Urbanization and the Geography of Development

J. Vernon Henderson

The World Bank
Sustainable Development Network
Urban and Disaster Risk Management Department
May 2014
Abstract

This paper focuses on three interrelated questions on urbanization and the geography of development. First, although we herald cities with their industrial bases as “engines of growth,” does industrialization in fact drive urbanization? While such relationships appear in the data, the process is not straightforward. Among developing countries, changes in income or industrialization correlate only weakly with changes in urbanization. This suggests that policy and institutional factors may also influence the urbanization process. In fact, the relationship between industrialization and urbanization is absent in Sub-Saharan Africa. Second, do development policies have a big-city bias and, if so, what does this imply for growth and inequality? Intelligent public infrastructure investment inevitably involves picking winners. One hopes that such choices are based on market indicators, such as where industry is starting to agglomerate and where there are clear needs. Yet governments seem to favor the biggest cities which in turn draw firms and migrants to these cities. To try to avoid excessive in-migration and oversized, congested cities, favored cities might adopt policies that make living conditions for migrants more unpleasant. This can result in increased inequality and social tension. Finally, the paper examines city sizes and city-size distributions. Factors determining both aspects are complex and poorly understood. It is hard to be prescriptive about either individual city sizes or overall city-size distributions. The best policies strengthen institutions in the relevant markets so that market forces can move the economy toward better outcomes.

This paper is a product of the Urban and Disaster Risk Management Department, Sustainable Development Network. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at j_henderson@brown.edu.
Urbanization and the Geography of Development

J. Vernon Henderson, Brown University and the National Bureau of Economic Research

Keywords: cities, urbanization, geography of development, industrialization, growth

JEL codes: R11, R58

This paper is part of a collection of papers prepared for the World Bank's Sixth Urban Research and Knowledge Symposium, October 2012. Financial support for this paper was provided by the Urban Development and Resilience Department and the Research Support Budget of the World Bank.
# Table of Contents

SPATIAL TRANSFORMATION AND CITIES AS ENGINES OF GROWTH ................................................................. 2

THE EMERGENT URBAN HIERARCHY .............................................................................................................. 5

  PRODUCT AND FUNCTIONAL SPECIALIZATION ............................................................................................... 5
  THE EVOLVING ROLE OF THE LARGEST CITIES ............................................................................................ 8
  ROLE OF TRANSPORT INVESTMENTS ............................................................................................................. 11

URBAN AND BIG-CITY BIAS ........................................................................................................................................ 14

  DOES BIAS EXIST? .................................................................................................................................................. 14
  WHY DO WE CARE ABOUT BIG-CITY BIAS? ....................................................................................................... 15

CITY SIZES AND THE CITY-SIZE DISTRIBUTION ......................................................................................... 17

  CITY SIZES .......................................................................................................................................................... 17
  CITY-SIZE DISTRIBUTIONS .................................................................................................................................. 20

CONCLUSIONS ..................................................................................................................................................... 22

NOTES ................................................................................................................................................................. 23

REFERENCES ...................................................................................................................................................... 24
This paper focuses on several interrelated key questions on the geography of development. Although we herald cities with their industrial bases as “engines of growth,” does industrialization in fact drive urbanization? What economic activities do cities of different sizes undertake? Does this change as countries develop? If so, what are the policy implications? Do development policies have a big-city bias? If so, what does this imply for growth and inequality, and what are appropriate place-based policies? Should countries have policies concerning optimal city sizes or city-size distributions?

Urbanization is central to the development process. Employment shifts out of agriculture into industry, and industrial production proceeds most effectively in cities, with their agglomeration economies. Cities are thus viewed as engines of growth. While such relationships appear in the data, the process is not straightforward. Among developing countries, changes in income or industrialization correlate only weakly with changes in urbanization. This suggests that policy and institutional factors may also influence the urbanization process, weakening the link between industrialization and urbanization. The relationship between industrialization and urbanization is absent in Sub-Saharan Africa.

Urban development involves a hierarchy: production patterns vary across cities by specialization and differences between bigger and smaller cities in manufacturing and service composition. What cities of different sizes produce differs over the development process. Two items are important. First, production patterns churn as cities change the goods they specialize in and produce for export, implying that cities need to be nimble in formulating policies that affect local industry competitiveness. Second, as development proceeds, industry decentralizes out of the biggest cities. Transport investment plays a key role in industrial relocation and big-city transformation, as well as in trade volumes and incomes of hinterland regions.

In many developing countries, especially predemocratization, policy makers exhibit a big-city bias. Picking “winners” is part of intelligent public infrastructure investment. Given limited resources, policy makers need to decide, for example, where to place highways, which will offer a competitive edge to at least nodal cities. One hopes that such choices are based on market indicators, such as where industry is starting to agglomerate and where there are clear needs. Still, however, deciding which places to link and where to run highways involves picking winners.

Here we focus on more explicit biases. These include offering some cities favorable access to and cost of credit in capital markets, favorable access to export licenses, better public schools, and the like. Such favoritism seems focused on the biggest cities—national capitals or homes of political elites. Favoritism tends to draw firms and migrants to these cities. To try to avoid excessive in-migration and oversized, congested cities, favored cities might adopt policies that make living conditions for migrants more unpleasant. This can result in increased inequality and social tension.
Finally, we will look at city sizes and city-size distributions. Factors determining both aspects are complex and poorly understood. It is hard to be prescriptive about either individual city sizes or overall city-size distributions. The best policies strengthen institutions in the relevant markets so that market forces can move the economy toward better outcomes. That said, the role of public-private interaction and the details of what institutional reforms different situations require are controversial.

**Spatial transformation and cities as engines of growth**

Urbanization is associated with economic growth and industrialization. According to the classic two-sector models of Harris and Todaro (1970) and Williamson (1986) and their more recent incarnations (Brueckner and Zenou [1999], for example), industrialization, perhaps export led, drives development. Such industrialization requires urbanization, drawing labor out of a passive rural sector. And such spatial reorganization of the population seems efficient: both manufacturing and services benefit with increased productivity from agglomeration economies and dense urban environments. In an endogenous growth context, human capital accumulation and rising effective technology levels can underlie this process as countries develop, releasing labor from agriculture into industry in cities. Learning and knowledge accumulation may also be more productive in cities.

