11

Fiscal Policy

Beginning in late 2007, what had been a booming market for housing in the United States evolved instead into a depressed market with house prices falling and many homeowners forced into foreclosure. Financial instruments created by packaging large numbers of mortgages that were rated as safe investments suddenly dropped sharply in value, triggering losses not only for individuals but for major financial institutions such as Bear Stearns, Lehman Brothers, and the American International Group (AIG). The failure of these financial institutions in turn threatened the viability not only of other financial institutions but of the entire U.S. economy. The potential failure of many American financial institutions and the related sharp decline in economic activity in the United States in turn led to similar threats to financial institutions in the rest of the world, notably in Europe, and initiated a global recession that was the most severe recession worldwide since the Great Depression of the 1930s. Developing countries had not caused the global recession. Many had gone through a similar financial crisis in 1997–98, which had led to a major recession in several large developing economies, particularly in Asia. That crisis in turn led many of those economies to reform their financial systems in ways that protected them from the kind of financial meltdown being experienced in the high-income European countries and the United States. Developing countries, however, did feel the impact of the global recession as they saw their exports to the United States and Europe fall.

The magnitude of the recession in the United States and globally was so large that the normal mechanisms used to reverse the downturn, mainly monetary policies that lowered interest rates and increased the money supply, were not sufficient to end the

decline. Instead, first the United States, together with several major European economies, had to revert to large expenditures through the government budget to reverse the sharply falling demand for goods and services in the economy. Soon thereafter a number of developing economies also introduced their own stimulus packages in an effort to reverse falling export demand that threatened to drive their economies into recession only a decade after their last crisis. The largest of these developing country stimulus packages was China's. China initiated a program of nearly US$600 billion to be spent mainly on infrastructure, such as a national high-speed highway system, eight new north-south and east-west high-speed rail lines, and new and modernized ports and airports. The stimulus package kept Chinese gross domestic product (GDP) growth from falling despite a large drop in the demand for Chinese exports in 2009. The average rate of GDP growth in China, thanks to a large degree to these government expenditures, was 9.8 percent a year in 2009 and 2010.

When a government raises its expenditures or cuts taxes on individuals and corporations so that they can raise their expenditures, these actions are referred to as **fiscal policy**. When government efforts instead rely on the central bank to adjust the money supply and interest rates to raise or lower individual and corporate expenditures, this is referred to as the exercise of **monetary policy**.

In this chapter we focus on fiscal policy and begin by noting that the main purpose of fiscal policy is not to deal with economic crises, whether crises of inadequate demand leading to recessions as in 2009–10 or of excess demand leading to inflation. The first function of fiscal policy is to manage government expenditures used to fund whatever activities government policy makers and the people they represent think the government should support. The next closely related function of fiscal policy is to find a way of financing those government expenditures, either through the collection of taxes, issuance of government bonds, or in more desperate cases by printing money.

This chapter starts with a discussion of what expenditures are typically undertaken by governments and follows with an analysis of the kinds of taxes commonly used to pay for those expenditures. With this as background, we then return to the question of what happens when something goes wrong either in the public or the private sphere of the economy. What is it that a government can do to repair what has gone wrong?

In many respects these issues are similar in high-, middle-, and low-income economies, but there are significant differences. Developing countries often have to rely more on fiscal policy to curb inflation or recover from a recession because many of the instruments of monetary policy are not as effective or do not even exist. Underdeveloped markets in developing countries are also often a major reason why they rely on government to take responsibility for investments and other expenditures that in a more advanced economy would be handled by the private sector.

The limitations of markets in developing countries would seem to suggest that government expenditures would be a higher share of GDP than in high income countries. But, as the data in Table 11-1 indicate, the opposite is the case. Governments in

TABLE 11-1 Central Government Expenditure as a Share of GDP, 2009

COUNTRY	PERCENT OF GDP
Low-income economies	
Bangladesh	11.3
Cambodia	11.0
Ghana	17.9
Lower-middle-income economies	
Indonesia	15.8
El Salvador	21.6
Morocco	27.9
Upper-middle-income economies	
Colombia	19.4
Thailand	19.7
Mauritius	20.9
High-income economies	
South Korea	21.9
Denmark	42.5
United Kingdom	46.4

Source: World Bank, "World Development Indicators," http://databank.worldbank.org.

developing countries on average spend a smaller share of GDP than do high-income countries. A major reason is that high-income countries typically spend a larger share of GDP on social welfare measures (old-age pensions, poverty reduction, healthcare for all) largely because the people of these countries demand such expenditures for reasons of fairness and feel that at their relatively high incomes they can afford to do so. Developing country governments are more apt to focus on measures to increase investment for overall growth. There is, however, considerable variation in the share of government expenditures in GDP among countries at both ends of the income spectrum.

GOVERNMENT EXPENDITURES

At the most general level all countries rely on government for certain activities because of public goods and market failures discussed in Chapter 5. For private goods, such as rice, saris, or TV sets, the signals provided by unfettered competitive market mechanisms guide producers to satisfy consumer demand efficiently. For pure public goods, the market fails entirely. There are few examples of pure public goods; national defense and basic scientific research generally are cited as illustrations. The term *public good* refers to a good or service that exhibits two traits: **nonrival consumption** and **nonexcludability**.

In *nonrival consumption,* one person's use of a good does not reduce the benefits available to others. That being the case, no one has any incentive to offer to pay for the good: If it is available to one, it is available to all. *Nonexcludability* means that it is either impossible or prohibitively expensive to exclude anyone from the benefits once the good is available. In either case, the private market cannot provide the good: Market failure is total. For most public goods, the characteristics of nonrival consumption and nonexcludability are present but less pronounced, and the market can function only in an inefficient way. Examples include vaccination against contagious diseases, police protection, and mosquito abatement. Therefore, it is evident that the appropriate role of the public sector, to a significant degree, is a technical issue.

In practice in all countries, government expenditures designed to offset market failures of this sort account for only a portion of government expenditures even in the most laissez faire, small government-oriented societies. In democratic societies the people through their elected representatives generally want government to do more, such as provide pensions for the elderly and healthcare for the sick regardless of income. In authoritarian societies, the leadership often does much the same because failure to do so can threaten the leaderships' ability to govern.

CATEGORIES OF GOVERNMENT EXPENDITURES

The principal distinction between types of government expenditure is between **capital** or **public investment** expenditures also called **development expenditures**, on the one hand, and **recurrent expenditures** on the other. Development or public investment expenditures are those that, like investment in general, generate a stream of income or services into the future. Government typically takes on the task of building a network of highways, and some governments, particularly in developing countries, take responsibility for providing electric power or exploitation of natural resource wealth (for example, petroleum). Education and health expenditures can also be thought of as a form of investment undertaken by government although in practice government budgets often treat them as recurrent consumption expenditures. The quality of government development investments in these areas should be judged by the same criteria that are used to judge private investments, as was discussed in Chapter 10.

Recurrent expenditures are those expenditures that a government must make year after year. They include the salaries of government employees, the purchase of office equipment, and annual payments to families living in poverty and for pensions for retirees. Most of these categories of recurrent expenditures are self-explanatory, but one category is of particular importance in developing countries largely because it is often ignored. Capital investments to build a highway are often given a high priority but the expenditures needed to maintain the highway once built are often neglected, and the highway is soon filled with potholes. This results not only from the neglect of maintenance by developing country governments but also because foreign aid agencies will frequently support a capital project but will seldom support the

recurrent costs needed to keep the capital project functioning. The problem, however, is not unique to developing countries. The United States has neglected bridge and road maintenance in many parts of the country for decades and, at the time of the 2008–09 economic recession, faced the need for large maintenance and replacement expenditures, some of which were built into the stimulus program designed to get the country out of the recession.

Other common recurrent expenditures in developing (and high-income) countries are interest payments on government debt, consumer subsidies, and subsidies to loss-making state-owned enterprises. Government debt crises are the subject of Chapter 13, and here we note only that the problem with government deficits in developing countries on average is not that they are typically very large, although that is the case in some countries. The problem with chronic large deficits is that interest payments rise over time taking up an increasing share of the government recurrent cost budget either crowding out other worthwhile expenditures or leading to even larger government deficits. If the deficits are financed by borrowing abroad, foreign debt, typically in convertible currencies such as the dollar or the euro, accumulates and eats into the ability of the country to import and may eventually lead to a debt crisis. If the government borrows money domestically to cover the deficit, the interest cost in the budget will also rise. In developing countries, however, borrowing domestically often means borrowing directly from the central bank, and that is essentially the same as printing money and increasing the money supply. The reason developing countries frequently have to borrow from the central bank is that capital markets are underdeveloped and it often is not possible to sell government bonds to the general public or even to private companies. The net result is that deficits lead to money creation, and excessive money creation leads to inflation, which usually proves detrimental to the productive investment so central to economic growth.

Consumer subsidies are also not unique to developing countries. The political logic behind consumer subsidies is that certain commodities are considered essential, and government subsidizes these commodities so as to lower their price to the consumer and thus make them available to all regardless of income. In high-income countries the government typically has the ability to find out the true income of families and then target the subsidy so that it applies only to those below a predetermined poverty line—food stamps in the United States is a good example. In developing countries, however, government usually lacks the administrative capacity to discover the true income of families, although some middle-income nations are gaining this ability (see the discussion of Mexico's Oportunidades program in Chapters 6 and 8). When targeting is difficult, one option is to subsidize all sales of the commodity even though most of the subsidy will go to families that are not poor and could afford to pay the market price. General subsidies of this sort can be very expensive. When a country runs into debt and balance of payments problems and has to call for help to the International Monetary Fund (IMF), one of the first measures the IMF sets as a condition for a loan is elimination of these subsidies. Subsidies, however, are popular

with the public and eliminating them typically leads to demonstrations against the government and even to major riots that threaten the existence of the government.

Even in China, where administrative capacities at the local level are quite high by developing country standards, targeting subsidies to poor families is not possible; subsidies in the 1990s and in the first decade of the twenty-first century have instead gone to designated poor counties and are available to all those in the county, including many families that are not poor. An alternative approach tried in the 1980s in Sri Lanka was to ask families to voluntarily state whether they fell below the poverty line or not. The result was that well over half the population of the country said that they fell below the poverty line even though objective data made it clear that the share below the poverty line was only a fraction of that self-reported figure. In some cases there is not even a pretense that the subsidy is designed to alleviate poverty. Countries with large oil reserves, such as Venezuela and Indonesia, have subsidized refined fuel costs to households even though the heaviest users of such fuel are the well off who can afford automobiles and motor bikes.

Subsidies to cover the losses of state owned enterprises (SOEs) grew rapidly in a great many developing countries when they first achieved independence. This was particularly so in Africa. In Tanzania alone, the number of state enterprises increased 10-fold from 1965 to 1985. By 1980, SOEs were common, and often dominant, in manufacturing, construction, banking services, natural resource industries, and agriculture. Although SOEs typically were small-scale undertakings in most developing countries before 1950, many by the beginning of the twenty-first century were among the largest firms in their countries, and some, notably in China, are among the largest enterprises in their fields anywhere in the world.

Deficits in state enterprises averaged 4 percent of the GDP across all developing countries in the mid-1970s. The problem worsened in the early 1980s, particularly in such countries as Brazil, Costa Rica, the Dominican Republic, Ecuador, Egypt, the Philippines, Turkey, and Venezuela, where SOE deficits reached between 3 and 12 percent of the GDP. In all these countries, the rest of the public sector would have generated a fiscal surplus, excluding the net transfers to the state enterprises.[1] By the late 1980s and through the first decade of the twenty-first century, the drain of SOE subsidies on government budgets has led to efforts by many developing countries to eliminate these subsidies by privatizing their SOEs. International aid agencies frequently have made SOE privatization a condition for further assistance. Privatization, however, has been a slow process, because it has been actively resisted by SOE managers and workers, not to mention politicians who use SOE employment as a patronage machine for their supporters. In countries where a large part of the private sector is controlled by minorities, as is the case with the Chinese minority in Southeast Asia or the South Asian minority in East Africa, the majority population often fears that privatization is simply a way of turning even more of modern business over to the minority.

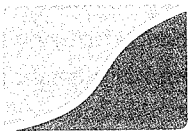
[1]World Bank, *World Development Report 1988* (New York: Oxford University Press, 1988), p. 171.

Finally transfers from central to subnational governments conventionally are treated as recurrent expenditures even though subnational governments (provinces, departments, counties, municipalities) may use the proceeds for capital formation, such as construction of schools and hospitals. Some countries, such as Indonesia, classify some transfers as capital spending and others as consumption, so international comparisons are difficult. In a few cases, subnational governments with small transfers have access to rich sources of revenue. This is true for Bolivia, where oil-producing provinces receive large oil royalties, and for Malaysia, where the state governments of Sabah and Sarawak have earned substantial tax revenues from the export of tropical timber. In most unitary states, such as Chile, government affairs at virtually all levels are run from the capital, and subnational units of governments have few responsibilities to go with their limited sources of local revenue.

There is often a compelling argument for moving decisions about expenditures to lower levels of government. Lower levels of government are typically closer to the people being served by both investment and recurrent expenditures and thus in principle are better able to design expenditures to best meet local needs. Problems arise, however, when local governments lacking adequate tax revenues under their control finance these project and recurrent expenditures through access to bank loans. The problem can be particularly acute where the local banks are owned or controlled by local politicians. The result can be projects driven more by the political needs of local politicians than the needs of the local population and an increasing often unsustainable local government debt. In the late 1990s local Argentinian governments borrowed heavily to cover local expenditures from both domestic and international sources running up debts neither they nor ultimately the central government had the capacity to repay. Argentina defaulted on this debt in 2002 and negotiated a large-scale debt restructuring in 2005–06. Similar although less severe problems existed in both Brazil in the 1990s, but the government took steps to strictly limit the ability of local governments to invest in this way (see Box 11–1).

In many places, including many federal countries, transfer of funds from the central government to subnational units is essential because the center has monopolized the most productive sources of tax revenue. In most cases, national governments impose income taxes because subnational governments lack the resources and skill required to administer such a complex tax and could not do so effectively anyway because the income tax base easily can migrate within a country. Similarly, in the past the major source of tax revenue for some developing countries, import duties, remained a central government resource because most countries have only a few serviceable ports. Very few developing countries today, however, receive a substantial portion of tax revenue from import duties, in part because import duties have been cut sharply over the years in the various international trade negotiations that have lowered tariffs for all countries.

We have just seen that there are few easy ways to reduce government recurrent expenditure to achieve higher public savings. The scope for doing so often was



BOX 11-1 REINING IN FISCAL DECENTRALIZATION IN BRAZIL AND CHINA

BRAZIL

Brazil has a federal system of government, and the states have a substantial role in capital and recurrent expenditures. The states in the past also had banks owned by their governments, and these banks contributed in a major way to large-scale profligate state spending over many decades. There were also problems involving borrowing by the central government from the central bank. In 1996, however, the central government took steps to increase the independence of the central bank and began a program to separate the banks from control by state governments. The first phase of this effort involved an effort to privatize state-owned banks in what was called the Program to Reduce State Involvement with Banking Activities. The central government helped with the privatization process, and many of the state banks were privatized and others turned into development agencies. Then in 2000 the government passed the Law of Fiscal Responsibility that prohibited government borrowing from banks it owned or from the central bank. All new borrowing by either the central government or state governments first had to be approved by the central bank and by the senate. In 2002 governments were further prohibited from borrowing within 180 days of the end of a government's term in office. This restrained the large expenditures during the final days of an administration often designed to pay off supporters knowing that the incumbent governor or president would not have to find ways of repaying the loans involved. There were also steps taken to reduce the banks' incentive for making such loans. These efforts have played an important role both in reducing wasteful government expenditures and in reining in the chronic inflation that had plagued Brazil for decades.

CHINA

China's provincial governments are not formally independent of the central government but they have had a substantial degree of autonomy in many economic areas. Formally the largest state-owned commercial banks are national not provincial banks, but in practice local governments and the local branches of the Communist Party have had a major say in appointments to head the local branches of these national banks. The result in the 1990s was that local governments borrowed heavily from the banking system to finance local projects and to support investments of provincial- and city-owned state industrial enterprises, which led to large-scale lending partly to prop up failing state-owned enterprises that fueled bursts of inflation. Inflation was stopped but only when the central government and the central bank imposed lending quotas on state-owned banks and their

branches. The government, however, did take steps to change the territory covered by the branches of the large commercial banks so that those territories did not coincide with that of the provincial and local governments, making it more difficult for local politicians to influence the selection of local bank managers.

During the large-scale investment surge that began with the 1997–98 financial crisis in Asia and elsewhere, the Chinese government began a massive state-led investment program that continued throughout the first decade of the twenty-first century, and local governments were pressured to carry out much of this investment program. Many of these investments were well thought out, but local governments lacked the tax base to pay for them and so borrowed heavily from the state-owned commercial banks and local state-owned banks. The result by 2011 was a local debt accumulation of over RMB 10 trillion, which was 25 percent of the gross domestic product (GDP). Given the weakness of the tax base of the local governments, there was a danger that the state-owned commercial banks would again be burdened by a large number of nonperforming loans requiring refinancing by the central government.[a]

[a]Anwar Shah, "Fiscal Decentralization and Macroeconomic Management," *International Tax and Public Finance* 13 (2006), 437–62.

greatest in the areas of military spending and subsidies to cover deficits of state-owned enterprises. Subsidies to state-owned enterprises, however, have been cut over the years either through privatization or by taking steps to make state enterprises competitive and profitable. Military expenditures in developing countries in the twenty-first century are still 2 percent of the GDP on average, with considerable variation across nations, which leaves some room for cutting but not a great deal. The main argument for cutting these latter expenditures is that military forces in developing countries are sometimes a greater danger to their own people than to potential enemies of the nation.

GOVERNMENT REVENUE AND TAXES

Government expenditures have to be funded or paid for. Most government expenditures are funded by revenue from taxes, including **direct taxes** (taxes on corporate or individual income and other payroll taxes such as Medicare and Social Security payments in the United States) or **indirect taxes** (sales taxes, import duties, and value-added taxes that are typically levied as a percentage of the value of a good being sold or imported), but there are other sources of funding. Mineral-rich countries often depend on royalties paid to the government by large petroleum and mineral mining companies, bonds are sometimes issued to cover the costs of a capital investment,

and fees are charged for many government-provided services. Revenue can also come through aid grants from other countries or in the form of loans from international lending agencies such as the World Bank.

As incomes per capita rise, the amount of taxes collected, expressed as a share of GDP, also tends to rise, although there is considerable variation among countries as the data in Table 11–2 indicate. The taxes reported in the table do not include what are often referred to as "social contributions"—that is, taxes paid for old-age support such as the Social Security and Medicare systems in the United States. Because high-income nations have more extensive transfer programs of this kind, they also collect more in taxes to finance them. If these taxes were included in the data, the share of taxes out of GDP would rise with country income level in an even more pronounced fashion.

The relationship between the kinds of taxes collected is also associated with per capita income. For countries with low incomes and in the early stages of development, there is typically much greater reliance on international trade taxes. For the three low-income economies in Table 11–2, at least one quarter of tax revenues is obtained from taxes on international trade. As incomes rise, reliance on taxes on international trade falls off significantly. Taxes on income, personal and corporate,

TABLE 11–2 Composition of Tax Systems by Type of Tax,* 2009

		COMPOSITION OF TAXES (PERCENT OF GDP)			
COUNTRY/ECONOMY	TOTAL TAXES	DOMESTIC INCOME TAXES	TAXES ON GOODS AND SERVICES	TAXES ON INTERNATIONAL TRADE	OTHER TAXES
Low-income economies					
Bangladesh	8.6	2.2	3.3	2.8	0.4
Cambodia	9.6	1.7	5.4	2.5	0.0
Ghana	12.5	4.2	5.4	2.9	0.0
Lower-middle-income economies					
Indonesia	11.4	5.7	4.8	0.3	0.6
El Salvador	12.5	4.7	6.8	0.9	0.0
Morocco	23.8	9.5	10.5	2.0	1.8
Upper-middle-income economies					
Colombia	11.8	4.5	5.5	0.9	0.8
Thailand	15.2	7.1	7.2	0.9	0.1
Mauritius	18.6	5.4	11.0	0.5	1.7
High-income economies					
South Korea	15.5	6.6	6.0	0.9	2.1
United Kingdom	26.0	13.3	10.1	0.0	2.5
Denmark	34.6	18.0	14.7	0.0	2.0

*Taxes refer to the central government only and are exclusive of social contributions—that is, to programs of old age support and so on.

Source: World Bank, "World Development Indicators," http://databank.worldbank.org.

BOX 11-2 TAX RATES AND SMUGGLING: COLOMBIA

In the classic case from Colombia, the import duty rate on cigarettes was over 100 percent, and it was virtually impossible to purchase duty-paid cigarettes. At such high rates, import duty collections on cigarettes were nil and the market was flooded with smuggled foreign brands. In 1969, the duty rate was reduced to 30 percent. Cigarette smuggling on the poorly policed Caribbean coast of that country continued, but duty-paid packages began to appear in the mountainous interior, and duty collections soared. Smuggling profits possible under a 30 percent duty were no longer high enough to compensate smugglers for the risks of arrest.

provide about half of the tax revenue of Denmark, Korea and the UK; far less in Bangladesh and Cambodia.

In the sections that follow, we review the different kinds of taxes and explain why some taxes are favored over others at different stages of development.

TAXES ON INTERNATIONAL TRADE

Although reliance on the taxes of foreign trade has fallen sharply in recent decades, particularly in middle-income countries, tax revenue structures in developing countries historically have depended heavily on import duties. Many low-income countries remain markedly dependent on import duties. This is particularly so for countries like the Gambia, Lesotho, Côte D'Ivoire, Togo, and Bangladesh, where as much as one quarter to one third of the total government revenue comes from taxes on imports.

For most countries, attempts to raise revenues through higher duties are unfeasible and undesirable on economic grounds. Higher import duties intensify the incentive for smuggling or evading tariffs. Various studies have shown that, for countries with already high duty rates, the incentive to smuggle increases disproportionately with further increases. Therefore, a 10 percent rise in duty rates can result in an increase in smuggling activity by more than 10 percent (Box 11-2). In mountainous countries, such as Afghanistan and Bolivia, or archipelago countries, such as Indonesia and the Philippines, borders are especially porous to smuggled imports. Singapore during the first four decades after independence did not even report its trade with Indonesia largely because it did not want Indonesia to know how much trade between the two countries went through informal (smuggling) rather than official channels.[2]

[2]Bill Guerin, "Indonesia-Singapore Gap More Than Numbers," *Asia Times,* June 27, 2003. Available online at www.atimes.com/atimes/Southeast_Asia/EF27Ae03.html, accessed February 29, 2012.

Export taxes are constitutionally prohibited in the United States and extremely rare in other industrial countries. But export taxes are not uncommon in developing countries, particularly in tropical Africa and Southeast Asia. Taxes ordinarily are imposed on exports of raw materials and have been applied on timber (Ivory Coast and Liberia), tin (Malaysia), jute (Pakistan), and diamonds (Botswana) as well as on foodstuffs, such as coffee (Colombia), peanuts (the Gambia and Senegal), cocoa (Ghana), and tea (Sri Lanka).

Over the past few decades, export taxes have declined in importance as various economic organizations, notably the World Trade Organization (WTO), have put pressure on new members and some old ones to eliminate such taxes or at least not use them to promote particular industries. Some economic groupings such as the European Union and Mercosur ban export taxes. Export taxes often are imposed in the belief that they are paid by foreign consumers. That is, the taxes themselves are thought to be exported to consumers abroad, along with the materials. But the conditions necessary for exporting taxes on exports to foreign consumers rarely are present.

Export taxes also are employed to promote nonrevenue goals, including increased processing of raw materials within natural resource–exporting developing countries. This is done by imposing high rates of export tax on unprocessed exports (cocoa beans or logs) and lower or no rates of tax on processed items fabricated from raw materials (chocolate and plywood). In principle, this use of export taxes should increase the local value added on natural resource exports and thereby generate greater employment and capital income for the local economy. Unfortunately, in many cases, the result has been that a government gives up more in export tax revenues than its country gains in additional local value added, particularly when processed raw materials are exported tax free. One study documents several instances in Southeast Asia and Africa in which the additional value added gained in heavily protected local processing of logs into plywood typically was less than half the amount of export tax revenue that would have been collected had timber been exported in the form of logs.[3]

SALES AND EXCISE TAXES

A much more promising source of government revenue is indirect taxes on domestic transactions, such as sales and excise taxes. **Sales taxes**, including **value-added taxes**, are broad-based consumption taxes imposed on all products except those specifically exempted, usually food, farm inputs, and medicine. **Excise taxes** also are taxes on consumption, but these levies are imposed only on specifically enumerated items, typically tobacco, alcoholic beverages, gambling, and motor fuel. On average, developing countries depend on domestic commodity taxes for a substantial share,

[3]Robert Repetto and Malcolm Gillis, eds., *Public Policies and the Misuse of Forest Resources* (New York: Cambridge University Press, 1988), chap. 10.

ranging from a bit over a quarter (the poorest countries plus the United States) to over half of total revenues (Table 11–2).

Virtually every developing country imposes some form of sales tax. In most, the tax is not applied to retail sales because of the burdensome administrative requirements of collecting taxes from thousands of small retailers. In the past, as in Chile before 1970 and in some Indian states, the tax was imposed as a gross turnover tax collected at all levels of production and distribution, with harmful implications for efficiency, income distribution, and virtually every objective of tax policy. In developing countries, administrative problems are more tractable when the sales tax is confined to the manufacturing level: A much smaller number of firms is involved, and the output of manufacturers is far more homogeneous than sales of retailers or wholesalers. For these reasons many low-income countries have used either the single-stage or the value-added form of manufacturers' tax, usually exempting very small producers. This kind of sales tax, however, involves more economic distortions than either a wholesale or a retail tax, and for that reason as well as for revenue motives, more and more middle-income countries have turned to taxes at the retail level.[4]

Increasingly, these taxes have taken the form of one or another variant of the value-added tax (VAT), widely seen as the most effective method of taxing consumption yet developed. More than 60 countries, including all members of the European Union, have adapted the VAT since it was first adopted in its comprehensive retail form in Brazil in 1967.[5] More than 40 developing countries had adopted the tax by 1994, either to replace older, outdated forms of sales taxes or to allow them to reduce reliance on harder-to-collect income and property taxes. The typical VAT tax rate varies from around 10 percent in some of the poorest countries to 15 to 20 percent in most of the others.

Excise taxes might appear to represent an ideal source of additional tax revenue. These typically are imposed on sumptuary items having relatively inelastic demand. When the price elasticity of demand for such products is very low (as for tobacco products) or relatively low (as for alcoholic beverages), an increase in excise tax rates will induce little reduction in consumption of the taxed good. If price elasticity is as low as –0.2, not uncommon for cigarettes in part because smoking tobacco has been shown to be addictive, then an additional 10 percent excise tax on this product would yield an 8 percent increase in tax revenues. Moreover, it is a hallowed theorem in optimal tax theory that taxes levied on items with inelastic demand and supply involve the smallest losses in economic efficiency, or what is the same thing, the least excess burden. **Excess burden** is defined as a loss in private welfare above

[4]For a full discussion of the distortions involved in different forms of sales tax, see John Due and Raymond Mikesell, *Sales Taxation: State and Local Structure and Administration* (Baltimore: Johns Hopkins Press, 1983).

[5]Carl Shoup, "Choosing among Types of VAT," in Malcolm Gillis, Carl Shoup, and Gerry Sicat, eds., *Value-Added Taxation in Developing Countries* (Washington, DC: World Bank, 1990).

the amount of government revenue collected from a tax.[6] Further, many agree, with much justification, that consumption of both tobacco and alcohol should be discouraged on health grounds.

PERSONAL AND CORPORATE INCOME TAXES

Harried ministers of finance, perceiving slack in personal and corporate income taxes, often resort to rate increases in these taxes, with no change in the tax base. The results usually are disappointing, particularly for the personal income tax. Even in middle-income countries, only a small proportion of the population is covered by the personal income tax; in contrast, the vast majority of the adult population in the United States files income tax returns each year (although just under half of all U.S. households were required to pay federal income taxes as of 2011).[7] Therefore, few developing countries can rely heavily on the personal income tax for revenues. Whereas the personal income tax accounts for over 40 percent of all federal taxation in the United States (not including Social Security and Medicare payments), rarely does the personal income tax account for close to that much of total central government revenues in developing countries. In low- and middle-income nations, when personal income taxes are applied, they are paid largely by urban elites. Not only are these groups usually the most vocal politically but they have developed such a variety of devices for tax evasion and avoidance that rate increases stand little chance of raising additional revenues. (The same is often true in high-income economies.)

