JOURNAL OF REGIONAL SCIENCE, VOL. 50, NO. 1, 2010, pp. 515–540

CITIES AND DEVELOPMENT

J. Vernon Henderson
Brown University, Department of Economics, Providence, RI 02912.
E-mail: J_Henderson@Brown.edu

ABSTRACT. This paper starts with a "primer" on what we know about the conceptual and empirical links between development and urbanization. While historical experience of developed countries is reviewed, today's rapid urbanization in developing countries offers an intense set of challenges. Rapid urbanization requires massive population movements and enormous local and inter-city infrastructure investments in a modern context of heavy government interventions in economies. This context raises under-researched issues, discussed in the second part of the paper. First concerns the spatial form of development. How much development should be focused in mega-cities, or huge urban clusters, as opposed to being more spatially dispersed, a critical question facing China and India today? How do we conceptualize and measure both the benefits and costs of increased urban concentration; and how are they linked to a country's evolving national industrial composition? Second, what is the evolution of spatial income inequality under massive rural-urban migration? Is inequality heightened today relative to the past by national government policies which "favor" certain cities and regions and by local government policies in those cities that may try to deflect migrants by offering them poor living conditions?

Urbanization and economic development are closely intertwined. While urbanization per se does not cause development, sustained economic development does not occur without urbanization. This paper starts with a "primer" on what we know about the conceptual and empirical links between development and urbanization. Empirical evidence is partly based on historical experience in developed countries such as the U.S, but I try to draw also from the experience in some countries in Latin America and East Asia, which have gone through the urbanization process and proceeded toward middle income status as countries. We then turn to an agenda that discusses key issues for which there is a very limited body of research and for which more research would inform the problems and policy issues facing today's rapidly urbanizing countries: the giants of China and India, and more generally South and South-East Asian counties, as well as those in sub-Saharan Africa.

In thinking about historical experience of today's developed countries versus the current experience in developing countries, a feature that distinguishes the two is the role that national government policy plays. Today's developing countries have either explicit national urbanization policies or policies governing resource allocation that are motivated by controlling and

Received: March 2009; revised: May 2009; accepted: June 2009.

 DOI: 10.1111/j.1467-9787.2009.00636.x

directing urbanization. Part of this arises because full urbanization—moving a country from a situation where 10–20 percent of the population is urbanized to one where 60–85 percent is urbanized—occurs now often in a span of about 30 years, as opposed to the more leisurely pace of urbanization in today's developed countries which played out over 100–150 years. Rapid urbanization is traumatic, requiring massive movement of population, replacement of traditional institutional and social structures with modern ones centered in a formal legal apparatus, and massive local and intercity infrastructure investments with the required financing mechanisms, all in a short time span of time. Having national urbanization policies is also a reflection of the worldwide increase over the last century in relative size of national governments, in terms of governments' share of national consumption and investment. There are major issues in today's urbanization process concerning the role of policy for which there seem few historical counterparts, as we will see.

Before turning to a partial review of the literature on urbanization and development, we outline the issues in the research agenda discussed in the second half of the paper, so that one can see how the points discussed in the first half set the stage for the second. In the second half, we focus on two issues. First concerns the spatial form of development. How spatially concentrated should urbanization be—how much development should be focused in megacities, or huge urban clusters, as opposed to spatially dispersed. This is critical question, facing China and India today, with fairly radical proposals on the table, which we will discuss for extraordinarily concentrated development, with huge portions of the population to potentially be housed just in megacities. The more concentrated the development the more rural populations have to uproot themselves and migrate longer distances to focal points of urbanization. Whatever an optimal pattern might be, how costly are deviations, even significant ones from that optimal pattern? How do we conceptualize and measure both the benefits and costs of increased urban concentration; and how are they linked to a country's evolving national industrial composition? Within large clusters, or megacities, what are the best forms of spatial development and types and spatial forms of infrastructure investments?

Second, what are the determinants of spatial income inequality and what is the evolution of such inequality under massive rural–urban migration? The natural evolution postulated in the Kuznets-Williamson (Williamson, 1965) hypothesis is one of first rising income inequality especially across regions in the earlier stages of development and urbanization, followed by declining inequality in the later stages. One question is whether the degree of inequality in this natural process tends to be heightened today relative to the past by national government policies, which "favor" certain cities and regions with investments at implicitly or explicitly highly subsidized rates. The idea that it is relative to the past arises again because of the potentially heightened role of governments in national markets over the last century. Does this favoritism in investment policies affect income inequality within and across regions and its evolution? A potential mechanism for increased income inequality with

national regional favoritism is local government policies in favored areas that try to restrain directly or indirectly in-migration.

This mechanism presents a related set of issues to analyze. Migration restrictions which are intended to make life difficult for migrants in favored areas to deflect them elsewhere may affect the form of urban development. For example, we will suggest that the "favela" or slum style development of Latin American cities in recent history is an example of one such restraint and one that is repeating today in China. Thus, while it is tempting to view slum development as an inevitable part of the urbanization process, due to the strain on evolving local fiscal and land market institutions in the face of rapid development, it may be in part intentional, driven by local policies which intend to restrain in-migration through offering very poor living conditions for migrants.