This move of people out of agriculture into cities engaged in industry and service production has played out historically. There are no major developed, nonurbanized countries. And the dynamics of development seem consistent with the ideas behind the two-sector model. As urbanization proceeds, we expect countries to start with a strong divergence between the urban and rural sectors. No matter what triggers the industrialization process nationally, industrializing cities with more modern technologies and better access to foreign markets offer higher incomes than rural areas. This income gap starts to draw migrants to cities, and countries go through an intergenerational adjustment phase of young people moving from rural to urban areas. Urbanization eventually improves conditions in rural areas through better labor-to-land ratios and the agricultural modernization that occurs with improved education and national exposure to modern technologies and business methods. Urban-rural incomes then start to converge. Urban-rural consumption differences start high, but as urbanization proceeds, they disappear (figure 1). At low levels of urbanization, the urban sector’s share of income can be 2–3 times its share of national population, while at high levels of urbanization, the two shares are equal.
Asia and Sub-Saharan Africa are the two major world regions yet to fully urbanize. Urbanization in Asia is following the traditional pattern discussed above, but Sub-Saharan Africa differs. It has experienced substantial urbanization in the face of modest income growth,5 with limited industrialization6 and little technological progress in agriculture (figure 2).7

Figure 2. Urbanization, income, technology, and sector change

a. Effective technology and urbanization level

b. Income and urbanization level

c. Changes in technology and changes in urbanization
d. Urbanization level and sector composition

\[ < \text{global mean(GDP pc)} \cdot \text{GDP pc} = (\text{PPP}, \text{PWT7.0}) \]

Source: Work in progress by J. Vernon Henderson, Mark Roberts, and Adam Storeygard.

The graphs in figure 2, for developing countries defined as having less than the mean real income per capita in 1970, distinguish Sub-Saharan Africa (SSA) from the rest of the world (NSSA). Part (a) concerns a measure of effective technology: the average years of high school and college of the adult population. There, the urbanization level in 2010 strongly correlates with effective technology, and Sub-Saharan Africa follows the same pattern as the rest of the world. Part (b) reveals a strong relationship between urbanization and income per capita for the rest of the world but not for Sub-Saharan Africa. Part (c) shows the long difference from 1970–2010 for changes in effective technology and urbanization. The change relationships are positive but weaker, and for income changes not shown, they are insignificant. These three graphs together suggest two things. First, effective technology levels correlate better with and potentially lead urbanization. Second, while changes over time corroborate that urbanization relates to development, many other factors influence changes in urbanization. We will discuss some of these later.

The biggest surprise concerns part (d). For the rest of the world, urbanization is associated with a reduced role of agriculture and increased role of industry (not shown). But for Sub-Saharan Africa, that is not the case. Some African countries are 40–45 percent urban (about China’s level of urbanization), with small shares in agriculture, and others are at the same level of urbanization, with large shares. Most of Sub-Saharan Africa has yet to develop modern industry. Cities are based on natural resource development—in particular, agriculture. Preindustry urbanization seems associated with consumer cities rather than producer cities. Cities provide retail, wholesale, and personal services, where resource rents are spent. Box 1 offers evidence.
Can growth in agriculture foster urban growth absent the development of manufacturing and tradable business and financial services? Can urbanization be driven by consumer cities rather than producer cities? Jedwab (2012) examines the relationship between cocoa booms in Ghana and Côte d’Ivoire and the development of cities. Sub-Saharan Africa has yet to undergo a green revolution; fertilizer and irrigation use are low, and yields are low compared with other developing regions. Cocoa farmers in Ghana and Côte d’Ivoire operate on a 25-year cycle. New cocoa production areas arise in recently deforested areas and last for about 25 years before depletion. Then farmers move production to the next forested frontier. In both countries, production has moved successively from east to west. Jedwab uses this staged movement with district-level data for Ghana (1891–2000) and Côte d’Ivoire (1948–98) to show at the frontier how cash-crop production accounts for 60–75 percent of urbanization. This urbanization is not accompanied by increased production of manufacturing and tradable services in the frontier areas. Urban growth comes from farmers spending their incomes in local towns, trading with each other, and spending on products imported from outside the region. Towns become wholesale and retail centers, as well as producers of personal and farm services. Surprisingly, prior urbanization in earlier frontier areas is maintained even after cocoa production moves on, and these areas face income losses.

In unpublished work, Henderson and Storeygard provide corroborating evidence suggesting the effects of increased agricultural incomes on nearby cities. For 541 cities in 18 Sub-Saharan African countries, they show that positive rainfall shocks in a city’s hinterland raise city income. The idea is that rainfall increases agricultural yields and leads to enhanced urban incomes as farmers spend their incomes in cities. A one standard deviation increase in rainfall increases the intensity of “night lights” 14 percent. City lights are night lights measured daily for all points on the globe by U.S. satellites in the late evening (on the dark half of the lunar cycle for days without cloud cover). Henderson, Storeygard, and Weil (2012) show that changes in lights are a reasonable proxy for changes in income in places with weak or no income measures, and a 14 percent increase in lights suggests about a 4 percent increase in local GDP.

The emergent urban hierarchy

In this section, we cover three topics. First is the urban hierarchy that emerges in fully urbanized economies and how and why the hierarchy differs somewhat in developing countries. Second is the role of a country’s largest cities. In developed countries, this role has changed with time, and in developing countries today, that evolution is playing out much faster. Third is the effect of transport infrastructure investments on the production patterns of the biggest cities and on regional development.

Product and functional specialization

A number of studies of the United States covering 1950–70 note a high degree of sector specialization (or product specialization) in metropolitan areas, especially smaller and medium-size ones. In a city, about 60–75 percent of employment is in nontraded good activities (personal, trade, restaurant, housing, construction, and utility services), and 25–40 percent is in the production of goods for export (sold outside the city). Historically, exports were mostly manufactured products. For example, using cluster analysis of similarity in production patterns based on 299 industries and 243 metro areas for the United States in 1970, Henderson grouped cities into 15 types. A typical cluster type is automobiles, with 12 metropolitan areas in 1970 heavily engaged in producing automobiles and their parts and components. The median city in the group had 13 percent of its employment in automobile production, with the highest (Flint,
having 36 percent. The analysis also included cities engaged in textiles, apparel, aircraft, communication, pulp and paper, shipbuilding, steel, petrochemicals, industrial machinery, and leather products. There were some specialized service cities, though in 1970 they were mostly state-college and state-capital cities (for example, Champaign-Urbana or Austin). Larger cities, such as Chicago and Dallas, were still classified as “diverse manufacturing.”

These studies demonstrate that automobile and textile manufacturing have few colocation benefits: they do not learn from each other or trade in intermediate inputs. So, to exploit within industry-scale externalities and limit city sizes to reduce urban congestion and other size disamenities, workers and firms partition into different types of cities specializing in production, with city sizes tailored to the degree of agglomeration benefits in their respective production types.

