CHAPTER 7

Urbanization and Rural-Urban Migration: Theory and Policy

Cities will increasingly become the main players in the global economy.
—Kofi Annan, former secretary general of the United Nations and Nobel laureate for Peace

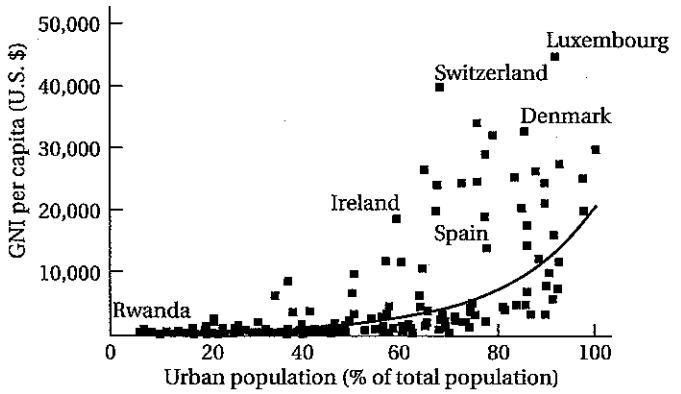
Poverty is becoming increasingly urbanized: Within twenty-five years, the number of urban poor will increase from an estimated 400 million to over 1 billion.
—James D. Wolfensohn, former president, World Bank

The Migration and Urbanization Dilemma

In this chapter, we focus on one of the most complex and nuanced dilemmas of the development process: the phenomenon of massive and historically unprecedented movements of people from the rural countryside to the burgeoning cities of Africa, Asia, and Latin America. In Chapter 6, we documented the extraordinary increase in world and especially developing-country population over the past few decades. By 2050, world population could exceed 9 billion people, and nowhere will population growth be more dramatic than in the major cities of the developing world. Indeed, according to United Nations estimates, in 2007, for the first time in human history, the world became more urban than rural.

After reviewing trends and prospects for overall urban population growth, we examine in this chapter the potential role of cities—both the modern sector and the urban informal sector—in fostering economic development. We then turn to a well-known theoretical model of rural-urban labor transfer in the context of rapid growth and high urban unemployment. In the final section, we evaluate various policy options that LDC governments may wish to pursue in their attempts to moderate the heavy flow of rural-to-urban migration and to ameliorate the serious unemployment problems that continue to plague their crowded cities. This chapter's case study looks at patterns of migration in India and Botswana.

FIGURE 7.1 Urbanization and Per Capita Income in Selected Countries



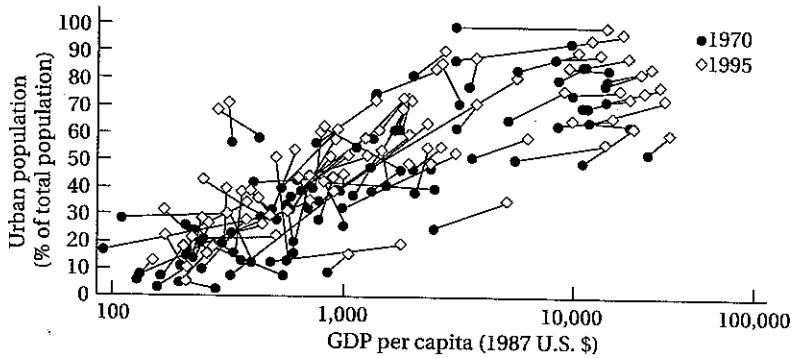
Source: UN-Habitat, "State of the World's Cities, 2001," http://www.unchs.org/Istanbul+5/86.pdf. Reprinted with permission.

Urbanization: Trends and Projections

The positive association between urbanization and per capita income is one of the most obvious and striking "stylized facts" of the development process. Generally, the more developed the country, measured by per capita income, the greater the share of population living in urban areas. Figure 7.1 shows urbanization versus GNI per capita; the highest-income countries, such as Denmark, are also among the most urbanized, while the very poorest countries, such as Rwanda, are among the least urbanized. At the same time, while individual countries become more urbanized as they develop, today's poorest countries are far more urbanized than today's developed countries were when they were at a comparable level of development, as measured by income per capita; and LDCs are urbanizing at a faster rate.

Figure 7.2 shows urbanization over time and across income levels over the quarter century from 1970 to 1995. The figure, from the World Bank's 1999–2000 *World Development Report*, gives real income per capita (in constant 1987 U.S. dollars) but does not adjust for purchasing power parity. Each line segment represents the trajectory of one country, starting from the solid dots, which represent the 1970 income and urbanization level for a given country, and ending at the end of the line segments (marked by a diamond), which represent the corresponding 1995 income and urbanization level for the same country. Although the World Bank caption to the figure stated that "urbanization is closely associated with economic growth," the figure may also be interpreted as showing that urbanization is occurring everywhere, at high and low levels of income and whether growth is positive or negative. Even when the lines point to the left, indicating shrinking incomes per capita over the period, they still generally point upward, indicating that urbanization continued. In short, urbanization is happening everywhere in the world, although at differing rates. So we need to consider urbanization

FIGURE 7.2 Urbanization across Time and Income Levels



Source: World Bank, *World Development Report, 1999–2000* (New York: Oxford University Press, 2000). Reprinted with permission.

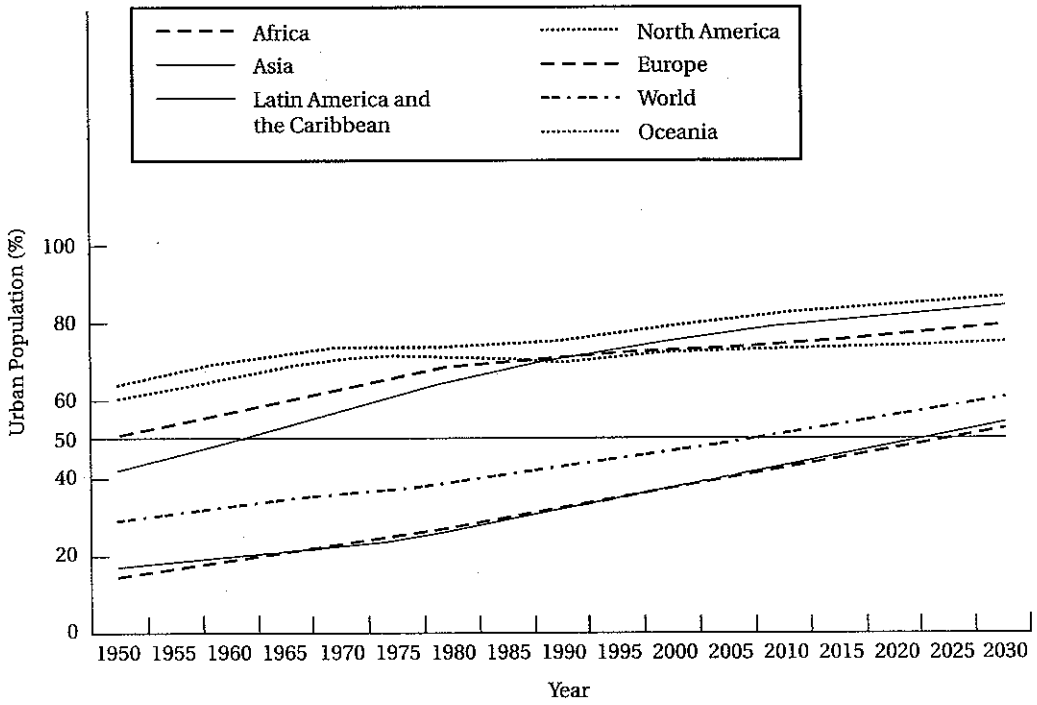
carefully—is it only correlated with economic development, or is causation also at work?