Rate increases for corporate taxes are usually not productive either. In only 16 of 82 developing countries did the corporation income tax account for more than 20 percent of total taxes in the 1980s. In 2006–08, corporate taxes and personal income taxes taken together accounted for more than 30 percent of all revenue in less than half of a sample of more than 60 developing countries. And, in many of those countries, corporate tax collections often originated with foreign natural-resource firms.

NEW SOURCES OF TAX REVENUES

If higher rates on existing taxes are unlikely to raise revenues much, a second option for increasing public savings is to tap entirely new sources of tax revenues. In many developing countries, whether by accident, design, or simply inertia, many sources of tax revenue may have been overlooked entirely. Many countries have not collected taxes on motor vehicle registrations; some have not used urban property taxes as a significant source of revenue; many have not applied corporation income taxes to the income of state-owned enterprises. Kenya, for example, did not seriously tax

[6]For a further discussion of optimal taxation and excess burden, see Figure 11-1 and Joseph Stiglitz, *Economics of the Public Sector* (New York: Norton, 1988).

[7]Rachel Johnson, Jim Nunns, Jeff Rohaly, Eric Toder, and Roberton Williams. "Why Some Tax Units Pay No Income Tax," Tax Policy Center, Urban Institute, July 2011.

farmland, and a few countries, such as Indonesia before its 1984 tax reform, did not collect personal income taxes on the salaries of civil servants. China began to experiment with the use of urban property taxes only in the first decade of this century.

CHANGES IN TAX ADMINISTRATION

A far more significant option for increasing tax revenues is to implement changes in tax administration that permit more taxes to be collected from existing tax sources, even at unchanged tax rates. The potential for increased revenues from such action is very large and seldom realized in virtually all developing countries. Shortages of well-trained tax administrators, excessively complex tax laws, light penalties for tax evasion, corruption, and outdated techniques of tax administration combine to make tax evasion one of the most intractable problems of economic policy in developing countries.

The case studies given in Box 11–3 discuss the magnitude of the problem in India and Bolivia, which is typical of many, perhaps most, other developing countries. These examples suggest that efforts to collect a greater share of the taxes due under the current law can increase revenues substantially. But the kinds of administrative reforms required are difficult to implement and especially difficult to sustain. Even Bolivia's successful reform raised its tax ratio to only 7 percent of GDP. Korea at the beginning of its rapid development period, a country noted for its efficient and determined administration, failed in the 1960s to reach its goal of increasing collections by 40 percent through more effective enforcement, although it was able to reduce underreporting of nonagricultural personal income from 75 percent to slightly under 50 percent. Administrative reform can help and is important at any stage of development; however, better tax administration tends to improve with economic development.

FUNDAMENTAL TAX REFORM

The final policy option available for increasing tax revenues is the most difficult to implement but the most effective when it can be done. Fundamental tax reform requires junking old tax systems and replacing them with completely new tax laws and regulations. Implementing tax reform engenders enormous technical and informational, not to mention political, difficulties in all countries. In general, governments resist genuine efforts to reform the tax structure until a fiscal crisis in the form of massive budgetary deficits threatens. Even during a fiscal crisis, it is difficult to mobilize a political consensus to allow unpopular tax measures to pass. Tax policies that protect favored groups and distort the allocation of resources did not just happen; more likely, they were enacted at the behest of someone, ordinarily the privileged and the powerful.

Probably, more has been said, to less effect, about tax reform than almost any topic in economic policy. This is no less true for the United States than for the scores of developing countries where major tax reform efforts have been mounted. That the process is painful and slow is evident from the experience of several countries: In the

BOX 11-3 TAX ADMINISTRATION IN INDIA AND BOLIVIA IN THE 1980S

INDIA

India during the 1980s provides one of the more egregious examples of poor tax administration. During the fiscal year 1981, for example, it was found that taxes were avoided on at least 40 and possibly 60 percent of potentially taxable income. Almost 70 percent of the taxpayers covered by a survey openly admitted bribing tax officials, and three quarters of the tax auditors admitted accepting bribes to reduce tax payments. The cost of a bribe was commonly known to be about 20 percent of the taxes avoided.

BOLIVIA

In Bolivia during the same period, the tax system was chaotic. More than 400 separate taxes were levied by the national, regional, and city governments. Taxpayer records were out of date, and tax collections were recorded more than a year after they had been paid. Of 120,000 registered taxpayers, one third paid no taxes at all, while 20,000 taxpayers who did pay were not on the register. The administration had all but ceased its data processing. Administration had gotten so bad that the government was helpless to collect taxes during the hyperinflation of the mid-1980s. Tax collections fell to only 1 percent of the GDP. Faced with this chaos, the Bolivian government reformed its tax administration, simplifying it drastically. This, together with price stabilization, enabled the government to collect revenues equivalent to over 7 percent of the GDP by 1990, a low ratio by world standards but a vast improvement for Bolivia.

Sources: Omkar Goswami, Amal Sanyal, and Ira N. Gang, "Taxes, Corruption and Bribes: A Model of Indian Public Finance" (201–13), in Michael Roemer and Christine Jones, eds., *Markets in Developing Countries: Parallel, Fragmented and Black* (San Francisco: ICS Press, 1991); Carlos A. Silvani and Alberto H. J. Radano, "Tax Administration Reform in Bolivia and Uruguay" (19–59) in Richard M. Bird and Milka Casanegra de Jantscher, eds., *Improving Tax Administration in Developing Countries* (Washington, DC: International Monetary Fund, 1992).

United States, the time lag between the birth of tax innovations (tax credit for child-care expenses, inflation proofing of the tax system) and their implementation usually is at least 15 years. If anything, the lag may be slightly shorter in developing countries.

In spite of the difficulties involved, some countries were able to carry through fundamental reforms in tax structure and administration before 1980. The classic example is Japan in the 1880s, when that society began its transformation to a modern industrial power. Korea implemented a major tax reform program in the early 1960s, as did Colombia and Chile in the 1970s and Indonesia in the 1980s (see Box 11-4).

BOX 11-4 INDONESIAN TAX REFORM

In the 1970s Indonesia was a major oil exporter and received more than half of its government revenue from royalties paid by the large oil companies. The increase in petroleum prices caused by the Organization of Petroleum Exporting Countries (OPEC) in 1973 and 1979 increased these royalties to an unprecedented share of the government budget. The Indonesian minister of finance at the time, however, knew that what goes up (in this case petroleum prices) can also come down and he wanted to put Indonesia's government finances on a more stable long term basis. Non-oil sources of tax revenue, however, were based on laws that in many cases went back to the Dutch colonial period and produced limited revenue for the government and did so inefficiently.

In 1981 the ministry of finance began a major effort to completely reform and modernize the Indonesian tax system. The goal of this effort was to create a broader base for the generation of tax revenue, increase the equity of the fiscal system, and generally bring the Indonesian system in line with international best practice. The process involved the complete rewriting of the tax laws, but it also involved fundamental changes in the way taxes were administrated and enforced together with the creation of a computerized tax information system among other major changes.

A team of 25 international experts, including economists, lawyers, accountants, and data processing experts, was put together to assist 10 Indonesian government officials in this endeavor. The role of the team was to help determine what aspects of best tax practice would work best in the Indonesian context. The experts then would present their findings for a full discussion of whether the expert recommendations would in fact work in the Indonesian context in meetings attended by the minister of finance and all of the ministry officials with relevant responsibilities in this area. When this process was completed and necessary adjustments to the recommendations made, the proposal was put first to the president of Indonesia and then to parliament, which passed the legislation in late 1983. Although Indonesia had an authoritarian government at that time, powerful rent-seeking individuals both inside and outside of the government could have had a major influence on the shaping of tax reform as happens in so many countries, but in this case the recommendations that were passed largely reflected international best practice, not the special interests of rent seekers.

The key reforms were the introduction of a low-rate simplified value-added tax together with major changes in the personal and corporate income tax codes. Nonpetroleum government revenues as a result of these reforms rose from

6 percent of Indonesia's GDP just before the passage of the reforms to 10 percent of GDP by the early 1990s. This increase in revenue had much to do with the fiscal stability that Indonesia enjoyed when petroleum prices and government royalties from petroleum fell in 1986.

Source: Michael Roemer and Joseph Stern, "Indonesia: Economic Policy Reform," (63–64), in Dwight H. Perkins, Richard Pagett, Michael Roemer, Donald Snodgrass, and Joseph Stern, eds., *Assisting Development in a Changing World: The Harvard Institute for International Development, 1980–1995* (Cambridge: Distributed by Harvard University Press, 1997).

After 1980, the pace of tax reform quickened notably, in both developing and industrial countries, with many similarities in the various reform programs. Throughout much of the postwar period, tax systems commonly were fine-tuned to achieve a wide variety of nonrevenue objectives. In particular, governments in developed and developing nations alike commonly sought substantial income redistribution through the use of steeply progressive tax rates. Also, complex and largely impossible to administer systems of tax incentives were widely used in attempts to redirect resources to high-priority economic sectors and promote foreign investment, regional development, and even stock exchanges.

While fine-tuning tax systems someday may yield the desired results, it requires, at a minimum, strong machinery for tax administration and traditions of taxpayer compliance. Within developing countries, at least, there has been growing recognition that such conditions seldom prevail. Governments increasingly have turned away from reliance on steeply progressive tax rates and complicated, costly tax incentive programs.

The 1980s through the beginning of the twenty-first century saw a worldwide movement toward an entirely different type of tax system, with a shift toward vastly simplified taxes imposed at much flatter rates and with much broader bases[8] and increasingly greater reliance on consumption rather than income taxation. Tax reform programs in Bolivia, Chile, Colombia, India, Indonesia, Jamaica, and Malawi exemplify most of these trends.[9]

Two aspects of this worldwide movement in tax reform are especially salient. First, the top marginal income tax rates of 60 to 70 percent were not uncommon from 1945 to 1979. But, since 1984, country after country has slashed the top marginal rate, often substantially. Table 11-3 shows this phenomenon in a selection of developing countries. During the same period, many industrial nations, ranging from Australia and Austria through the United Kingdom and the United States also cut the top rate sharply.

[8]Joseph A. Pechman, ed., *World Tax Reform: A Progress Report* (Washington, DC: Brookings Institute, 1988), p. 13.

[9]A number of these cases are discussed in Malcolm Gillis, ed., *Tax Reform in Developing Countries* (Durham, NC: Duke University Press, 1989).

TABLE 11-3 Countries Reducing Highest Rates of Income Tax, 1984–2009

COUNTRY	1984–85	2009
Low-income economies		
Tanzania	75	30
Uganda	70	30
India	62	30
Ghana	65	25
Lower-middle-income economies		
Papua New Guinea	50	42
Pakistan	60	20
Philippines	60	32
Sri Lanka	55	35
Indonesia	45	35
Peru	50	30
Guatemala	42	31
Jamaica	58	25
Upper-middle-income economies		
Botswana	75	25
Brazil	60	27.5
Costa Rica	50	15
Colombia	49	33
Thailand	65	37
Argentina	45	35
Mexico	55	28
Malaysia	55	27
Chile	56	40

Sources: George J. Yost III, ed., *1994 International Tax Summaries, Coopers & Lybrand International Tax Network* (New York: Wiley, 1994); Glenn P. Jenkins, "Tax Reform: Lessons Learned," in Dwight H. Perkins and Michael Roemer, eds., *Reforming Economic Systems in Developing Countries* (Cambridge: Harvard Institute for International Development, 1991); Denise Bedell, "Personal Income Tax Rates," *Global Finance*, available at www.gfmag.com/tools/global-database/economic-data/10442-personal-income-tax-rates.html#axzz1YDeg5nMb, accessed February 2012.

In many of these cases, deep cutbacks in the highest tax rates were accompanied by reforms involving a very substantial broadening of the income tax base, through the reduction of special tax incentives, abolition of tax shelters, and the like. This pattern was especially notable in Bolivia, Colombia, Indonesia, Jamaica, and Sri Lanka, so that, even with a rate reduction, higher-income groups often ended up paying a higher proportion of total taxes than before.

Reasons for the worldwide shift toward lower tax rates on broader income tax bases are not difficult to find. First, income taxes imposed at high marginal rates have proven difficult or impossible to administrate, even in wealthy countries such as the United States. With high marginal tax rates, the incentives to evade taxes (through concealment of income) or avoid taxes (by hiring expensive legal talent to devise tax shelters) are very high. Second, the growing mobility of capital across international boundaries has meant that the risk of capital flight from a particular country increases

when that country's top rates of income tax exceed those prevailing in industrial nations, where tax rates have been falling.[10] Third, the operation of the income tax systems of such developed nations as the United States and Japan has placed downward pressure on the tax rates everywhere. This is because of the *foreign income tax credit,* wherein a country like the United States allows foreign income taxes to be credited (subtracted) from U.S. taxes due on income repatriated from abroad. This credit could be used, however, up to only the amount of tax payable at U.S. rates, which was reduced from 50 to 28 percent in 1986. Finally, high marginal rates of income tax did not prove to be particularly efficacious in correcting severe inequalities in income distribution in either rich or poor countries.

The second striking feature of recent tax reforms worldwide has been the steadily growing number of countries adopting the value-added tax. Several factors account for the popularity of the VAT: The two most important are its reputation as a money machine and its administrative advantages, relative to other forms of sales taxes and income taxes.[11] The record of the VAT in generating large amounts of revenue quickly and in a comparatively painless fashion has given it a reputation as an efficient way to generate money. Although this reputation stems largely from the experience in European countries, the record in developing countries does lend some support to the alleged revenue advantages of the VAT. In Indonesia, the 4 percent share of the value-added tax in GDP in 1987 was nearly three times the share garnered in 1983 by the taxes it replaced. Notwithstanding the marked revenue success of the VAT in nations such as Brazil, Chile, and Indonesia, its reputation as a money machine for many other countries appears to have been at least slightly overstated.

In the 1980s, it was common to hear the claim that the VAT was largely self-administering. This is not so, but the tax-credit type of VAT has three principal advantages over single-stage retail and nonretail sales taxes in limiting the scope for evasion. First, the VAT is self-policing to some extent because underpayment of the tax by a seller (except, of course, a retail firm) reduces the tax credit available to the buying firm. Even so, firms that also are subject to income taxes have incentives to suppress information on purchases and sales to avoid both the value-added and income taxes. Also, this possible advantage of the VAT is diminished when evasion at the final (retail) stage of distribution is endemic. Second, cross-checking of invoices enables the tax administration to match invoices received by purchasers against those retained by sellers. The cross-check feature is a valuable aid in audit activities but no substitute for a true, systematic audit. Third, that a large share of the VAT is collected before the retail level is an advantage particularly because, in most developing countries, an abundance of small-scale retail firms do not keep adequate

[10]For discussion of the implications for taxation of growing international mobility of financial and physical capital, see Dwight R. Lee and Richard B. McKenzie, "The International Political Economy of Declining Tax Rates," *National Tax Journal* 42, no. 2 (March 1989), 79–87.

[11]For a full discussion of these reasons, see Alan A. Tait, *Value-Added Tax* (Washington, DC: International Monetary Fund, 1988), chap. 1.

records. In sum, the administrative advantages of the VAT are very real, if sometimes exaggerated by enthusiastic proponents.[12]

TAXES AND INCOME DISTRIBUTION

For decades, developed and developing countries alike have sought to use the fiscal system, particularly taxation, to redress income inequalities generated by the operation of the private market. Social philosophers from John Stuart Mill and eminent nineteenth-century Chilean historian Francisco Encina to twentieth-century philosophers such as John Rawls have sought to establish a philosophical basis for income redistribution, primarily through progressive taxes. No scientific basis is used to determine the optimal degree of income redistribution in any society. And across developing countries different views prevail as to the ideal distribution of income. But, in virtually all countries, the notion of **fiscal equity** permeates discussions of budgetary operations. In the overwhelming majority of countries, fiscal equity is typically defined in terms of the impact of tax and expenditure policy on the distribution of economic well-being. Progressive taxes, those that bear more heavily on better-off citizens than on poor ones, and expenditures whose benefits are concentrated on the least advantaged are viewed as more equitable than regressive taxes and expenditures.

On the tax side of the budget, the materialistic conception of equity requires that most taxes be based on the **ability to pay**. The ability to pay can be measured by income, consumption, wealth, or some combination of all three. Clearly, individuals with higher incomes over their life spans have a greater ability to pay taxes, quite apart from the moral question of whether they should do so. Indeed, the redistribution impact of taxation almost always is expressed in terms of its effects on income. However, philosophers since the time of Thomas Hobbes, a seventeenth-century English philosopher, have argued that consumption furnishes a better index of ability to pay than income; in this view, tax obligations are best geared to what people take out of society (consume) rather than what they put into society (as measured by income).

In practice, developing countries have relied heavily on these two measures of ability in fashioning tax systems. Personal and corporate income taxes employ income as the indicator; sales taxes and customs duties are indirect assessments of taxes on consumption. At a minimum, equity usually is assumed to require the avoidance of regressive taxes whenever possible. A number of tax instruments have been employed to secure greater progressivity in principle, if not in practice; all suffer from limitations to one degree or another.

[12]For a succinct summary of some of these issues, see John F. Due, "Some Unresolved Issues in Design and Implementation of Value-Added Taxes," *National Tax Journal* 42, no. 4 (December 1990), 383–98.

PERSONAL INCOME TAXES

The most widely used device for securing greater progressivity has been steeply progressive rates under the personal income tax. In some countries in some periods, nominal or legal marginal income tax rates have reached very high levels, even for relatively low incomes. Thus, for example, tax rates applicable to any income in excess of $1,000 in Indonesia in 1967 reached 75 percent, largely because tax rates were not indexed to rapid inflation; in Algeria in the 1960s, all income in excess of $10,000 was subject to marginal tax rates of nearly 100 percent; Tanzania imposed top marginal rates of 95 percent as late as 1981. Although, in most developing countries, marginal income tax rates are considerably lower than the preceding examples and, as is apparent from Table 11-3, have been falling, some countries still attempt to impose rates in excess of 50 percent, such as Angola, Gabon, and Cuba, but most countries with marginal rates at this level in the twenty-first century are high-income European states with large social welfare programs.

If the tax administration machinery functioned well and capital were immobile among countries, the pattern of actual tax payments of high-income taxpayers would resemble the legal, or theoretical, patterns just described. In fact, in most countries, **effective taxes** (the taxes actually collected as a percent of income) fall well short of theoretical liabilities. Faced with high income tax rates, taxpayers everywhere tend to react in three ways: (1) They evade taxes by concealing income, particularly capital income not subject to withholding arrangements; (2) they avoid taxes by altering economic behavior to reduce tax liability, whether by supplying fewer labor services, shipping capital to tax havens abroad, or hiring lawyers to find loopholes in the tax law; and (3) they bribe tax assessors to accept false returns.

For all these reasons, the achievement of substantial income redistribution through progressive income taxes has proven difficult in all countries, including the United States and the three Scandinavian nations where tax rates were long among the world's most progressive. Tax avoidance is the favored avenue for reducing tax liability in the United States, where use of the other methods can result in imprisonment. But, where tax enforcement is relatively weak, particularly where criminal penalties for evasion are absent and tax officials deeply underpaid, tax evasion and bribery are used more commonly. The scope for substantial redistribution through the income tax is, therefore, even more limited in developing countries than in the United States or Sweden.

Notwithstanding these problems, a significant share of the income of the wealthiest members of society is caught in the income tax net in many developing countries. Revenues from personal income tax collections in countries such as Colombia, South Korea, and Chile have been as high as 15 percent of total taxes and in a few others have run between 5 and 10 percent of the total. In virtually all developing countries, the entirety of such taxes is collected from the top 20 percent of the income distribution. This means, of course, that the very presence of an income tax, even one imposed at proportional rather than progressive rates, tends to reduce income inequality.

TAXES ON LUXURY CONSUMPTION

In view of the difficulties of securing a significant redistribution through income taxes, countries may employ heavy indirect taxes on luxury consumption as a means of enhancing the progressivity of the tax system. Efforts to achieve this goal usually center on internal indirect taxes, such as sales, and on customs duties on imports, but not excises on tobacco and alcohol.

Several developing countries have found that, provided tax rates are kept to enforceable levels, high rates of internal indirect taxes on luxury goods and services, coupled with lower taxes on less income-elastic items, can contribute to greater progressivity in the tax system. For revenue purposes, countries typically impose basic rates of sales taxes on nonluxuries at between 4 and 8 percent of manufacturers' values. This is equivalent to retail taxes of between 2 and 4 percent because taxes imposed at this level exclude wholesale and retail margins. Food, except that consumed in restaurants, almost always is exempted from any sales tax intended to promote redistributive goals. In developing countries, the exemption of food by itself renders most sales taxes at least faintly progressive, given the high proportion (up to 40 percent in many middle-income countries and over 50 percent in low income countries) of income of poor households spent on food. Sales taxes involving a limited number of luxury rates of between 20 and 30 percent at the manufacturers' level have been found to be workable.

The redistributive potential of sales tax rates differentiated in this way, however, is limited by the same administrative and compliance constraints standing in the way of the heavier use of income taxation in developing countries. While sales taxes are not as difficult to administer as income taxes, they do not collect themselves. A manufacturer's sales tax system employing three or even four rates may be administratively feasible in most countries, even when the highest rate approaches 40 percent. Rates much higher than that or reliance on a profusion of rates in an attempt to fine-tune the tax lead to substantial incentives and opportunities for tax evasion. Jamaica had over 15 rates before 1986, and Chile had over 20 from 1960 to 1970. In recognition of these problems, Indonesia adopted a flat-rate manufacturers' tax in 1985: The tax applied at a rate of 10 percent on *all* manufactured items and imports. The tax nevertheless was slightly progressive because it did not apply to items that did not go through a manufacturing process, including most foodstuffs consumed by low-income families.

Although the use of internal indirect taxes, such as sales taxes, can contribute to income redistribution goals without causing serious misallocation of resources, the same cannot be said for the use of customs duties. Sales taxes are imposed on all taxable goods without regard to national origin, including goods produced domestically as well as abroad. Tariffs apply only to imported goods. Virtually all countries, developed and developing, use customs duties to protect existing domestic industry. Developing countries in particular employ customs duties as the principal means of

encouraging domestic industry to produce goods that formerly were imported. This strategy, called **import substitution**, is examined at length in Chapter 18.

Deliberate policies to encourage import substitution through the use of high protective tariffs, under certain conditions, might lead to results sought by policy makers. But accidental import substitution arises when tariffs are used for purposes other than protection, and this is unlikely to have positive results. Many countries, as already pointed out, use high tariffs to achieve heavier taxation of luxury consumption. Often heavy tariffs are imposed on imported luxury items for which there is no intention of encouraging domestic production. Therefore, many Latin American and some Asian countries have levied customs tariffs of 100 to 150 percent of value on such appliances as electric knives, hair dryers, sporting goods, and mechanical toys. For most countries, these items are clearly highly income elastic and apt candidates for luxury taxation.

But efforts to tax luxuries through high customs duties lead to unintended, and almost irresistible, incentives for domestic production or assembly of such products. In virtually all countries, save the very poorest, alert domestic and foreign entrepreneurs have been quick to seize on such opportunities. By the time local assembly operations are established, they usually can make a politically convincing case that the duties should be retained to enable local production to continue, even when the value added domestically is as low as 10 percent of the value of the product. Such operations, if subject to any local sales taxes, usually succeed in being taxed at the basic tax rate, usually 5 to 10 percent. By relying on tariffs for luxury taxation, the government ultimately forgoes the revenues it previously collected from duties on luxury goods, as well as severely undermining the very aims of luxury taxation.

If, instead, higher luxury rates on imports are imposed under a sales tax collected on both imports and any domestic production that may develop, unintended import substitution can be avoided. The use of import tariffs for luxury taxation—indeed, for any purpose other than providing protection to domestic industry—is one illustration of the general problem of using one economic policy instrument (tariffs) to achieve more than one purpose (protection, luxury taxation, and revenue). Reliance on import duties for revenue is subject to the same pitfalls just discussed. If it is desired to increase government revenue from imports, a 10 percent sales tax applied both to imports and any future domestic production yields at least as much revenue as a 10 percent import duty, without leading to accidental protection.

CORPORATE INCOME AND PROPERTY TAXES: THE INCIDENCE PROBLEM

Income taxes on domestic corporations and property taxes often are mentioned as possible methods for securing income redistribution through the budget. Corporate income ultimately is received, through dividends and capital gains, almost exclusively by the upper 5 to 10 percent of the income distribution. Also, ownership of

wealth, which in many lower-income countries largely takes the form of land, tends to be even more concentrated than income. But, to a greater extent in developing than in developed countries, efforts to secure significant fiscal redistribution through heavier taxes on domestic corporations and property are limited both by administrative and economic realities.

Administrative problems bedevil efforts to collect income taxes from domestic firms to at least as great an extent as for income taxes on individuals. Hence, in many countries where corporate taxes on local firms have been important, as much as two thirds to three quarters of nonoil corporate taxes flow from state-owned firms, not from private firms owned by high-income individuals. Taxes on land should be subject to less severe administrative problems because it is an asset that cannot be hidden easily. However, land valuation for tax purposes has proven difficult even in Canada and the United States. It is more difficult in developing countries. Few developing countries have been able to assess property at anything approaching its true value.

Economic realities hinder efforts to achieve greater progressivity in the tax system through heavier use of corporate and land taxes because of the tendency for taxation to unintentionally burden groups other than those directly taxed. This is the **incidence problem**. The incidence of a tax is its ultimate impact, not who actually pays the tax to the government but whose income finally is affected by the tax when all economic agents have adjusted in response to the tax. The point of incidence is not always the point of initial impact. Taxes on domestic corporations may reduce the income of capitalists, who in turn might shift their investment patterns to reduce taxation. The income of the workers they employ and the prices charged to consumers may be affected as well. Ultimately, all taxes are paid by people, not by things such as corporations and property parcels.

The implications of incidence issues may be illuminated by a simple application of incidence analysis to the corporation income tax. Consider a profit-maximizing company that has no significant monopoly power in the domestic market. If taxes on the company's income are increased in a given year, then after-tax returns to its shareholders in that year are reduced by the full amount of the tax. In the short term, the incidence of the tax clearly is on shareholders. Because shareholders everywhere are concentrated in higher-income groups, the tax is progressive in the short run. If capital were immobile, unable to leave the corporate sector, the long-term incidence of the tax also would rest on shareholders and the tax would be progressive in the long run as well.

But, in the long run, capital can move out of the corporate sector. To the extent that capital is mobile domestically but not internationally, the corporate tax also is progressive in the long run. Returns on capital remaining in the corporate sector are reduced by the tax. Untaxed capital owners employed outside the corporate sector also suffer a reduction in returns, because movement of capital from the taxed corporate sector drives down the rate of return in the untaxed sector. Because the corporate tax reduces returns to capital throughout the economy, all capital owners suffer,

including owners of housing assets, and in a closed economy, the long-run incidence again is progressive.

However, few if any developing economies are completely closed; indeed, capital in recent years has become much more mobile internationally. To the extent that capital can move across national borders and higher returns are available in other countries, domestic capital migrates to escape higher corporate taxes. But as capital leaves an economy, ultimately output is curtailed and the marginal productivity of workers falls. Prices of items produced with domestic capital rise. In this way, an increase in corporate taxes may be borne by domestic consumers, who pay higher prices for the reduced supply of corporate-sector goods. Domestic workers, whose income is reduced when production is curtailed, may also bear a part of the burden of the corporate tax. Hence the corporate tax may be regressive (worsen the income distribution) in the long run. Although under other plausible conditions, an increase in the corporation income tax may result in greater relative burdens on capitalists, this scenario is sufficient to illustrate that often the intentions of a redistributive tax policy may be thwarted by all the workings of the economy. Policy makers cannot be sure that all taxes imposed on wealthy capital owners ultimately are paid by them.

The foregoing discussion suggests that, whereas some tax instruments may achieve income redistribution in developing countries, the opportunities for doing so are limited in most countries, a conclusion supported by a large number of empirical studies. With few exceptions, these studies show that the failure to administer personal income taxes effectively—the failure to utilize the limited opportunities for heavier taxes on luxury consumption, overreliance on revenue-productive but regressive excise taxes, and inclusion of food in sales taxes—significantly reduces the redistribution impact of tax systems. Tax systems in developing countries tend to produce a burden roughly proportional across income groups, with some tendency for progressivity at the very top of the income scale. The very wealthy pay a somewhat greater proportion of their income in taxes than the poor, but the poor still pay substantial taxes: at least 10 percent of their income in many cases studied. Of course, in the absence of such efforts, the after-tax distribution of income may have been even more unequal. This suggests that, although difficult to implement and often disappointing in results, tax reforms intended to reduce income inequality are not futile exercises and they may prevent taxes from making the poor worse off.