Why then do we have big, diverse manufacturing cities? These could be centers of experimentation in an urban product-cycle model. According to Duranton and Puga (2001), while there are specialized cities, there are also diversified cities where potential manufacturers experiment with different technologies and products until they find what they are best suited to produce. Then, they decentralize to a specialized city. More generally, bigger cities are centers of research and development and higher technology development, where diversity and cross-sector fertilization in a Jacobs (1969) framework aids innovation and new product development. Fujita and others (2004), on Japanese electronics firms, show an innovation and production hierarchy. Branches of electronics firms doing research and development and trial production are found in larger metropolitan areas, while mass production of standard items is found in nonmetropolitan areas or offshore in developing countries. This pattern coincides with the higher return to learning in larger cities observed in the United States, even after accounting for the sorting of high-skill and high-ability workers into larger cities.¹⁰

There are many complications to this paradigm. First are changes in developed countries, such as the United States since 1970. In a study based on 2000 data rather than 1970 data, Black and Henderson (2003) show that U.S. deindustrialization has reduced the share of manufacturing cities and resulted in the development of many more service cities. These include health, entertainment, transport services, insurance, and finance cities, in addition to the traditional state-college and state-capital cities. A number of cities are transitioning out of manufacturing, and former large, diverse manufacturing cities are now more market and service oriented. Duranton and Puga (2005) show that, as cities have deindustrialized in recent years, sector specialization has declined. And as Kolko (1999) shows, the most industrialized parts of the United States are now rural (really nonurbanized) areas and smaller cities.

To some extent, larger cities now display what Duranton and Puga (2005) call “functional specialization.” Larger firms now separate production from administrative and service functions, including the widespread development of sales office networks.¹⁰ Firms now have multiple “headquarters,” mostly administrative offices working in law, advertising, marketing, or finance.
The authors show that functional specialization, or concentration in cities, has increased dramatically. This increased concentration of functions occurs only in cities with more than 1.5 million people, and especially in the key market centers with more than 5 million. As Kolko (1999) shows, tradable business and financial services in the United States concentrate in the largest metropolitan areas, and their roles decline sharply going down the urban hierarchy.

Another feature is churning—the turnover of export production activities of cities. As established in Eaton and Eckstein (1997) for France and Japan and by Black and Henderson (1999) and Duranton (2007) for the United States, there is little turbulence in city-size ranking over periods of 100 years or more. Big cities stay big, even if their economic bases change. This might be due partly to their immobile housing capital stocks, which help retain employment even in the face of large negative economic shocks. Lower house prices keep people in depressed locations.11 Contributing to the stability of city-size rankings may be the accumulated stock of location-specific general knowledge.12 However, there is enormous churning in cities’ industrial bases, as documented in Black and Henderson (1999) and Duranton (2007). As an example, Beardsell and Henderson (1999) examine the computer industry in the United States, where, with the invention of the personal computer, computer industry employment doubled over 1977–85. But by 1992, the offshoring of production had caused employment to fall back to 1977 levels. Apart from the dominant producer, San Jose, there was widespread change in dominant cities, with big declines in cities like Los Angeles, Minneapolis, and Poughkeepsie, New York (IBM), and big gains in new producers like Austin and San Diego. More generally, over a period of 50 years or so, many cities change industrial bases. Over 1963–92, the dominant U.S. cities in computers, aircraft, instruments, and electronic components all changed.

Duranton (2007) argues that churning is critical to innovation and urban growth. With national and regional cross-fertilization of ideas, a city producing an electronic auto component may invent a process relevant to a different industry—say, communications equipment. As that development takes hold and producers move to that city to learn the new process, the city transforms from an auto component city to a communications equipment city. For developing countries, there is a critical lesson, based on related literature on firms.13 Churning among firms whereby less productive, stale producers go out of business and are replaced by higher productivity start-ups is key to an industry maintaining a competitive edge.14 Policies that prop up stagnant state-owned firms and industries retard growth, as do those that try to force or induce noncompetitive private firms to stay in business.

The analogy to cities is clear. While cities need to be aware of their industrial bases to properly serve them, policies that prop up or subsidize growth of declining or stagnant local industries spell future unemployment. Rather, cities need to be nimble, ready to change public service and infrastructure provision to match the needs of new producers, especially in countries with export-led growth. A city specializing in the production of certain apparel items might find that another, lower cost country is more competitive, that tastes change, or that buyers experience economic downturns, reducing demand. The city will lose its business, and its entrepreneurs will need to
focus on another set of products. The city may thus need to alter the composition and level of services it provides to meet the needs of the new producers.

Cities in developing countries, even ones that are industrializing, experience less product and functional specialization. The basis can be history, such as Maoist planning in China, whereby major cities and provinces operated almost in autarky and, by plan, produced all products. One can also apply the Duranton and Puga (2001) product-cycle model to argue that a country’s early industrialization stages of experimentation with products and imported technologies proceed most effectively in diverse environments, especially larger cities. Also, as Donaldson (2010) points out, high transport costs lead cities to produce some consumer products inefficiently with backward technology rather than import them. As developing countries invest in transport infrastructure, the decline in transport costs makes trade cheaper. Cities can then specialize more in products for which they have a comparative advantage and import and stop producing items for which they have less advantage. Specialization will thus increase with transport investments. In the Republic of Korea, industry decentralized from the three largest cities to smaller cities and hinterland areas following massive transport and communications infrastructure investments in the early 1980s. Specialization or spatial concentration of employment across cities increased overall for manufacturing—and especially in such key sectors as heavy industry and machinery.

**The evolving role of the largest cities**

The biggest cities in a developing country tend to start off as manufacturing centers. Box 2 traces New York City’s development as a major manufacturing center in the 19th and early 20th centuries to its role today as a deindustrialized major service producer and exporter. Calcutta, Jakarta, Shanghai, and Seoul all started as major manufacturing centers for their nations. Even in the mid-1990s, Shanghai remained a key heavy manufacturer in China. As noted above, this might have occurred because these megacities are key centers for importation and adaptation of knowledge, which are best carried out in a large, diverse environment. But also, as we will discuss, public services, capital markets, and allocation of licenses may be biased toward these major centers.

**Box 2. New York City’s transformation**

New York is and has been throughout U.S. history America’s dominant city. As detailed by Glaeser (2005), New York historically had considerable natural advantages over competitor cities. It has a superb natural deepwater harbor better than Boston’s and is directly on the coast unlike Philadelphia. It is centrally located along the Eastern seaboard and historically was the center of a hub-and-spoke system including water links to the Great Lakes. Its hinterland had fairly rich agricultural land. New York grew initially as a shipping center, especially for cotton and textiles and as part of trade with the West Indies. With its harbor and central location, New York was also a main entry point for European immigrants. Many of these immigrants settled in New York’s ethnic neighborhoods, where people of like origin traded job and housing information in common cultural and linguistic neighborhoods.