Indeed, one of the most significant of all modern demographic phenomena and the one that promises to loom even larger in the future is the rapid growth of cities in developing countries. In 1950, some 275 million people were living in cities in the developing world, 38% of the 724 million total urban population. According to UN estimates, the world's urban population had reached nearly 3 billion by 2003, with well over two-thirds living in metropolitan areas of developing countries.

Figure 7.3 shows the growth of the proportion of the population living in urban areas by region. From 2005 to 2030, the UN projects that world population will grow at a 1.78% average annual rate. Accordingly, there will be almost 5 billion urban dwellers by 2030, nearly five-eighths of the projected 8.1 billion world population in that year. In fact, after 2015, the number of people living in rural areas in the world is projected to actually begin to decrease, by some 155 million people from 2015 to 2030, or an annual rate of −0.32%. The most rapid urbanization in now occurring in Asia and Africa; well before 2030, more than half of all people in these regions will live in urban areas. More than half the world's urban population will live in Asia, and the projected 2030 urban population of Africa of 748 million will be larger than the entire projected 685 million total population of Europe.[1]

Although a majority of urban growth will be found in small to medium-size cities, the developing world is also coming to dominate the world's largest cities, including the megacities with over 10 million inhabitants. Figure 7.4 provides a map locating megacities, the largest cities in the world containing a population of at least 10 million people. As the figure shows, in 1975, there were only 5 megacities, but by 2000, there were 19 such metropolises. Of these 19, all but 2,

FIGURE 7.3 Proportion of Urban Population by Region, 1950–2030

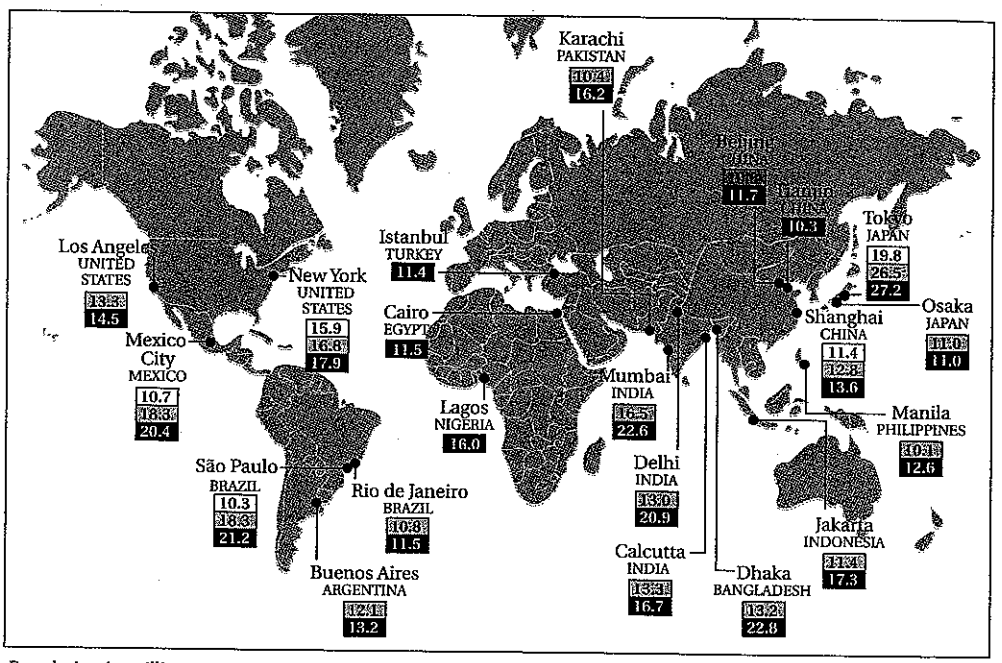


Source: Adapted from UN-Habitat, "State of the World's Cities, 2006–07," http://www.unhabitat.org.

New York and Tokyo, were located in the developing world. By 2015, at least 4 more giant cities are projected to join their ranks: Istanbul in Turkey, Hyderabad in Pakistan, Bangkok in Thailand, and Tianjin in China. (Note that lists of the largest cities, even across UN agencies, may differ, owing to alternative definitions of a metropolitan area, but any way they are counted, the trend toward an increasing number of urban giants in the developing world is unmistakably clear.) Moreover, as Figure 7.5 shows, almost all of the increments to the world's population will be accounted for by the growth of urban areas as migrants continue to stream into the cities from rural areas and as urbanization rates in the developing world continue to approach those of the developed world.

A central question related to the unprecedented size of these urban agglomerations is how these LDC cities will cope—economically, environmentally, and politically—with such acute concentrations of people. While it is true that cities offer the cost-reducing advantages of agglomeration economies and economies

FIGURE 7.4 Megacities: Cities with Ten Million or More Inhabitants



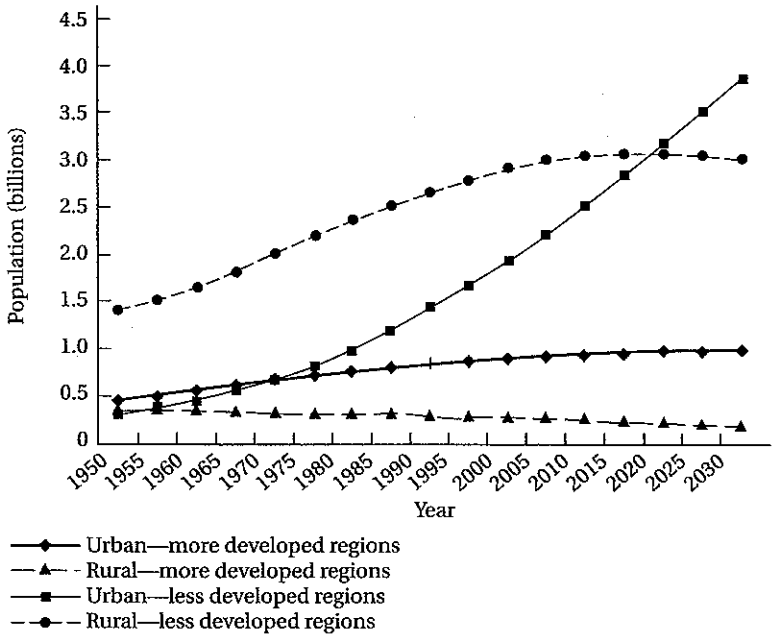
Source: Data from United Nations Population Division, March 2002.

of scale and proximity as well as numerous economic and social externalities (e.g., skilled workers, cheap transport, social and cultural amenities), the social costs of a progressive overloading of housing and social services, not to mention increased crime, pollution, and congestion, can outweigh these historical urban advantages. Former World Bank president Robert McNamara expressed his skepticism that huge urban agglomerations could be made to work at all:

> These sizes are such that any economies of location are dwarfed by costs of congestion. The rapid population growth that has produced them will have far outpaced the growth of human and physical infrastructure needed for even moderately efficient economic life and orderly political and social relationships, let alone amenity for their residents.[2]

Along with the rapid spread of urbanization and the **urban bias** in development strategies has come this prolific growth of huge slums and shantytowns. From the *favelas* of Rio de Janeiro and the *pueblos jovenes* of Lima to the *bustees* of Calcutta (Kolkata) and the *bidonvilles* of Dakar, such makeshift communities have been growing rapidly. Today, slum settlements represent over one-third of the urban population in all developing countries.