The limits of tax policy suggest that, if the budget is to serve redistribution purposes, the primary emphasis must be on expenditure policy. Where redistribution through expenditures has been a high priority of governments, the results generally have been encouraging. The effects of government expenditure on income distribution are even more difficult to measure than those of taxes. But both the qualitative and the quantitative evidence available strongly indicate that, in developing countries, budget expenditures may transfer very substantial resources to lower-income households, in some cases as much as 50 percent of their income. Some of the measures using government expenditures to attack poverty were discussed in Chapter 6.

ECONOMIC EFFICIENCY AND THE BUDGET

SOURCES OF INEFFICIENCY

On the expenditure side of the budget, *social cost-benefit analysis* can be deployed to enhance efficiency (reduce waste) in government spending. On the tax side, promotion of economic efficiency is more problematic.

All taxes lead to inefficiencies to one degree or another. The objective, therefore, is to minimize tax-induced inefficiencies consistent with other goals of tax policy. In most developing societies, this objective largely reduces the necessity of identifying examples of waste engendered by taxes and purging them from the system. If a particular feature of a tax system involves large efficiency losses, called *excess burden* in public economics, and at the same time contributes little or nothing to such other policy goals as income redistribution, then that feature is an obvious candidate for abolition. A full discussion of those elements of tax systems that qualify for such treatment is properly the subject of an extended public finance monograph. We can do little more here than indicate some of the principal examples.

A major source of inefficiency in taxation is excessive costs of tax administration. In some countries and for some taxes, these costs have been so high that they call into question the desirability of using certain taxes for any purpose. This is true for certain kinds of narrow-based stamp taxes that have been used in Latin America to collect government revenue on the documentation of transfer of assets, rental agreements, checks, and ordinary business transactions. Many stamp taxes cost more to administer than they collect in revenue.

In some countries, even broad-based taxes have had inordinately high costs of collection. Sales taxes in Chile and Ecuador in the 1960s before the reform of their tax systems cost $1 in administration for every $4. Administrative costs to collect net taxes of $100 in most Organization for Economic Co-Operation and Development (OECD) countries in the early twenty-first century in contrast range from $0.42 to $2.08, with the United States at $0.47 at the low end and Japan at $1.50 at the higher end.[13] And, because taxes on capital gains are so difficult to administer everywhere, including North America, the cost of collecting this component of income taxes often exceeds the revenue in developing countries.

Many developing countries offer substantial tax incentives to encourage investment in particular activities and regions. Many of these, particularly income tax holidays for approved firms, have proven very difficult to administer and few have

[13]These are the average rates for 2000–02 and are taken from OECD, *Tax Administration in OECD Countries: Comparative Information Series* (Paris: Organization for Economic Cooperation and Development, 2004).

led to the desired result.[14] Given persistently pressing revenue requirements in most countries, granting liberal tax incentives may have no effect other than requiring higher rates of tax on taxpayers who do not qualify for incentives. It is a dictum of fiscal theory that economic waste (inefficiency) arising from taxation increases by the square of the tax rate employed, not proportionately. Therefore, it is not difficult to see that unsuccessful tax incentive programs involve inefficiencies for the economy as a whole that are not compensated for by any significant benefits. Largely for this reason, Indonesia abolished all forms of tax incentives in a sweeping tax reform in 1984.

Finally, some features of major tax sources involve needless waste. From the earlier discussion of the use of import duties for luxury tax purposes, it is clear that this is often a major source of inefficiency. The use of progressive tax rates in a corporation income tax is another example. Progressive rates of corporate tax, where they cannot be enforced, do little to contribute to income redistribution; and where they can be enforced, they lead to several kinds of waste. Two of the most important are fragmentation of business firms and inefficiency in business operations. The incentive for fragmentation is evident: Rather than be subject to high marginal rates of taxes, firms tend to split into smaller units and lose any cost advantages of size. Where high progressive rates are employed for company income, the tax takes a high proportion (say, 70 percent) of each additional dollar of earnings, so the incentive to control costs within the firm is reduced. For example, for a firm facing a marginal tax rate of 70 percent, an additional outlay of $1,000 for materials involves a net cost to the firm of only $300, because at the same time taxes are reduced by $700.

NEUTRALITY AND EFFICIENCY: LESSONS FROM EXPERIENCE

Experience around the world, both in developed and developing countries, seems to indicate that, in societies where efficiency in taxation matters, this objective is best pursued by reliance on taxes that are as neutral as possible. A **neutral tax** is one that does not lead to a material change in the structure of private incentives that would prevail in the absence of the tax. A neutral tax system, then, is one that relies, to the extent possible, on uniform rates: a tax on all income at a flat rate or a sales tax with the same rate applied to all goods and services. A neutral tax system cannot be an efficient tax system.

An **efficient tax system** involves a minimum amount of excess burden for raising a required amount of revenue, where the *excess burden* of a tax is the loss in total welfare, over and above the amount of tax revenues collected by the government. Figure 11-1 demonstrates how the excess burden of, say, a commodity tax is greater the more elastic the demand or the supply of the taxed item. Figure 11-1a depicts the inelastic case, good A, whereas panel b shows the elastic case, good B. Constant

[14]See, for example, Arnold C. Harberger, "Principles of Taxation Applied to Developing Countries: What Have We Learned?" in Michael Boskin and Charles E. McLure Jr., eds., *World Tax Reform: Case Studies of Developed and Developing Countries* (San Francisco: ICS Press, 1990).

marginal costs (MC) are assumed in both cases and at the same level for both goods to portray more starkly the contrasting results achieved in those cases. Before the tax is imposed on either good, the equilibrium price and quantity are P_a and Q_a for good A and P_b and Q_b for good B. Now, we impose a tax rate (t) on both goods. The new equilibrium (post-tax) magnitudes are P_{at} and Q_{at} for good A and P_{bt} and Q_{bt} for good B. For good A, the total amount of government revenue is the rectangle $P_aP_{at}cd$. The total loss in consumer surplus arising from the tax is the trapezoid $P_aP_{at}ce$. The excess of the loss in consumer surplus over the amount of government revenue is the conventional measure of efficiency loss from a tax, or excess burden. For good A, the excess burden is the small triangle *cde*. By similar reasoning, the excess burden in the case of good B is the larger triangle *fgh*. We see that taxes of equivalent rates involve more excess burden when imposed on goods with elastic demand.

We can see that efficient taxation requires neither uniformity nor neutrality but many different tax rates on different goods, with tax rates lower for goods with elastic demand and higher for goods with inelastic demand. This is known as the **Ramsey rule**, or the **inverse elasticity rule**. The problem is that a tax system under this rule would be decidedly regressive: The highest taxes would be required on foodstuffs, drinking water, and sumptuary items. Taxes would be lower on items such as clothing, services, and foreign travel, which tend to be both price and income elastic.

The principle of neutrality in taxation is not nearly as intellectually satisfying a guide to tax policy as efficient taxation. Nevertheless, neutral taxation is to be preferred as one of the underlying principles of taxation, along with equity, until such

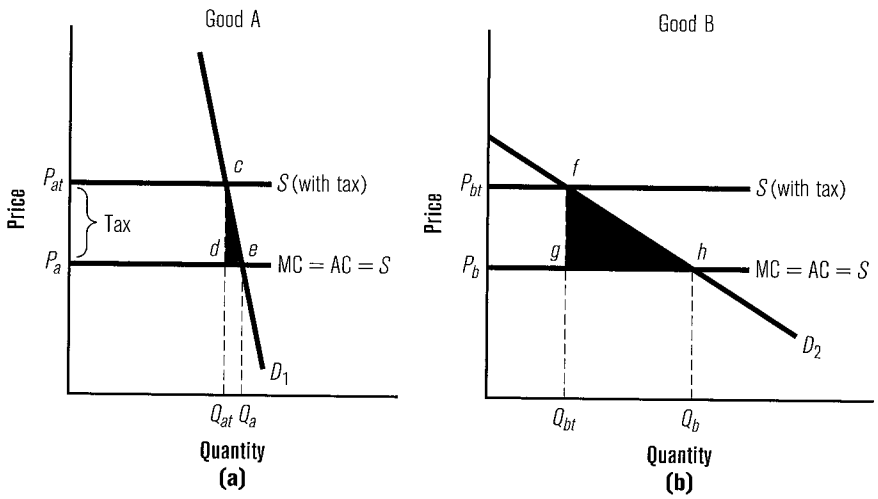


FIGURE 11-1 Taxation and Efficiency: Excess Burden of Commodity Taxes with Constant Marginal (MC) and Average Costs (AC), under Competition

(a) Case 1: Inelastic demand. (b) Case 2: Elastic demand. The shaded area in each case represents the excess burden of equal tax rates imposed on different goods. The greater the elasticity, the greater the excess burden.

time as analysts are able to identify optimal departures from neutrality—and uniformity in tax rates—in real-world settings. More intellectually satisfying tax policies must also wait until such time as administrative capacities are equal to the task of operating necessarily complicated structures of efficient or optimal taxes.

There is a paradox here. Neutral, uniform-rate taxes are less suited for efficiency goals than perfectly administered efficient taxes. Yet neutral tax systems are more likely to enhance efficiency in the economy than are efficient or optimal systems because neutral systems with uniform rates can be administered most easily and are much less vulnerable to evasion. This is not to say that neutrality has ever been or should be the overriding goal of tax policy. Governments often undertake very deliberate departures from neutral tax treatment of certain sectors or groups of society to achieve other policy goals. But, in real-world settings, these departures involve costs, not only in terms of tax administration but often in both equity and efficiency terms as well.

SUMMARY

- The expenditure side of the government budget is made up of capital or development items together with a wide variety of current use expenditures that are called *recurrent costs*. Developing country governments and donor aid programs frequently waste capital by neglecting the recurrent costs needed to maintain that capital.
- Developing countries rely mainly on indirect taxes (sales or value-added taxes and customs duties) rather than income and capital gains taxes because the latter are extremely difficult and expensive to administer in developing countries and thus produce less revenue.
- Tax reform in developing countries generally involves simplifying the tax system (one or a small number of different tax rates) and introducing taxes such as the value-added tax that have some self-enforcement properties. Greater complexity in the tax code leads to greater difficulty and greater corruption in administering the code. Simplified tax codes also reduce the distortions in the economy introduced by taxes.
- Taxation on the basis of an individual's ability to pay (higher-income people pay a larger share of their income in taxes) is a fundamental principle of equity. A good tax system is thus progressive rather than regressive, but progressive tax systems are extremely difficult to administer, particularly in developing countries, except in limited areas such as the exemption of food from sales taxes. When one looks at the actual incidence of taxes on high- and low-income earners, taxes that are designed to be progressive often turn out to be the opposite, and tax reform needs to take this into account.

12 Financial Development and Inflation

A country's **financial system** consists of a variety of interconnected financial institutions, both formal and informal. A central bank lies at the core of the organized financial system and is responsible for the control of the money supply and the general supervision of organized financial activity. Virtually everywhere, and particularly in developing countries, the commercial banking system is the most visible and vital component of the formal financial system, as acceptor of deposits and grantor of shorter-term credit. Other elements of the formal financial system include savings banks; insurance companies; and in a growing number of middle-income countries, pension funds and investment banks specializing in long-term credit as well as emerging stock exchanges. In high-income countries since the early 1990s there was the creation of a new set of institutions called *hedge funds* that played an important role in the global economic crisis of 2008–09.

Standing behind these financial institutions are a variety of government and sometimes private institutions whose role is to regulate the financial system. Typically separate agencies exist to regulate the commercial banks and the stock market, but the nature of regulation of the financial system is constantly changing. The need for regulation arises first and foremost from the fact that financial institutions exist primarily to manage the financial assets of people and companies other than the owners and managers of the financial institutions themselves. Regulation is designed to ensure that the financial assets of the public are managed responsibly without excessive risk. In the 1990s and during the first years of the twenty-first century the trend in high-income countries in North America and Europe was to reduce and even eliminate many of the regulations governing the financial system in the belief that the

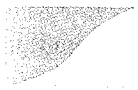
market would be an adequate regulator on its own. Many financial companies also developed new financial instruments that were designed to provide a service, such as insurance, that would not be subject to the standard regulations governing insurance contracts (these new instruments were called *credit default swaps*). This trend toward further liberalization of financial markets came to an abrupt end in 2009 and 2010 as the global financial crisis made it clear that markets on their own could not adequately control risky behavior of financial institutions, and a wide variety of new regulations were put in place.

The financial crisis of 2008–09 started in the United States and spread quickly to Europe, but developing countries were soon enveloped in the consequences of the global recession. The primary difference between developing and high-income countries in this context was that the developing country financial systems were simpler and made much less use of the more exotic financial instruments than did high-income countries. Commercial banks dominated the financial system in developing countries, investment banks if they existed at all were smaller and weaker, and insurance companies were in their infancy. The existence of financial systems depends on trust that they will manage a depositor's or an investor's funds well and that trust is built on a history of responsible behavior, on an effective regulatory system, and on the transparency of the operations of the financial organizations themselves. It takes decades or longer to develop trustworthy institutions and trust can be lost quickly as happened in 2008–09, but in developing countries this trust and the related effective regulation was largely confined to the commercial banks.

The limited scope of the formal financial sector in developing countries, however, left many people and small companies without ready access to credit or even to a reliable place to put their savings. In recent decades this has led to the creation of new **micro-credit institutions** designed to reach these left-out individuals, but long before formal micro-credit institutions were created, there was in virtually all developing countries an informal financial sector existing alongside the formal one. Informal financial institutions include pawnshops, local moneylenders, trade credit, and other arrangements involving the borrowing and lending of money such as intrafamily transfers and cooperative credit. In very low income countries, or even in some middle-income countries, the informal financial sector may rival the formal system in size. In almost all cases the informal sector is unregulated, although governments do interfere in its operation from time to time.

Financial policy embraces all measures intended to affect the growth, utilization, efficiency, and diversification of the financial system. In North America and western Europe, the term *financial policy* ordinarily is used as a synonym for *monetary policy*, the use of monetary instruments to reduce price instability caused by fluctuations in either internal or external markets. In the United States, these instruments include open-market operations, changes in legal reserve requirements of commercial banks, and shifts in central bank (Federal Reserve System) lending (rediscount) rates to commercial banks (these terms are explained later in this chapter). In developing

countries, the term *financial policy* typically has a much broader meaning. Monetary policy is part of financial policy but so are measures intended to encourage the growth of savings in the form of financial assets, develop money and capital markets, and allocate credit among different economic sectors. Regulation to control risk, as in the 1997–98 and the 2008–09 financial crises, is also a component of financial policy.



THE FUNCTIONS OF A FINANCIAL SYSTEM

The financial system provides four basic services essential for the smooth functioning of an economy. First, it provides a medium of exchange and a store of value, called *money*, which also serves as a unit of account to measure the value of the transactions. Second, it provides channels for mobilizing savings from numerous sources and channeling them to investors, a process called *financial intermediation*. Third, it provides a means of transferring and distributing risk across the economy. Fourth, it provides a set of policy instruments for the stabilization of economic activity. When these functions are managed poorly the result on a macro level can be a financial crisis leading to a global recession, inflation, or a debt crisis. The management of inflationary pressures will be discussed in this chapter and debt crises are the subject of Chapter 13.

MONEY AND THE MONEY SUPPLY

An economy without money as a **medium of exchange** is primitive. Trade between individuals must take the form of high-cost, inefficient barter transactions. In a barter economy, goods have prices, but they are expressed in the relative prices of physical commodities: so many kilos of rice for so many liters of kerosene, so many meters of rope for so many pairs of sandals, and so on. Trading under such circumstances involves onerous information costs.

Few societies have ever relied heavily on barter because of the high costs implicit in this means of exchange. At some point, prices of goods and services begin to be expressed in terms of one or more universally accepted and durable commodities, like gold and silver or even beads and cowrie shells. The rise of commodity money diminishes the transaction and storage costs of trade but still involves problems of making exchanges across space and time. Gold and silver prices fluctuate, and the commodities, therefore, are not fully reliable as **units of account**. As specialization within an economy increases, financial instruments backed by commodities appear. In the last century, with the rise of central banking all over the world, currency evolved into *fiat* money, debt issued by central banks that is legal tender. It is backed not by commodities of equivalent value but only by the full faith and credit of the central bank.

As markets widen and specialization proceeds apace, a need arises for still another financial instrument, **transferable deposits**. In the normal course of development, *checking* or *demand deposits* (deposits that may be transferred to any economic agent at the demand of the depositor) appear first and ordinarily bear little or no interest. Rising levels of economic activity, however, create increasing needs for transaction balances; individuals will always maintain some balances in demand deposits to meet these needs but tend to economize on the levels of such deposits if no interest is paid on them. With further monetization still another financial instrument begins to grow in importance, *time deposits*, which also are legally transferable on demand but sometimes only after stated periods. Time deposits involve contractual interest payments; higher interest rates induce people to hold greater amounts of deposits in this form.

While checking (demand) and time deposits are **liabilities** (or debts) of commercial banks, they are **financial assets** for the individuals who hold them. Both demand and time deposits are known as *liquid financial assets*. Unlike *nonfinancial assets* that also can be held by households and businesses (inventories, gold, and land), demand and time deposits can be quickly and conveniently converted into their currency equivalents. *Currency*, by definition, is the most liquid of all assets. The concept of liquid financial assets is an important one in any discussion of financial policy in developing countries. For most developing countries, the movement of savers in and out of liquid financial assets may be the prime factor behind the success or failure of a financial policy. We will see that long-term shifts from tangible, or nonfinancial, physical assets to financial assets, particularly liquid assets, bodes well not only for economic growth but also for economic stability.

A country's **money supply** may be defined as the sum of all liquid assets in the financial system. While not all economists agree about what constitutes a liquid financial asset, most vastly prefer this money supply concept to those commonly employed in early postwar monetary analysis. Formerly, the money supply was conventionally defined as the sum of only two liquid financial assets: currency in circulation outside banks (C) plus demand deposits (D), which together are known as M1 (*narrow money*). However, it later became clear that, because depositors tend to view time and savings deposits (T) as almost as liquid as demand deposits, the former also should be included in any workable concept of money supply, called M2 (*broad money*). Finally, for high-income countries, specialized deposit-taking financial institutions have arisen and offer an array of options to savers other than those available in commercial and savings banks. The liabilities of these specialized institutions (O) are included in M3 (*total liquid liabilities*), the broadest measure of money. Thus

$$\text{M1} = C + D \qquad [12\text{—}1]$$

$$\text{M2} = \text{M1} + T \qquad [12\text{—}2]$$

$$\text{M3} = \text{M2} + O \qquad [12\text{—}3]$$

For most low-income countries and many middle-income countries, liquid financial assets constitute by far the greatest share of outstanding financial assets. But, as income growth continues and the financial system matures, less-liquid financial assets assume progressively greater importance. These include primary securities such as stocks, bonds (issued both by government and firms), and other financial claims on tangible (physical) assets that are convertible into currency equivalents with only some risk of loss to the asset holder and are hence less liquid than demand or time deposits.

The evolution of financial activity follows no set pattern across countries. Differing economic conditions and policies may result in widely divergent patterns of financial growth. Nevertheless, as per capita income rises, money increases as a ratio to gross domestic product (GDP). Table 12–1 shows these patterns for broad money (M2).

TABLE 12–1 Broad Money (M2) as a Percentage of Gross Domestic Product (GDP), 1980 and 2008

	PERCENT OF GDP	
COUNTRY/ECONOMY	1980	2008
Low-income countries	19	36
Ethiopia	–	31
Tanzania	–	28
India	32	70
Bangladesh	13	55
Kenya	30	40
Nigeria	24	30
Ghana	16	24
Honduras	23	51
Lower-middle-income countries	28	54
Pakistan	39	44
Bolivia	18	55
Cameroon	21	19
Philippines	22	59
Sri Lanka	28	33
Indonesia	13	36
Peru	17	31
Egypt	52	84
Upper-middle-income countries	29	79
Colombia	24	33
Argentina	29	25
Mexico	27	26
Malaysia	71	118
Korea, Republic	29	62
High-income countries	72	113
Saudi Arabia	13	48
Japan	138	207
United States	68	83

Source: World Bank, *World Development Indicators 2011* (Washington, DC: World Bank, 2011), available at http://databank.worldbank.org/ddp/home.do.

actively advocate inflation but tended to have a higher threshold of tolerance for a steadily rising general price level than now is common.

INFLATION EPISODES

Inflationary experiences vary widely among developing countries and generalizations are difficult to make. Nevertheless, postwar economic history offers some interesting national and regional contrasts in both susceptibility to and tolerance for different levels of inflation. The period before the early 1970s was one of relative price stability in developing countries. In the southern cone of Latin America, however, particularly in Argentina, Brazil, and Chile, **chronic inflation** (prices rising 20 to 50 percent per year for three years or more) was an enduring fact of economic life for much of the 1950s through the 1980s and even beyond. The experience of these countries indicates that a long period of double-digit inflation does not necessarily lead to national economic calamity in all societies. In all, about two dozen countries experienced chronic inflation from 1950 to 2008, as shown in Table 12–2, but only a few, including Tajikistan, Venezuela, and Uzbekistan, have experienced chronic inflation throughout much of the first decade of the twenty-first century.

However, a tolerable rate of inflation in one country may constitute economic trauma in another. This may be seen more readily by considering the often progressive inflationary disease, acute inflation. **Acute inflation**, defined here as inflation between 50 and 200 percent for three or more consecutive years, was experienced by 18 countries listed in Table 12–2 over the postwar period, in some cases more than once per country. For Brazil, the progression from chronic to acute inflation did not result in any noticeable slowing of that country's relatively robust economic growth, whatever it may have meant for income distribution. In Ghana, on the other hand, a decade of acute inflation coincided with a decade of decline in GDP per capita. Although it may be tempting to attribute economic retrogression in Ghana at the time to acute inflation, it is more likely that the same policies that led to sustained inflation, not the inflation itself, were responsible for declines in living standards there.[1] Since the middle of the 1990s acute inflation has been rare.

Although acute inflation has proven toxic to economic development in some settings and only bothersome in others, **runaway (hyper) inflation** almost always has had a devastating effect. Inflation rates in excess of 200 percent per year represent an inflationary process that is clearly out of control; 22 countries have undergone this traumatic experience since 1950. One major bout of runaway inflationary experience occurred in three Latin American countries: Bolivia and Argentina in 1985, Argentina again in 1988–90, and Peru in 1988–91. In Bolivia, the annual rate of inflation over

[1]For a diagnosis of the causes of the Ghanian economic decline after 1962, see Michael Roemer, "Ghana, 1950 to 1980: Missed Opportunities," and Yaw Ansu, "Comments" (201–30), both in Arnold C. Haberger, ed., *World Economic Growth* (San Francisco: ICS Press, 1984).

TABLE 12-2 **Inflation Outliers: Episodes of Chronic, Acute, and Runaway Inflation among Developing Countries, 1948–2008 (average annual rates)**

	CHRONIC INFLATION (20–50 PERCENT, 3 YEARS)		ACUTE INFLATION (50–200 PERCENT, 3 YEARS)		RUNAWAY INFLATION (200+ PERCENT, 1 YEAR)	
COUNTRY	YEAR(S)	RATE	YEAR(S)	RATE	YEAR(S)	RATE
Angola			1997–2003	143	1993–96	858
Argentina	1950–74	27	1977–82	147	1976	443
			1986–87	111	1983–85	529
					1988–91	1,400
Armenia					1990–97	700
Azerbaijan					1990–96*	590
Belarus			1997–2002	87	1990–96*	715
Bolivia	1979–81	33	1952–59	117	1983–86	1,132
Brazil	1957–78	36	1979–84	108	1985	227
					1986	145
					1987–93	831
Chile	1952–71	29			1973–76	308
	1978–80	36				
	1983–85	26				
Colombia	1979–82	26				
	1988–92	28				
Dominican Republic			1988–92	51		
Ecuador	1997–2001	49				
Georgia					1990–96*	2,279
Ghana	1986–90	32	1976–83	73		
			1995–97	75		
Indonesia					1965–68	306
Kazakhstan					1990–96*	605
Kyrgyz Republic					1990–96*	256
Latvia					1990–96*	111
Lithuania					1990–96*	179
Malawi	1998–2002	40				
Mexico			1982–88	70		
Myanmar	2001–03	29	1992	116		
Nicaragua	1979–84	33			1985–91	2,130
Paraguay			1951–53	81		
Peru	1975–77	32	1950–55	102	1988–91	1,694
			1978–87	85		
Romania			1991–93	132		
			1997–92	56		
Russia	1996–99	42			1993–95	376
Sierra Leone			1983–92	81		
South Korea			1950–55	95		
Tajikistan	2000–08	22			1990–96*	394
Turkmenistan					1990–96*	1,074
Tanzania	1980–89	27				
Turkey	1981–87	38	1978–80	69		
			1988–2002	70		

(*Continued*)

TABLE 12-2 ***Continued***

Uganda	1990–92	37	1981–89	101		
Ukraine					1993–95	1,167
Uruguay	1948–65	26	1965–68	83		
	1981–83	34	1972–80	68		
Uzbekistan	2000-08	29	1984–92	76		
Venezuela	1987–92	40				
	1993–98	43				
	2002–08	28				
Zaire/Congo	1981–82	36	1976–80	68	1991–92	2,987
			1983–90	68	1992–95	3,206
					1996–2002	218
Zambia	1985–87	44	1988–92	113		
			1993–96	71		
Zimbabwe					2002–07	5,305†

*For a number of the former Soviet republics, the gross domestic product (GDP) price deflator data are given for the entire period 1990–96, not broken down by year.
†Data for Zimbabwe are not available after 2007. Some estimates place Zimbabwe's annual inflation at over 30,000 in the late 2000s.

Sources: International Monetary Fund, *International Financial Statistics Yearbook 1994* (Washington, DC: International Monetary Fund, 2004) and International Monetary Fund, *International Financial Statistics Yearbook*, 2003 (Washington, DC: International Monetary Fund, 2003); and World Bank, *World Development Indicators 2010*, at http://databank.worldbank.org/ddp/home.do accessed June 2011.

a period of several months in 1985 accelerated to a rate of nearly 4,000 percent. In Argentina, the monthly rate of price increases was 30.5 percent in June 1985 alone; on an annual basis that would have been an inflation rate of 2,340 percent. In both Bolivia and Argentina for much of 1985, workers had little choice but to spend their paychecks within days of receipt, for fear that prices would double or triple over the next week. In Peru, hyperinflation in 1989 gave birth to publications devoted only to the tracking of inflation (see Box 12-1).

In the 1990s, hyperinflation was experienced by virtually all the new republics formed after the breakup of the Soviet Union. In the first part of the 1990s, inflation reached a rate of nearly 400 percent a year in Russia over a period of several years and over 1,000 percent a year in the Republic of Georgia and Turkmenistan. In the latter two countries, civil war put heavy demands on government expenditures while reducing the ability of those governments to collect taxes, leading them to finance their activities by printing money. In Africa in the 1990s, civil war in Angola and the Congo also led to inflation rates of over 1,000 percent per year. Since the year 2000, however only the Congo (briefly) and Zimbabwe (throughout the first decade of the twenty-first century) have experienced runaway inflation. In the Congo the ongoing war in the eastern part of the country is the main cause. In Zimbabwe runaway inflation has been the result of across the board mismanagement of the economy throughout that decade.

BOX 12-1 HYPERINFLATION IN PERU, 1988-90

The economic and social havoc wrought by hyperinflation is difficult to comprehend for those who have not lived through the experience. The Peruvian hyperinflation, which began in 1988 and continued until 1991, provides some rueful examples.

The inflationary process was triggered by large budgetary deficits and sustained by subsequent ongoing deficits, virtual economic collapse, and steadily rising inflationary expectations. Peru already had experienced two serious bouts with acute inflation since 1950, but the pace of inflation in 1989 was the highest in the nation's history: 28 percent per month, or about 2,000 percent per year. Moreover, inflation accelerated in the first six weeks of 1990, as prices rose by 6 percent per week, or about 1 percent per day. From January 1989 to December 1990, the value of the Peruvian currency (the inti) on the free market fell from 1,200 intis per dollar to 436,000 intis per dollar.