In the latter half of the 19th century, a rapidly growing New York emerged as a manufacturing center. Its history of trade in textiles made it natural for developing a garment industry. Its ties to the West Indies, its natural harbor, and its central location made it the point for refining and distributing that region’s unrefined sugar. Finally, New York was the central importer of books from the United Kingdom and produced pirated versions for distribution in the
United States. Such focus on manufacturing was not unique to New York. Kolko (1995) shows that, in 1910, the four largest U.S. cities had 35 percent of their employment in manufacturing, with shares moving down the urban hierarchy to the rural sector (at 25 percent). Big cities were manufacturing engines of growth.

During this period, before rail commuting and cars and continuous process production (on a single floor), New York’s manufacturing developed in dense, five- and six-story buildings, and its population lived in similarly dense neighborhoods. Over the last 60–70 years, New York has deindustrialized and transformed into a service center. In 2002, 24 percent of employment in Manhattan was in professional, scientific, or technical services or in security and commodity contracts. From a different perspective, in 1997, while Manhattan had 1.8 percent of the nation’s total employment, it had 12 percent of national employment in financial headquarters, 12 percent in financial services, and 8 percent in business services, with 15 percent of advertising. Manhattan’s historical high density has been central to its role as a service center. Arzaghi and Henderson (2008) estimate the high productivity benefits of very close spatial proximity, or density, of producers trading information. Close spatial proximity seems to most benefit services like advertising, which then out-compete manufacturers for these locations with their expensive land and high-priced labor.

Note:

As development proceeds, manufacturing decentralizes from these large cities in two stages (as in the United States). First is decentralization to periurban or suburban areas of these cities. Second is decentralization from these metropolitan regions to hinterland areas, small cities, and rural areas. Henderson (2010) presents this process for the Republic of Korea and Japan. World Bank (2012a) also looks at the Korean example and links decentralization to the widespread transport and communications infrastructure investments that the Republic of Korea made in the early 1980s.

As another example here, I consider 108 Chinese metropolitan areas for which there are good 1990 GDP data. For the first stage, in 1990, the nine largest metropolitan areas had about 85 percent of industrial GDP produced in their center cities as opposed to periurban areas (the rest of the prefecture); by 2005, this share had fallen to just over 60 percent, in a period during which the population share of these center cities actually increased modestly. For the second, now emerging stage, in 1990 the ratio of the share of industrial GDP of the nine largest cities in all 108 cities to the share of total GDP of the nine largest in all 108 was 1.05. By 2005, this figure had fallen to about .87. That is, the nine largest are changing from being more industrial than other cities to being substantially less industrial.

Box 3 considers recent decentralization in India and consequently raises another issue. We have painted decentralization here as a positive development. However, there may be premature decentralization in India and a loss of agglomeration benefits. Firms are driven from cities because of poor environments: poorly allocated infrastructure investments, a lack of public utilities, and inappropriate land market regulations. Bertaud and Brueckner (2005) discuss how land market regulations (limiting floor-area ratios) in Mumbai have led to sprawl and inefficiently low densities near the city center. This may result in a costly lessening of agglomeration benefits.
Box 3. Industrial decentralization in India

Two studies look at recent patterns of production in Indian cities and how the patterns differ between the urban and rural sectors (box figure 1). Desmet and others (2012) study manufacturing employment growth in districts where it is concentrated and in districts where it is not. They do the same for services. The time period is short, 2000–05, but the patterns are striking. The authors fit locally a trend with error bands, and as the data move to higher intensity districts with fewer observations, the error bands widen. For manufacturing, there is strong mean reversion, whereby districts with high concentrations in 2000 grow much more slowly over 2000–05 than districts with low concentration in 2000. In fact, on average, high-concentration districts have negative growth rates and lose employment. The pattern for services is quite different. High-concentration districts have higher average growth rates than at least the middle-concentration districts. At the upper end, growth rises with concentration. Overall, this suggests decentralization of manufacturing, which we explore below, with services concentrating even more in high-intensity districts, found in the biggest cities.

Box figure 1. Indian manufacturing density, 2000–05, and Indian services density, 2001–06

Ghani, Goswami, and Kerr (2012) look more specifically at manufacturing and document the move away from cities to the rural sector. In box figure 2, the first part shows the urban sector’s decline in share of industrial output over 1989–2005, though its share of employment stabilized after 1995. The second part shows data just for the organized sector, as opposed to the informal sector. For bigger firms that dominate national output, the urban shares of plants, employment, and output of the organized sector all declined after 1995. The authors argue that the organized sector is ruralizing. But as World Bank (2012a) points out, a lot of this movement is really suburbanization—the first stage of decentralization in which firms move to rural areas on the outskirts of major metropolitan areas, with vastly cheaper land and somewhat cheaper labor. While the urban share of employment in the organized sector declined, the urban share of informal sector employment rose. The informal sector may contain smaller experimental firms, for which city location is important, as well as traditional parts suppliers, which may now face spatial mismatch as their buyers in the organized sector move away.
What drives this change in major cities from manufacturing leaders to service leaders? Several forces are at work. First, imported and adapted technologies learned in big cities become standardized, and producers benefit less from big-city knowledge spillovers. They then want to move production to less congested locations, with lower land and labor costs, especially if big cities have costly land market regulations and are poorly managed. Second, national production of traded business and financial services grows compared with other sectors, as economies mature with more developed legal, financial, and marketing sectors and with more functional specialization. Big cities are more competitive for producing services and facilitating functional specialization. Third, whether rationally (conservation of resources) or due to big-city bias, big cities tend to start with better public infrastructure provision, important for firms. When such investments become more widespread nationally, decentralization becomes possible. We illustrate next with the key investment—transport infrastructure—and discuss location issues more generally.

Role of transport investments

Transport investments play two key roles in developing countries. First, they reduce commuting costs and the costs of moving goods within cities. That reduces the need to cluster activities in and around the city center, allows city populations to spread out, and enables industry to move out of the central business district. The second is the role of intercity, interregional, and hinterland investment in transport links. Since transport infrastructure investments are costly and make up a large portion of national budgets, understanding their consequences is important for making better investment decisions.