FIGURE 7.5 Estimated and Projected Urban and Rural Population of the More and Less Developed Regions, 1950–2030

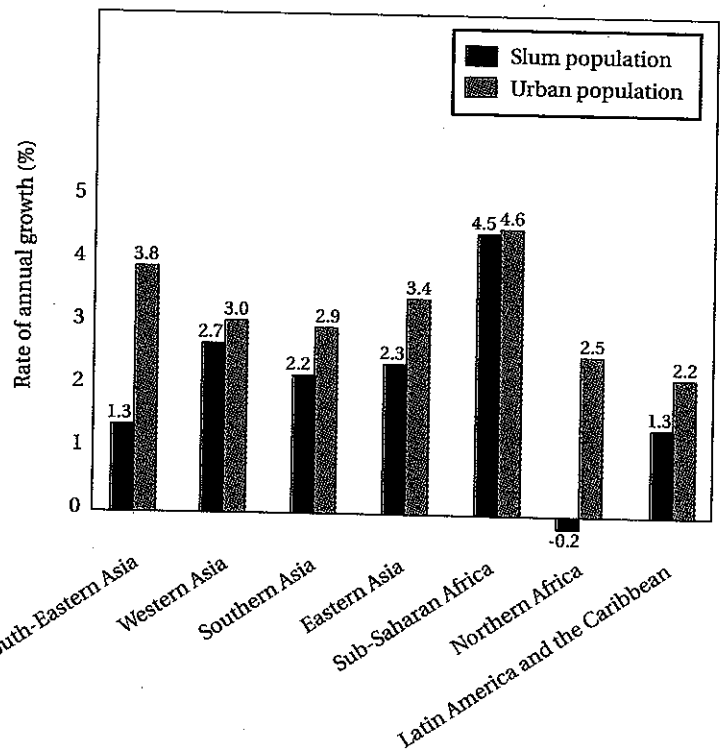


Source: United Nations Population Division, *World Urbanization Prospects: The 1999 Revision* (New York: United Nations, 2000). Reprinted with permission.

Figure 7.6 shows the annual growth of urban and slum populations in the 1990–2001 period, drawn from the 2006 United Nations *Millennium Development Goals Report.* As the *Report* summarized, "Sub-Saharan Africa is the world's most rapidly urbanizing region, and almost all of this growth has been in slums, where new city residents face overcrowding, inadequate housing, and a lack of water and sanitation. In Western Asia, as well, most of the urban growth is occurring in slums. The rapid expansion of urban areas in Southern and Eastern Asia is creating cities of unprecedented size and complexity and new challenges for providing a decent environment for the poor. Northern Africa is the only developing region where the quality of urban life is improving: In this region, the proportion of city dwellers living in slums has decreased by 0.15 per cent annually."

Although population growth and accelerated **rural-urban migration** are chiefly responsible for the explosion in urban shantytowns, part of the blame rests with LDC governments. Their misguided urban-planning policies and outmoded building codes often means that 80% to 90% of new urban housing is "illegal." For example, colonial-era building codes in Nairobi, Kenya, have made it impossible to build an "official" house for less than $3,500. The law has also required every

FIGURE 7.6 Annual Growth of Urban and Slum Populations, 1990–2001



Source: Adapted from United Nations, *Millennium Development Goals Report, 2006* (New York: United Nations, 2006), p. 20. Used with permission.

dwelling to be accessible by car. As a result, two-thirds of Nairobi's land has been occupied by 10% of the population, while many slum dwellings cannot legally be improved. Similarly, in Manila, Philippines, a large majority of the population has historically been too poor to be able to buy or rent an officially "legal" house.[3]

Statistics show that rural migrants constitute anywhere from 35% to 60% of recorded urban population growth. Accordingly, 90 out of 116 developing countries responding to a UN survey indicated that they had initiated policies to slow down or reverse their accelerating trends in rural-urban migration.[4]

Given the widespread dissatisfaction with the experience of rapid urban growth in developing countries, the critical issue that needs to be addressed is the extent to which national governments can formulate development policies that can have a definite impact on trends in and the character of urban growth. It is clear that the emphasis on industrial modernization, technological sophistication, and metropolitan growth created a substantial geographic imbalance in economic opportunities and contributed significantly to the accelerating influx of rural migrants into urban areas. Is it possible and or even desirable now to attempt to reverse these

trends by pursuing a different set of population and development policies? With birth rates declining in many LDCs, rapid urban growth and accelerated rural-urban migration will undoubtedly be one of the most important development and demographic issues of the coming decades. And in urban areas, the growth and development of the informal sector as well as its role and limitations for labor absorption and economic progress will assume increasing importance.

Before examining conditions in developing-country cities more closely, let us first consider the potential advantages offered by cities. Urban areas have played a highly constructive role in the economies of today's developed countries, and they offer huge and still largely untapped potential to do the same for developing countries. A detailed look at the informal sector in developing cities will give an idea of its potential as an engine of growth. We also consider in more detail what has been different—and what has gone wrong—with urban development in LDCs and the excessively rapid pace of rural-urban migration. We conclude with a look at constructive policies to help cities foster successful urban development while at the same time giving more balanced treatment to development in rural areas.

The Role of Cities

What explains the strong association between urbanization and development? To a large degree, cities are formed because they provide cost advantages to producers and consumers through what are called **agglomeration economies**. As noted by the pioneering regional economist Walter Isard, these agglomeration economies come in two forms. **Urbanization economies** are effects associated with the general growth of a concentrated geographic region. **Localization economies** are effects captured by particular sectors of the economy, such as finance or automobiles, as they grow within an area. Localization economies often take the form of backward and forward linkages of the type introduced in Chapter 4. When transportation costs are significant, users of the outputs of an industry may benefit from a nearby location to save on these costs. This benefit is a type of forward linkage. In addition, firms of the same or related industries may benefit from being located in the same city, so they can all draw on a large pool of workers with the specific skills used in that sector or from specialized infrastructure. This is a type of backward linkage. Workers with specialized skills appropriate to the industry prefer to be located there as well so that they can easily find a new job or be in a position to take advantage of better opportunities.

Industrial Districts

An economic definition of a city is "an area with relatively high population density that contains a set of closely related activities." Firms often also prefer to be located where they can learn from other firms doing similar work. Learning takes place in both formal relationships, such as joint ventures, and informal ones, such as from tips learned in evening social clubs or over lunch. These spillovers are also agglomeration economies, part of the benefits of what Alfred Marshall

called "industrial districts," and they play a big role in Michael Porter's "clusters" theory of competitive advantage.[5] Firms located in such industrial districts also benefit from the opportunity to contract out work easily when an unusually large order materializes. Thus a firm of modest size does not have to turn down a big job due to lack of capacity, an arrangement that provides "flexible specialization."[6] Further, firms may wish to operate in well-known districts for the marketing advantages of locating where consumers of their goods know to shop to get the best selection.