This hyperinflation may turn out to be one of the best documented in history: In 1988, Richard Webb, an internationally respected Peruvian economist, began to publish a magazine devoted essentially to helping producers and consumers cope with the chaos associated with runaway inflation. The magazine, called *Cuanto?* (*How Much?*) appeared monthly. It not only provided details on price developments for a large number of commodities and services but managed to extract what little humor there is in a situation in which the price of a movie ticket rises while people are waiting in line to buy it or taxi drivers must carry fare money in a burlap bag because the domestic currency collected in fares each evening is too bulky to fit in a wallet.

But precious little is funny about hyperinflation. In Peru, it was a story of government employees going without pay for weeks at a time, of indices of poverty nearly doubling from 1987 to 1989, further impoverishing the poorest 40 percent of the population. It was a story of precipitous decline in gross domestic product (GDP) and the rise of pervasive black markets in everything from dollars to gasoline to cement. It was a time when, on each payday, laborers rushed to the market to buy their weekly food supplies before they were marked up overnight. It was a tale of wide variations in price rises, in which prices for such items as pencils and chicken increased by more than 25 times from February 1989 to February 1990, but prices of light bulbs and telephone services increased "only" 10-fold.

Imagine life in Lima, the capital city, in the first few weeks of 1990, for a middle-income family trying to survive. For the first 40 days of the year, increases in the price of dying outpaced the price of living: The cost of funerals rose 79 percent, while house rent rose by 56 percent and the price of restaurant meals and haircuts increased by 44 percent. By mid-1990, the economic paralysis of Peru was virtually complete. Peru's hyperinflation ended in 1992 (although inflation remained acute at 75 percent) as government reduced its budget deficit to 1 percent of GDP.

The countries listed in Table 12–2 brought on inflation in three different ways. In one group (including Argentina, Chile, Ghana, Indonesia, Peru, Russia, and Ukraine), large budget deficits relative to GDP were financed by borrowing from the central bank. In a second group (Paraguay in the early 1950s and Brazil and Uruguay before 1974), inflation was caused by a massive expansion of credit to the private sector. And in Nicaragua, Sierra Leone, Uganda, Angola, Congo, Zimbabwe, and several of the nations created out of the former Soviet Union, political strife or civil war exacerbated the fiscal and monetary causes of inflation. Whatever the initial impetus to inflation, once it begins to accelerate, the public begins to expect inflation to continue, and this leads to even-higher, more-sustained price increases.

For developing countries as a group, Table 12–3 shows that, on average, inflation was only 13 percent a year until the oil crisis began in 1973 but jumped to 21 percent a year during the period of rising oil prices until 1981. Inflation then accelerated to more than 30 percent a year even as oil prices began falling and exceeded 50 percent a year after oil prices collapsed in 1986. In the 1990s, these high rates of inflation continued through the first half of the decade and then slowed markedly after 1995. The principal exceptions were the states of the former Soviet Union, where triple-digit inflation in many cases continued to the end of the century and in a few cases a few years more. The industrial countries, in contrast, had much lower inflation throughout and the highest price increases coincided with the rise in oil prices.

MONETARY POLICY AND PRICE STABILITY

We saw from Table 12–3 that inflation largely was conquered in Asia during the 1970s but accelerated in Latin America and Africa during and especially after the oil-price boom of the 1970s. In the 1990s, however, attempts to control inflation were more serious and widespread with the notable exception of the republics formed out of the collapse of the Soviet Union. Even in these former Soviet republics, inflation by the beginning of the twenty-first century was generally down to single digits. The only case of runaway inflation during the first decade of the twenty-first century, as already noted, was Zimbabwe, where inflation was the result of a general collapse of the economy and of a government fiscal policy in the context of a ruling elite desperately trying to hold onto power. Controlling inflation in the midst of a civil war and the collapse of government functions requires first ending the war and reestablishing a functioning government. To control inflation under more normal circumstances, monetary policy is the principal instrument used to achieve price stability.

TABLE 12-3 Inflation by Regional Groupings, 1963–2008 (percent per annum)

REGION	1963–73	1973–81	1981–92	1992–95	1995–98	1998–01	2001–08
World	26.3	13.8	17.9	19.2	6.8	4.2	4.5
Industrial countries	24.6	10.3	13.0	2.5	1.9	1.8	2.1
Developing countries	12.9	20.7	44.1	47.9	12.6	7.1	8.5
Africa	24.9	17.3	27.0	40.2	12.1	8.2	9.1
Asia	13.5	28.8	16.8	11.6	6.9	2.4	3.9
Middle East*	24.2	16.6	16.0	15.0	9.5	5.5	12.6*
Latin America and Caribbean	18.4	43.7	162.8	216.5	15.5	8.4	9.6

*The Middle East gross domestic product deflator series ends in 2007.

Sources: International Monetary Fund, *International Financial Statistics Yearbook 1994* (Washington, DC: International Monetary Fund, 1994); International Monetary Fund, *International Financial Statistics, 1998* (Washington, DC: International Monetary Fund, 1998); International Monetary Fund, *International Financial Statistics, 2003* (Washington, DC: International Monetary Fund, 2003); and International Monetary Fund, *International Financial Statistics 2010* (Washington, DC: International Monetary Fund, 2010).

MONETARY POLICY AND EXCHANGE-RATE REGIMES

Appropriate use of monetary policy in controlling inflation depends critically on the type of exchange-rate regime used by a country. Exchange-rate regimes form a continuum with fixed (pegged) exchange rates at one end and floating (flexible) exchange rates at the other. Under a **fixed-exchange-rate** system, a country attempts to maintain the value of its currency in a fixed relation to another currency, say, the U.S. dollar: The value of the local currency is **pegged** to the dollar. This is done through intervention by the country's monetary authorities in the market for foreign exchange and requires the maintenance of substantial **international reserves** (reserves of foreign currencies), usually equivalent to the value of four or more months' worth of imports.

Consider the case of Thailand, where from 1987 to early 1997 the Thai currency, the baht, was fixed at an exchange rate close to 25 baht to US$1. Because the exchange rate, if left to its own devices, would change from day to day to reflect changes in both the demand for and supply of exports and imports and capital flows, to defend the peg, the government must be prepared to use the country's international reserves to buy or sell dollars at an exchange rate of 25 to 1 to keep the exchange rate from moving. If, for example, a poor domestic rice harvest caused the nation to increase its food imports, the baht-dollar exchange rate would tend to rise (the baht would depreciate, as its dollar value falls) in the absence of any net sales of dollars from Thailand's international reserves. To sustain a fixed exchange rate, of course, the country must have sufficient foreign exchange reserves to keep on buying baht at that fixed rate. In 1997, because of large capital outflows, Thailand actually ran out of foreign exchange and had to abandon its support of the pegged rate.

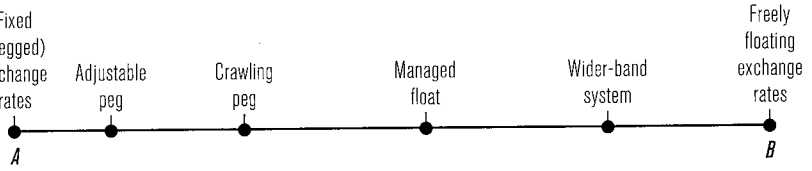


FIGURE 12-1 Continuum of Prototypes of Exchange-Rate Regimes

As one moves from point *A* to point *B*, both the frequency of intervention by domestic monetary authorities and the required level of international reserves tend to be lower. Under a pure fixed-exchange-rate regime (point *A*) authorities intervene so that the value of the currency vis-à-vis another, say, the U.S. dollar, is maintained at a constant rate.

Under freely **floating rates**, the authorities simply allow the value of local currency vis-à-vis foreign ones to be determined by market forces. Between the two ends of this continuum lie a number of intermediate options (Figure 12-1).[2] Closest to the floating-exchange-rate option is the *wider band* system, by which the exchange rate of a country is allowed to float or fluctuate within a predefined band of values, say, between 23 and 27 baht to US$1. But when conditions threaten to push the value of the currency beyond the band, the authorities intervene by buying or selling local currency as appropriate to stay within the band. Further along the continuum away from floating rates is the *managed float*, where the authorities are committed to defend no particular exchange rate, but they nevertheless intervene continuously at their discretion. A country with steadily shrinking international reserves, for example, might allow the value of its currency to depreciate against the value of other currencies—that is, allow the exchange rate to rise against other currencies.

Two other systems are closely related hybrids of fixed and floating rules. The *crawling peg*, used over a long period by Brazil, Colombia, and Indonesia, involves pegging the local currency against some other currency but changing this in gradual, periodic steps to adjust for any differential between the country's inflation rate and the world inflation rate. Closest to a fixed-exchange-rate system is the *adjustable peg*, involving a commitment by the monetary authorities to defend the local exchange rate at a fixed parity (peg), while reserving the right to change that rate when circumstances require.

Two very rigid forms of pegged exchange rates that a small number of countries have adopted are currency boards and dollarization. With a *currency board*, the government issues domestic currency only when it is fully backed by available foreign exchange reserves at the given exchange rate. Currency in circulation increases when additional foreign exchange becomes available (say, through increased export receipts) and decreases when foreign exchange becomes scarcer (say, through an

[2]For a full discussion of these and other types of exchange-rate regimes, see John Williamson, *The Open Economy and the World Economy* (New York: Basic Books, 1983), pp. 238–41; or Anne O. Krueger, *Exchange Rate Determination* (New York: Cambridge University Press, 1983), pp. 123–36.

increase in imports or capital outflows). This system ensures that the country will not run out of foreign exchange. However, the main instrument of adjustment becomes domestic interest rates, which increase when foreign exchange (and domestic currency) becomes scarcer, and decline when foreign exchange becomes more available. Hong Kong, Bulgaria, Argentina, Brunei, Djibouti, Estonia, and Lithuania all have or have had currency boards.

With *dollarization*, one country adopts another country's currency, as Panama did many years ago when it adopted the U.S. dollar as its currency. Most economists believe that currency boards and dollarization are appropriate in only a very limited number of developing countries that are small, very open to trade, and not vulnerable to large commodity price swings. In Hong Kong, the currency board system works reasonably well, although high interest rates and hence slower growth were required to prevent large currency outflows during and immediately after the Asian financial crisis. In Argentina at the beginning of the twenty-first century, the inability of the central government to rein in excessive local government spending caused pressures on reserves that ultimately made it impossible to maintain the peg to the dollar that was essential to the continuance of its currency board.

The currencies of all the major industrial countries have floated vis-à-vis one another since the early 1970s, with occasional intervention by national monetary authorities to prevent very sharp swings in rates. Most developing countries have adhered to either the adjustable-peg or the crawling-peg system, although an increasing number, particularly in Africa, have been adopting floating-rate systems as part of stabilization programs. Because, in practice, both pegged systems operate for particular periods like fixed-rate regimes, for our analysis, we focus most of our attention on monetary policy issues arising under fixed exchange rates in small, open economies.

SOURCES OF INFLATION

In open developing economies with fixed exchange rates, the rate of monetary expansion no longer is under the complete control of domestic monetary authorities. Rather, countries with fixed exchange rates may be viewed as sharing essentially the same money supply, because the money of each can be converted into that of the others at a fixed parity.[3] Under such circumstances, the stock of money (M), by definition, is the sum of two components: the amount of domestic credit of the banking system that is outstanding (DC) and the stock of international reserves of that country (IR), measured in terms of domestic currency. The money supply, therefore, has a domestic component and an international component, so we have

$$M = DC + IR \qquad [12\text{–}4]$$

[3]This section draws substantially on syntheses of monetary and international economics by Arnold C. Harberger. See his "A Primer on Inflation" and "The Inflation Syndrome," papers presented in the Political Economy Lecture Series, Harvard University, March 19, 1981.

Changes in the domestic money stock can occur either through expansion of domestic credit or by monetary movements that lead to changes in international reserves. That is,

$$\Delta M = \Delta DC + \Delta IR \qquad [12\text{–}5]$$

Under fixed exchange rates, a central bank of any small country can control *DC*, the domestic component, but has only very limited control over *IR*, the international component. Under such circumstances, developing countries that attempt to keep the rate of domestic inflation below the world inflation rate (through restrictive policies on domestic credit) are unable to realize this goal. If, fueled by monetary expansion abroad (growth in the world money supply), world inflation initially is running in excess of domestic inflation, the prices of internationally traded goods rise relative to those of domestic, nontraded goods.[4] Imports fall, exports rise, and the balance of payments moves toward surplus and causes a rise in international reserves. Therefore, the foreign components of the money stock rise. This is tantamount to an "importation of money" and eventually undoes the effort to prevent importation of world inflation. Again, a small country on fixed exchange rates can do little to maintain its inflation rate below that of the rest of the world. For very open countries with few restrictions on the movement of goods and capital into and out of the country, the adjustment to world inflation can be very rapid (less than a year). For less-open countries with substantial restrictions on international trade and payments, the process takes longer, but the outcome is inevitable under fixed exchange rates.

The fact that financial policy for stabilization in countries with fixed exchange rates is heavily constrained by international developments is sometimes taken to mean that changes in the domestic component of the money stock have no impact on prices in economies adhering to fixed exchange rates. On the contrary, excessive expansion in money and credit surely will result in domestically generated inflation that, depending on the rate of expansion, for a time, can be well in excess of world inflation rates. However, such a situation cannot continue for long, as excess money creation spills over into the balance of payments via increased imports and leads to a drain on international reserves and, ultimately, an inability to maintain the fixed exchange rate. As reserves dwindle, the country no longer can defend its exchange rate and devaluation becomes inevitable.[5] Inflation, therefore, can be transmitted to small, open economies through the working of the world economy or generated by domestic developments.

A growing number of developing countries have floating exchange rates (point *B* on the continuum in Figure 12–1). A floating exchange-rate regime allows countries

[4]This is but one of several mechanisms that led to changes in international reserves sufficient to thwart efforts by developing countries to insulate themselves from world inflation.

[5]Import controls are frequently used to stem the drain of reserves and avoid devaluation for a time. But import controls engender another set of distortions and inefficiencies, explored in Chapter 18, that eventually require more drastic measures, including devaluation.

to insulate themselves from world inflation. Under such a system, the rise in world prices attendant on world inflation would initially favor exports from the country and discourage imports. As a consequence, the current account of the country's balance of payments improves, international reserves rise, and the exchange rate soon appreciates (fewer baht are required to buy dollars, for instance). The appreciation in the country's exchange rate cancels out external price increases and prevents the importation of world inflation.

Under any exchange-rate regime, domestically generated inflation may result from excessive increases in domestic credit from the banking system to either the public or the private sector. Budget deficits of the central government, for example, must be financed by borrowing, but the embryonic nature of money and capital markets in most developing countries generally means that governments facing deficits ordinarily must resort to borrowing from the central bank. Borrowing from the central bank is equivalent to direct money creation via the printing press. The result is a direct addition to the reserve base of the monetary system, an increase in so-called high-powered money. It is important, however, to recognize that not all budgetary deficits are inflationary. We have seen that a growing economy is characterized by a growing demand for liquid assets, including money. Moderate budgetary deficits year after year, financed by the central bank, can help satisfy this requirement without leading to inflation. In general, the money stock may expand at least as fast as the growth in real income, with little or no inflationary consequences.

Earlier we saw that liquid assets normally are between 40 and 60 percent of GDP in low- and lower-middle income economies (with wide variations), equivalent to roughly three to six months of income. Therefore, the public generally is willing to hold this much in money balances. A deficit of 2 percent of GDP financed by money creation adds only marginally to the money supply and easily may be accepted by the public. But a deficit of 8 percent of GDP increases the stock of money by an amount equal to one more month of income, an amount the public may be unwilling to hold (unless nominal interest rates on deposits are greatly increased). The excess spills over into higher prices.

Use of bank credit to finance government deficits has not been the only source of inflationary monetary expansion in developing countries. Sometimes excessive growth of credit to the private sector has played a more significant role in domestically generated inflationary processes. Nevertheless, as a general rule, inflation rates that are much in excess of world inflation usually have been traceable to budgetary deficits.

For countries attempting to maintain fixed exchange rates, efforts to avoid price increases in excess of world inflation must be a matter primarily of fiscal policy, not monetary policy. If budget deficits are not held to levels consistent with world inflation, even very deft deployment of monetary policy instruments are unable to prevent rapid inflation, devaluation, or both. There still is a role for monetary policy in developing countries, but that role must be largely passive. Resourceful use of monetary policy can help by not making things worse and moderating strong

inflationary pressures until the budget can be brought under control, provided the latter is done fairly quickly.

Monetary factors are causes of inflation in both fixed- and floating-exchange-rate countries. In the case of fixed exchange rates, both world monetary expansion and domestic monetary expansion generate inflation; in flexible-exchange-rate countries, inflation arises from domestic monetary sources. But, thus far, no mention has been made of so-called nonmonetary causes of inflation. It seems plausible that internal and external shocks, such as those arising from widespread crop failure in the domestic economy or a drastic increase in prices of imported energy, could have important effects on inflation in countries suffering such shocks. This is true, but the mechanism whereby nonmonetary factors may initiate or worsen inflation needs to be clearly portrayed.

Nonmonetary disturbances indeed may precipitate policy reactions that lead to domestic monetary expansion large enough to accommodate higher relative prices of food or oil and large enough to cause inflation. In the absence of accommodating monetary expansion in the face of such shocks, inflation can be contained, but at some cost. Failure to allow the money supply to expand to accommodate higher relative prices of important goods leads to increases in unemployment that most governments find unacceptable. Governments in such cases usually attempt to allow monetary expansion sufficient to avoid unwanted consequences for employment. But it is important to remember that, however advisable monetary accommodation may be on social and employment grounds, expansion in the money stock is required to fuel inflation, whatever external or internal factors may precipitate the expansion. This truth, known for centuries, is often incorrectly interpreted to mean that nonmonetary factors cannot "cause" inflation. They can, but only through an expansion of the national or international stock of money or both.

CONTROLLING INFLATION THROUGH MONETARY POLICY

The array of available instruments for anti-inflationary monetary policy in developed countries include (1) open-market operations, in which the central bank can directly contract bank reserves by sales of government securities;[6] (2) increases in legal reserve requirements of banks, so that a given volume of reserves support a lower stock of money (and reduce the credit expansion multiplier as well); (3) increases in rediscount rates, so that commercial bank borrowing from the central bank becomes

[6]Open-market operations are used as an instrument of monetary policy in countries with well-developed financial markets. When the Federal Reserve System in the United States or a central bank in Europe wants to curtail the growth of the money supply, it sells government securities (bonds, bills) on the open market. When a buyer pays for the securities, the effect is to reduce directly the reserves of the banking system because the funds are transferred from commercial bank deposits or household cash holdings to the account of the Federal Reserve. When the Federal Reserve wants to expand the money supply, it buys securities on the open market and thus directly adds to bank reserves.

less attractive; and (4) moral suasion, by which the exhortations of monetary authorities are expected to lead to restraint in bank lending policies.

For virtually all developing countries, the first instrument (open market operations) is not available for inflation control. Securities markets typically are absent or not sufficiently well developed to allow the exercise of this powerful and flexible instrument, although some emerging market economies can now use this tool. The other three monetary policy instruments are employed, with varying degrees of success, in developing countries. In addition, developing countries often resort to two other tools employed only infrequently in developed countries: (5) credit ceilings imposed by the central bank on the banking system and (6) adjustments in allowable nominal rates of interest on deposits and loans.

RESERVE REQUIREMENTS

All central banks require commercial banks to immobilize a portion of their deposits in the form of legal reserves that may not be lent to prospective customers. For example, legal reserve requirements for Indonesian and Malaysian banks in the late 1970s were expressed as 30 percent of deposits in domestic currency in the former and 20 percent of all deposits in the latter. Banks in Malaysia were required to add 20 units of currency to reserves for every 100 units of deposits. These figures are not too far out of line with legal reserve requirements in many industrial nations, where reserve ratios of 15 percent for demand deposits and 5 percent for time deposits are common.

Increases in reserve requirements can be used to help moderate inflation. An upward adjustment in reserve requirements works in two ways: It reduces the stock of money that can be supported by a given amount of reserves, and it reduces the money multiplier. The first effect induces banks to contract credit outstanding; the second reduces the growth in the money stock possible from any future increment to reserves.[7] Changes in legal reserve requirements are usually employed only as a last-ditch measure, although China raised reserve requirements several times during the 2000–10 period, when price increases were still modest but rising. Even small changes in the required ratio of reserves to deposits can have a very disruptive impact on commercial bank operations, unless banks are given sufficient time to adjust.

CREDIT CEILINGS

In some countries, such as Indonesia from 1947 to 1983; China in 1994 to 1996 and during the first part of the twenty-first century; and at various other times in Malaysia, Sri Lanka, and Chile, credit ceilings have been used as supplementary instruments

[7]In its simplest form the money multiplier (m) can be expressed as

$$m = (c + 1)/(c + k)$$

where c = the ratio of currency outside banks to deposits and k = the ratio of reserves to deposits. If k is raised, then m falls.

of inflation control. The International Monetary Fund (IMF) often requires countries seeking balance-of-payments support to adopt credit ceilings as a prerequisite for assistance. General ceilings of domestic credit expansion represent a useful method of controlling growth in the domestic components of the money supply. Credit ceilings, however, do not allow full control of money-supply growth in developing countries operating under fixed-exchange-rate regimes because the monetary authorities have no control over foreign components of the money supply. Nevertheless, general credit ceilings sometimes can be usefully deployed in combating inflation in countries not experiencing major imbalances in external payments. Unfortunately, ceilings work the least well where they are needed the most because countries attempting to deal with chronic inflation are usually those experiencing the most destabilizing changes in their international reserve positions. General credit ceilings are unlikely to have much effect on inflation unless the government simultaneously takes steps to reduce the budgetary deficits that, except in major oil-exporting countries, typically are the root causes of chronic, acute, and especially runaway inflation.

Countries often supplement general credit ceilings with specific ceilings on lending to particular sectors of the economy. Indonesia attempted to fine-tune credit controls in this way from 1974 to 1983, with poor results. The system of ceilings was so detailed and cumbersome that domestic banks were unable to come close to exhausting the ceilings. Excess reserves rose. The banks had little choice but to place their excess reserves in deposits overseas, primarily in banks in Singapore. Many domestic firms in Jakarta were then forced to seek credit from Singapore banks, which held well over a billion dollars of deposits from Jakarta banks that might have been lent to domestic firms at a lower rate in the absence of credit ceilings. China, in the first part of the twenty-first century, placed ceilings on credit to real estate developers in what was believed to be an overheated real estate market but with only limited impact on investments in this area.

INTEREST-RATE REGULATION AND MORAL SUASION

In most industrial countries, the central bank can influence interest rates by varying the *rediscount rate* charged on central bank loans to commercial banks that require additional liquidity. Because the rediscount rate is central to commercial banks' operations, it is important in determining the market rate of interest on both deposits and loans. As more developing countries adopt financial reforms that free interest rates from central bank control, they are better able to use the rediscount rate as a tool for influencing market interest rates.

In developing countries that controlled rates on loans and deposits, the controlled rates have been instruments of anti-inflationary packages. The use of such interest rate adjustments have been common in Latin America, and increases in deposit rates and loan rates were major elements in the stabilization programs of South Korea and Taiwan in the mid-1960s and Indonesia in both 1968 and 1974. The

objective in each case was twofold: to stimulate the demand for liquid assets and to discourage the loan demand for marginal investment projects by private-sector borrowers. The extent to which such measures can be successful depends on the interest elasticity of the demand for liquid assets and the interest elasticity of the demand for loans. In most of the cases just cited, particularly in the three Asian countries, both sets of elasticities evidently were sufficiently high because the stabilization packages succeeded to a large degree. Interest rate increases, however, were less effective in China in the 1990s and during 2000–09, largely because many firms, particularly state-owned or -controlled enterprises, are not very sensitive to the impact of higher interest costs on their investment plans.

Moral suasion by the monetary authorities, sometimes called *open-mouth operations* or *jawbone control*, is practiced no less extensively in developing than in developed countries. Warnings and exhortations to commercial banks to restrict lending or to encourage them to focus lending on particular activities have been quite common in Ghana. They also were used at various times in Malaysia, Singapore, Brazil, and elsewhere, sometimes before the imposition of credit ceilings and often to reinforce pressure on banks to adhere to ceilings. In both developed and developing countries, however, moral suasion has proven credible only when accompanied by forceful use of more tangible instruments of monetary control.

INTERNATIONAL DEBT AND COMBATING RECESSIONS

Our discussion to this point has focused on the role of monetary instruments in controlling domestic inflation because that was the dominant problem facing monetary and fiscal policy makers in developing countries since the end of World War II. In this chapter, however, we have not dealt with another major problem that has played an important role in developing countries, the role of accumulating debt obligations by many countries, particularly in Latin America and Africa. Borrowing from other countries and from international institutions such as the World Bank is also a way of financing a development program when domestic financial resources are inadequate, but it can lead to accumulating international debt to a level that the borrowing country can no longer afford. Because these international debt issues are both important and complex, we devote the next chapter to understanding the origins and solutions to developing country debt problems.

Until recently most developing countries did not have to face a problem common in high-income countries in recessions: the lack of sufficient aggregate demand to keep a country's resources of capital and labor fully employed. Developing countries frequently experienced recessions, but they were typically the result of a fall in the price of their natural resource exports or some other balance-of-payments problem. Increasing demand in these circumstances led not to the employment of unused resources but to either inflation or a worsening balance-of-payments deficit. But in 2008–09, countries that had come to rely on exports of manufactures to Europe and

North America suddenly experienced a sharp decline in the demand for their products, as a severe global recession led to a fall in incomes in high-income countries.

In the United States a standard method for getting out of a recession in previous years had been to rely on the central bank to lower interest rates, which would then lead to an increase in investment in a number of sectors, notably in housing. Even in mild recessions, however, interest rates in developing countries are seldom effective in generating increases in investment demand, largely because financial markets in these areas are undeveloped and many borrowers are not very sensitive to small changes in interest charges. In the case of the global recession of 2008–09 lowering interest rates was not very effective in either high- or middle-income countries. The only alternative for the countries affected was to turn to fiscal policy and to introduce what were called *stimulus packages,* large government expenditure programs paid for by borrowing mostly from the central bank. Among developing countries China had the largest stimulus package of well over US$500 billion if all of the components are included, and similar smaller packages were tried in Vietnam and elsewhere in Southeast Asia. These packages were made possible in part because the countries involved had built up large reserves of foreign exchange and thus did not have to fear that the increased demand would lead to large balance-of-payments deficits that could not be covered from existing foreign exchange reserves. China had enormous reserves of over $2 trillion (over $3 trillion in 2011). Many other countries in Asia, in the aftermath of the 1997–98 financial crisis that had been due in part to a lack of sufficient reserves, had built up their foreign exchange reserves as well. In the 1997–98 financial crisis, the most affected countries could not have introduced large stimulus packages to offset the impact on incomes and employment because their reserves were inadequate for even existing balance-of-payments requirements.

FINANCIAL DEVELOPMENT

Up to this point we have discussed the general role of financial systems in economic development and then turned to how financial systems can be mismanaged in a way that undermines stability in the economy and threatens growth in incomes and employment. High inflation by itself has been a major reason why the economic performance of many developing countries has suffered, and preventing inflation has often been the major focus of policy makers both within developing countries and in the international financial agencies. Controlling inflation is a major tenet of Washington Consensus–style reforms. But there is more to financial development than controlling inflation. The development of a financial system that carries out its major functions efficiently and fairly is an important and integral part of the development process.

SHALLOW FINANCE AND DEEP FINANCE

Policies for **financial deepening** seek to promote growth in the real size of the financial system, the growth of financial assets at a pace faster than income growth. In all but the highest-income developing countries, private-sector financial savings predominantly take the form of currency and deposits in commercial banks, savings and loan associations, postal savings accounts, and in some countries, mortgage banks. For most developing countries, growth in the real size of the financial system is reflected primarily in growth in the share of liquid assets in GDP. In contrast, under **shallow finance**, the ratio of liquid assets to GDP grows slowly or not at all over time and typically falls: The real size of the financial system shrinks. Countries able to mobilize large volumes of government or foreign savings can sustain high growth rates even under shallow finance policies, although even these countries may find financial deepening attractive for reasons of employment and income distribution. But, for countries where mobilization of government savings is difficult and foreign savings scarce or unwanted, deep finance may be essential for sustained income growth. This is because growth in the share of liquid assets in GDP provides an approximate indication of the banking system's ability to increase its lending for investment purposes. We will see that the hallmark of a deep financial strategy is avoidance of negative real interest rates; shallow finance, on the other hand, typically involves sharply negative real interest rates.