Baum-Snow and others (2012) look at the effects of intracity transport investment in China on the move of industry and population to periurban areas. The authors econometrically estimate the impact on decentralization of having more highway rays (radial rays going out from the city center), more ring roads, and more railway rays. They examine the growth of activities in the
1990 city center (urban districts as of 1990) from 1990 to either 2005 or 2010 and find that each additional highway ray causally displaced 4 percent of the city center population to the suburbs over 1990–2010. They also find that, in China, where until very recently there was little long-distance highway shipping, rail rays play a central role in production decentralization. Each additional rail ray leads to a 26 percent reduction in city center industrial production, because suburban firms then have better access to rail nodes and sidings. More speculatively, they find that ring roads lead to major decentralization of both population and industry. Ring roads for industry link to suburban rail sidings and terminals and provide support for trucking as well.

Clearly, transport investments within cities cause them to spread out, perhaps leading to urban sprawl and a large ecological footprint. This reflects market-reoriented spatial redistribution of activities, generated by people’s demand for space and firms’ demand for cheaper land and labor on the urban fringe. But sprawl can be inefficient, especially if fuel is subsidized, encouraging more commuting, or if land is taken on the urban fringe at below-market prices (as in China), encouraging excessive taking.

The role of intercity transport investments is key but still poorly understood. It helps to provide some context. Where does economic activity—urban or rural—locate in a country? As Rappaport and Sachs (2003) indicate, most economic activity occurs on a coast or on navigable waterways, due to the important role of water transport and water’s amenity value. However, many countries (India and China, for example) developed with centers far from the coast and huge interior hinterland populations. In countries with huge interior populations, per capita income clearly declines moving away from the coast. As Bosker and others (2012) argue for China, countries with large interior populations can sustain huge hinterland cities not really served by a waterway. Still, access to domestic and international markets has a big impact on the location of economic activity and income produced by hinterland regions. Transport access raises incomes and induces in-migration in the areas it serves. Thus, there is a tradeoff of cost for maintenance of hinterland populations. If a country wishes to maintain large hinterland populations, it must invest in expensive transport infrastructure to keep these locations competitive and connected to international and coastal markets.

Over the last few years, economists have produced several studies that document the role of transport in these situations. Duranton and Turner (2012) show that transport investments increase city populations in the United States. Duranton, Morrow, and Turner (2011) show that, for the United States, improved transport networks facilitate trade: cities today that are better connected to the national highway system have larger trade flows and tend to produce heavier items. In an impressive paper, Donaldson (2010) argues that improved transport from British investment in railways in the 19th and early 20th centuries enabled Indian districts to stop producing the goods they produced inefficiently and to import them instead. These districts could then focus production on goods for which they had a comparative advantage. Donaldson estimates that the advent of a rail passing through a district raised its real income per unit of land about 17 percent. In a differently conceived paper, Banerjee, Duflo, and Qian (2012) argue that
better rail transport connections in China after the mid-19th-century advent of treaty ports have had output effects on counties that got rails similar to the output effects on those farther off the grid. Rather than try to make inferences based on historical rails in India or China, box 4 looks at trucking costs in Sub-Saharan Africa.

**Box 4. Hinterland income and the cost of transport in Sub-Saharan Africa**

Storeygard (2012) examines Sub-Saharan African countries where the primary city is on the coast (Tanzania, Mozambique, Angola, Gabon, Nigeria, Cameroon, Benin, Togo, Ghana, Cote d’Ivoire, Liberia, Sierra Leone, Guinea, Senegal, and Mauritania). He traces the distance along the best road network connecting 287 interior cities to their respective primary cities. Because data are not available to trace the time development of roads, he examines another dimension, the travel costs on those roads, inspired by two considerations. First, per-mile vehicle trucking costs have a high fuel component; the elasticity of vehicle travel costs per unit of distance with respect to fuel prices is 0.35. Second, over 2002–08, the price of oil nearly quadrupled, from $25 a barrel to $97, with a commensurate rise in local diesel fuel prices.

What is the impact of the rise in fuel prices on city incomes, as differentiated by distance? We expect places far from the coast to experience greater income losses, since their shipping costs per unit rise more. Unfortunately, we do not have measures of city income for Sub-Saharan African countries for any point in time, let alone as they vary over time. Storeygard measures income changes by changes in the intensity of city lights over 1992–2007. What does he find?

As box table 1 shows, cities with greater road distance from the coastal primary city suffered greater income losses than those closer to the coast, with substantial effects. To interpret the magnitude of the estimated coefficients, we give an example. For maximum oil price change in the time period (from $25 a barrel to $97), a one standard deviation increase in distance from the primary city lowers city lights by 23 percent. That translates roughly into a 6 percent lower GDP. Storeygard shows that the effect on connections to the primary city is greater for cities (endogenously) served more by paved roads. Paving occurred for cities more likely to be better connected to international markets, so these cities are most sensitive to changes in cost of access to the coast. For cities served more by unpaved roads, access to a local regional center is more important.

**Box table 1. Log (night lights per city) annually, 1992–2008**

<table>
<thead>
<tr>
<th>Coefficient: Road distance to primary city X oil price</th>
<th>−0.683***</th>
<th>−0.520**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation method and controls</td>
<td>Tobit; year and city fixed effects, linear city specific time trend</td>
<td>ordinary least squares; year and city fixed effects, city-specific splines on growth</td>
</tr>
<tr>
<td>n</td>
<td>263</td>
<td>263</td>
</tr>
</tbody>
</table>

*** significant at 1 percent or better; ** significant at 5 percent or better
Urban and big-city bias

Does bias exist?

Literature on urban bias argues that developing countries tend to favor the urban sector with improved terms of trade and better access to capital markets. However, the urban sector includes a diverse set of large cities and small towns. A more refined argument holds that the biggest cities—national capitals especially—receive the most favoritism. As noted, infrastructure investments in transportation and communications are necessarily spatially targeted, based as much as possible on evolving private market demand for such investments. Placement of such investments may or may not involve bias. Here we focus on bias from public intervention in capital and export markets that favor certain locations—particularly the politically connected or very biggest cities. Bias can also involve the quality of public services, such as schooling. Here, if intervention is based on correcting market failures, there is little rationale for favoring big cities, as failure in the relevant markets may be more pronounced in the rural sector than in big cities. This section presumes bias is economically inefficient and focuses on the unintended consequences of such bias for both efficiency and equity.

First, is there evidence of bias? Evidence is usually indirect, restricting lending or access to export licenses in hinterland areas, as argued for Indonesia by Henderson and Kuncoro (1996). More tellingly, Jefferson and Singh (1999) and Au and Henderson (2006a) estimate that in China in the 1990s, the rate of return to capital in the rural sector exceeded that in the urban sector (especially in bigger cities) by at least 25 percent. Of course, some government capital market programs may be geared toward the rural sector, so the assessment is difficult. While Brazil had subsidized agricultural lending programs, it likely strongly favored the biggest cities in capital markets in the nondemocratic era under state capitalism. Another issue is the degree to which public service provision is slanted toward the biggest cities, whether in showcase investments in Beijing and Shanghai or in the gap in quality of education facilities between rural and urban areas in Brazil or China.