It may not matter so much where such industrial districts are located as that they somehow got an early start there, perhaps because of a historical accident. For example, in the United States, many innovative computer firms located in Silicon Valley, California, simply because other such firms were already located there. Analogously, suppliers to shoe firms located in the Sinos Valley in southern Brazil and in Guadalajara in Mexico because so many shoe firms located in those regions. Some of the benefits are gained simply by the fact of location—Khalid Nadvi has termed this "passive collective efficiency"—but other benefits must be achieved through collective action, such as developing training facilities or lobbying government for needed infrastructure as an industry rather than as individual firms ("active collective efficiency").

A growing body of evidence shows that industrial clusters are now common in developing countries, at stages of industrial development ranging from cottage industry to advanced manufacturing techniques, and appear to be significant factors in emerging industrial competitiveness.[7] Nevertheless, the dynamism of these clusters has varied widely. Some of the identified districts are traditional clusters of artisans that have shown little ability to innovate, export, or expand. Traditional cottage industries are often grouped together by village, a phenomenon found throughout the developing world that is particularly prominent in Java. But such groupings often remain one-family microenterprises with little division of labor or use of modern techniques. Producers in a village are better off sharing a common specialization than producing a random assortment of goods, in part because intermediaries work with villages with a high concentration of producers in their sector. But such traditional producers benefit little from "internal" divisions of labor within the firm, generally producing a largely complete product within the household and remaining at very low productivity and incomes. For example, a small town in Kenya may have a dozen or more families fabricating wheelbarrows, each family starting with timber and a few simple purchased metal inputs and producing a final product for sale.

In other cases, such traditional township specializations have evolved into clusters, with still modest-size but somewhat larger firms using a more detailed division of labor, such as a group of wheelbarrow producers with some specialization, each employing a few workers. Eventually, the cluster might expand in scope and become a low-tech metal products industrial district selling products throughout the country as the town grows into a small city. These clusters are reminiscent of the industrial districts of developed countries but require that sufficient financing be gathered to invest in core firms using somewhat larger-scale capital goods.

As Hermine Weijland found in her study of Java, Indonesia, "It needs only a few fortunate years of market expansion to create gains from externalities and joint action."[8] She cites as examples local clusters that have upgraded and now competitively produce such goods as roof tiles, rattan furniture, cast metal, and textiles. Similarly, Dorothy McCormick concluded from a study of six representative clusters in Africa that "groundwork clusters prepare the way; industrializing clusters begin the process of specialization, differentiation, and technological development; and complex industrial clusters produce competitively for wider markets."[9] In some cases, the evidence suggests that coordination failures are not overcome, and so there may be a role for government policy in encouraging the upgrading of clusters. In other cases, it is the government itself that shares blame for cluster stagnation when it enforces irrational and stifling regulations, which are far more damaging than the usual policy of benign neglect toward nascent clusters in the informal sector. Examples of clusters in developing countries that are widely considered successful include surgical instruments in Sialkot, Pakistan; software in the Bangalore (Bengaluru) area in India; and footwear in the Sinos Valley, Brazil (although this last industry is also known for its use of child labor). Clusters of all kinds, however, face substantial challenges from globalization and trade liberalization.

Again, not all of the collective efficiency advantages of an industrial district are realized through passive location. Others are actively created by joint investments and promotional activities of the firms in the district. One factor determining the dynamism of a district is the ability of its firms to find a mechanism for such collective action. While the government can provide financial and other important services to facilitate cluster development, **social capital** is also critical, especially group trust and a shared history of successful collective action, which requires time to develop. Government can help by bringing parties together and helping them gain experience cooperating on more modest goals before tackling larger ones, but social capital normally grows organically in an economic community and cannot be created forcibly. Even with collective action to supplement passive benefits of agglomeration, traditional clusters may not survive in their current form into more advanced stages of industrialization. Nonetheless, as Hubert Schmitz and Khalid Nadvi note, even if transitional, districts in the informal sector may still play a crucial role in mobilizing underused human and financial resources.[10]

Statistical estimates show that benefits of agglomeration can be quite substantial in practice. For example, studies have demonstrated that "if a plant moves from a location shared by 1,000 workers employed by firms in the same industry to one with 10,000 such workers, output will increase an average of 15%, largely because the pool of specialized workers and inputs deepens." Moreover, "productivity rises with city size, so much so that a typical firm will see its productivity climb 5% to 10% if city size and the scale of local industry double."[11]

Efficient Urban Scale

Localization economies do not imply that it would be efficient for all of a country's industries to be located together in a single city! These economies extend across

closely related industries, such as those with strong backward and forward linkages, but there are fewer productivity benefits for unrelated industries to locate together. One notable exception is the potential spillover from technological progress in one industry to its adaptation for different uses in another industry. But there are also some important **congestion** costs. The higher the urban density, the higher the costs of real estate. It is much more expensive to build vertically than horizontally, increasingly so as skyscraper scale is reached, so that when market forces work properly, tall buildings are built primarily when urban land costs become high. (Note that skyscrapers and other buildings of monumental scale are sometimes built for political show rather than for economic efficiency, such as the world's tallest buildings in Dubai, United Arab Emirates; Taipei, Taiwan; and Kuala Lumpur, Malaysia.) In large urban areas, workers may find themselves with longer and longer commutes and greater transportation costs and may demand higher wages to cover these costs. In addition, the costs of infrastructure such as water and sewer systems are higher in concentrated urban areas. In theory, if costs of transportation of finished goods are high and consumers wish to be located in the largest city to avoid paying those transportation costs as much as possible, economic activity could become indefinitely concentrated within a city (called the "black hole" effect), but it is generally much less costly to improve the transportation system of a country than to pay the costs of maintaining a gargantuan urban complex. Under competitive forces and other things being equal, if workers are mobile, a worker in a large city with higher wages but higher costs of living (such as higher housing prices) is no better off in real material terms than a worker with comparable education, experience, ability, and health in a small city who has lower wages and lower costs of living.[12]

Thus the concentrating, or "centripetal," forces of urban agglomeration economies are opposed by the dispersing, or "centrifugal," forces of diseconomies featuring increasing costs with greater concentration, because some of the factors of production, most obviously land, are not mobile. We can "create" more central city land by building skyscrapers, but only to a certain scale and only at substantial cost. Thus it is normal for an economy to have a range of cities, with sizes dependent on the scale of the industries it sponsors and the extent of agglomeration economies found for that industry or cluster of industries.

Two well-known theories of city size are the urban hierarchy model (central place theory) and the differentiated plane model.[13] In the urban hierarchy model, originated by August Losch and Walter Christaller, plants in various industries have a characteristic market radius that results from the interplay of three factors: economies of scale in production, transportation costs, and the way the demand for land is spread over space. The larger the economies of scale in production and the lower the transportation costs, the larger the radius of territory that will be served by that industry to minimize costs. In contrast, if the price of real estate is bid up to high levels in the resulting cities, this will tend to create smaller radii. As a result, small cities contain activities with short market radii, while large cities emerge to contain activities of both small and large radii. Generally speaking, activities of a national scope, such as government and finance, will be located in a single city (though not necessarily the same large city because of the effect of congestion costs). Clearly, the urban hierarchy approach applies better to nonexport industries than to export

industries. When countries have different specializations in the international market or are at different stages of economic development, the size distribution of cities may potentially differ. For example, a developing country that still overwhelmingly specializes in agriculture might reasonably have one or two large cities serving national industries such as finance and government and many smaller towns serving local agricultural areas. A country with a highly differentiated manufacturing and service base might have a large number of medium-size cities.