Growth in the real size of the financial system enhances its capacity for intermediation, gathering savings from diverse private sources and channeling these savings into productive investment. The need for financial intermediation arises because savings endowments do not necessarily correspond to investment opportunities. Individuals with the greatest capacity to save are usually not those with the entrepreneurial talents required for mounting new investment projects. Except in very simple, rudimentary economies, mechanisms are required to channel savings efficiently from savers to entrepreneurs.

If a country is restricted to self-financing by individuals or companies, many productive opportunities involving high private and social payoffs will never be seized because the resources of even the small number of very wealthy are not unlimited. Innovative, smaller-scale investors are not the only groups that fare poorly where financial intermediation is poorly developed; savers are penalized as well. Let us first examine the case in which even the most basic financial intermediaries, commercial banks, are absent. Under these circumstances, the domestic options open to savers are limited to forms of savings, such as acquisition of gold and jewelry, purchase of land and consumer durable goods, or other relatively sterile forms of investment in physical assets. Alternatively, wealthier savers may ship their savings abroad. The common feature of all such investments is that the resources devoted to them are inaccessible to domestic entrepreneurs who would adopt new technology, begin new firms, or expand production in existing enterprises.

All developing countries, therefore, have financial institutions, however embryonic, to serve as intermediaries between savers and investors, even where these intermediaries are limited to commercial banks that accept checking (demand) and time (savings) deposits from savers, to relend to prospective investors for a short term. Intermediation flourishes under deep finance, but under strategies of shallow finance intermediation is constricted and the financial system can contribute little to further the goals of economic growth. Later, we see that shallow finance may have unintended effects on employment and income distribution as well.

SHALLOW FINANCIAL STRATEGY

Shallow financial policies have a number of earmarks: high legal reserve requirements on commercial banks, pervasive nonprice rationing of credit, and most of all, sharply negative real interest rates. Countries rarely, if ever, have consciously and deliberately adopted strategies of shallow finance. Rather, the repression of the financial system flows logically from certain policies intended to encourage, not hinder, investment.

In developed and developing countries alike, policy makers often have viewed low nominal rates of interest as essential for the expansion of investment and controlled interest rate levels tightly. So long as the supply of investment funds is unlimited, low interest rates foster all types of investment activities, as even projects with low returns appear more attractive to investors. In accordance with that observation and in the belief that low interest rates are particularly essential to assist small enterprises and small farmers, governments often have placed low ceilings on nominal interest rates charged on all types of loans. These low ceilings are quite apart from special credit programs involving subsidized credit for special classes of borrowers. Because financial institutions ultimately must cover costs (or else be subsidized by governments), low legal ceilings on nominal loan rates mean low nominal interest rates on deposits as well.

As long as inflation is held in check, low ceilings on nominal loan and deposit interest rates may not retard growth, even when these ceilings are set below the opportunity cost of capital. The United States over the period 1800–1979 managed rather respectable rates of income growth, even in the presence of a set of archaic usury laws and other interest rate controls that (particularly before 1970) often involved artificially low, administered ceilings on interest rates. Even so, throughout most of the period before 1979, real interest rates in the United States remained positive; periods in which real interest rates were sharply negative were intermittent and confined to wartime (1812, 1861, 1917–18, and 1940–46).[8]

Usury laws and other forms of interest rate ceilings also have been common in developing countries. Financial officials in many developing countries, observing

[8]Steven C. Leuthold, "Interest Rates, Inflation and Deflation," *Financial Analysis Journal* (January-February 1981), 28–51.

gross imperfections in financial markets, have concluded that the market should not be permitted to determine interest rates. Monopoly (or oligopoly) power in financial markets, particularly in commercial banking, provides ample scope for the banks and other lenders to exercise market power in setting interest rates on loans at levels higher than the opportunity cost of capital.

There are ample observations of gross imperfections in financial systems in developing countries. Barriers to entry into banking and finance often allow a few large banks and other financial institutions to possess an inordinate degree of control over financial markets and thus exercise monopoly power in setting interest rates. Often these barriers are a direct result of government policies. Governments have prohibited new entrants into the field, adopted such stringent financial requirements for entry that only the very wealthy could amass the needed capital, or reserved permission for entry to political favorites who were attracted to banking and finance largely by the monopoly returns available when entry was restricted.

In this way, one set of government policies, entry restrictions, helps give rise to the need for extensive controls on prices charged by financial institutions. Typically, these controls take the form of interest rate ceilings imposed to limit the scope of monopoly power in the financial system. Controls by themselves do not necessarily lead to shallow finance. Rather, a combination of rigid ceilings on nominal interest rates and inflation impedes financial development and ultimately retards income growth.

Few economists believe that steeply positive real interest rates are essential for healthy growth in the real size of the financial system. The Chilean experience with very high real interest rates from 1981 to 1983 strongly suggests the opposite. There is no widely accepted answer to the question of what the required level of real interest rates is for a steady development of the financial system. The required real rate differs across countries in different circumstances. In some, financial growth may continue even at 0 or mildly negative real interest rates; for others, moderately high positive real rates of between 3 and 5 percent may be essential.

Apart from a few Latin American countries and Indonesia, most developing countries were able to keep rates of inflation at or below 5 to 6 percent before 1973. Inasmuch as nominal deposit rates typically were between 3 and 5 percent, real interest rates tended to be slightly positive or only mildly negative. When inflation accelerated in many developing countries after 1973, because few countries made more than marginal adjustments in nominal deposit rates, real interest rates turned significantly negative in a number of nations, as Table 12–4 shows for 1980. Negative interest rates endured in a few African and Latin American countries throughout the 1980s but, as the table suggests, are much less common today.

When real interest rates turn significantly negative, the maintenance of low nominal rates for promoting investment and income growth becomes counterproductive. Inflation taxes on liquid financial assets bring real growth in the financial system to a halt. Sharply negative real rates lead to shrinkage in the system, as the demand for liquid assets contracts. Contraction in the financial system results in a reduction in

TABLE 12-4 Real Lending Interest Rates, 1980s and 2008–09

COUNTRY	1980S	2008–09
Low-income countries		
Ethiopia	−18	−17
Tanzania	−18	7
India	4	4
Bangladesh	−5	8
Kenya	1	8
Nigeria	−4	19
Honduras	12	14
Lower-middle-income countries		
Bolivia	2	15
Philippines	0	6
Sri Lanka	−1	10
Indonesia	22	6
Peru	−19	18
Egypt	−4	1
Upper-middle-income countries		
Brazil	–	37
Hungary	1	6
Colombia	9	8
Malaysia	5	13
Korea	−5	2
High-income countries		
United Kingdom	−3	−1
Japan	2	3
United States	6	2

Source: World Bank, *World Development Indicators 2011* http://databank.worldbank.org/ddp/home.do, accessed June 2011.

the real supply of credit and thus constricts investment in productive assets. Under such circumstances, nonprice rationing of investible resources must occur and can take many forms. In most developing countries, only those borrowers with either the highest-quality collateral or the "soundest" social and political connections or those willing to make the largest side payments (bribes) to bank officers are successful in securing finance from the organized financial system. These criteria do not yield allocations of credit to the most productive investment opportunities.

Negative real interest rates make marginal, low-yielding, traditional types of investment appear attractive to investors. Banks and financial institutions find such projects attractive as well because they may be the safest and the simplest to finance and involve the most creditworthy borrowers. Satisfying the financial requirements of such investors constricts the pool of resources available to firms, with riskier projects offering greater possibilities for high yields. In addition, in the presence of substantial inflation, interest rate ceilings discourage risk taking by the financial institutions

themselves because under such circumstances they cannot charge higher interest rates (risk premia) on promising but risky projects. Also, negative real interest rates are inimical to employment growth because they make projects with relatively high capital-output ratios appear more attractive than if real interest rates were positive. This implicit subsidy to capital-intensive methods of production reduces the jobs created for each dollar of investment, even as the ability of the financial system to finance investment is shrinking.

Negative real rates of interest tend to lower the marginal efficiency of investment in all the ways described. In terms of the Harrod-Domar model described in Chapter 4, shallow financial strategies cause higher capital-output ratios. Consequently, growth in national income and, therefore, growth in savings tend to be lower than when real rates are positive. Therefore, shallow finance retards income and employment growth even if the interest elasticity of savings is 0. And if savings decisions are responsive to real interest rates, then shallow finance has even more serious implications for income growth, as the ratio of private savings to the GDP also contracts.

DEEP FINANCIAL STRATEGY

Deep finance as a strategy has several objectives: (1) mobilizing a larger volume of savings from the domestic economy—that is, increasing the ratio of national savings to the GDP (where the interest elasticity of savings is thought to be positive and significant); (2) enhancing the accessibility of savings for all types of domestic investors; (3) securing a more efficient allocation of investment throughout the economy; and (4) permitting the financial process to mobilize and allocate savings to reduce reliance on the fiscal process, foreign aid, and inflation.

A permanent move toward policies involving positive real interest rates or, at a minimum, avoidance of sharply negative real rates is the essence of deep finance. In turn, this requires either financial liberalization that allows higher nominal rates on deposits and loans, curbing the rate of inflation, or some combination of both.

Given the difficulties involved in securing quick results in reducing inflation to levels consistent with positive real rates of interest, the first step involved in a shift from shallow to deep financial strategies ordinarily is to raise the ceilings on nominal rates for both deposits and loans. In extreme cases of acute inflation, the initial step has involved raising ceilings on nominal deposit rates to as much as 50 percent in Argentina and Uruguay in 1976 and to nearly 200 percent in Chile in 1974 (where real interest rates nevertheless remained negative until 1976). As the real rate moves toward positive levels, savers strongly tend to increase their holdings of liquid assets; this allows a real expansion in the supply of credit to investors. Available evidence suggests that countries that attempt to maintain modestly positive real interest rates over long periods tend to be among those with the highest rates of financial growth.

Where finance is deep, inflation tends to be moderate; therefore, savers are not subject to persistently high inflation taxes on liquid-asset holdings. That being the case, they are less inclined to shift their savings into much more lightly taxed domestic assets such as gold, land, or durable goods and foreign assets such as currencies or land and securities. Rather, financial resources that otherwise may have been used for these purposes flow to the financial system, where they are more accessible to prospective investors. Nonprice rationing of credit, inevitable under shallow finance, diminishes as well, and the capacity of the financial system to identify and support socially profitable investment opportunities expands. Higher-risk, higher-yielding investment projects stand a far better chance of securing finance under deep than shallow finance. Growth prospects are enhanced accordingly.

Investment finance problems, however, do not end with the provision of a growing real flow of short-term credit. As economies move to higher levels of per capita income, the pattern of investment shifts toward longer horizons. Longer-term investment requires longer-term finance. Commercial banks everywhere are ill-suited for providing substantial amounts of long-term finance, given that their deposits primarily are of a short-term nature.

Therefore, as financial and economic development proceeds, the need for institutions specializing in longer-term finance rises accordingly: insurance companies, investment banks, and ultimately equity markets (stock exchanges) become important elements in financial intermediation. The type of well-functioning commercial bank system that tends to develop under deep finance almost always is a necessary condition for the successful emergence and long-term vitality of institutions specializing in longer-term investment finance.

Earlier we observed that entry into the financial field is rarely easy, and other factors also often lead to gross imperfections in financial markets. Under such circumstances, many developing-country governments have found intervention essential to develop financial institutions specializing in longer-term finance. Intervention may take the form of establishment of government-owned development banks and other specialized institutions to act as distributors of government funds intended as a source of longer-term finance, as in Indonesia and Pakistan. In Mexico and Colombia, governments provided strong incentives for private-sector establishment of long-term financial institutions. Other governments sought to create conditions favorable for the emergence of primary securities (stocks and bonds) markets, the source par excellence for long-term finance. In cases in which these measures have been undertaken in the context of financial markets with strong commercial banking systems (Hong Kong, Singapore, Brazil, and Mexico), efforts to encourage long-term finance have met with some success. In cases in which commercial banking has been poorly developed as a consequence of shallow finance (Ghana, Uruguay before 1976) or the government has sought to "force feed" embryonic securities markets through tax incentives and other subsidies (Indonesia before 1988, Kenya, Turkey), the promotional policies have been less effective.

INFORMAL CREDIT MARKETS AND MICRO CREDIT

The discussion of financial development to this point has dealt with modern credit institutions, the formal market. But in many developing countries, as noted at the beginning of this chapter, **informal credit markets** coexist with modern financial institutions. These markets arise in many forms. In rural India, village moneylenders make loans to local farmers who have no access to commercial banks. In Ghana and other West African countries, market women give credit to farmers by paying for crops in advance of harvest, and they assist their customers by selling finished goods on credit. In South Korea, established lenders actually made loans on the street outside modern banks; this justifies their designation as the *curb* (or *kerb*) *market.* In much of rural Africa wealthy family members make loans to less fortunate kin, and all over the developing world there are cooperative arrangements to raise funds and share credit among members. Even in modern economies, pawnbrokers and others give credit outside the formal credit system.

Informal credit generally is financed by the savings of relatively wealthy individuals, such as local landowners, traders, family members who have moved into lucrative jobs or businesses, and the pooled efforts of cooperative societies. But informal lenders also may have access to the formal banking system and borrow there, to relend to customers with no access to banks. How can they do this if the banks cannot? First, because they know their borrowers so well and may have familial, social, or other ties to them, informal lenders face lower risks than distant, large banks that might loan to the same borrowers. Loan recovery rates are higher (usually much higher than found in large banks in developing countries) because those who borrow in informal markets know that the availability of loans in the future depends on repaying current loans. Second, they also face lower administrative costs in making loans. Of course, moneylenders charge very high interest rates, and this is a third reason they coexist with banks, which are often prevented by law from charging rates high enough to cover the risks and costs of loans in small amounts to very small firms and low-income borrowers.

Beginning in the 1970s, micro-credit institutions and micro-finance institutions more broadly were developed to reach populations in developing countries that had no access to credit, except in informal markets in which interest charges frequently reached annual rates of 100 percent. The best known of these new institutions was the Grameen Bank, founded in 1976 by Muhammad Yunus. Building on innovations such as Grameen, the movement to provide micro credit gained momentum and spread around the world to Africa, Latin America, and Asia where billions of people had no access to reasonable credit. The expansion in micro credit was then supplemented with programs to provide institutions designed to facilitate savings and other financial transactions by the rural poor in developing countries. **Micro finance** was the term applied to this broader set of financial transactions.

The basic idea of micro credit and micro finance was that the poor could be safe credit risks if they were organized appropriately. The Grameen Bank began by

providing micro loans to poor women in the village of Jobra, Bangladesh. Today, Grameen covers virtually all of Bangladesh, with more than 2,500 branch offices and more than 20,000 employees. The Grameen Bank reports that they have over 8 million borrowers in Bangladesh; about 97 percent of whom are women, most are landless, and all are poor. Loans are made to individual women, but only through local groups that provide social pressure for repayment. Loans for micro enterprises average 28,000 taka (about US$400) and carry an annual interest rate of 20 percent. Loans are also given for other purposes, including for housing and education, at 8 and 5 percent, respectively. There is even a special program targeted at beggars, which offers very small loans at 0 interest. Today, Grameen maintains that its loans are not heavily subsidized, neither by donors nor the government of Bangladesh. The recovery rate on Grameen loans is impressive; it exceeds 97 percent and the bank reports it has earned a profit every year since it was founded. The Grameen Bank is more than a financial institution. Its loans require recipients to accept certain "social disciplines," such as cleanliness and family planning, and the bank provides such services as advice on home construction and access to education for some borrowers.[9]

A major problem with many efforts to promote micro credit, however, was that many of the institutions set up to provide such credit depended on subsidies, often quite large subsidies, either from international aid agencies or from private donations from the well to do. This was true of the Grameen Bank earlier in its history. In Indonesia in the 1980s, a government bank, the Bank Rakyat Indonesia (BRI), followed a different path. It began to provide full banking services, both loans and savings deposit facilities, to farmers, traders, and other small-scale borrowers, through their branches in over 3,000 villages. Loans are provided for either working capital or investment purposes. Today, loans range from 25,000 to 25 million rupiahs (from about $3 to $3,000), with repayment periods as short as three months and as long as two years. BRI programs—*kupedes* for the lending program and *simpedes* for the savings program—replaced a government-promoted effort that depended on heavy government subsidies and had a high rate of loan default. These new loans, in contrast, were commercially profitable without any subsidies right from the beginning.

BRI charged market rates of interest, around 30 percent a year, on its loans and paid attractive rates, about 12 percent, on deposits. These rates changed during the financial crisis of 1997–98, largely because of the high inflation in that period, but the basic rate-setting principles remained the same. In the absence of inflation, rates of 30 percent on loans may seem high, but they were considerably below the rates charged by informal money lenders. At these rates, the BRI was able to attract sufficient savings deposits to more than finance its loan program. Several million savers and borrowers mostly in rural areas gained access to the formal banking system for the first time through these BRI programs. Roughly 97 percent of the loans were repaid on time. In addition, the program became a major generator of profits for the

[9]World Bank, *World Development Report 1989*, p. 117; (New York: Oxford University Press, 1989) A. Wahid, *The Grameen Bank: Poverty Relief* (Boulder, CO: Westview Press, 1993).

BRI, something that could not be said about many of its credit programs for large-scale producers and borrowers during the 1997–98 financial crisis.[10] These loans have continued to be a major source of BRI profits in the years since the crisis.

Micro credit is sometimes promoted by its supporters as the answer to rural and urban poverty in developing countries, but as we have emphasized throughout this book, there is no single answer to how to lift the world's poor out of poverty. Micro credit and micro finance have helped poor households in Bangladesh, Indonesia, and elsewhere, but there is considerable debate over their impact on poverty (Box 12-2). While we know that some programs have done well, we do not know whether

[10]Richard H. Patten and Jay Rosengard, *Progress with Profits: The Development of Rural Banking in Indonesia* (San Francisco: ICS Press, 1991); Marguerite Robinson, "Rural Financial Intermediation: Lessons from Indonesia," Development Discussion Paper 434, Harvard Institute for International Development, October 1992.

BOX 12-2 DOES MICRO CREDIT REDUCE POVERTY?

To determine whether micro-credit loans reduce poverty, the experience of borrowers might be compared with nonborrowers. If borrowers, on average, have a lower poverty rate than do nonborrowers, we might conclude that micro credit was an effective instrument for reducing poverty, and the impact of micro credit on poverty reduction could be quantified. But this approach is problematic. We face a classic problem of *identification*. What if borrowers are systematically different from nonborrowers in some unobservable way? Maybe those who obtain micro-credit loans are more entrepreneurial or have a higher tolerance for risk. Their lower poverty rates may have more to do with these attributes than with the loans they received. Borrowers of micro credit might have sought out other types of informal credit, making it difficult to isolate the marginal impact of micro credit on poverty reduction.

Problems of identification like these can be addressed by randomized controlled trials (RCTs), which compare the impact of an intervention between an otherwise similar treatment and control group. Researchers at the Abdul Latif Jameel Poverty Action Lab (J-PAL) located at the Massachusetts Institute of Technology (MIT) conducted an RCT on micro finance in 2005 in the city of Hyderabad, India. Working with an Indian micro-finance firm, Spandana, branches were opened in half of 104 randomly selected slum neighborhoods; in the other half, branches were not opened. Spandana loans are similar to those pioneered by the Grameen Bank. They are offered to women who form groups with 6 to 10 members who are jointly responsible for the loans. Initial loans are for 10,000 rupees (about $200 at market exchange rates and $1,000 in terms of purchasing power parity [PPP]), with an annual percentage rate of 24 percent. Then 15

to 18 months after the introduction of the new branches, detailed household surveys were administered to almost 7,000 households living in the treatment and control neighborhoods to determine the impact of the micro-credit loans.

Because micro-finance firms other than Spandana were available throughout Hyderabad, residents of both treated and untreated neighborhoods had access to micro-credit loans. In neighborhoods where Spandana opened a new branch, 27 percent of residents took out a loan compared to 19 percent in untreated areas. This suggests that the treatment had an effect on borrowing rates. But it also indicates that between 70 and 80 percent of households did not take advantage of Spandana or any other micro-finance loans. The overwhelming majority of households either were not credit constrained or were unwilling to assume the risks associated with taking out a loan. This suggests that it may be a lack of profitable opportunities to invest borrowed funds rather than a lack of cheaper credit that constrains most poor households.

In comparing treated and untreated neighborhoods, the J-PAL study found that 7 percent of households started a new business in neighborhoods where Spandana opened a new branch compared to only 5.3 percent of households in the control areas. Access to micro credit therefore had a significant impact on the formation of new businesses. About one third of Spandana borrowers started a new business, and about one fifth spent their loans on equipment for existing businesses. The other half of borrowers used their loans for current consumption, including on food or medicine, to buy a durable for household use (for example, a television set), or to pay down other debts. In the period covered by the survey, there was no statistically significant increase in total consumption expenditures between treatment and control areas. In other words, micro credit did not reduce poverty rates over this time period.

As the J-PAL study notes, 15 to 18 months may be too short a time to assess the long-term impact of micro credit. The new businesses started as a result of Spandana loans may need more time to earn profits and lift families above poverty. In the short term micro credit did not reduce poverty, nor was it a recipe for improving the schooling or health of children or for empowering women. Of course, the J-PAL study is only one study, and as with all RCTs we must be careful in drawing general conclusions from any one experiment. Evidence from a variety of nonexperimental approaches finds many benefits of micro-credit loans and other micro-finance products. The authors of the Hyderabad study themselves are measured in their conclusions, noting that micro credit "may not be the 'miracle' that is sometimes claimed on its behalf, but it does allow households to borrow, invest and create and expand businesses."

Sources: Abhijit Banerjee, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan, "The Miracle of Microfinance? Evidence for a Randomized Evaluation," Massachusetts Institute of Technology, June 2010; Kathleen Odell, "Measuring the Impact of Microfinance: Taking Another Look," Grameen Foundation Publication Series, 2010.

the hundreds and thousands of such programs around the world justify the large outlays of outside aid that they have used. About half of the estimated $12 billion committed to micro-finance organizations worldwide came as grants or concessional loans from aid agencies and donors. These resources might have been spent in other ways and had more of a poverty impact.

Micro credit often has its biggest impact when it helps individuals with limited means to start a new business and gradually expand that business. But there are many other barriers in addition to the lack of credit that can inhibit the development of new businesses and undermine the value of the credit. Many individuals borrow saying the loan is for a small business but use the money for consumption that leaves them with a debt that they may have difficulty repaying. Some borrow funds from one micro-finance organization to pay off loans they secured from another micro-finance organization. Such practices, clearly, are not sustainable for either the borrower or the lender.

Research indicates that profitable micro finance on the BRI model is not a complete substitute for subsidized micro finance in reaching the poorest parts of the population. Profitable micro finance reaches large numbers of people who previously had no access to credit or appropriate vehicles for savings, but many of the programs designed to reach the poorest parts of the population in poor countries and provide other nonfinancial services along with credit also reach people that the for-profit programs do not. There is clearly a role for both kinds of credit facilities.[11]

[11]For a lengthy analysis of these and many other related issues see Beatriz Armendariz and Jonathan Morduch, *The Economics of Microfinance* (Cambridge: MIT Press, 2005); Robert Cull, Asli Demirguc-Kunt, and Jonathan Morduch, "Microfinance Meets the Market," *Journal of Economic Perspectives*, 23, no. 1 (2009), 167–92. Another source that evaluates micro finance is *David Roodman's Microfinance Open Book Blog* (http://blogs.cgdev.org/open_book).

SUMMARY

- The money supply is made up of the liquid assets of an economy but the degree of liquidity of particular assets varies, leading to different, more precise definitions of the money supply.
- The rate of inflation varies greatly among developing economies. Inflation, in effect, is a tax on those who hold money balances, and a moderate rate of inflation can sometimes increase government savings and investment and not harm growth, but the higher the rate of inflation, the more people shift away from liquid assets, thus undermining the development of the financial system and harming economic growth.
- Exchange rate management is an essential component of any effort to control inflation. Exchange rate systems range from fixed exchange rates,

by which the local currency is pegged to the dollar or some other currency, to floating rates that move with market forces. In small countries with fixed-rate systems, worldwide inflation is transferred rapidly to the country with the fixed exchange rate. With floating rates, inflation arises mainly from domestic, not international, sources.

- A variety of additional mechanisms for controlling inflation are at the disposal of the central bank, all of which must, in one way or another, reduce the growth rate of the money supply. Mechanisms such as open-market operations are generally more efficient than mechanisms such as credit ceilings, but most developing countries are not in a position to conduct open-market sales and purchases of government bonds. Credit ceilings and increased bank reserve requirements are less efficient but are effective in lowering the growth of the money supply and inflation.
- Financial panics have been present throughout the last several centuries, but countries have learned how to eliminate some of the reasons for financial panic. As the experience in the late 1990s and 2007–09 of several countries around the world demonstrated, however, there are reasons why panics can still occur.
- The nominal rate of interest is the agreed-on rate between lenders and borrowers, but the real rate of interest, the nominal rate adjusted for inflation, most influences whether individuals are willing to hold liquid assets or not.
- Positive real interest rates are necessary for financial deepening, defined as the rising ratio of liquid assets to GDP. Negative real interest rates have the opposite effect. Financial deepening generally supports growth, although growth sometimes still occurs in its absence.
- As development proceeds and the economy becomes more complex and the need for long-term investment financing increases, economies require a wider variety of financial institutions. Stock markets and bond markets are one source of such long-term finance but so are insurance companies and development banks, and the latter are often supported by government for the explicit purpose of providing longer term finance.
- The formal financial system particularly in developing countries tends not to reach small businesses and individuals with limited financial resources. Informal financial markets exist to provide financing for such potential borrowers, but the interest charges on such loans tend to be very high. To counter this problem a variety of more formal efforts in the form of micro-credit institutions have arisen, providing more reasonable credit terms. The expansion of micro finance since the 1970s has been rapid, but there is still much we do not know about its overall impact on poverty.

15 Managing Short-Run Crises in an Open Economy

Economic development takes place in the long term. Most of the processes discussed in the previous chapters, whether improving human welfare, increasing saving, or shifting toward manufactured exports, take years and even decades to bear significant results. If policy makers in developing countries gaze only at the far horizon, however, they are unlikely ever to reach it. Much happens in the short term, within a few months or a couple of years, to throw an economy off balance and make pursuing long-term strategies difficult and sometimes impossible. Policy makers need to emulate a ship's captain, who, always steering toward the port of destination, nevertheless must deal decisively with any storms at sea.

Among the most dangerous and likely of these storms are changes in world prices that throw the balance of payments into deficit, excessive spending that fuels inflation or unsustainable debt, and droughts or other natural disasters that disrupt production. Unless a government counteracts these economic shocks, they create greater uncertainty and higher risk for private producers and investors, who take evasive actions that reduce future investment, worsen the crisis, and cause development efforts to flounder.

During the 1970s and 1980s, the late 1990s, and in Europe in 2010–11, as pointed out in Chapters 12 and 13, many economies became unbalanced because of unstable world market conditions and their own macroeconomic mismanagement. In Chapters 5 and 11 through 13, we discussed the consequences of such **macroeconomic instability**. Countries with overvalued exchange rates and rapid inflation were unable to grow rapidly. *Stabilization programs*, many funded by the International Monetary Fund (IMF), were intended to correct these macroeconomic imbalances.

In this chapter, we develop a mechanism for analyzing the macroeconomic policies that a developing country should pursue to stabilize its economy and create a climate for faster economic growth. The model developed here incorporates the two main policy approaches for correcting macroeconomic imbalances: changing the level of domestic expenditures and adjusting relative prices. In many cases, managing economic crises specifically requires expenditure reduction (by lower government budget deficits and slower creation of money) and exchange-rate devaluation.

EQUILIBRIUM IN A SMALL, OPEN ECONOMY

Developing economies[1] have two features central to understanding how macroeconomic imbalances occur and can be corrected. First, they are **open economies**, in that trade and capital flow across their borders in sufficient quantities to influence the domestic economy, particularly prices and the money supply. Most economies are open in this sense, especially because of economic reforms in China beginning in the late 1970s and in Eastern Europe and the states of the former Soviet Union beginning in the early 1990s. Today only a few economies, such as Cuba, North Korea, and Burma, are so heavily protected and regulated (and subject to foreign embargoes) that they might not qualify as open to trade and finance.