Favoritism may result from policies based on economic beliefs that the biggest cities are the engines of economic growth, incorrectly suggesting that government-influenced investments should be directed to the biggest cities. Or it may have an element of nepotism and corruption, whereby national leaders favor the city where they live and where their family and friends have businesses.

Davis and Henderson (2003) make two important points about factors that can help offset bias. First is the role of transport investments in helping hinterland development. The authors show that the degree of primacy—the share of the urban population in a country’s largest city—is influenced by transport investments. They argue that, conditional on population and land area, a larger national road network and a larger waterway system lead to reduced primacy, as hinterland cities have better access to markets. Second, and very critically, the ability to favor the biggest cities is constrained by institutions as they develop. Conditional on population and land
area, countries that have a federal constitution show less primacy—a smaller share of the urban population in the largest city of a country. Federalism helps level the playing field across cities, giving smaller hinterland cities greater freedom to finance and provide their own public services and to have regional banks. Similarly, improvements in the democratic environment lead to decreased primacy. Democratization allows elected candidates to represent hinterland region interests in a national legislature.

Figure 3 shows how a more general measure of urban concentration is lower in more democratized states and declines as democratization proceeds. Of course, institutions are endogenous, so causal inference is tough. Shocks that affect urbanization may also affect democratization. For example, Bruckner and Ciccone (2011) argue that rainfall shocks in Sub-Saharan Africa affect trade and income, which in turn affect the opportunity costs of conflict and hence democratization. But rainfall directly affects urban-rural incomes and urbanization.

**Figure 3. The spatial Gini and political structures**

(a) 1970 degree of democracy versus 2000 Gini  


Note: The graphs look at the partial relationship between democratization and a measure of urban concentration, the spatial Gini, after factoring out national population, national income, and planned economic and federal structure influences. The spatial Gini is like an income Gini. Rank all cities in a country from smallest to largest. The spatial Lorenz curve is the cumulative share of the national urban population as we move from the smallest city to the largest. The Gini is the area between the Lorenz curve and the 45-degree line. On the 45-degree line, all cities would be of equal size in a graph of size rank against cumulative share. Democratization is measured by the polity index.

Graph (a) shows the relationship between 1970 democratization and 2000 spatial concentration. Graph (b) shows the relationship between changes in democracy and changes in spatial concentration.

**Why do we care about big-city bias?**

Favoring big cities in national markets and public service provision makes them attractive for firms—and hence for workers looking for jobs. Favored cities thus face increased in-migration,
potentially leading to excessive city sizes, as in Bangkok, Jakarta, and Mexico City. This relates to the aforementioned idea that favoritism leads to increased primacy and urban concentration and raises two questions. First, should we be concerned about excessive city sizes? Second, because excessive size increases congestion and lowers the quality of city life, residents and elites of favored cities may try to limit in-migration. Collectively, they would like to keep the favors without the added population; of course, individual firms may want to hire cheap migrant labor. What policies could residents or elites follow to restrain in-migration?

In urban models, modest increases in city size beyond an “optimal level” do not necessarily lead to substantial decreases in resident welfare. While there are increases in diseconomies, there are still scale benefits, and the net losses may be quite low. However, Henderson’s (2003) panel productivity study, the only study to look at this issue closely, implements an econometric model whereby one can calculate the best primacy level for each country (varying by country size). Best primacy maximizes productivity growth, as reflected in growth in real GDP per capita. A one standard deviation in primacy above its best level reduces a country’s annual growth 1.4 percentage points, a very large effect. Excessive size leads to wasted resources and reduced growth.

Cai’s (2006) description of Chinese policy on urbanization addresses whether residents and local policy makers resist in-migration. China had strong formal migration restrictions enforced through the household registration (hukou) system, but they were lifted in 2001. The Chinese were concerned that cities like Beijing and Shanghai would be flooded with rural immigrants. Beijing and Shanghai are certainly favored with massive infrastructure investments and in capital markets, as the role of state-owned enterprises has increased over the past decade. The articulated policy for the largest cities relative to the free-migration policy for small cities is to “raise the doorsill” to stem in-migration. “Raise the doorsill” means to provide migrants with poor living conditions and restricted access to public services. Migrants live either in dorms or, in Beijing, in about 300 urban villages scattered throughout the city and on the urban fringe, with little or no access to the formal housing market. Urban villages are still under rural governance, where the local rural residents “farm the housing”—provide essentially slum housing for migrants. These are units in which 4–6 migrants live in a single room with all cooking, washing, and toilet facilities outside. Access to state schools, despite many policy initiatives, remains severely restricted for migrants. Most either leave their children back home or send them to quasi-legal private schools that face repeated attempts to close them and have unqualified teachers and very limited facilities. Migrants are not part of social security systems in cities and have no access to local health insurance. The policy debates about the resulting increases in inequality between residents and migrants are muted by the recognition that improving conditions would only attract more migrants.

Inspired by this articulated policy for China, Feler and Henderson (2011) look at Brazil in the 1980s, before democratic reforms. National zoning regulations on formal sector housing created sizes potentially larger than what lower income residents wanted. The result was the
development of illegal (favelas) and quasi-legal (loteamentos) settlements. Loteamentos were “illegal” developments built on private land, violating national zoning laws. Cities were thus not required to service them—in particular, with piped water—forcing households to buy from expensive private vendors. Some cities were more willing and had a lower cost for servicing these migrants than others. City districts (the units largely responsible for providing public piped water) that provided services had more growth in small houses that migrants live in than did districts that denied services. A one standard deviation reduction in small houses served in the 1980s led to a 15 percentage point decline in growth in numbers of households in a district over 1991–2000, where average growth was 40 percentage points. The data suggest that higher income and already more populated districts are more likely to underservice small houses.

All this presents a dilemma. If national policy makers favor the largest cities with subsidized access to capital markets and inefficiently high comparative public investments, these cities draw more migrants, resulting in excessively large and crowded cities. But acting to stem in-migration by either making living conditions very poor or acquiescing to poor conditions increases inequality and potentially social tension.

Of course, favoritism is the heart of the problem: lack of a level playing field across cities as they compete for resources. Leveling the playing field means providing equal access to capital markets, fiscal resources, and export permissions and other licenses. But interpreting what equal access means is more difficult, as big cities have levels of development and degrees of sophistication different from rural areas. But markets can provide effective equal access. Banks may employ different screening mechanisms, information requirements, and repayment inducements in different areas. The issue is whether in the end lending across areas is geared so that rates of return on investments are equal.