In the differentiated plane model, originated by Alfred Weber, Walter Isard, and Leon Moses, the limited number of transportation routes linking the industries within an economy plays a key role. The model predicts urban concentrations at the points where the scarce transportation routes cross, called "internal nodes." The hierarchy of urban sizes depends on the pattern of nodes and the industrial mix. Primary processing industries have few inputs and are usually located near the source of the primary resource. However, there will also be incentives for industries with strong backward or forward linkages to locate in the same city.

The Urban Giantism Problem

In the case of developing countries, the main transportation routes are often a legacy of colonialism. Theorists of the dependency school (see Chapter 3) have compared colonial transportation networks to drainage systems, emphasizing ease of extraction of the country's natural resources. In many cases, the capital city will be located near the outlet of this system on the seacost. This type of transportation system is also called a "hub-and-spoke system," which is especially visible when the capital city is located in the interior of the country. Many nations inherited a hub-and-spoke system from ancient colonial times, such as that of Paris and London, as well as more recent colonial experiences, such as cities in Africa and Latin America. In the case of London and Paris, these transportation systems apparently were designed in the time of the Roman Empire to facilitate movement of troops from the capital to the outlying towns to suppress revolts. A similar motivation was likely present in the more recent designs of the African colonies of Britain and France and the Central and South American colonies of Spain.

The differentiated plane approach emphasizes the lasting impact of historical accidents. In this case, it helps explain where the most oversized cities are found in the developing world and suggests where policies of urban decentralization may be most helpful. Note that not all countries inherited such a hub-and-spoke system; Germany did not; the United States did not, in part because it is the result of the merger of 13 separate British colonies, which retained some measure of local autonomy, as do the federal states of Germany. The recent development of the United States makes the emergence of cities such as Atlanta from the crossing of transportation routes especially clear, but the same principle has applied elsewhere over longer historical periods. Of course, as nations become wealthy, they generally build better transportation systems.

Sometimes one urban core becomes too large to keep the costs of the industries located there to a minimum. In developed countries, other cores are often developed

within the broad metropolitan region, enabling the region as a whole to continue to receive benefits of agglomeration while lowering some of the costs; or new cities may develop in entirely different parts of the country. But this creation of new urban cores does not happen automatically if there are advantages to locating where other firms and residents are already present. This is another chicken-and-egg coordination problem of the type described in Chapter 4. Who will be the pioneer if it is less costly to stay where you are and wait for other pioneers to settle in the new city first? In economic terms, the agglomeration economies of cities are externalities, which must somehow be internalized or the market will fail. How can this be done?

In the United States, developers frequently internalize the externality by creating a new "edge city" within a metropolitan area, financing and building a new center where land is still relatively inexpensive, perhaps 10 to 50 kilometers from the original urban core. This takes place within a context of public oversight in the form of zoning regulations and inducements such as tax breaks. In LDCs, however, capital markets generally do not (yet) work well enough for this process of development to take place. In Europe, the public sector plays a much larger role in coordinating new towns and large developments.

In developing countries, however, governments are less involved in the dispersal of economic activity to more manageable sizes or, if they are involved, are often less effective. For example, government may seek to disperse industry without regard to the nature of agglomeration economies, giving incentives for dispersal but no attention to clustering relevant industries together, a problem seen in industrial parks in Pakistan. And all too often, the incentives are for firms to concentrate in the capital city or other "urban giants." A key problem of countries such as Peru and Argentina is that their giant capitals suffer from enormous levels of congestion, but adequate midsize cities that might provide alternative locations for growth are lacking. A well-designed infrastructure development program, including more efficient links between medium-size cities and better roads, utilities, and telecommunications within these cities, can help alleviate this problem.

A more detailed comparison of North and South America is instructive. The largest urban area in the United States, the New York metropolitan area, has about 6% of the national population (3% in New York City itself and about the same percentage in its suburbs). Toronto, the largest metropolitan area in Canada, has 4.3 million residents, some 14% of the Canadian population. But Mexico City holds more than one-fifth of the population of Mexico, Montevideo two-fifths of the population of Uruguay, Lima about one-fifth of the population of Peru, and Buenos Aires and Santiago more than a third of the populations of Argentina and Chile, respectively.[14]

First-City Bias

A form of urban bias that has often caused considerable distortions might be termed *first-city bias*. The country's largest or "first-place" city receives a disproportionately large share of public investment and incentives for private investment, in relation to the country's second-largest city and other smaller cities. As a result, the first city receives a disproportionately, and inefficiently, large share of population and economic activity.

TABLE 7.1 Population of the Largest and Second-Largest Cities in Selected Countries (millions)

Country	Largest-City Population	Second-Largest-City Population	Ratio
Canada	Toronto, 4.3	Montreal, 3.3	1.3
United States	New York, 19.7	Los Angeles, 15.3	1.3
Argentina	Buenos Aires, 10.7	Rosario, 1.1	9.7
Brazil	São Paulo, 9.8	Rio de Janeiro, 5.5	1.8
Chile	Santiago, 4.3	Concepción, 0.3	14.3
Mexico	Mexico City, 15.0	Guadalajara, 2.9	5.2
Peru	Lima, 6.4	Arequipa, 0.6	10.7

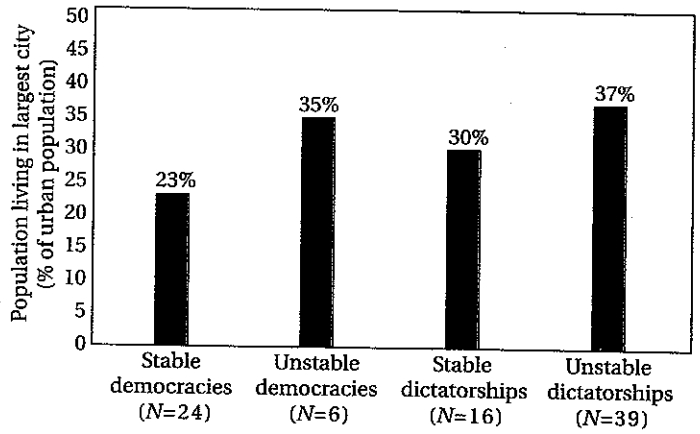
Source: Data from United Nations demography Web page, http://www.un.org/Depts/unsd/demog/city.htm.
Note: Definitions of city size differ across studies.

Table 7.1 shows the largest and second-largest cities in the United States, Canada, and major Latin American countries. Notice that in all of the outsized capital cities—Buenos Aires, Santiago, Mexico City, and Lima—the first city also serves as the capital. Patterns are similar in many other developed and developing countries. Although two European countries, the former colonial powers of the United Kingdom and France, have exceptionally large first cities, some seven times the size of the second city in both cases, in virtually all other European countries, the second city is about half, or more, the size of the first city.[15] Some other developing countries have remarkably outsized first cities, notably Thailand, where Bangkok has a population nearly 20 times the size of the second city of Nakhon Ratchasima. Further examples can be found in Egypt and Iran.