Second, these are **small economies**, meaning that neither their supply of exports nor their demand for imports has a noticeable impact on the world prices of these commodities and services. Economists call these countries *price takers* in world markets. A number of developing countries can exert some influence over the price of one or two primary exports in world markets: Brazil in coffee, Saudi Arabia in oil, Zambia in copper, South Africa in diamonds, for example. But they almost never affect the price of goods they import, and for macroeconomic purposes, it usually is adequate to model even these countries as price takers.[2]

These two qualities, smallness and openness, are the basis for the **Australian model** of a developing economy.[3] Countries typically trade both importable and

[1]In developing this and the next two sections, we acknowledge an intellectual debt to Shantayanan Devarajan and Dani Rodrik, who wrote an excellent set of notes for their class on macroeconomics for developing countries at Harvard's John F. Kennedy School of Government in the late 1980s and to Richard E. Caves, Jeffrey A. Frankel, and Ronald W. Jones, who develop the open economy model in Chapter 19 of *World Trade and Payments: An Introduction* (Glenview, IL: Scott, Foresman, Little, Brown, 1990).

[2]Among developing countries, China and India are large enough that they could become exceptions to the small country rule, given continued growth in China and both greater growth and openness in India.

[3]So called because it was developed by Australian economists, including W. E. G. Salter, "Internal Balance and External Balance: The Role of Price and Expenditure Effects," *Economic Record* 35 (1959), 226–38; Trevor W. Swan, "Economic Control in a Dependent Economy," *Economic Record* 36 (March 1960), 51–66; W. Max Corden, *Inflation, Exchange Rates and the World Economy* (Chicago: University of Chicago Press, 1977). Australia also is a small, open economy.

exportable goods and services. The Australian model lumps importables and exportables together as *tradables* and distinguishes these from all other goods and services, called *nontradables.* (We use this specification again in Chapter 18's discussion of Dutch disease.)

Tradable goods and services are those whose prices within the country are determined by supply and demand on world markets. Under the small economy assumption, these world market prices cannot be influenced by anything that happens within the country and so are *exogenous* to the model (determined outside the model). The domestic (local currency) price of a tradable good is given by $P_t = eP_t^*$ where e is the nominal exchange rate in local currency per dollar (pesos per dollar for Mexico or rupees per dollar for Pakistan) and P_t^* is the world price of the tradable in dollars. Even if the supply of and demand for tradables change within an economy, the local price will not change because domestic supply and demand have a negligible influence on the world price. Yet, changes in the nominal exchange rate change the domestic price of tradables commodities. If a country **devalues** its nominal exchange rate, it increases the amount of local currency required to purchase a dollar (e increases); if a country **revalues** its nominal exchange rate it decreases the amount of local currency required to purchase a dollar (e decreases). Devaluations thus increase the local currency price of tradables (all else equal), while revaluations tend to decrease the local currency price of tradables. Because this model simplifies all tradables into one composite good, the price of tradables P_t is best thought of as an index, a weighted average of the prices of all tradables, much like a consumer price index.

Tradables include exportables, such as coffee in Kenya and Colombia, rice in Thailand, beef in Argentina, cattle in West Africa, palm oil in Malaysia and Indonesia, copper in Peru and Zambia, oil in the Middle East, and textiles and electronics in East Asia, and importables, such as rice in West Africa, oil in Brazil or Korea, and intermediate chemicals and machinery in many developing countries.

Nontradables are goods and services, such as transportation, construction, retail trade, and household services that are not easily or conventionally bought or sold outside the country, usually because the costs of transporting them from one country to another are prohibitive or local custom inhibits trade. Prices of nontradables, designated P_n, therefore, are determined by market forces within the economy; any shift in supply or demand changes the price of nontradables. Nontradable prices thus are *endogenous* to the model (determined within the model). The term P_n, like P_t, is a composite or weighted average price incorporating all prices of nontradable goods and services.

INTERNAL AND EXTERNAL BALANCE

Figure 15–1 depicts equilibrium under the Australian model. The vertical axis represents nontradables (N); the horizontal axis takes both exportables and importables and treats them together as tradables (T). The production possibility frontier shows

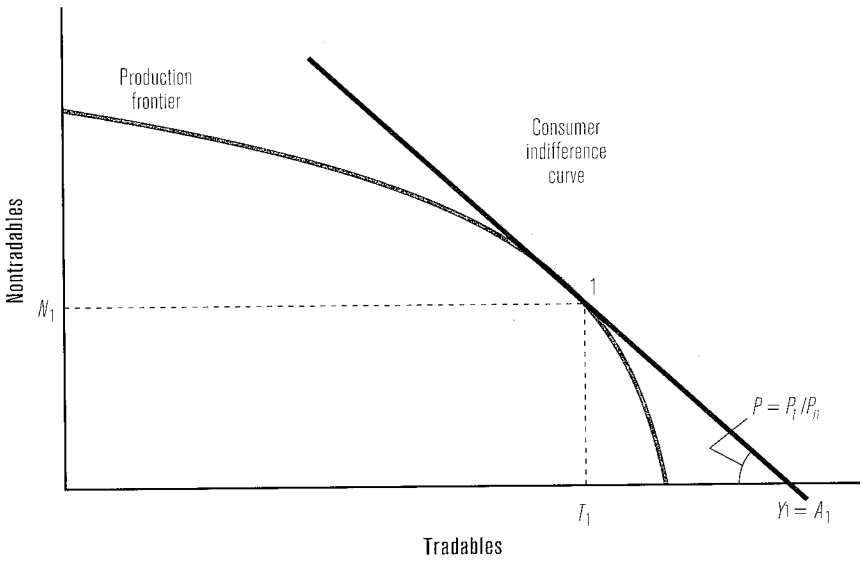


FIGURE 15-1 Equilibrium in the Australian Model

With equilibrium at point 1, the tangency of the production frontier and a community indifference curve, the country produces and consumes T_1 of tradables and N_1 of nontradables. The relative price, P, is a measure of the real exchange rate (see text).

Y_1, national income measured in tradable prices.

the menu of possible combinations of outputs of the two kinds of goods, N and T. The community indifference curves show consumer preferences between consumption of tradables and nontradables.

Equilibrium is at point 1, the tangency of a consumer indifference curve and the production possibilities frontier. At this point, the production of tradables, determined by the production frontier at point 1, is T_1, equal to the demand for tradables, determined by the indifference curve at 1, and similarly, for nontradables, supply equals demand at N_1. This is a defining characteristic of equilibrium in the Australian model: At point 1, the markets for both goods are in balance. Put another way, there is **external balance (EB)**, because the supply of tradables equals demand, and **internal balance (IB)**, because the supply of nontradables equals demand.

Point 1 simultaneously indicates the optimal (profit maximizing) combination of tradables and nontradables for producers, the optimal (utility maximizing) combination of tradable and nontradables for consumers. For both producers and consumers, this optimum occurs with respect to relative prices—in this case the relative price of tradables to nontradables. This relative price is indicated by the slope of the price line in Figure 15-1. This joint equilibrium for producers and consumers is indicated by the tangency of the indifference curve and the production possibility frontier. The tangency of the indifference curve and production frontier is jointly determined with the relative price of tradables in terms of nontradables, $P = P_t/P_n$. This relative

price, P, is one way to define the **real exchange rate (RER)**, one of the important innovations of the Australian model.[4] (Box 15–1 provides a more detailed explanation of the real exchange rate.) This formulation separates out prices that are under the influence of monetary and fiscal policy and domestic market forces, P_n, from prices that can be changed only by adjustments of the nominal exchange rate, $P_t = eP_t^*$. Note that the slope of the price line that is tangent to the production possibility curve and the consumer indifference curve is the only real exchange rate consistent with equilibrium in the model.

If P rises (the price line becomes steeper in the diagram), tradables become more expensive relative to nontradables. Producers then attempt to switch along the production frontier away from N goods, toward T goods. Consumers attempt to switch in the opposite direction, up along the indifference curve to consume fewer T goods and more N goods. Therefore, a rise in P should increase the surplus of T-good production over consumption.

If the production of T goods exceeds consumption of T goods, there is an external surplus, which is identical to a surplus in the balance of trade. To see this, start with the definition of the trade balance as

$$B_t = X - M \qquad [15\text{–}1]$$

where X and M are exports and imports. Because exports are the surplus of supply over demand for exportable goods, while imports are the opposite, a surplus of demand over supply, we can write the balance of trade as

B_t = value of X-goods supply − value of X-goods demand − (value of M-goods demand − value of M-goods supply),
= value of X-goods supply + value of M-goods supply − (value of X-goods demand + value of M-goods demand),
= value of tradables supply − value of tradables demand;

or if we let the supply of tradables be S_t and demand be D_t,

$$B_t = P_tS_t - P_tD_t = P_t(S_t - D_t) \qquad [15\text{–}2]$$

In Figure 15–1, with the economy in equilibrium, consumption of tradables is equal to production, so the balance of trade is 0.

The value of income (GDP) also can be found in Figure 15–1. It is the sum of the value of output of N goods (N_1) and T goods (T_1). This value is given by Y_1, the

[4]Chapters 18 and 19 will define the real exchange rate index as RER = R_oP_w/P_d. The term R_o is an index of the nominal exchange rate; in this chapter we use e, the nominal exchange rate itself. The term P_w is an index of world prices, often the U.S. consumer or wholesale price index and is similar or identical to P^* as measured in practice. But P_d is a domestic consumer or wholesale price index that includes both tradable and nontradable prices, whereas P_n is an index of nontradable prices only. Thus, the Australian formulation of the real exchange rate is a more-precise definition than those given in the later chapters.

BOX 15-1 REAL VERSUS NOMINAL EXCHANGE RATES

Most anyone who has traveled outside their home country has experience with nominal exchange rates. The nominal exchange rate is simply the number of units of local currency you can buy from the country you visit with one unit of your own currency. For instance, in August 2010, you could buy just over 46 Indian rupees, or 12.6 Mexican pesos, for US$1. (Note, this implies that you could also have bought 3.65 Indian rupees with one Mexican peso.) But when you bring a dollar into India or Mexico what you really care about is not how many rupees or pesos you can buy with that dollar but rather the quantity of actual goods and services that you can buy with that dollar. This depends on the prices of goods and services in the host country in addition to the price of its currency. This distinction is the key idea underlying the concept of the real exchange rate (RER).

In the most general sense, RER is the relative price of foreign goods in terms of domestic goods. In practice, economists have developed a range of approaches to quantifying this idea. Economists working on developed economies typically measure the RER as the relative price of domestic and foreign goods. In contrast, economists working on developing countries typically measure the RER as the relative price of tradables and nontradables. Tradable commodities are goods that are *or could be* traded internationally, in contrast to nontradable goods (such as housing and many services), which are not traded internationally. A key practical distinction between these categories of goods is that there are world market prices for tradable goods, but the prices of nontradables are determined purely by local supply and demand conditions in each country.

The nominal and real exchange rates are linked together by the requirement that the relative prices of tradable and nontradable goods in the RER be expressed in the same currency units. The prices of tradables are typically expressed in U.S. dollars, whereas the prices of nontradables are expressed in units of the local currency. Thus we need to use the nominal exchange rate to convert the dollar-denominated price of tradables into units of the local currency to calculate the RER.

We can construct this RER as follows. Expressing the nominal exchange rate, e, in terms of the number of local currency units per dollar, and expressing the prices of tradables and nontradables as P_t and P_n, respectively, we can construct the RER as

$$\text{RER} = \frac{eP_t^*}{P_n} = \frac{P_t}{P_n}$$

In this equation, the asterisk on P_t in the second term indicates that the price of tradables is expressed in dollars. Multiplying this dollar-denominated price by e, the nominal exchange rate as defined above, converts the price of tradables into the same local currency units as the price of nontradables.

It is important to consider several issues related to this concept of the RER. First, note that the RER is based on price indices rather than actual nominal price levels. The RER is thus expressed relative to a base year, and changes in the relative price of tradables to nontradables, say from 1 to 1.2, would indicate percentage changes (in this case, 20 percent) relative to the base year. We refer to an increase in the RER as a *depreciation* of the local currency, and a decrease in the RER as an *appreciation* of the local currency. While this may sound counterintuitive, the rationale for these terms is that when a currency depreciates in real terms, a given quantity of foreign goods can be exchanged for a greater quantity of that country's domestic goods (and vice versa in the case of an appreciation). It is for this reason that the RER is often thought of an as an indicator of a country's international competitiveness: When a country's currency depreciates in real terms relative to its trading partners' currencies, that country's goods become less expensive to foreigners.

An important practical challenge in constructing an RER lies in the need to choose price indices for tradables and nontradables. P_t and P_n are indices of the prices of entire categories of goods. Thus constructing these price indices first requires deciding which goods (and services) belong in which category. Specific price data may also be lacking. One short cut for addressing these challenges may be to use the U.S. consumer price index in place of P_t and a similar indicator from the home country (with the same base year) in place of P_n.

Although the availability of such price indicators as the consumer price index makes them convenient, their use in constructing RERs is problematic. Theory calls for an index of nontradables prices, but the consumer price index (CPI) is typically constructed to reflect the price of a basket of consumption goods that includes both tradable and nontradable goods. The larger the share of tradable goods in that basket, the greater the divergence between what the RER tells us in theory and what we actually measure if we construct an RER using those broad price indices. In practice, aggregate price indices purely for nontradables rarely exist. Similar problems exist in choosing a price index to represent tradables prices (for use in the numerator of the RER for a given country). In this case, in which the goal is to choose a price index based to the greatest extent possible on tradables, many authors use the wholesale price index (WPI) from the United States or from a given country's trading partners. Yet this approach too is problematic because (as Lawrence Hinkle and Peter Montiel note[a]) for-

[a]Lawrence E. Hinkle and Peter Montiel, *Exchange rate misalignment: concepts and measurement for developing countries.* (Washington, DC: World Bank, 1999).

eign WPIs may not provide a very close indication of the tradables prices actually faced by consumers in the home country. There is no perfect match between the theoretical requirements and practical data availability in constructing empirical RERs. A common compromise is to use the foreign WPI to represent P_t and the domestic CPI to represent P_n.

An additional question is whether the relevant RER is purely between two specific countries (the home country and a single trading partner—that is, the bilateral RER) or between the home country and multiple trading partner countries. In general, policy makers in a given country will be more concerned with how their currency relates in real terms with all of their trading partners. In that case, it is necessary to take a (trade-weighted) average of all the bilateral RERs between the home country and its trading partners. This average is called the **real effective exchange rate (REER)**.

The central challenge for policy makers concerned with their country's international competitiveness is whether the level of the REER at any given time reflects its equilibrium value or whether it is overvalued or undervalued relative to that equilibrium. Equlibrium in this setting generally refers to the level of the REER at which a country's internal market (that is, its supply and demand for nontrable goods and labor) and its external market (that is, its supply and demand for tradable goods) are in balance.

intersection of price line P from point 1 to the T axis.[3] In national income accounting, we distinguish two concepts. Gross domestic *product*, a measure of the value of output, is given by

$$\text{GDP} = C + I + X - M \qquad [15\text{-}3]$$

where C and I are consumption and investment by both the government and the private sector. Gross domestic *expenditure*, often called **absorption**, is

$$A = C + I = \text{GDP} + M - X \qquad [15\text{-}4]$$

When, as in Figure 15-1, the economy is in equilibrium, $X = M$ and income equals absorption. Indeed, this is the other condition for equilibrium in the Australian model. From equation 15-4, we can also see that $A - \text{GDP} = M - X$, indicating

[3]Along the T axis, Y_1 is measured in prices of the T good, so $P_t Y_1 = P_t T_1 + P_n N_1$ or $Y_1 = T_1 + (P_n/P_t)N_1$. But $P_n/P_t = \Delta T/\Delta N$, with $\Delta N = N_1$ and $\Delta T = Y_1 - T_1$, the distance along the T axis from T_1 to Y_1. Thus the value of both goods in T prices is $T_1 + Y_1 - T_1 = Y_1$.

that any excess of expenditures over income implies a negative trade balance of equal amount. (The appendix to this chapter provides a more detailed review of balance of payments accounting.)

This exploration of the Australian model yields three results. First, macroeconomic equilibrium is defined as a balance between supply and demand in two markets: nontradable goods (internal balance) and tradable goods (external balance). Second, to achieve equilibrium in both markets, two conditions must be satisfied: Expenditure (absorption) must equal income, and the relative price of tradables (the real exchange rate) must be at a level that equates demand and supply in both markets (the slope of P in Figure 15–1). Third, this also suggests two remedies for an economy that is out of balance: A government can achieve equilibrium (stabilize the economy) by adjusting absorption, the nominal exchange rate, or both. Generally, both instruments must be used to achieve internal and external balance.

THE PHASE DIAGRAM

Using the perspective of trade theory, we tie the small, open economy model of macroeconomic management to the tools of analysis already used in this text. But the principles of stabilization can be explored from a more-useful perspective, the **phase diagram**. To develop this approach, consider the markets for tradables and nontradables from the perspective of conventional supply and demand diagrams, as in Figure 15–2.

In these diagrams, we use the real exchange rate, which is the relative price of T goods in terms of N goods (P_t/P_n), as the price in both markets. For tradable goods, that gives a conventional supply and demand diagram: As the price rises, supply increases and demand decreases. But in the nontradables market, a rise in P means a fall in the relative price of N goods, so supply decreases and demand increases. Note that, in both markets, any increase in expenditure, or absorption, A, causes an outward shift of the demand curve: At any price, consumers buy more of both goods.

To use these diagrams as a basis for macroeconomic analysis, we need to change the interpretation of the supply curve for tradables. Until now, we have assumed that all tradables are produced within the home country. But foreign investment and foreign aid can add to the supply of tradables by financing additional imports. Therefore, the supply curve should not be S_t, but $S_t + F$, where F is the inflow of long-term foreign capital in the form of aid, commercial loans, and investment.

Figure 15–2 constitutes a simple model of the small, open economy that is based on two variables: The real exchange rate, P, on the vertical axis and absorption, A, which determines the position of the demand curves. These, of course, are the conventional variables of microeconomics, price and income. But in this model, they also are the two main macroeconomic policy tools of government: The exchange rate and the level of expenditure. Because these two variables are central to macroeconomic management, it would be helpful to develop a diagram that uses them explicitly on the axes.

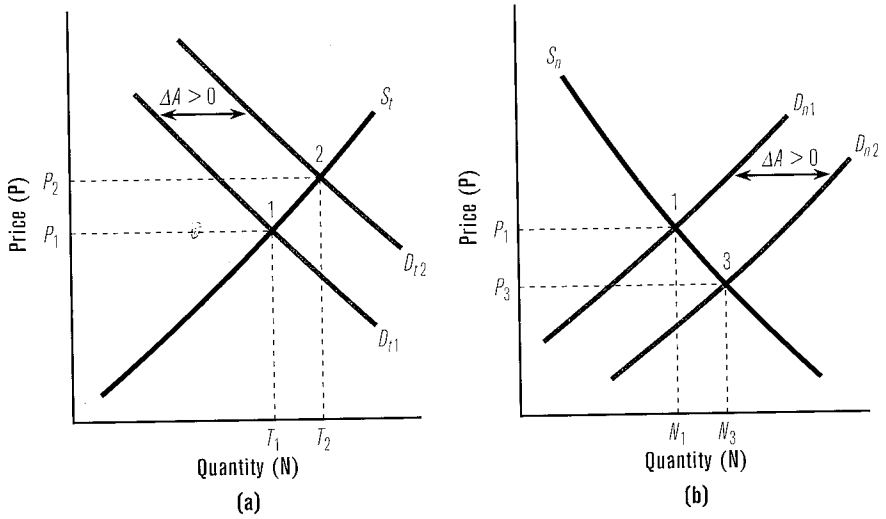


FIGURE 15-2 Tradables and Nontradables Markets

(a) Tradables market: The demand and supply curves for tradables S_t and D_t have the conventional slopes. (b) Nontradables market: The slopes are reversed. S_n falls as P rises (because the relative price of N is falling) and D_n rises as P rises. In both markets, demand increases when absorption (expenditure) increases, shown by and outward shift of D_t and D_n.

S, supply; D, demand; P (price) $= P_t/P_n$.

Figure 15-3 does this. It puts the real exchange rate, $P_t = eP_t^*/P_n$, on the vertical axis and real absorption, A, on the horizontal axis. The diagram also contains two curves, each representing equilibrium in one of the markets. Along the EB, or external balance, curve, the T-goods market is in balance ($S_t = D_t$). Along the IB, or internal balance, curve, the N-goods market is in balance ($S_n = D_n$).

The slopes of the two curves, EB and IB, can be derived from Figure 15-2. In the tradables market, when absorption is A_1, equilibrium is at P_1, where T_1 is produced and consumed. This equilibrium point 1 also is shown in Figure 15-3a. If absorption increases to A_2 in Figure 15-2a, the demand curve moves outward and shifts equilibrium to point 2. Note that with higher absorption, A_2, the real exchange rate, P_2, must be higher to restore equilibrium in the T-goods market. Increased absorption raises the demand for T goods. To meet this demand, it is necessary to raise output, which can be achieved only through a higher relative price of T goods, P_2. This higher price also helps regain balance by reducing the demand for T goods along the new demand curve. Point 2 is transferred to Figure 15-3a at (P_2, A_2).

In the nontradables market, when absorption is A_1, equilibrium is at P_1, where N_1 is produced and consumed. This equilibrium point 1 also is shown in Figure 15-3b. If absorption increases to A_3 in Figure 15-2b, the demand curve moves outward and shifts equilibrium to point 3. In the N-goods market, higher absorption, A_3,

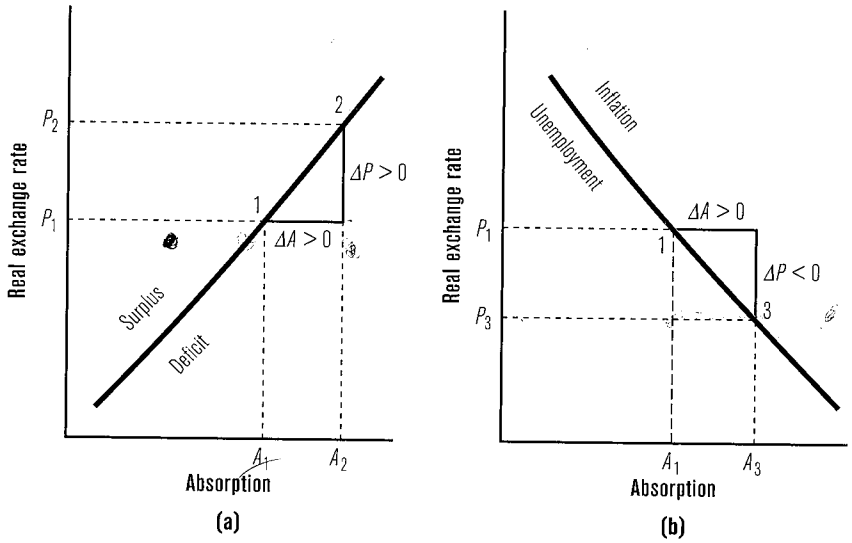


FIGURE 15-3 The Phase Diagram

(a) External balance (EB). (b) Internal balance (IB). The axes are the main policy variables: the real exchange rate, P, and the real absorption, A. The curves show equilibrium in the T-goods market (EB) and N-goods market (IB).

requires a lower, or appreciated, real exchange rate to restore equilibrium. Increased absorption raises demand for N goods, met by raising output, which can be achieved only through a lower relative price of T goods, P_3. This lower real exchange rate, or higher price of N goods, also helps regain balance by reducing the demand for N goods along the new demand curve. Point 3 is transferred to Figure 15-3b at (P_3, A_3).

Figure 15-3 also shows the **zones of imbalance**. In the T-goods market (panel a) for any given level of absorption, say, A_1, any real exchange rate greater than P_1 causes external surplus: The production of tradables exceeds the demand for tradables because the relative price, P, is at a more depreciated level than required for equilibrium. Any real exchange rate below (more appreciated than) P_1 causes an external deficit and the demand exceeds the supply of tradables. Therefore, the zone of surplus is northwest of EB and the zone of deficit is southeast.

In the N-goods market (Figure 15-3b), inflation is to the right of the IB curve, where the demand for N goods exceeds the supply. In that region, for any given real exchange rate, such as P_1, absorption is too high, say, A_3. To the left is the zone of unemployment, where there is an excess supply of N goods. In that region, for any given real exchange rate, say, P_3, absorption is too low, say, A_1.

The meanings of *inflation* and *unemployment* are precise in our model but not in the real world. It is best to think of inflation as being an increase in prices faster than is customary in the country in question. That rate would be quite low in Germany, Japan, or China, probably less than 5 percent a year, but quite high in

Brazil or Argentina. Unemployment implies not only jobless workers but also idle capital and other factors of production. In other words, there is unemployment when an economy is inside the production frontier in Figure 15–1. A country may have high levels of labor unemployment but be unable to increase output because it is fully utilizing its capital or land.

EQUILIBRIUM AND DISEQUILIBRIUM

The two balance curves are put together in Figure 15–4. All along the external balance curve, the demand for T goods equals the supply produced at home plus any net foreign capital inflow. All along the internal balance curve, the demand for N goods equals the supply of N goods. The only point at which there is both internal and external balance (equilibrium in both the T- and N-goods markets) is the intersection of the two curves. This is sometimes called the *bliss point*. It is the same as the tangency of the indifference curve to the production frontier in Figure 15–1 at point 1. The objective of macroeconomic policy is to adjust the exchange rate and absorption to keep an economy stable, in both external and internal balance.

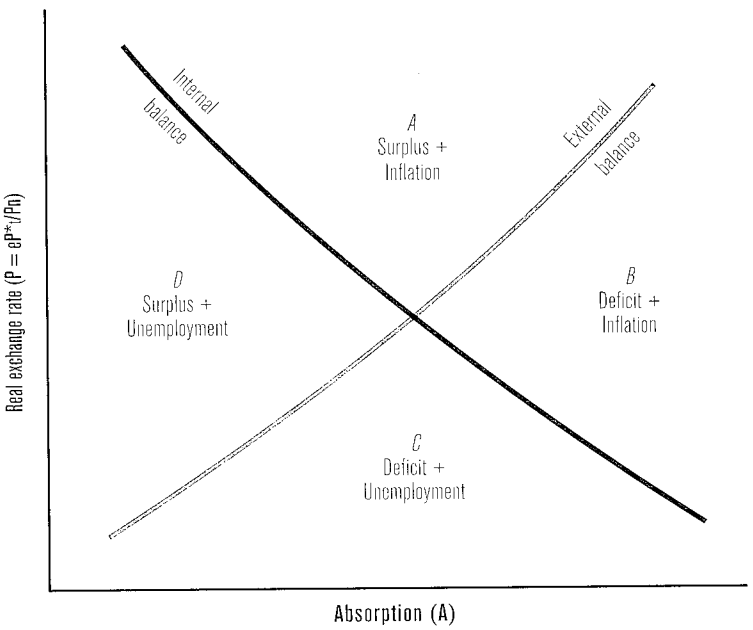


FIGURE 15-4 **Zones of Imbalance**

The economy is in equilibrium only at the intersection of the external balance (EB) and internal balance (IB) curves. Zones of imbalance are labeled. For example in zone A, the supply of T goods exceeds demand, so there is a surplus, and the demand for N goods exceeds supply, so there is inflation.

Economies spend considerable time in one of the four zones of imbalance shown in Figure 15-4. Zone *A* to the north is a region of external surplus and inflation, where the exchange rate is *undervalued*. In zone *B* to the east of equilibrium, the economy faces inflation and a foreign deficit, due principally to excessive expenditure (absorption is greater than income). To the south is zone *C*, where the exchange rate is *overvalued* (too appreciated) and there is both unemployment and an external deficit. And west of the bliss point the economy is in zone *D*, where, because of insufficient absorption, there is unemployment of all resources but a foreign surplus.