City sizes and the city-size distribution

This section turns to two topics. First is the extent to which we expect individual cities, given what they produce and where they are located, to be of reasonably efficient size—and what policies that depends on. Second is the overall size distribution of cities in their respective countries and the relevant policy issues, if any.

City sizes

We have yet to talk about city-size determination, other than to suggest there may be a big-city bias that tends to draw migrants to favored cities. What institutions and market forces determine city sizes in general, regardless of where cities are in the hierarchy? The urban economic literature identifies two traditional settings, both based on how new cities are likely to form. One is self-organization, whereby individual firms and people mill about and cluster into settlements in an undirected and uncoordinated fashion—no real estate developers, no local governments investing to attract residents or subsidize businesses, and no company towns. The other is a national market in city land development, with the entrepreneurs being large real estate
developers or local governments. For example, town or village governments may decide that they want to become a more major urban center. As formulated in Henderson (1974), Helsley and Strange (2012) detail how formation of cities by self-organization of residents and firms without large-scale developers and institutions leads to cities of very inefficient sizes. They also show that industrial compositions may be very inefficient. Similarly, Behrens, Duranton, and Robert-Nicoud (2012) model significantly oversized cities under self-organization, though Henderson and Venables (2009) show that with sophisticated contracting and capital markets, in a dynamic context, cities will tend to be efficient sizes under self-organization.

Henderson and Becker (2000), building on Henderson (1974), argue that the potential poor outcomes under self-organization can be largely avoided with “free market” formation of new cities or with conversion of towns to cities led by private developers that can finance large-scale settlements. If existing cities become oversized or operate inefficiently, the threat of new, better-sized, more efficient settlements maintains market discipline. To establish new cities, small town governments must have the fiscal autonomy and ability to finance expansionary public infrastructure, and large land developers, in conjunction with private employers and housing providers, can effectively start new, large-scale settlements as demand requires. Garreau’s (1991) book on the U.S. development of “edge cities” details a large variety of partnerships between the public and private sectors, as well as purely private sector developments. Behind all these mechanisms lies the need for good local institutions—well-defined and enforceable private property rights (without risk of expropriation) and disciplined access to capital markets. Box 5 discusses aspects of the requirements as argued in World Bank (2012a,b).

Box 5. Essentials for timely city formation and expansion

World Bank (2012a,b) analyzes issues of new city formation and city expansion and discuss the history of “planned cities.” Two general issues are important for a reasonable process: how cities can finance urban infrastructure, and how real estate markets should be structured to encourage private real estate development of urban land for residential or commercial use.

When cities borrow, they offer no collateral for loans: private banks or bondholders cannot seize public property. Ensuring repayment requires a strongly regulatory and sophisticated credit rating system whereby each city borrower is subject to regulation that requires transparency and full accounting and assurance of repayment. For example, if a local government is failing to pay its debts, a higher level of government may be able to intercept monies coming into the local government and force repayment of the debts. In many countries, capital markets and regulatory systems are not mature enough for the widespread issuance of local public debt. Instead, the national government can create substitutes.

World Bank (2012b, spotlight 4) discusses the FINDETER initiative in Colombia. FINDETER is a quasi-public financial institution that intermediates between commercial banks and borrowing localities. It offers subsidized loans to banks as an inducement to lend to localities and monitors and approves applications with an eye to the proposed venture’s quality. When the bank lends, it absorbs all the risk but has the right to intercept certain city revenues (in this case, intergovernmental transfers) in event of nonrepayment. FINDETER has a AAA local credit rating from Duff & Phelps. Also discussed is an initiative in Tamil Nadu, India, through which a public-private financial intermediary issues its own bonds and lends to localities with creative ways to limit risk, involve the private sector, and ensure that projects are producing revenue to help repay loans. Its loan recovery rate is 98
percent. Of course, in some countries like China, the national government intervenes to fund local city development directly. While people point to the creation of Shenzhen as a success, there are also spectacular failures, like Ordos.

World Bank (2012a) analyzes property rights problems in India that inhibit large-scale developers. The country has limits on private assembly of land, and public taking (through eminent domain) is an opaque system tainted by years of corruption. More generally, it has poorly defined property rights and a judiciary system that operates too slowly to effectively enforce rights and decide disputes. This creates insecurity in tenure rights and increased risk to development projects. For eminent domain to work and for the local public sector to use property tax revenues, a proper system for land evaluation is needed. That requires trained appraisers, public assessments (to deter corruption), and transparency and consistency in evaluation (to get public acceptance). For large-scale developers and city land markets to work, India needs comprehensive reforms focused on well-defined and enforceable property rights and local governments that offer transparency, expertise, and limited corruption.

For developing countries, the issue is whether market formation of new cities works well enough to ensure reasonable outcomes and, if not, whether the national or regional governments can fill the gap by planning new cities. In this context, there is the need not just for existing cities to grow but for towns to expand into major urban centers. Over 1960–2000, new city formation accounted for more than 30 percent of urban population growth in developing countries. National governments do not have a great record in economic planning (hence the collapse of planned economies), and the scant evidence on new planned cities suggests that they do no better in planning city formation.

World Bank (2012b) illustrates Egypt’s poor performance in new city planning. For market formation of cities, two issues emerge. First is the national government’s willingness to allow local hinterland towns working with the private sector to develop, and thus compete with, say, the primary city. Henderson and Wang (2007) suggest that this could be linked to democratization. They show that a one standard deviation increase in the polity measure of democracy noted above increases the growth in numbers of cities in a country 37 percent more than the typical increase. The effects are robust, linking growth in numbers of cities to democratization and implied regional representation.

Second, even if national governments are willing to encourage the development of hinterland cities, there is the needed timely development of local fiscal autonomy and access to capital markets for small towns ripe for expansion and development. Perhaps as critical, land markets in developing countries are often largely informal and with limited property rights, discouraging the evolution of large-scale real estate developers. We know property rights are important on a micro scale in house and labor markets, but on a larger scale, there is no research on their effects on timely city formation and the achievement of desirable sizes. Aspects of property rights issues are also noted in box 5.

To be clear, we have little idea what the optimal size of any type of city should be. Not knowing optimal sizes, we also do not know corresponding optimal numbers. Au and Henderson (2006b), using a unique dataset and unusual context, look at the sizes of cities that maximize GDP per
employee in China in 1997, where this so-called efficient size depends on place in the urban hierarchy.