Causes of Urban Giantism

Why have first cities often swelled to such a large multiple of second cities in developing countries? Overall, urban giantism probably results from a combination of a hub-and-spoke transportation system and the location of the political capital in the largest city, thus combining the effects of the urban hierarchy model with the differentiated plane model. This is further reinforced by a political culture of rent seeking and the capital market failures that make the creation of new urban centers a task that markets cannot complete. Other more detailed explanations also generally involve unfortunate consequences of political economy (see Chapter 11). One argument, featured in the work of Paul Krugman, stresses that under import substitution industrialization (see Chapter 12), with a high level of protection, there is much less international trade, and population and economic activity have an incentive to concentrate in a single city, largely to avoid transportation costs. Thus firms wish to set up operations in the city where the most consumers already live, which attracts more people to the region in search of jobs and perhaps lower prices (made possible because there are fewer transport costs to be passed on to consumers); this concentration in turn attracts still more firms and consumers in a circle of causation. However, when trade barriers are reduced, the incentive to focus production on the home market is also reduced, and exporters and their suppliers have much less incentive to be located in the country's biggest

FIGURE 7.7 Politics and Urban Concentration



Source: Data from Alberto F. Ades and Edward L. Glaeser, "Trade and circuses: Explaining urban giants," *Quarterly Journal of Economics* 110 (1995): 196. Copyright © 1995 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.
Note: N = number of countries in group.

population center. This moves production toward ports and borders, or elsewhere in the country, to escape the excessive congestion costs of the largest city.[16]

Another recent explanation for urban giants focuses on the consequences of dictators' efforts to remain in power. As Figure 7.7 shows, on average, a much larger share of a country's urbanized population (37%) lives in the first city in unstable dictatorships than in stable democracies (23%). In interpreting this finding, Alberto Ades and Edward Glaeser argue that unstable dictatorships (fearing overthrow) must provide "bread and circuses" for the first city (usually the capital) to prevent unrest; this extreme urban bias in turn attracts more migrants to the favored city and a still larger need for bread and circuses. It should be noted that although the authors attempt to control for reverse causality, it may still be the case that unstable dictatorships also tend to emerge in countries with high first-city concentrations.[17]

In the developing world, until recently, relatively few countries were effective democracies. In the 1970s, almost all developing countries had authoritarian governments of one form or another. To remain in power and prevent popular uprisings and coups, which were generally thought to be most threatening when launched from the capital city, governments had an incentive to "buy off" the population of the largest city. This focus of national government spending on the capital city is the bread-and-circuses effect, recalling the phrasing of "rent-sharing" policies in ancient Rome in its period of expansion. The availability of better opportunities, whether the equivalent of the grain handouts in ancient Rome or jobs, wages, infrastructure, and other government services concentrated in the capital city of many of today's developing countries, attracts an ever-growing migrant population, in turn leading to larger precautionary government spending as the fear of political instability grows.

A final political economy factor contributes to capital city giantism: It becomes advantageous for firms to be located where they have easy access to government officials, to curry political favor from a regime that can be induced to give companies special favors for a price or that simply demands bribes to function at all. The resulting first-city giantism may be viewed as a form of underdevelopment trap, which may be escaped fully only with a return to democratic rule together with a better balance of incentives to compete for exports as well as home consumption. Democracy does not eliminate political benefits of location in the national capital, but while lobbyists still congregate in the political capital, there may be less incentive for production to become overconcentrated there. Moreover, a free press tends to expose corruption and generate public pressure to root it out, as recent experience in many democratizing countries in Latin America and East Asia makes clear.

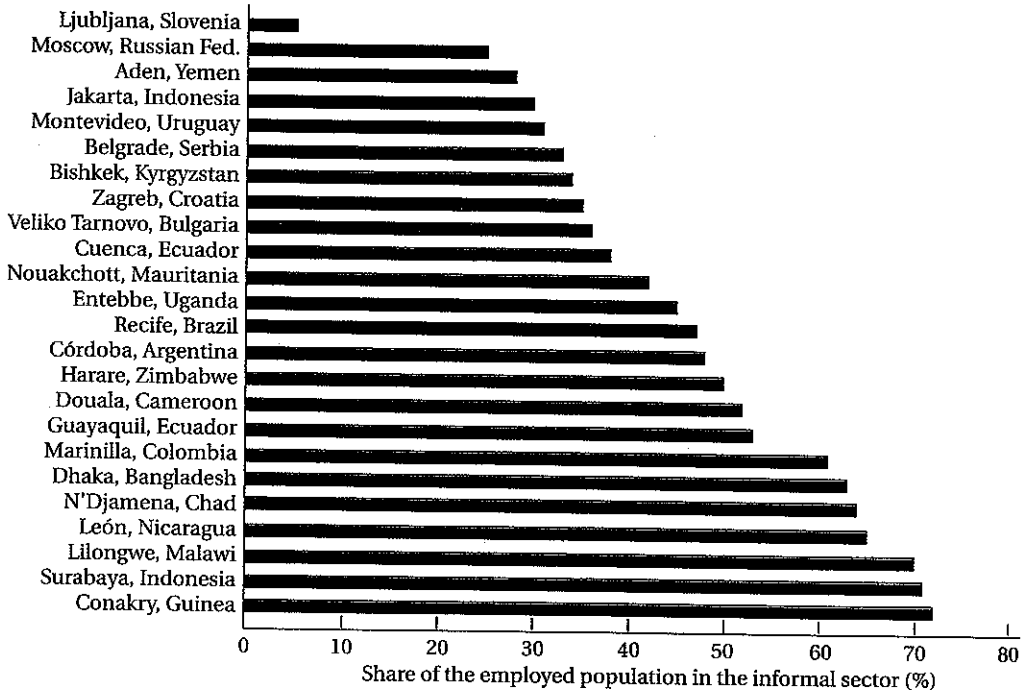
The four explanations for urban giantism—production for the home market in the face of high protection and transport costs, few adequate smaller cities as alternative locations for firms reflecting infrastructure patterns, location of the capital in the largest city, and the political logic of unstable dictatorships—are complementary and help explain some of the advantages of democracies with more balanced economic policies, including well-planned investments in infrastructure. Such countries are able to avoid some of the costs of urban giantism.

The Urban Informal Sector

As noted in Chapter 3, a focus of development theory has been on the dualistic nature of developing countries' national economies—the existence of a modern urban capitalist sector geared toward capital-intensive, large-scale production and a traditional rural subsistence sector geared toward labor-intensive, small-scale production. This dualistic analysis has also been applied specifically to the urban economy, which has been decomposed into a formal and an informal sector.

The existence of an unorganized, unregulated, and mostly legal but unregistered **informal sector** was recognized in the 1970s, following observations in several developing countries that massive additions to the urban labor force failed to show up in formal modern-sector unemployment statistics. The bulk of new entrants to the urban labor force seemed to create their own employment or to work for small-scale family-owned enterprises. The self-employed were engaged in a remarkable array of activities, ranging from hawking, street vending, letter writing, knife sharpening, and junk collecting to selling fireworks, prostitution, drug peddling, and snake charming. Others found jobs as mechanics, carpenters, small artisans, barbers, and personal servants. Still others were highly successful small-scale entrepreneurs with several employees (mostly relatives) and higher incomes. Some could even eventually graduate to the formal sector, where they become legally registered, licensed, and subject to government labor regulations. With the unprecedented rate of growth of the urban population in developing countries expected to continue and with the increasing failure of the rural and urban formal sectors to absorb additions to the labor force, more attention is being devoted to the role of the informal sector in serving as a panacea for the growing unemployment problem.