Once in disequilibrium, economies have built-in tendencies to escape back into balance. Figure 15-5 describes them separately for external balance (panel a) and internal balance (panel b). Start with an external surplus, point 1 (Figure 15-5a). The excess supply of tradables generates two self-correcting tendencies. First, the net inflow of foreign exchange adds to international reserves. If the central bank takes no countermeasures, the money supply increases and interest rates fall and induce both consumers and investors to spend more. The increase in absorption moves the economy rightward, back toward external balance. Second, the inflow of foreign exchange creates more demand for the local currency and, if the exchange rate is free to float, forces an appreciation. This is a move downward in the diagram, also toward the EB line. The net result of these two tendencies is the resultant, shown as a solid line in the diagram, heading toward external balance. If, instead, the economy starts in

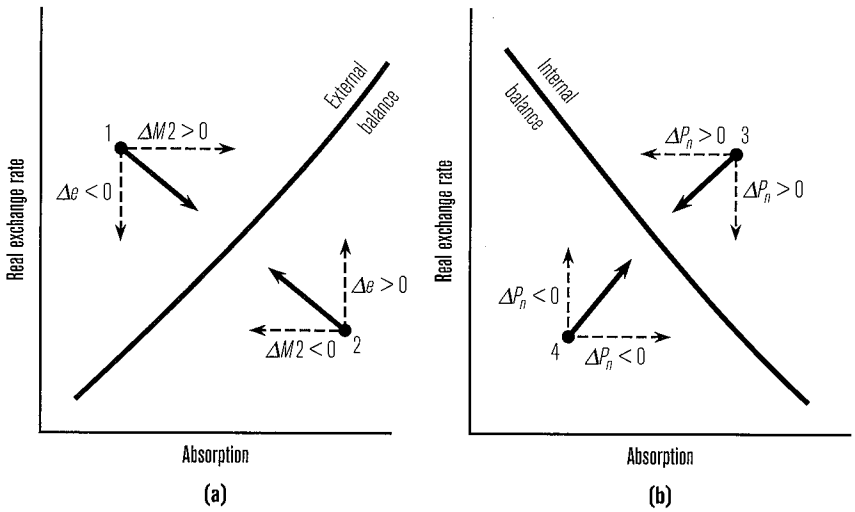


FIGURE 15-5 Tendencies toward Equilibrium

(a) External balance (EB). If the economy faces an external surplus (point 1), reserves and the money supply tend to rise while the exchange rate tends to appreciate; this drives the economy toward EB. (b) Internal balance (IB). For a deficit, if the economy faces inflation (point 3), the rise in prices leads to real appreciation of the exchange rate and a reduction in the real value of absorption; this moves conditions toward IB. Conversely for unemployment at point 4, but only if prices can fall flexibly.

external deficit at point 2, the tendencies are the opposite but the result is the same, a tendency to regain external balance.

The tendency to regain internal balance is shown in Figure 15-5b. When there is inflation (point 3), it affects both the real exchange rate and real absorption. If the nominal exchange rate remains fixed (or is not allowed to depreciate as fast as inflation), the rise in P_n causes a real appreciation. At the same time, the rise in prices can cause a fall in the real value of absorption, assuming that the central bank does not take steps to increase the money supply to compensate for inflation. Under these assumptions, the economy would move from inflation at point 3 back toward internal balance. Unemployment (point 4) would be self-correcting also if prices are able to fall as easily as they rise, but this seldom is the case.

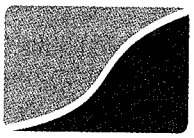
Despite these self-correcting tendencies, in practice, they often fail to work smoothly or quickly enough because of *structural rigidities* in the economy. For instance, exchange-rate changes may take time to affect actual imports and exports, perhaps as long as two years to have a full impact. In economies like Ghana and Zambia, dominated by one or two export products such as cocoa, oil, and copper, with long gestation periods for new investment, supply elasticities for tradables may be especially low, and foreign deficits can persist for a time despite real devaluations.

Nontradables prices probably rise very quickly when demand exceeds supply, as in Figure 15-5b. But in many developing economies, inflation, once started, may resist corrective policies, and prices do not fall so easily when there is unemployment: Unions strike wage bargains that try to maintain real wages by continually raising nominal wages; banks use their market power to keep interest rates high; producers depend on imports, the prices of which are responsive only to exchange rate adjustments; and large firms with monopoly or oligopoly power keep prices up to cover costs that resist downward pressures. Such rigidities have frequently been cited to explain chronic trade deficits and inflation in Latin America, especially in Argentina and Brazil.

However, arguments about structural rigidities can be overstated. There is some flexibility in production for most export industries, even in the short term. And many producer prices are quite flexible, including those of most farm products, those in the large informal sector, and even those of some modern manufacturing firms. Nevertheless, the automatic tendencies toward external and internal balance depicted in Figure 15-5 are likely to be too slow and politically painful to satisfy most governments.

Not all the barriers to adjustment are structural. Sometimes, policies work against adjustment. When foreign reserves fall, for example, the money supply also falls automatically unless the central bank's policy is to *sterilize* these shifts by expanding domestic credit to compensate for the fall in reserves and keep the money supply from falling. Sterilization prevents the move from points 1 or 2 of Figure 15-5a toward external balance. And nominal exchange rates respond to changing market conditions only if the exchange rate is allowed to float or the government makes frequent adjustments in the nominal exchange rate to match changing economic conditions.

However, the opposite policy, a fixed nominal exchange rate, is needed if inflation in nontradables prices is to cause a real exchange-rate appreciation, as depicted at point 3 of Figure 15-5b. This fixed nominal rate is called an exchange rate *anchor* because the fixed rate alone can halt the upward drift of prices as the economy moves due south from point 3. Chile used such an anchor to slow inflation during the late 1970s (Box 15-2). If government devalues the rate to keep up with inflation,



BOX 15-2 PIONEERING STABILIZATION: CHILE, 1973-84

In the last year of the Salvadore Allende regime in Chile, when the public sector deficit soared to 30 percent of the gross domestic product (GDP) and was financed mostly by printing money, inflation exceeded 500 percent a year. In 1973, General Augusto Pinochet overthrew Allende and established an autocratic regime. An early goal of his government was to stabilize the economy. It proved to be a difficult task of many years, with important lessons for later stabilizations in Latin America.

Faced by rapid inflation and unsustainable external deficits, the government imposed a fiscal and monetary shock on the economy. The budget deficit was cut to 10.6 percent of GDP in 1974 and again to 2.7 percent in 1975. Monetary policy was tight: From the second quarter of 1975 through the middle of 1976, it has since been estimated, households and firms were willing to hold more money than was in circulation. But inflation persisted; consumer prices nearly doubled in 1977.

Despite draconian measures, prices continued to rise for two reasons. First, the peso was aggressively devalued to improve the foreign balance, the more so because of the 40 percent fall in copper prices in 1975. In 1977, the peso was worth about one-80th its 1973 value against the dollar. Second, wages in the formal sector were determined by rules that permitted adjustments based on the previous year's rate of inflation, a rule that helped perpetuate the higher rates of earlier years. It also was argued by some that the monetary policy was not stringent enough.

In 1978, the government switched gears and began using the exchange rate as its main anti-inflation weapon. At first a crawling peg was adopted with preannounced rates, the *tablita*, that did not fully adjust to domestic inflation. In 1979, the rate was fixed at 39 pesos to the dollar for three years. The appreciating real exchange rate, or *anchor*, helped control inflation, which was down to 10 percent by 1982. But it also discouraged export growth and contributed to a growing

current-account deficit. At the same time, Chile liberalized its controls over foreign capital flows and attracted large inflows of loans: Net long-term capital rose from negligible amounts before 1978 to average over $2 billion a year in the following five years, equivalent to 8 percent of GDP in 1980. This inflow not only financed the growing current deficit but contributed to the real appreciation of the exchange rate.

Not until after 1984 did Chile finally achieve a semblance of both internal and external balance. It did so through a large real devaluation, approaching 50 percent, supported by tighter fiscal and monetary policies. After a decade and a half of falling income per capita, Chilean incomes grew by 5.8 percent a year from 1985 to 1991.

Source: Based on the account by Vittorio Corbo and Andrés Solimano, "Chile's Experience with Stabilization Revisited," in Michael Bruno et al., eds., *Lessons of Economic Stabilization and Its Aftermath* (Cambridge, MA: MIT Press, 1991).

Brazil's practice for many years, then real appreciation is thwarted and there is no anchor. Similarly, real absorption falls with inflation only if the government fixes its expenditure and its deficit in nominal terms and allows inflation to erode the real value of the expenditure and if the central bank restrains the money supply to grow more slowly than inflation. More typically, the fiscal authorities adjust the expenditure, while the monetary authorities adjust both the money supply and the nominal exchange rate, to fully compensate for inflation. In that case, rising prices have no impact on the real exchange rate or real absorption and an inflationary economy remains at point 3 in Figure 15-5b.

STABILIZATION POLICIES

Whether the barriers to rapid automatic adjustment are inherent in the economic structure or created by policy contradictions, in most cases, governments need to take an active role to stabilize their economies. They have three basic instruments for doing so: exchange-rate management, fiscal policy, and monetary policy.

Alternative **exchange-rate regimes** were introduced in Chapter 12. Governments can vary the exchange rate by having the central bank offer to buy and sell foreign currency at a predetermined or *fixed* official exchange rate (e in our nomenclature) that nevertheless can be changed from time to time or by allowing the rate to *float* in the currency market, although the central bank sometimes may intervene to influence the price. An intermediate case is the *crawling peg*, under which the central

bank determines the rate but changes it frequently, as often as daily, to ensure that the official rate stays in line with domestic and world inflation; this results in a constant or slowly adjusting real exchange rate (P).

Governments have two policies that can influence the level of absorption. Fiscal policy, adjusting levels of government expenditure and taxation, directly affects the government's components of consumption and investment. It also influences private expenditure, especially consumption, which depends on *disposable income*, or income net of taxes. Monetary policy also affects private expenditure. If the central bank acts to increase the money supply, as described in Chapter 12, it increases the liquidity of households and firms, lowers interest rates, and stimulates private consumption and investment.

The power of the phase diagram is that it indicates the necessary directions for these policies, depending on the state of the economy. Figure 15-6 provides such a policy map. It shows the same external and internal balance lines as in the previous diagrams but adds a new element: four policy quadrants, I to IV, within which the policy prescription always is the same.

Take, for example, point 1, which has been placed on the external balance line but in the inflationary zone. For many years, Brazil was in this situation, with buoyant exports and balance in foreign payments but chronic inflation running from 40 to well over 100 percent a year. Because the demand for nontradables exceeds supply, we know that one necessary correction is a reduction in real absorption, monetary and fiscal *austerity*, that would reduce demand and move the economy due west from point 1. But, if that is the only policy taken, the economy would not reach internal balance until point 4, in the zone of external surplus. One imbalance is exchanged for another. To avoid generating a surplus, reduced absorption needs to be accompanied by an appreciation of the exchange rate, a move due south from point 1. The result would be a move approximately toward the equilibrium or bliss point, 0.

Note three things about this result. First, this combination of policies, austerity and appreciation, would work from any point within quadrant I to return the economy to equilibrium. That is, the same combination is needed whether the economy had inflation with a moderate external surplus or inflation with a moderate deficit, either just above or just below the EB line. If the economy starts just below external balance, with a moderate deficit, it may seem strange (*counterintuitive*) to recommend an appreciation that, on its own, would worsen the deficit. But the reduction in absorption, needed to reduce inflation, also reduces the deficit because it also lowers the demand for tradables. Indeed, it reduces the demand for tradables too much and throws the economy into surplus; this is the reason an appreciation is needed. Of course, the relative intensity of each policy is different, depending where in quadrant I the economy starts. But the basic principle holds: Anywhere in quadrant I, the right combination of policies is austerity and appreciation, the combination that moves the economy toward point 0.

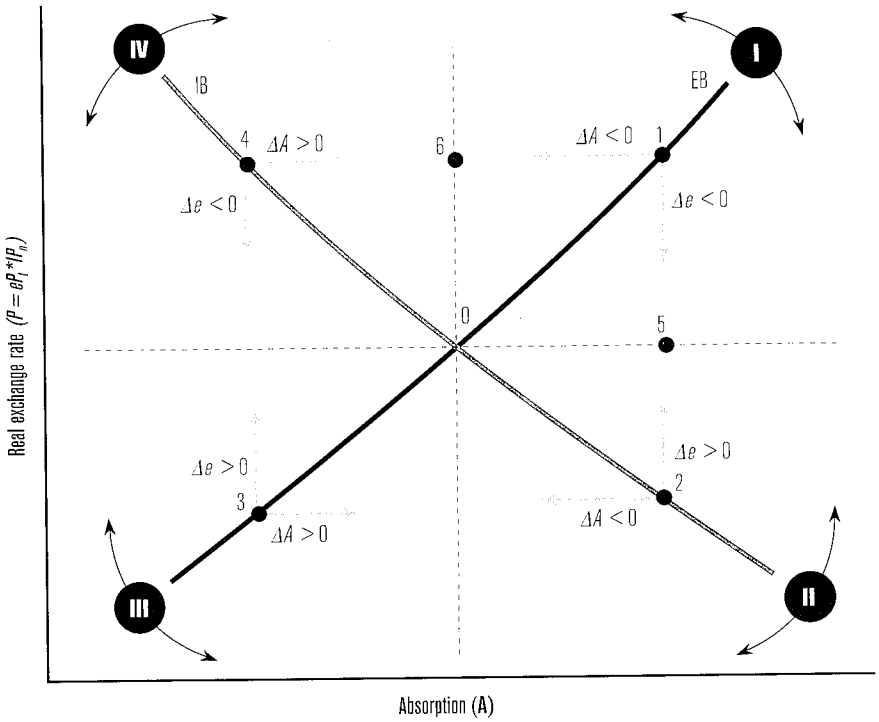


FIGURE 15-6 **Policy Zones**

From any position of disequilibrium, two policy adjustments generally are needed to restore internal and external balance. In each quadrant (I–IV) a particular combination of exchange rate and absorption policy is prescribed.

Second, in general, two policy adjustments are required to move toward equilibrium. This is a simple example of the general rule enunciated by Dutch economist Jan Tinbergen: To achieve a given number of policy goals, it generally is necessary to employ the same number of policy instruments. Here we have two goals, internal and external balance, and need adjustments in both absorption (austerity) and the real exchange rate (appreciation) to reach them both. It is not always necessary to use two goals, however. If the economy lies just to the east of equilibrium at point 5, then a reduction in absorption achieves internal and external balance simultaneously. And, if the initial situation is point 6, due north of 0, then appreciation alone does the job.

Third, we could view the policy prescription in either of two ways. Austerity is needed to reduce inflation (move west) and appreciation is used to avoid surplus (move south). Or appreciation can be targeted on internal balance (move south toward point 2) but alone would cause a deficit, so that austerity then is required to restore external balance. Therefore, no logic in macroeconomics suggests that one particular policy should be assigned to one particular goal. Economic institutions often do this anyway. In practice, the central bank might use the exchange rate to achieve external

balance while the finance ministry uses the budget for internal balance. But if these two approaches are not coordinated, they may well fail to reach equilibrium.

With these principles established for quadrant I, it is fairly routine to go around the map in Figure 15–6 and see what policy responses are required:

- In quadrant II at a point like 2, with an external deficit but internal balance, exchange-rate devaluation is needed to restore foreign balance but, taken alone, would push the economy into inflation. Fiscal and monetary austerity also are needed to avoid inflation and reach equilibrium. We could reverse this assignment of policies and use austerity to achieve external balance and devaluation to stimulate the economy. Many African countries have been in this situation right up to the present, with low inflation but an insufficiency of export earnings and foreign investment to pay for the imports required for economic development.
- In quadrant III at point 3, an expansionary fiscal or monetary policy eliminates unemployment but at the cost of a foreign deficit, so devaluation is needed to reach equilibrium. Or devaluation stimulates employment and so requires expansion to eliminate the resulting surplus. This is the situation of a mature industrialized economy during a recession, with unemployed labor and capital, but it is not so common in developing countries.
- In quadrant IV at point 4, exchange-rate appreciation can eliminate the external surplus while fiscal expansion prevents unemployment. Or fiscal expansion can end the surplus while appreciation prevents a resulting inflation. A few countries in Asia, such as Taiwan and Malaysia in the 1980s, have been in this situation.

So the principles of macroeconomic stabilization are simple: If policy makers know where to place their economy on this map, they know how to move toward equilibrium. But how do policy makers know where they are? The answer lies partly in measurement, partly in art. Regularly available data on the balance of payments, changes in reserves, and inflation can help locate an economy with respect to the external and internal balance lines. Data on the nominal and real exchange rates, the budget deficit, and the money supply can indicate movements from one policy quadrant to another. Some kinds of data, such as private sector short-term borrowing abroad, however, may not be readily available to policy makers. Such was the case in Korea at the beginning of that country's financial crisis in 1997. In principle, barring such surprises as an unknown large short-term foreign debt that has to be repaid immediately, econometric models can locate the economy and indicate the policies needed to balance it. In practice, especially but not only for developing economies, such models can be too imprecise and too unstable to be wholly dependable. The art of stabilization policy comes in knowing just how hard to push on each component of policy and how long to keep pushing. In this, experience in managing a particular economy is as important a guide as the models estimated by economists.

APPLICATIONS OF THE AUSTRALIAN MODEL

Throughout this book we refer to different kinds of economic problems that are associated with developing countries, including the Dutch disease, debt crises, terms-of-trade shocks, foreign-exchange shortages, destructive inflation, and droughts or other natural catastrophes. The Australian model and its phase diagram can be used to show how these and other shocks affect macroeconomic balance and how they should be handled.

DUTCH DISEASE

In Chapter 18 we will discuss the strange phenomenon of the Dutch disease, in which a country that receives higher export prices or a larger inflow of foreign capital may end up worse off than without the windfall. The Dutch disease was first analyzed by Australian economists Max Corden and Peter Neary, using a version of the open-economy model.[6] Figure 15-7 traces the impact of a windfall gain using the phase diagram. (Box 18-1 provides an alternative exposition of Dutch disease.)

An economy in equilibrium at point 1 suddenly begins to receive higher prices for its major export or is favored by foreign aid donors or foreign investors. All the oil producers, from Saudi Arabia to Indonesia to Mexico, were in this position in the 1970s, as were coffee (and many other commodities) exporters during the boom of the mid-1970s. Egypt and Israel were rewarded with large aid programs by the United States after the Camp David accord of 1978, as was Ghana by the World Bank and others during its stabilization of the 1980s (Box 15-3). Both Chile in the late 1970s and Mexico after its stabilization in the late 1980s received large inflows of private capital, much of it a return of previous flight capital. Foreign exchange windfalls are more frequent than sometimes is supposed. In some cases, these windfalls result from new discoveries of natural resources, such as the major offshore oil reserves discovered by Ghana in 2007.

When the windfall occurs, the supply of tradable goods rises at any given price. This can be shown as a rightward shift in the supply curve in Figure 15-2a. In the phase diagram of Figure 15-7, there is a rightward shift in the EB curve. At point 1, for example, which had been in external equilibrium along EB_1, the economy now is in surplus, so the new EB curve must be to the right—for example, at EB_2. The economy cannot remain at point 1 because the inflow of reserves increases the money supply; this adds to demand and, because the windfall increases private income and government revenue, leads to greater expenditure. So absorption rises, a move

[6] W. Max Corden and J. Peter Neary, "Booming Sector and Deindustrialisation in a Small Open Economy," *Economic Journal* 92 (1982), 825–48.

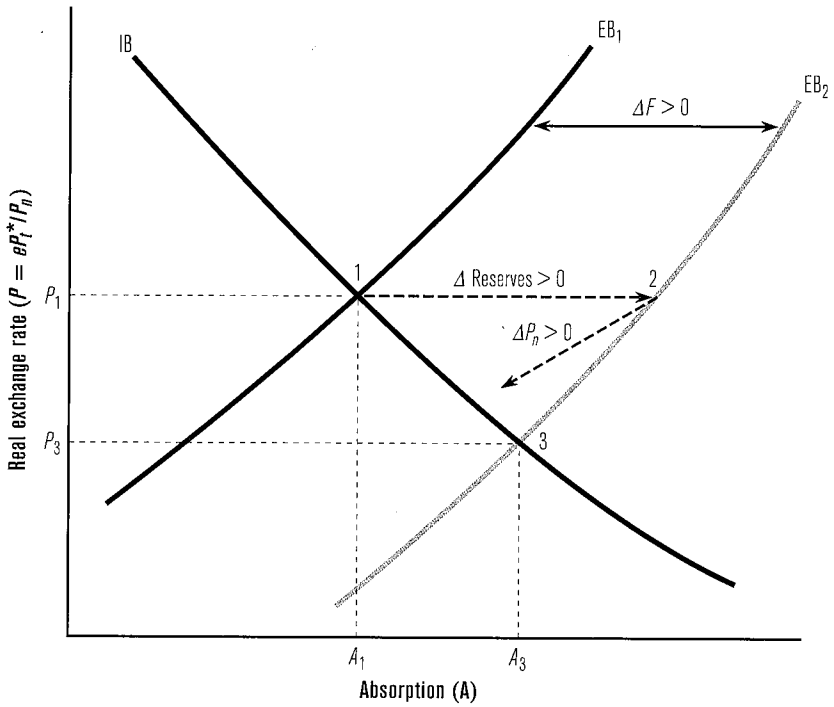


FIGURE 15-7 The Dutch Disease

An export boom or capital inflow shifts the EB curve rightward and leaves the economy at point 1 in surplus. As reserves accumulate and the money supply rises (or as the government and consumers spend the windfall), absorption rises and the economy moves eastward, into inflation. As nontradable prices rise, the real exchange appreciates. At the new equilibrium (point 3), because P is lower, the supply and demand is balanced with less production of T goods and more output of N goods than before. The loss of tradable output is what makes this a disease.

from point 1 toward point 2. This moves the economy off its internal balance, into inflation.[7]

The resulting rise in P_n has two effects: a reduction in real absorption that partially corrects the initial rise in A and, assuming the official rate is fixed, a real appreciation of the exchange rate. (The real rate also appreciates if the nominal rate is floating because the greater supply of foreign currency drives down the price of foreign currency.) Therefore, the economy first moves from point 1 toward point 2 in Figure 15-7, then begins to head in the general direction of the new equilibrium, point 3. In this case, market forces are likely to be sufficient to reach the new equilibrium, unless the

[7]If the windfall is an inflow of capital, this treatment is precise. In the case of a rise in export prices, however, the move from point 1 to point 2 is an approximation. Strictly speaking, a rise in export prices should raise P_t^*, a depreciation of the real exchange rate that moves the economy upward from point 1, after which the economy moves east toward EB_2.

BOX 15-3 RECOVERING FROM MISMANAGEMENT: GHANA, 1983-91

In 1983, after a decade of economic mismanagement, Ghana's gross domestic product (GDP) was 20 percent below its 1974 peak, investment was only 4 percent of GDP, exports had sunk to 6 percent of GDP, and inflation rocketed to 120 percent for the year. After a decade of economic decline, Ghana's military government, headed by Flight Lieutenant Jerry Rawlings, was ready to undertake drastic measures to stabilize the economy and restart economic development.

Working closely with the International Monetary Fund (IMF), Ghana focused on three deep-seated problems: exchange-rate reform, fiscal adjustment, and monetary policy. At first, the government maintained its fixed exchange rate but drastically devalued the cedi from 2.75 to the dollar in 1983 to 90 to the dollar by 1986. In 1986, Ghana adopted a restricted floating currency, using periodic auctions to determine the rate. The official exchange market was broadened in 1988, when many foreign exchange bureaus were authorized to trade currencies and virtually absorbed the parallel market in currency; by 1990, the banks were empowered to trade in an interbank currency market. This completed the move to a floating rate regime. By the end of 1992, the cedi traded at 520 per dollar.

In 1983, with fiscal revenues less than 6 percent of GDP, the urgent need was to restore revenues and control expenditures. The deficit was cut from 6.2 to 2.7 percent of GDP in the first year of austerity, and by 1985, the government had begun a major public investment program to stimulate growth. By 1988, the government had restored total expenditures to 15 percent of GDP, 20 percent of which was investment, and was running a surplus of nearly 4 percent of GDP.

Throughout the period, the money supply was constrained but inflation remained stubbornly above 20 percent a year until 1991, when it was reduced to 16 percent and real interest rates finally became positive. Because food prices play a large role in the consumer price index, investment in food production was seen as an important component of any long-run attack on inflation.

The aid donors responded handsomely to Ghana's stabilization and the accompanying economic reforms: The sum of net official transfers and net long-term capital rose from just over $100 million in 1983 to $585 million in 1991.

Stabilization helped restore economic growth. From the depression of 1983–91, GDP grew by 5.1 percent a year and investment rose to 17 percent of GDP. The improvement, although dramatic in relation to the early 1980s, still left Ghana with a lot to be done: In 1991, income per capita remained 25 percent below its 1973 level.

Source: This account is based on Ishan Kapur et al., *Ghana: Adjustment and Growth, 1983–91* (Washington, DC: International Monetary Fund, 1991).

authorities prevent appreciation and maintain real absorption and so keep the economy in an inflationary posture like point 2.

What, then, is the problem? The economy is at a new equilibrium, its terms of trade improved, its currency appreciated and so citizens have more command over foreign resources, people spending and consuming more without having to work any harder. There are two flaws in this otherwise idyllic picture. First, such windfalls generally are temporary. When export prices fall or the capital inflow dries up, the EB curve shifts back and a costly adjustment is necessary. We analyze that process in the next section.

The second problem is that, in shifting from the old to the new equilibrium, adjustments in the economy must be made. The real exchange rate P is lower, so S_t has fallen, while S_n has risen. Because the booming export sector does not retrench, nonboom tradables bear the brunt of the adjustment. Frictions in the labor market are likely to mean at least temporary unemployment as workers switch from tradable to nontradable production. If the tradable sector includes modern manufacturing, then long-term development may be set back because manufacturing is the sector likely to yield the most rapid productivity growth in the future. And if tradable industries close, it is more difficult to make the inevitable adjustment back toward point 1 when the windfall is over. This decline in nonboom-tradable production turns a foreign exchange windfall into a "disease."

What can be done to cure the disease? The government could try to move the economy back toward the old (and probably future) equilibrium at point 1. Its tools are the official exchange rate, which would have to be devalued against the tendencies of market forces, and expenditure, which would have to be reduced through restrictive fiscal and monetary policies that also reduce inflation (lower P_n or at least its growth). The resulting buildup of reserves and bank balances have to be sterilized through monetary policy so they are held as assets and not spent. It is a neat political trick to manage an austere macroeconomic policy in the face of a boom because all the popular pressures are for more spending. Not too many countries have managed it. Indonesia is among the few that have.

DEBT REPAYMENT CRISIS

When Mexico announced in 1982 that it no longer could service the debt it acquired during the oil boom of the 1970s, many other developing countries followed Mexico's lead, and the financial world entered a decade of debt crisis (Chapter 13). Most Latin American countries largely have overcome their debt problems, but many African countries continue to struggle to repay the money they borrowed, mostly from aid agencies. Although debt service insolvency encroaches gradually on an economy and can be foreseen, it often appears as a national crisis because economic management has been inept.

The formal analysis of a debt crisis is similar to that of another common phenomenon, a **decline in the terms of trade** that leads to a foreign exchange shortage, which in turn is simply the reverse of the Dutch disease. Therefore, the oil exporters, such

as Indonesia, Nigeria, and Venezuela, faced a similar kind of crisis once oil prices began falling in the 1980s. We can understand the similarity between a debt crisis and a decline in the terms of trade more clearly by seeing them in their common context in the balance of payments. A decline in the terms of trade implies deterioration in the balance of trade (in which the excess of imports over exports increases). All else equal, a declining trade balance adds directly to the current account deficit. As detailed in the appendix to this chapter, one of the few ways in which countries can finance current account deficits is by borrowing abroad.[8] Indeed, many developing countries financed chronic current account deficits by borrowing abroad, in the process accumulating enormous stocks of debt (often to levels greater than their GDP). Debt crises ensue (as discussed in Chapter 13) when current account deficits become unsustainable and lenders want to be repaid.

Figure 15–8 captures this process. An economy in balance at point 1 needs to find additional resources to repay its foreign debt or needs to adjust to falling terms of trade. The supply of tradables therefore shifts to the left in Figure 15–2a; in the phase diagram, the EB curve also shifts leftward to EB_2.[9] If the crisis leads to debt relief or additional foreign aid, the curve moves less far and might settle at EB_3.

Now in foreign deficit, the economy begins losing reserves. If the government has to repay some of the debt or falling export prices cut into its revenues, the government needs to reduce its expenditures as well. Both cause a reduction in absorption. These actions move the economy toward external balance but also into unemployment. To gain the new equilibrium at point 3, it is also necessary to devalue the currency. This could be done by the central bank under a fixed rate or by the foreign exchange market under a floating rate. At the new equilibrium, the country produces more and consumes fewer tradables because P has risen. This, of course, is a loss of welfare for the populace. The surplus of S_t over D_t is used to repay the debt or simply compensates for reduced export prices.

