticities in the Real World

In this section, we examine how the price elasticity of demand for imports and exports is measured and present some real-world estimates, discuss the J-curve effect, and examine the "pass-through" of exchange rate changes to domestic prices.

Elasticity Estimates

The Marshall–Lerner condition postulates a stable foreign exchange market if the sum of the price elasticities of the demand for imports and the demand for exports exceeds 1 in absolute value. However, the sum of these two elasticities will have to be substantially greater than 1 for the nation's demand and supply curves of foreign exchange to be sufficiently elastic to make a depreciation or devaluation feasible (i.e., not excessively inflationary) as a method of correcting a deficit in the nation's balance of payments. Thus, it is very important to determine the real-world value of the price elasticity of the demand for imports and exports.

Before World War II, it was widely believed not only that the foreign exchange market was stable but that the demand for and the supply of foreign exchange were very elastic. *Marshall,* among others, advanced this view in his *Money, Credit and Commerce,* published in 1923, but offered no empirical support for his belief.

During the 1940s, a number of econometric studies were undertaken to measure price elasticities in international trade. Two representative studies were undertaken by *Chang,* one in 1945 to measure the price elasticity of the demand for imports in 21 nations for which data existed from 1924 to 1938, and the other in 1948, to measure the price elasticity of the demand for exports of 22 nations over the same period. Chang found that the sum of the demand elasticities on the average barely exceeded 1, so that while the foreign exchange market was stable, the demand and supply curves of foreign exchange were probably fairly steep and inelastic (i.e., as $D^*_€$ and $S^*_€$ rather than as $D_€$ and $S_€$ in Figure 16.1). Other studies reached similar conclusions, confirming that the sum of the elasticities of the demand for imports and the demand for exports was either below or very close to 1 in absolute value. Thus, the prewar elasticity optimism was replaced by postwar **elasticity pessimism.**

However, writing in 1950, *Orcutt* provided some convincing reasons for the view that the regression technique used to estimate elasticities led to gross underestimation of the true elasticities in international trade. In short, it was likely that Marshall had been broadly correct, while the new econometric estimates, though seemingly more precise, were in fact likely to be far off the mark.

One reason advanced by Orcutt for the belief that the early econometric studies of the 1940s grossly underestimated the price elasticity of the demand for imports and exports results from the **identification problem** in estimation. This is explained with the aid of Figure 16.4. This figure is similar to the right panel of Figure 16.2 in that it shows the effect of a depreciation or devaluation of the dollar on the U.S. export market when the foreign demand curve and the U.S. supply curve of exports are expressed in terms of the foreign currency (euros). Suppose that points E and E^* are, respectively, the equilibrium points actually observed before and after the United States devalues its currency or allows it to depreciate (with none of the curves in Figure 16.4 being observed). The downward shift from S_X to S^*_X in Figure 16.4 is due to the depreciation or devaluation of the dollar (as in the right panel of Figure 16.2). The depreciation or devaluation of the dollar does not affect the foreign demand for U.S. exports.

If no other change (such as a change in tastes for U.S. exports) occurs, then the estimated foreign demand curve of U.S. exports is inelastic, as shown by D_X in

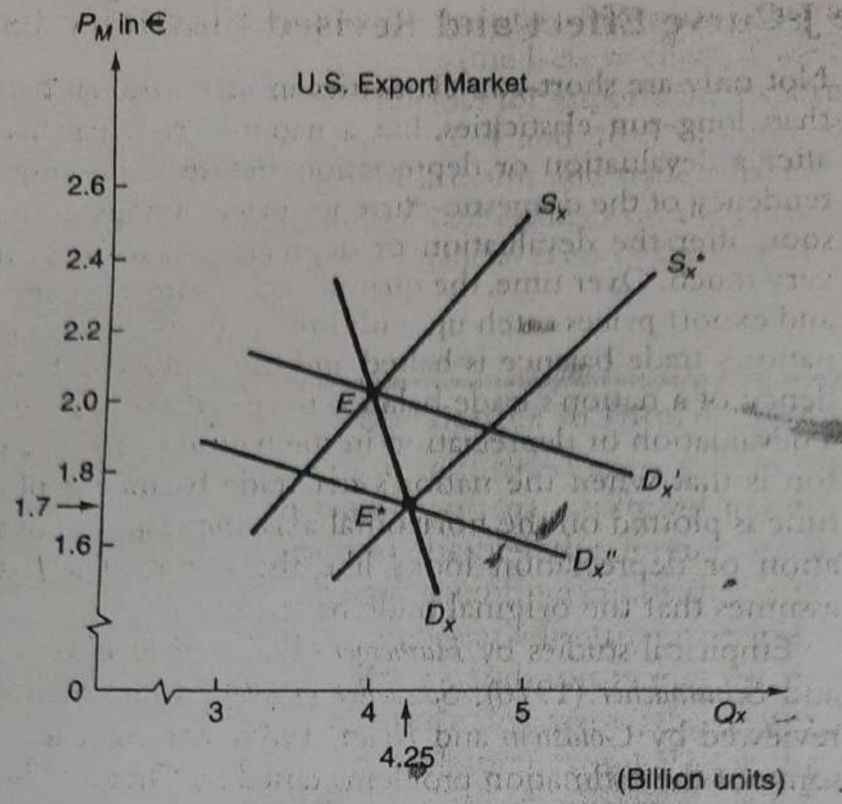

FIGURE 16.4. The Identification Problem. Observed equilibrium points E and E^* are consistent either with nonshifting inelastic demand curve D_X or with elastic demand curve D'_X shifting down to D''_X. The estimation techniques used in the 1940s ended up measuring the elasticity of (inelastic) demand curve D_X even when the relevant demand curve was elastic D'_X.