FIGURE 7.8 Importance of Informal Employment in Selected Cities



Source: UN-Habitat, "State of the World's Cities, 2001," http://www.unchs.org/Istanbul+5/statereport.htm. Reprinted with permission.

The informal sector continues to play an important role in developing countries, despite decades of benign neglect at best and outright hostility at worst. In many developing countries, about half of the employed urban population works in the informal sector. Figure 7.8 shows the relative importance of informal unemployment in selected cities. Most of these cities reflect the typical range of informal-sector employment share, from about 30% to 70%. (The only exception is Ljubljana, a virtually developed city near Austria and Italy.) We find a similar pattern of high informal-sector employment in cities throughout the developing world. For example, in India, the urban informal sector comprises 28.5% of employment in Calcutta (Kolkata), 46.5% in Ahmedabad, 49.5% in Bombay (Mumbai), 53.8% in Madras (Chennai), 61.4% in Delhi, and 65.5% in Bangalore (Bangaluru).

The informal sector is characterized by a large number of small-scale production and service activities that are individually or family-owned and use simple, labor-intensive technology. They tend to operate like monopolistically competitive firms with ease of entry, excess capacity, and competition driving profits (incomes) down to the average supply price of labor of potential new entrants. The usually self-employed workers in this sector have little formal education, are generally

unskilled, and lack access to financial capital. As a result, worker productivity and income tend to be lower in the informal sector than in the formal sector. Moreover, workers in the informal sector do not enjoy the measure of protection afforded by the formal modern sector in terms of job security, decent working conditions, and old-age pensions. Many workers entering this sector are recent migrants from rural areas unable to find employment in the formal sector. Their motivation is often to obtain sufficient income for survival, relying on their own indigenous resources to create work. As many members of the household as possible are involved in income-generating activities, including women and children, and they often work very long hours. A large fraction inhabit shacks that they themselves have built in slums and squatter settlements, which generally lack minimal public services such as electricity, water, drainage, transportation, and educational and health services. Others are even less fortunate. Many are homeless, living on the pavements of Calcutta (Kolkata), Manila, Dakar, Nairobi, Rio de Janeiro, and Bogotá—to mention just a few major developing-country cities. They find sporadic temporary employment in the informal sector as day laborers and hawkers, but their incomes are insufficient to provide even the most rudimentary shelter.

Policies for the Urban Informal Sector

In terms of its relationship with other sectors, the informal sector is linked with the rural sector in that it allows excess labor to escape from extreme rural poverty and underemployment, although under living and working conditions and for incomes that are often not much better. It is closely connected with the formal urban sector: The formal sector depends on the informal sector for cheap inputs and wage goods for its workers, and the informal sector in turn depends on the growth of the formal sector for a good portion of its income and clientele.

Informal-sector incomes have remained persistently higher than those in the poorest rural regions despite continued flow of rural-urban migration. The Nobel laureate Sir Arthur Lewis in the 1950s viewed traditional-sector workers, the petty traders like newspaper hawkers, as unproductive and essentially engaged in distractions from the main urban work of industrialization. But if wages are persistently higher in very competitive activities such as informal work, this must reflect higher productivities as well. Consequently, a revisionist view espousing the constructive role of cities in economic development has taken hold. This approach has been championed by the Dar es Salaam–based UN-Habitat, in its "State of the World's Cities" reports.[18] The 2001 report systematically criticized what it termed the "anti-urban bias of the development agencies." Acting on the strong development tradition beginning with the Lewis skepticism of the urban informal sector, developed with the Todaro migration model (examined later in this chapter) emphasizing the negative consequences of urban bias for both efficiency and equity, continuing with the influential work of the integrated rural development school of the 1970s and recast and reemphasized in recent years under the Wolfensohn presidency at the World Bank, development agencies have indeed stressed rural development rhetorically. Many scholars have concluded, however, that this rhetoric often goes untranslated into real resources for the rural areas. Thus any

pro-rural bias of development agencies is typically little more than a partial correction to the overriding forces for urban bias. However, the renewed focus on the development role of cities is an important trend. Besides UN-Habitat, the World Bank and other agencies have placed increasing emphasis on improved urban development.[19] The new focus is on how to make cities in developing countries more dynamic engines of growth and more livable environments, and it promises to be one of the more important streams of emerging research and policy analysis in economic development in coming years. In any case, while medium-size cities undoubtedly deserve greater attention for the constructive role they play in the development process, this does not obviate the problem of overconcentration of activities in first-city urban giantism.

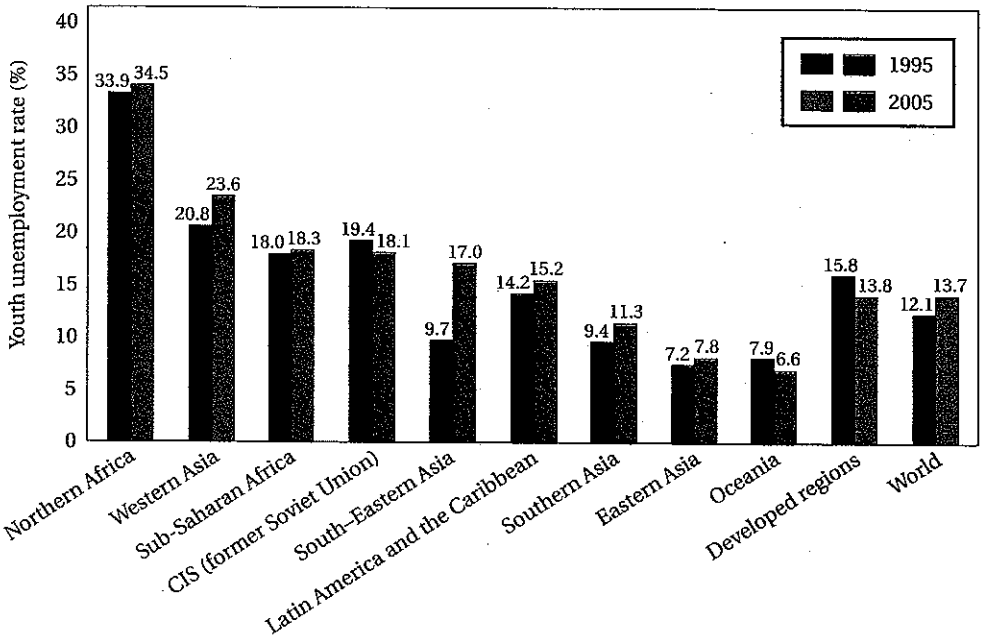
The important role that the informal sector plays in providing income opportunities for the poor is clear. There is some question, however, as to whether the informal sector is merely a holding ground for people awaiting entry into the formal sector and as such is a transitional phase that must be made as comfortable as possible without perpetuating its existence until it is itself absorbed by the formal sector or whether it is here to stay and should in fact be promoted as a major source of employment and income for the urban labor force.[20]

There seems to be a good argument in support of the latter view. The formal sector in developing countries has a small base in terms of output and employment. To absorb future additions to the urban labor force, the formal sector must be able to generate employment at a very high rate. This means that output must grow at an even faster rate, since employment in this sector increases less than proportionately in relation to output. This sort of growth seems highly unlikely in view of current trends. Thus the burden on the informal sector to absorb more labor will continue to grow unless other solutions to the urban unemployment problem are provided. But young people face increasingly difficult job prospects, as can be seen in Figure 7.9.