Debt crises and the hardships they cause are not an inevitable consequence of borrowing to finance development, as was discussed in Chapter 13. If the borrowed resources are invested productively, they increase the potential output of both tradables and nontradables. Added production increases income and generates the capacity to repay the debt out of additional income, without a crisis and an austerity program. Countries such as Korea and Indonesia have been large international borrowers, but before the financial crisis of 1997–99, they escaped debt crises.

[8]The only other ways (beyond foreign borrowing) to finance a current account deficit are to attract foreign investment and/or to run down the central bank's stock of foreign reserves. See the appendix to this chapter for a concise summary of balance of payments accounting.

[9]Strictly speaking, we cannot analyze the fall in export prices this way, but it is a reasonable approximation for many situations. See note 7.

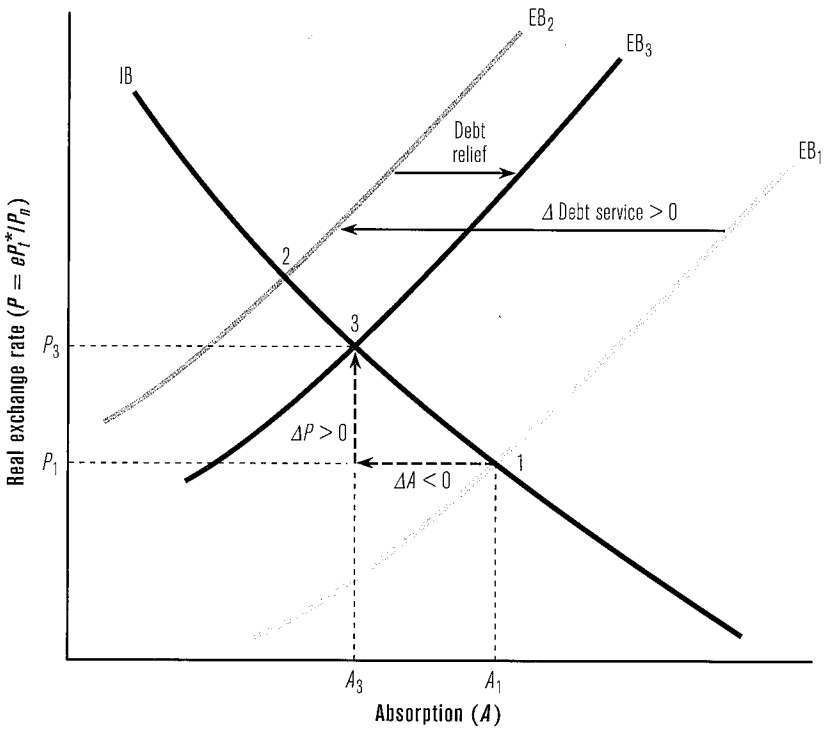


FIGURE 15-8 Debt Crisis or Declining Terms of Trade

An economy in equilibrium at point 1 suddenly needs to repay its debt (or faces falling export prices). External balance shifts from EB_1 to EB_2, although debt relief or increased foreign assistance might reshift the balance line back to EB_3. If policies accommodate the fall in reserves and income, absorption declines. A devaluing exchange rate, via central bank action or market forces, helps the economy move to its new equilibrium at point 3. With more tradables produced and less consumed, the surpluses can be used to repay the debt.

STABILIZATION PACKAGE: INFLATION AND A DEFICIT

External shock is not the only way an economy gets into trouble. Reckless or misguided government policies often are to blame. Impatient with sluggish development or intent on benefiting its constituencies, a government expands its spending and incurs a budget deficit. Unable to finance the deficit by borrowing from the public, the ministry of finance sells short-term bills to the central bank; this adds to the money supply. The economy drifts into inflation and a foreign deficit, at a point like point 1 in Figure 15–9, far from equilibrium at point 2 on the economy's original external balance curve EB_1. When economies become unstable in this way, private investors get skittish and try to invest in nonproductive assets like land or, more often, invest abroad; this deepens the external deficit. The government, recognizing the error of its ways or just hoping for some outside help to avoid painful adjustment, calls in the IMF.

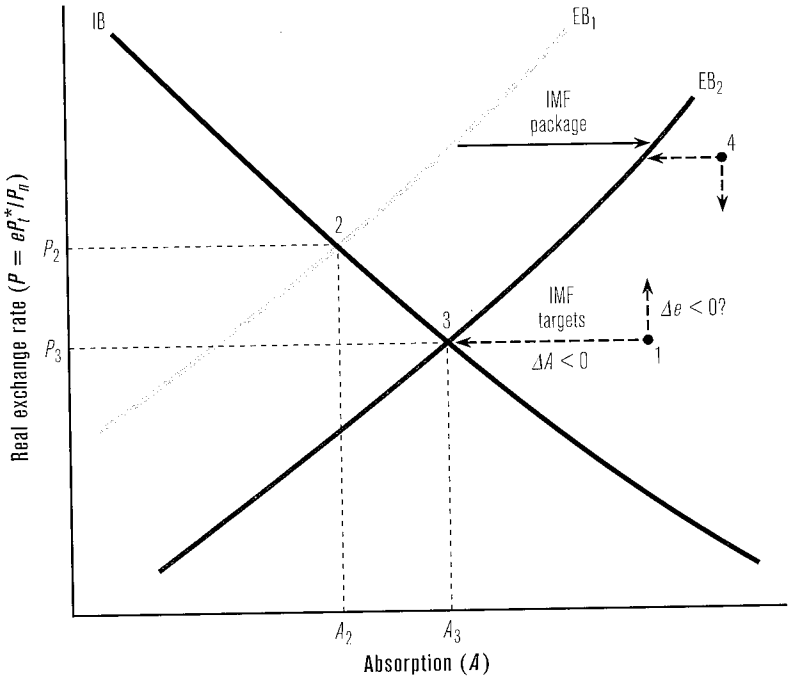


FIGURE 15-9 Stabilization from Inflation and a Deficit

An economy at point 1, far from equilibrium at point 2, above all needs to reduce absorption through austerity: reduced budgetary deficits and slower growth of the money supply. An International Money Fund (IMF) and donor package of aid might bring equilibrium closer by shifting the external balance to EB_2, but the aid package is conditional on the austerity program. Whether any exchange rate action is required depends on the precise initial position, point 1.

The core IMF stabilization program consists of a reduction in the government's budget deficit and programmed targets for domestic credit that, in effect, cap the growth of the money supply. Together, these measures reduce absorption in the economy and move it westward from point 1, closer to external and internal balance. IMF packages frequently include an exchange-rate devaluation as well. Whether this is needed or not depends on the precise location of the economy (point 1) relative to equilibrium (point 2). In some cases, the reduction in absorption is sufficient to reach both internal and external balance. As pictured in Figure 15–9, a small devaluation is needed to reach point 2 and avoid unemployment.

However, IMF programs usually come with substantial aid attached, not only from the fund, but from the World Bank and bilateral donors. The aid package, by adding to the economy's capacity to buy tradables, shifts the EB curve to the right, to EB_2 in the diagram, and moves equilibrium to point 3. Note two things about this aid package. First, it reduces the need for austerity to some extent, as A_3 is greater than A_2. Second, it reduces the need for devaluation of the exchange rate. Indeed, as shown,

there is little or no need to devalue to move from 1 to 3. Donors and the IMF nevertheless frequently insist on devaluation. Sometimes, that may be a requirement just to reach a point like 3. In other cases, donors and the IMF may have in mind a self-sustaining stabilization that will be valid even after aid is reduced and the external balance curve moves back toward EB_1. Whatever the motive, it is important to realize that aid itself is a partial substitute for both devaluation and austerity. In essence, the aid does what higher production of tradables otherwise must do and it finances expenditures that otherwise must be cut. Ghana's experience, which fits this description, is discussed in Box 15-3. More recently, Greece has suffered a major debt crisis, the analysis of which draws on elements of both Figures 15-8 and 15-9 (Box 15-4).

BOX 15-4 THE GREEK DEBT CRISIS OF 2010-12

Problems of unsustainable debt are not limited to the poorest countries. Starting in 2010, Greece faced a debt crisis so severe as to threaten the stability of the European Union (EU). As of September 2011, it remained uncertain whether Greece would be forced to default on its foreign debts or whether other members of the EU would provide its second bailout package for Greece in two years. In May 2010, the EU and the International Monetary Fund (IMF) provided a package of loans and balance of payments support worth €750 billion (approximately $938 billion). In return, the government of Greece committed itself to an austerity program that included severe reductions in expenditures and wages, termination of tens of thousands of government jobs, and extensive tax increases. This agreement sparked widespread civil unrest in Greece, which contributed to a lack of confidence in the country's ability to implement the promised reforms and service its foreign debts.

Greece's debt problem accumulated over many years of current account deficits, the magnitude of which ballooned during the late 2000s. Greece's current account deficits as a share of GDP began mounting in the mid-1990s. Deficits on the order of 3 percent of GDP in the late 1990s more than doubled relative to GDP during 2000–05. By 2005, Greece's current account deficit was equivalent to 7.5 percent of GDP, a proportion that doubled again to nearly 15 percent by 2008. Between 2008 and 2010, Greece succeeded in reducing its current account deficit to 10.5 percent of GDP, mainly by reducing imports. Borrowing by the Greek government to finance its deficits had also grown rapidly, as the government budget deficit as a share of GDP increased from 5.7 percent in 2006 to 15.4 percent in 2009. The government's total debt shot up from €183.2 billion in 2004 (equivalent to 99 percent of GDP, or approximately $228 billion)

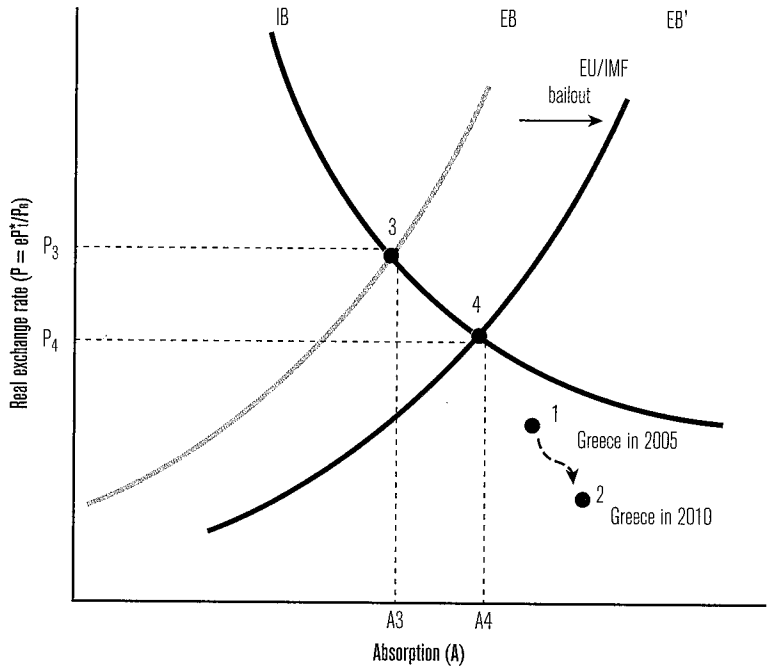
to €330.4 billion in 2010 (equivalent to 144 percent of GDP, or approximately $438 billion). The country's general economic stagnation was further reflected in its unemployment rate, which after fluctuating around 10 percent for a decade, fell to 7.5 percent in May 2008, only to double by March 2011. Greece's GDP shrank by 6.6 percent in 2010.

By the fall of 2011, the government of Greece found itself between a rock and a hard place, as the EU debt relief package was imperiled by a threatened cutoff in response to Greece's apparent inability to meet its austerity commitments. Yet daily strikes and mounting political pressure within Greece were preventing the government (led by a socialist party, with a slim parliamentary majority) from fully implementing the promised cuts. A disorderly default was a distinct possibility, along with an exit by Greece from the euro zone.

The Australian model lends itself well to a depiction of the Greek debt crisis. The analysis shown in the figure combines elements of Figures 15–8 and 15–9. By 2005, Greece was suffering from both a balance of payments deficit and unemployment, conditions indicated by point 1. By 2010, both unemployment and the balance of payments deficit had substantially worsened, moving Greece toward point 2 (farther from both internal and external balance and far from equilibrium at point 3). It seems straightforward, based on the phase diagram, that Greece needed to reduce absorption and depreciate its real exchange rate, but there's a catch. As a member of the EU, Greece was a member of the euro zone and was thus unable to devalue its currency as a means of reducing its balance of payments deficit. The only way for Greece to induce the necessary real depreciation was to reduce P_n sufficiently. In short, austerity presented itself as virtually the only tool available to the Greek government. Pushing down wages—a central element of the Government's austerity program—would reduce P_n (because labor can be counted as a key nontradable) and help depreciate the real exchange rate. Austerity was also critical to restoring external balance because borrowing to finance the current account deficit was no longer an option for Greece.

The bailout package offered by EU member states in concert with the IMF would effectively shift the EB curve out to *EB'*, thus reducing the amount of both real depreciation and spending cuts required to reach the new equilibrium at point 4. Herein lay the impasse encountered in the fall of 2011. Greece was widely seen as incapable of making the cuts necessary to reach equilibrium at point 3. Without the bailout package, default seemed unavoidable. Yet even the degree of austerity required to reach the postbailout equilibrium at point 4 was becoming politically impossible for the Greek government. Absent a credible commitment by the Greek government to impose the necessary spending and public employment cuts (along with substantial tax increases), the EU was

unwilling to provide the bailout. The brinksmanship between Greece's government and its debtors continued into early 2012. Greece's next round of debt repayments were due in on March 20, 2012, and, absent a new bailout agreement, default was inevitable. In late February, just weeks before the prospective default, EU and Greek negotiators agreed on a second bailout program worth €130 billion ($169 billion) in return for renewed promises of severe austerity and greater EU influence over Greek budgets. The deal was intended to keep the Greek government solvent until 2014. Yet, Greece's continuing deep recession and political uncertainty suggested that the crisis was far from over.



Another kind of stabilization also can be illustrated with Figure 15-9, **rapid inflation** (or hyperinflation). In Bolivia's hyperinflation of the mid-1980s or the chronic inflations in the past in Brazil and Argentina, external balance is a secondary consideration or not a major problem. Point 4 in the diagram depicts this situation. Austerity still is required to move toward equilibrium at point 2 or 3 (if there is an aid package), but devaluation only intensifies inflation. Instead, the currency must be appreciated, which also dampens inflation. One way to achieve this would be to fix the nominal rate and let the continuing (if decreasing) inflation in nontradable prices (P_n) appreciate the real rate P. This is the *exchange-rate anchor,* a device used often in Latin America, especially in Chile during the late 1970s, in Bolivia during the

mid-1980s, and in Argentina during the 1990s. It has the disadvantage that a lower real rate discourages export growth. Yet investment in new exports may be part of a strategy to open the economy, diversify exports, and move the external balance curve to the right.

DROUGHT, HURRICANES, AND EARTHQUAKES

The human tragedy of drought, earthquakes, and other natural disasters in places such as Ethiopia, the West African Sahel, and Haiti in 2010 dwarfs issues of macroeconomic management. But the adept management of an economy racked by natural disaster is essential to reduce the misery of starving or displaced people. Drought, for example, reduces a country's capacity to produce food, export crops, and in some countries, generate electricity from hydropower. At the same time, income is lower because farmers and others have less product to sell. Government then needs to provide social safety nets; this means spending more on the provision of food, transportation, health services, and sometimes shelter. Foreign governments often provide financial, food, and technical aid under these situations.

The macroeconomic reflection of a drought or earthquake is depicted in Figure 15–10. The economy begins in equilibrium at point 1. Drought or earthquake reduces the economy's capacity to produce both nontradables (some foods, hydroelectricity, water supplies) and tradables (export crops, importable foods, some manufactures). We show this as a leftward shift in both the IB and EB lines: Reduced output of S_n at any given price means a larger zone in which D_n exceeds S_n; this is inflationary. Similarly for S_t, and this enlarges the area of deficit. The new external balance curve, EB_2, may be augmented (shifted back to the right) by foreign aid to EB_3, in which case, the new equilibrium is point 3.

The economy, still at point 1, is inflationary. The fall in incomes creates a tendency for absorption to shrink on its own and move the economy leftward toward the new equilibrium. At the same time, the government tries to spend more to relieve hunger, disease, and other problems. The outcome depends on the relative force of these tendencies. The impact of most natural disasters is temporary, typically lasting a year, although some African droughts have been much longer, and the Haiti earthquake of 2010 is likely to last for several years or more. It is appropriate to try to ride out such shocks with minimal adjustment, especially if foreign aid can bear much of the burden. Therefore, for example, even if an exchange-rate adjustment is called for to reach equilibrium, it is unlikely to work very well during the natural disaster and probably should be resisted. This could be said for fiscal austerity too, except that the rise in prices can deepen the suffering of those already hurt by the disaster. If the government is able to shift its expenditures so that a greater portion goes into alleviating the impact of the disaster, it may be able to relieve the worst suffering while restricting the rise of total expenditures and containing inflation.

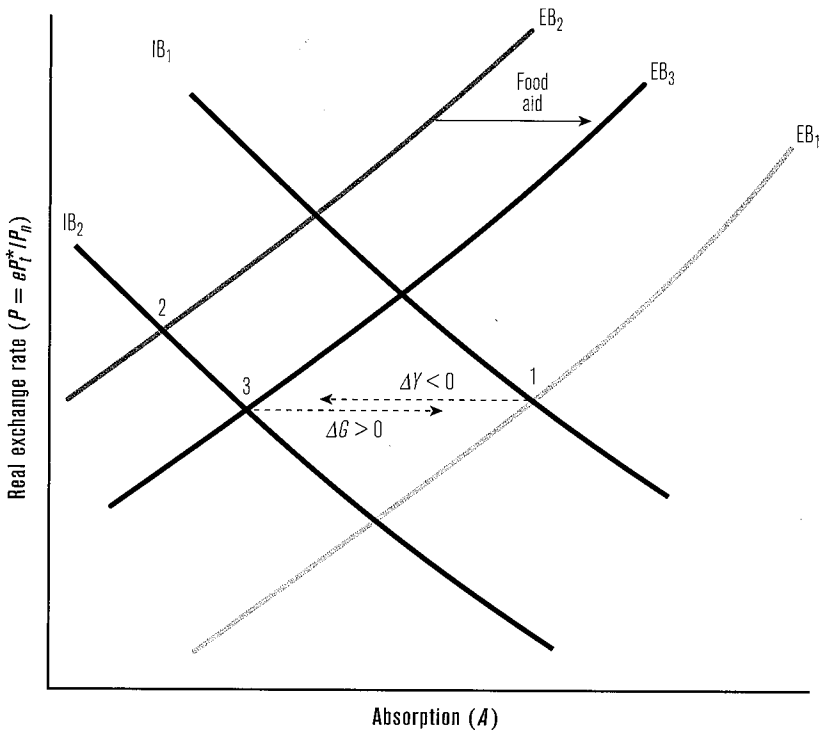


FIGURE 15-10 Drought

Drought or another natural disaster reduces the capacity to produce both nontradables and tradables, so the curves shift to the west. Disaster relief from abroad augments the external balance curve and shifts it to EB_3, with equilibrium at point 3. Remaining temporarily at point 1, the economy becomes inflationary. The reduction in output and therefore in income reduces absorption, but the government's need to spend more on relief tends to offset this move toward equilibrium. The outcome could be continued inflation.

SUMMARY

- Open economies may be exposed to a variety of shocks, both positive and negative, either of which may require policy responses to balance potential negative effects. In many cases, we have seen that a country's vulnerability to external shocks may be increased by its own history of poor policy choices.
- One common problem among developing countries arises from extended periods of spending more than they earn and financing the resulting current account deficits with foreign borrowing. For many developing countries, these persistent current account deficits become unsustainable, typically requiring both spending reductions and currency devaluation.

- The Australian model provides a convenient and versatile framework within which to diagnose common macroeconomic imbalances. Equilibrium in the Australian model requires balance in the markets for both tradable and nontradable commodities (the relative price of which is one expression of the real exchange rate). Equilibrium in the tradables markets is known as external balance; equilibrium in the nontradables market is known as internal balance.
- Disequilibrium in the tradables market implies either a balance of payments deficit (in the case of excess demand) or a balance of payments surplus (in the case of excess supply). Disequilibrium in the nontradables market implies either a higher rate of inflation (in the case of excess demand) or higher unemployment (in the case of excess supply).
- Considered together, these disequilibrium conditions may arise in four combinations: balance of payments deficit with inflation, balance of payments deficit with unemployment, balance of payments surplus with inflation, and balance of payments surplus with unemployment.
- The main policy tools available to regain macroeconomic equilibrium are fiscal and monetary policy interventions to change the level of domestic spending (absorption) and changes in nominal exchange rate policy aimed at altering the real exchange rate.
- Recommendations to use some combination of these policy tools to regain equilibrium must take account of the economy's natural tendencies toward equilibrium. The functionality of these natural tendencies will vary with each country's economic and policy circumstances.

APPENDIX TO CHAPTER 15: NATIONAL INCOME AND THE BALANCE OF PAYMENTS

Many macroeconomic problems and crises in developing countries have their roots in the balance of payments. Chapter 10 opens the discussion of balance of payments by describing the role of foreign savings in financing investment. Chapter 13 discusses debt crises in developing countries, and Chapter 15 provides a model for diagnosing and responding to macroeconomic crises—including debt crises. Balance of payments accounting provides a framework within which to understand many of the opportunities and risks created by opening economies to trade and financial flows.

This appendix provides a brief overview of some key aspects of balance of payments accounting.[10]

We begin by placing the balance of payments in the broader context of national income accounting. The value added from all goods and services produced domestically sums to Gross Domestic Product (GDP). In a closed economy, this exactly equals gross national expenditure, which consists of private consumption on final goods and services (C), government consumption and investment in final goods and services (G), and private investments to expand the country's capital stock (I). That is, GDP = C + I + G in a closed economy. In Chapter 15, we refer to gross national expenditure as **absorption** (A), or A = C + I + G. Adopting this notation, we can say that in a closed economy, GDP = A. In addition, in a closed economy all of gross national expenditures (or absorption) is paid to domestic entities, so GDP also equals gross national income (GNI).

When we allow for trade with other countries, it is no longer required that gross domestic product exactly equals absorption or gross national income. This is where the balance of payments comes in. Balance of payments accounting records all of the (home) country's transactions with the rest of the world, including both trade and financial flows. The main components of the balance of payments are the **current account**, the **financial account**, and the **capital account**.

We begin by incorporating the current account into national income. The current account records all foreign trade in goods (the trade balance) and services (net factor income from abroad), along with net unilateral transfers (such as foreign aid or remittances from residents working abroad). The current account also includes interest payments on foreign debts (debt service payments, but not repayment of principal on the debt). In an open economy, part of absorption may be spent on imports (M), and must be subtracted from A; conversely, some domestic production may be exported abroad (X), and the resulting payments to domestic firms are added to A. In this case, we can write the national income accounting identity as

$$GDP = C + I + G + X - M \qquad \text{[A15-1]}$$

Substituting absorption (A) in for gross national expenditure, we can re-write equation A15-1 as

$$GDP = A + X - M \qquad \text{[A15-2]}$$

[10]This discussion draws on Robert C. Feenstra and Alan M. Taylor, *International Macroeconomics*, 2nd edition (New York: Worth Publishers, 2012), Chapter 5 (to which the reader is referred for a more detailed treatment), as well as the IMF's *Balance of Payments and International Investment Position Manual*, 6th edition (Washington, DC: International Monetary Fund, 2011).

That is, gross domestic product equals domestic expenditure plus net exports. Rearranging the terms of equation A15-2, we can see that A − GDP = M − X. In other words, if domestic spending exceeds GDP, then our imports exceed our exports and we run a negative trade balance. The trade balance is typically the major component of the current account. As noted above, net factor income from abroad (NFIA), or income from the net export of services (e.g., wages paid to a home country resident working abroad for a foreign firm), is also part of the current account and is included in gross national income (GNI). Thus, GNI = GDP + NFIA. The difference between gross national income (GNI) and gross national disposable income (Y) is that the latter also includes the final component of the current account, net unilateral transfers (NUT). Putting these steps together, we can write

$$Y = A + (X - M) + NFIA + NUT = A + CA \qquad \text{[A15-3]}$$

where CA is the current account, which equals the sum of net exports, net factor income from abroad, and net unilateral transfers. Equation A15-3 is important. It implies that if the home country's total expenditures exceed its total receipts, the difference must be provided by running a deficit on the current account. From equation A15-3, we can see that $A - Y = -CA$. The current account balance thus serves as an indicator of whether a country's spending exceeds its income.

As discussed in Chapter 10, we can also approach the current account from the perspective of national savings. National savings is what is left over from total receipts net of private and government consumption, or $S = Y - C - G$. If we expand A in equation A15-3 to write $Y = C + I + G + CA$ and subtract $C + G$ from both sides, we arrive at the **current account identity**,

$$S = I + CA \qquad \text{[A15-4]}$$

which says that we can only save more than we invest if there is a CA surplus, and we can only invest more than we save if there is a CA deficit.

Domestic savings has two components—private savings (S_p) and government savings (S_g). Income in the private sector must be consumed, saved, or paid to the government as taxes. Thus,

$$S_p = Y - C - T \qquad \text{[A15-5]}$$

Similarly, government savings is the difference between the tax revenue collected by the government (T) and its expenditures (G), or

$$S_g = T - G \qquad \text{[A15-6]}$$

which is the government's fiscal surplus (or deficit, if $G > T$). The sum of these two equations equals our previous observation that $S = Y - C - G = S_p + S_g$, or total national savings is the sum of private savings and government savings. Substituting this idea into equation A15-4 and rearranging, we can see that

$$CA = S_p + S_g - I \qquad \text{[A15-7]}$$

Stated slightly differently, if investment is to exceed national savings, the difference must be financed from abroad, and that difference equals the current account deficit. In this sense, we can think of the current account as foreign savings. In addition, if we substitute equation A15–6 into equation A15–7, we see that

$$CA = (S_p - I) + (T - G) \qquad \text{[A15–8]}$$

which tells us that if investment exceeds private savings and/or the government runs a fiscal deficit, the current account must be in deficit by an equivalent amount.

The rest of the balance of payments story is about how the current account is financed. As noted above, in addition to the current account, the balance of payments includes the financial account (FA) and the capital account (KA). The financial account records foreign transactions involving financial assets such as stocks, bonds, and real estate ownership. The major components of the financial account (using the IMF definition) are foreign direct investment, portfolio investment, and (importantly) reserve assets (which is the stock of foreign currency—usually U.S. dollars, euros, or yen) held by the central bank. The capital account refers to the transfer of assets such as gifts to and from abroad.[11]

A fundamental principle of balance of payments accounting is that bills must be paid. This implies that the balance of payments must sum to zero, or

$$CA + FA + KA = 0 \qquad \text{[A15–9]}$$

This is the **balance of payments identity**. It implies that there are only three ways in which a country can finance a current account deficit: 1) it can borrow from abroad (from foreign governments, banks, or international lending bodies), 2) it can attract capital inflows (in the form of direct or portfolio investment), or 3) it can run down its stock of foreign reserves. If the current account is in deficit, then it must be balanced by a surplus in the financial and capital accounts (which potentially includes running down the country's stock of foreign reserves). If the balance of payments must also equal zero, what is meant when we hear about a "balance of payments deficit?" In common use, a balance of payments deficit refers to a situation in which a CA deficit is not fully balanced by a surplus in FA + KA *where FA is defined narrowly to exclude the foreign reserve assets of the central bank.* In such a case, the total balance of payments must be brought to zero by a change in foreign reserves (running them down, in this instance, to finance the remaining CA deficit).

Running a surplus in the (non-reserve) financial and capital accounts means that the home country is either borrowing from abroad or selling domestic assets to foreigners. In either case, the home country is increasing the claims by foreigners

[11]This narrow definition of the capital account (and broad definition of the financial account) is in accordance with the approach now taken by the IMF, along with OECD and the UN System of National Accounts. Previously, the accepted tradition was to define the capital account broadly enough to include everything that the IMF now assigns to the financial account, with the exception of the reserve assets of the central bank.

on domestic residents. Conversely, if the home country is running a current account surplus, it must also be running a deficit on the financial and capital accounts, thus increasing the claims by residents on foreigners. Since the home country's stock of net foreign assets equals the difference between claims by residents of the home country on foreigners and claims by foreigners on home country residents, the current account equals the change in net foreign assets.