Why can we not use investigations like this to plan? First, it is not clear nationally that all cities should be at the size that maximizes GDP per worker. For example, at the sizes that maximize welfare per worker, cities with better amenities can offer higher welfare per worker than cities with poorer amenities. Basic welfare analysis tells us that more attractive cities should expand beyond the size that maximizes welfare per worker to accommodate more people than their sister cities with poorer amenities. Second, the best sizes are measured statistically with considerable error. Third, we believe that these efficient sizes are increasing with technological advances, meaning that what is efficient in 1997 is too small in 2007. In hybrid outcomes of market and planned sizes, world city sizes more than doubled over 1960–2000.26 One wants not to proscribe sizes but to provide environments for the market to drive sizes toward efficient levels.

**City-size distributions**

In 2009, McKinsey issued a report on urbanization in China calling for the development of an urban network housing up to 500 million people in cities with 20–40 million people each. That would effectively mean placing most of the urban population in megacities, as opposed to the typical worldwide fraction of well under 10 percent for metropolitan areas of more than 12 million people. Our discussion of the urban hierarchy suggests that the McKinsey report is misguided. Most production is carried out efficiently in smaller, fairly specialized cities, and the role for megacities is strictly limited, as they specialize in high-profile items with limited national demand, such as high-end advertising; theater; high-fashion apparel; and financial, legal, and other business services. For example, the United States, an already heavily service-oriented country, has only two megacities, New York and Los Angeles.

But there is a fascination with city-size distribution. Part of this focuses on two empirical regularities. First, the city-size distribution seems to closely follow a Pareto distribution; second, the relative size distribution is remarkably stable over time. Black and Henderson (2003) document this for the United States for 1910–2000. For the world, figure 4 illustrates both the Pareto shape and the stability for 1960–2000 of the relative size distribution of metropolitan areas. Given the radical changes in production patterns and the structure of the urban hierarchy in the United States over 1910–2000 or in the world since 1960, how can the relative size distribution remain the same? According to our notions of urban hierarchy, changes in economic structure should bring changes in the size distribution; indeed, as Henderson and Wang (2007) show, they do at the margin. But we rarely see massive shifts in relative urban size structure within countries over time. Why?
Figure 4. The world size distribution of cities


Note: There are 20 cells on the horizontal axis, each spaced apart by the same percentage increase in size. The vertical axis shows the share of each cell in the number of world cities for the respective year. Sizes are normalized (as is the cutoff point) by the mean size in each year to obtain a relative size distribution. The absolute size distribution is continually shifting right with technological improvements. The intent of normalizing is to compare shapes of the relative size distributions over time. The shapes are consistent with a Pareto distribution. Inspection reveals that the relative size distribution in 1960 is the same as in 2000.

A controversial paper by Gabaix (1999) sparked papers by Eeckhout (2004), Rossi-Hansberg and Wright (2007), and Duranton (2007), among others. Gabaix formulates a stochastic process whereby cities receive repeated draws (uncorrelated with size) on overall productivity or quality of life. Cities that get repeated good draws by luck emerge as big, others are more middling, and the unlucky cities are small. Under certain conditions, a Pareto size distribution theoretically emerges (“Zipf’s Law”), at least in the upper tail of the distribution. Rossi-Hansberg and Wright (2007) show that this formulation can be applied to a specialized hierarchy model, and Duranton (2007) shows it is consistent with ongoing urban innovation and churning. One disconcerting feature is that these models require cities to also churn in the long term, with cities moving up and down the hierarchy of sizes, something we do not observe in the data, at least among bigger cities. One can also formulate a simple, nonstochastic model whereby city productivity depends on a permanent draw of a production amenity (like quality of access to water transport). If those amenities are Pareto distributed, the city-size distribution will approximate a Pareto distribution as well.

All these exercises focus on a single dimension, a one-time or repeated draw of a productivity parameter. Recent work by Desmet and Rossi-Hansberg (2012) and Albouy (2008) emphasizes
the complex nature of the process. Cities get differential long-term draws on production amenities (like access to natural resources) and consumer amenities (like weather and natural recreation). They also get draws on their “culture” or how well, at least in the medium term, their public sector and institutions operate. Desmet and Rossi-Hansberg (2012) argue that these differences in the draws across cities in China profoundly influence productivity and population distribution, while in the United States their impacts are more muted.

What is the bottom line for policy? First is not to proscribe. We simply do not know enough about what is most efficient and how costly market inefficiency is compared with a bad policy error. Instead, countries need to provide the institutions and decentralized government structure to allow new cities to form in a timely manner and to allow city industries to decline and emerge. They need to facilitate local public investment to provide needed infrastructure. The size distribution that emerges will reflect underlying market forces and more likely approximate a reasonably efficient outcome than a proscriptive one.

Conclusions

Urbanization is complex. The link in developing countries between urbanization and income growth or industrialization is weaker than might be expected, suggesting that many other factors influence a country’s industrialization. In the urban hierarchy, there is ongoing churning whereby cities gain and lose comparative advantage in producing different items, perhaps because of innovation. Over time, the biggest cities deindustrialize and become service centers. Manufacturing moves to hinterland areas and smaller cities. It is difficult to evaluate whether individual cities are oversized or have the right industrial composition—or whether the city-size distribution in a country is relatively efficient. Theory provides limited direction, and solid empirics with good verification are mostly absent.

It seems that the best a country can do is create a level playing field in capital and export markets and provide institutions to foster competition and promote better outcomes. Creating a level playing field means eliminating spatial biases in policies and not targeting specific cities with favorable lending and market conditions. Providing institutions to foster competition and promote better outcomes requires well-defined property rights for developers and reasonable and structured access to capital markets for public infrastructure investments by local governments. But the development of institutions is endogenous, promoted by human capital accumulation and political events. As with firm technical efficiency, institutions advance and mature as human capital accumulates. Market-determined urban outcomes will thus improve with development. For example, after democratization in the late 1980s, Brazil worked to clean up urban messes—regularizing property rights in loterimentos and favelas, developing the institutions for disciplined access to capital markets by local governments, and removing biases in capital markets.
Notes

1 World Bank 2009.
4 Baum-Snow and Pavin 2012; Moretti 2004.
7 Binswanger and Townsend 2000.
8 For example, Alexandersson (1959) and Henderson (1988); Henderson also looks at city types in Brazil.
9 Baum-Snow and Pavin 2012; Wang 2012.
10 Holmes 2005.
12 Rauch 1993.
13 Duranton 2012.
15 Duranton 2012.
17 World Bank 2012b.
18 Kahn 2006.
19 Naughton 2007.
20 Renaud 1981.
22 See Feler (2012) on Brazil.
23 Kwong 2004.
References