The informal sector has demonstrated its ability to generate employment and income for the urban labor force. As pointed out earlier, it is already absorbing an average of 50% of the urban labor force. Some studies have shown the informal sector generating almost one-third of urban income.

Several other arguments can be made in favor of promoting the informal sector. First, scattered evidence indicates that the informal sector generates surpluses even in a hostile policy environment that denies it access to the advantages offered to the formal sector, such as credit, foreign exchange, and tax concessions. Thus the informal sector's surplus could provide an impetus to growth in the urban economy. Second, as a result of its low capital intensity, only a fraction of the capital needed in the formal sector is required to employ a worker in the informal sector, offering considerable savings to developing countries so often plagued with capital shortages. Third, by providing access to training and apprenticeships at substantially lower costs than that provided by formal institutions and the formal sector, the informal sector can play an important role in the formation of human capital. Fourth, the informal sector generates demand for semiskilled and unskilled labor whose supply is increasing in both relative and absolute terms and is unlikely to be absorbed by the formal sector with its increasing demands for a skilled labor force. Fifth, the informal sector is more likely to adopt appropriate

FIGURE 7.9 Youth Unemployment Rates, 1995 and 2005



Source: Adapted from United Nations, *Millennium Development Goals Report, 2006* (New York: United Nations, 2006), p. 24. Used with permission.

technologies and make use of local resources, allowing for a more efficient allocation of resources. Sixth, the informal sector plays an important role in recycling waste materials, engaging in the collection of goods ranging from scrap metals to cigarette butts, many of which find their way to the industrial sector or provide basic commodities for the poor. Finally, promotion of the informal sector would ensure an increased distribution of the benefits of development to the poor, many of whom are concentrated in the informal sector.

Promotion of the informal sector is not, however, without its disadvantages. One of the major disadvantages in promoting the informal sector lies in the strong relationship between rural-urban migration and labor absorption in the informal sector. Migrants from the rural sector have both a lower unemployment rate and a shorter waiting period before obtaining a job in the informal sector. Promoting income and employment opportunities in the informal sector could therefore aggravate the urban unemployment problem by attracting more labor than either the desirable parts of the informal or the formal sector could absorb. Furthermore, there is concern over the environmental consequences of a highly concentrated informal sector in the urban areas. Many informal-sector activities cause pollution and congestion (e.g., pedicabs) or inconvenience to pedestrians (e.g., hawkers and vendors). Moreover, increased densities in slums and low-income neighborhoods, coupled with poor

urban services, could cause enormous problems for urban areas. Any policy measures designed to promote the informal sector must be able to cope with these various problems. Finally, it is an almost universal observation that when regular formal-sector employment becomes available, many informal-sector microentrepreneurs switch sectors to take these jobs—clear evidence of "revealed preference."

There has been little discussion in the literature as to what sorts of measures might be adopted to promote the informal sector. The International Labor Organization has made some general suggestions. To begin with, governments will have to abandon their hostility toward the informal sector and adopt a more positive and sympathetic posture. For example, in Latin America, although improving in many cases, bureaucratic red tape and an inordinate number of administrative procedures needed to register a new business result in delays of up to 240 days in Ecuador, 310 days in Venezuela, and 525 days in Guatemala. Until recently, Brazil, Mexico, and Chile all required more than 20 applications before a company could be approved to do business. Such procedures not only cause excessive delays but can also inflate the costs of doing business by up to 70% annually. So informal-sector businesses simply skirt the law.

Because access to skills plays an important role in determining the structure of the informal sector, governments should facilitate training in the areas that are most beneficial to the urban economy. In this way, the government can play a role in shaping the informal sector so that it contains production and service activities that provide the most value to society. Specifically, such measures might promote legal activities and discourage illegal ones by providing proper skills and other incentives. It could also generate taxes that now go unpaid.

The lack of capital is a major constraint on activities in the informal sector. The provision of credit would therefore permit these enterprises to expand, produce more profit, and hence generate more income and employment. Microfinance institutions have been leading the way in providing enhanced credit access. Access to improved technology would have similar effects. Providing infrastructure and suitable locations for work (e.g., designating specific areas for stalls) could help alleviate some of the environmental consequences of an expanded informal sector. Most important, better living conditions must be provided, if not directly, then by promoting growth of the sector on the fringes of urban areas or in smaller towns where the population will settle close to its new area of work, away from the urban density. Promotion of the informal sector outside the urban areas may also help redirect the flow of rural-urban migration, especially if carried out in conjunction with the policies discussed later in this chapter.

Women in the Informal Sector

In some regions of the world, women predominate among rural-urban migrants and may even comprise the majority of the urban population. Though historically, many of these women were simply accompanying their spouses, a growing number of unattached women in Latin America, Asia, and Africa migrate to seek economic opportunity. With the exception of the export enclaves of East Asia and a few other cities, where everything from computers to running shoes are manufactured, few

of these migrants are able to find employment in the formal sector, which is generally dominated by men. As a consequence, women often represent the bulk of the informal-sector labor supply, working for low wages at unstable jobs with no employee or social security benefits. The increase in the number of single female migrants has also contributed to the rising proportion of urban households headed by women, which tend to be poorer, experience tighter resource constraints, and retain relatively high fertility rates. The changing composition of migration flows has important economic and demographic implications for many urban areas of the developing world.

Because members of female-headed households are generally restricted to low-productivity informal-sector employment and experience higher dependency burdens, they are more likely to be poor and malnourished and less likely to obtain formal education, health care, or clean water and sanitation. Among the Brazilian poor, for example, male-headed households are four times as likely as female-headed households to have access to government-sponsored health services. Dropout rates among children from households headed by women are much higher because they are more likely to be working to contribute to household income.

Many women run small business ventures or microenterprises that require little or no start-up capital and often involve the marketing of homemade foodstuffs and handicrafts. Though women's restricted access to capital leads to high rates of return on their tiny investments, the extremely low capital-labor ratios confine women to low-productivity undertakings. Studies in Latin America and Asia have found that where credit is available to women with informal-sector microenterprises, repayment rates have equaled or exceeded those for men (see Chapter 15). And because women are able to make more productive use of capital and start from a much lower investment base, their rates of return on investments often surpass those for men.

Despite the impressive record of these credit programs, they remain limited. The majority of institutional credit is still channeled through formal-sector agencies, and as a result, women generally find themselves ineligible for even small loans. Government programs to enhance income in poor households will inevitably neglect the neediest households so long as governments continue to focus on formal-sector employment of men and allocation of resources through formal-sector institutions. To solve the plight of poor urban women and their children, it is imperative that efforts be made to integrate women into the economic mainstream. Ensuring that women benefit from development programs will require that women's special circumstances be considered in policy design.

The legalization and economic promotion of informal-sector activities, where the majority of the urban female labor force is employed, could greatly improve women's financial flexibility and the productivity of their ventures. However, to enable women to reap these benefits, governments must repeal laws that restrict women's rights to own property, conduct financial transactions, or limit their fertility. Likewise, barriers to women's direct involvement in technical training programs and extension services must be eradicated. Finally, the provision of affordable child care and family-planning services would lighten the burden of women's reproductive roles and permit them a greater degree of economic participation.