Figure 16.4. However, equilibrium points E and E^* are also consistent with elastic demand curve D'_X, which shifts down to D''_X as a result, for example, of reduced foreign tastes for U.S. exports. Regression analysis will always measure the low elasticity of demand D_X even if the true demand is elastic and given by D'_X and D''_X (i.e., regression techniques fail to identify demand curves D'_X and D''_X). Since shifts in demand due to changes in tastes or other unaccounted forces frequently occur over time, estimated elasticities are likely to greatly underestimate true elasticities.

The estimated elasticities of the 1940s also measured short-run elasticities in that they were based on quantity responses to price changes over a period of one year or less. *Junz* and *Rhomberg* (1973) have identified five possible lags in the quantity response to price changes in international trade. These are the *recognition* lag before the price change becomes evident, the *decision* lag to take advantage of the change in prices, the *delivery* lag of new orders placed as a result of price changes, the *replacement* lag to use up available inventories before new orders are placed, and finally the *production* lag to change the output mix as a result of price changes. Junz and Rhomberg estimated that it takes about three years for 50 percent of the final long-run quantity response to take place and five years for 90 percent to occur. By measuring the quantity response only during the year of the price change, the early econometric studies of the 1940s greatly underestimated long-run elasticities.

16.5B The J-Curve Effect and Revised Elasticity Estimates

Not only are short-run elasticities in international trade likely to be much smaller than long-run elasticities, but a nation's trade balance may actually worsen soon after a devaluation or depreciation, before improving later on. This is due to the tendency of the domestic-currency price of imports to rise faster than export prices soon after the devaluation or depreciation, with quantities initially not changing very much. Over time, the quantity of exports rises and the quantity of imports falls, and export prices catch up with import prices, so that the initial deterioration in the nation's trade balance is halted and then reversed. Economists have called this tendency of a nation's trade balance to first deteriorate before improving as a result of a devaluation or depreciation in the nation's currency the **J-curve effect.** The reason is that when the nation's net trade balance is plotted on the vertical axis and time is plotted on the horizontal axis, the response of the trade balance to a devaluation or depreciation looks like the curve of a *J* (see Figure 16.5). The figure assumes that the original trade balance was zero.

Empirical studies by *Harberger* (1957), *Houthakker* and *Magee* (1969), *Stern, Francis,* and *Schumacher* (1976), *Spitaeller* (1980), *Artus* and *Knight* (1984) (summarized and reviewed by *Goldstein* and *Khan,* 1985), and *Marquez* (1990) attempted to overcome some of the estimation problems raised by Orcutt. These studies generally confirmed the existence of a J-curve effect but also came up with *long-run* elasticities about twice as high as those found in empirical studies of the 1940s. The upshot of all of this is that real-world elasticities are likely to be high enough to ensure stability of the foreign exchange market in the short run and also to result in fairly elastic demand and supply schedules for foreign exchange in the long run. In the very short run (i.e., during the first six months), however, the so-called *impact elasticities* are small enough to result

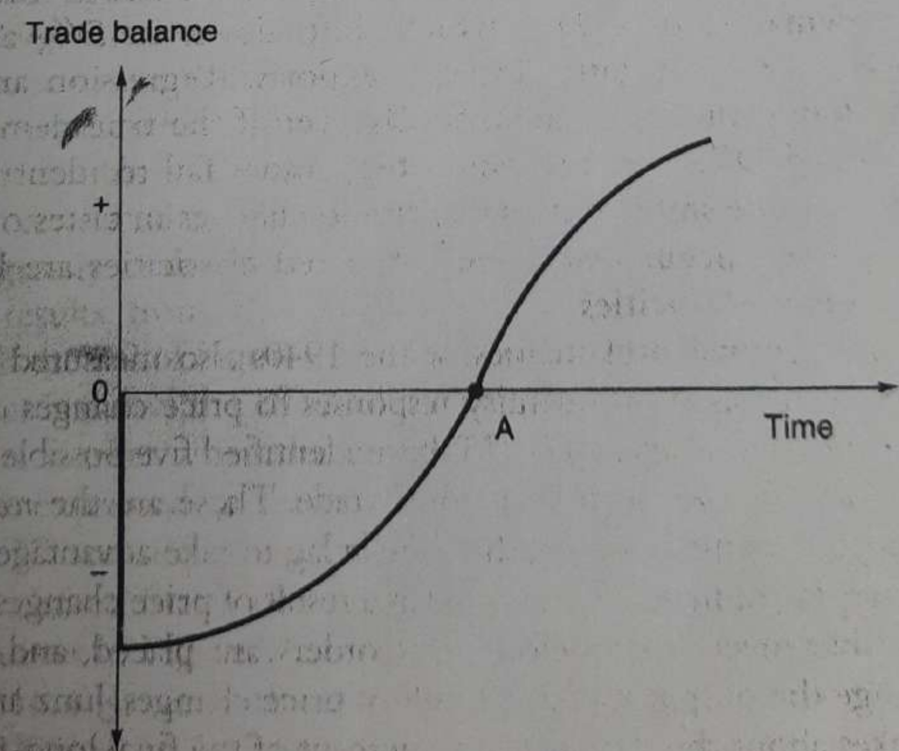

FIGURE 16.5. The J-Curve. Starting from the origin and a given trade balance, a devaluation or depreciation of the nation's currency will first result in a deterioration of the nation's trade balance before showing a net improvement (after time *A*).