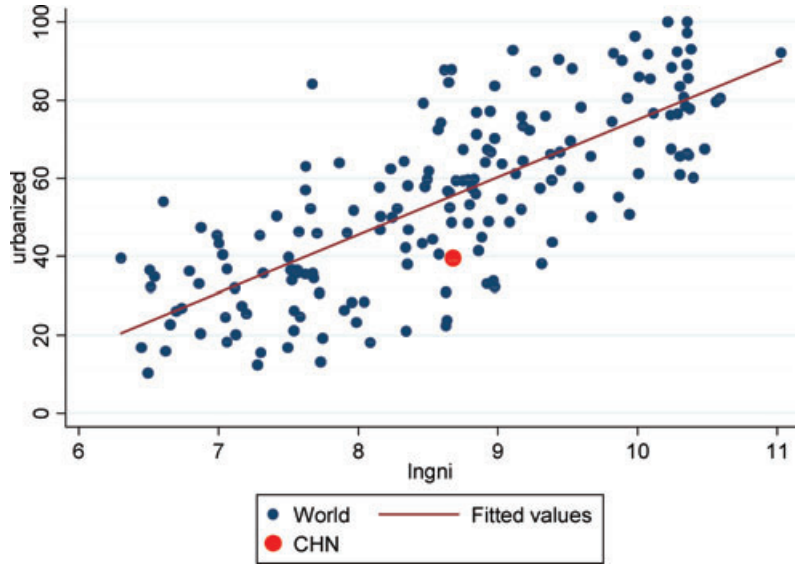
1. WHAT WE KNOW ABOUT URBANIZATION AND DEVELOPMENT

This section is a brief primer on what we know about urbanization and development, with some reference to related modeling. The primer is focused on urbanization at the national level, across the system of cites, and not on internal city form and its evolution in the development process. A more extensive treatment of both modeling and some facts may be found in Henderson (2005) for example.

The Urbanization-Growth Link

Figure 1 illustrates a relationship, which is well understood: higher levels of urbanization, as measured by the percent of the national population that lives in urban areas, is associated with increased income per capita. Figure 1 plots the percent urbanized in 2004 for countries around the world against PPP GNI p.c. (purchasing power parity GNI per capita). The level of urbanization and income per capita are highly correlated [$R^2 = 0.57$], and much of the variation around the regression line in the graph is explained by differences in definitions of urban across countries. China is noted separately for later reference.

The connection between income levels and urbanization is modeled in two interrelated ways. First are the "static" models in Harris-Todaro (1970) and the new economic geography (Krugman, 1991; Fujita, Krugman, and Venables, 1999), which stress the idea of either a technology change or productivity differential which leads to a shift of resources out of agriculture or a hinterland region into an urban sector or core region. Second are growth models, with an urban version in Henderson and Wang (2005), where in an endogenous economic growth context, a country can start off as agricultural, hit a tipping point at which urbanization starts and, then as urbanization proceeds, develop a system of cities. Growth models have as an underlying premise that development involves improvements in national technology such that labor is



Data Source: WDI, 2006.

FIGURE 1: Percent Urbanized vs. Log (PPP GNI per Capita): Countries of the World 2004.

released from agriculture to work in manufacturing and service sectors. That release occurs for two reasons. One is a demand shift toward manufacturing spurred by technological improvements in manufacturing and world demand patterns where investors are always looking for new low cost countries to house standardized, labor-intensive production of parts and components. The other is labor saving improvements in technology in agriculture. Second, manufacturing and services enjoy economies of agglomeration discussed below, which requires efficient production to be in high density locations, or cities.

While there is this association between urbanization and development, it is an equilibrium not causal relationship. There are examples of rapid urbanization, such as in sub-Saharan countries from 1970–2000 (World Development Report, 2000 (WDR)), where urbanization occurs in the face of little or no per capita income growth. The WDR, 2000, notes that urbanization without growth in some African counties in this time period may in part have been fueled by a push out of rural areas due to war and lack of upkeep of rural infrastructure. Correspondingly, econometric studies indicate that, while the form of urbanization causally affects growth, urbanization per se does not (Henderson, 2003).

Agglomeration and Development

There are two interrelated conceptual issues as to why development is linked to industrial agglomeration in cities. First is the vast conceptual and

empirical literature on localized external economies of scale, as reviewed in Duranton and Puga (2004) and Rosenthal and Strange (2004). Duranton and Puga review some of the micro-foundations underlying localized scale externalities, although the basic Marshallian notion (1890) of "mysteries. . .in the air," or very localized information spillovers (Fujita and Ogawa, 1982; Lucas and Rossi-Hansberg, 2002) remains a dominant consideration. Rosenthal and Strange suggest that, for a typical industry, doubling local own industry size leads to a 2–10 percent increase in productivity of worker; in addition, doubling city size for the same local industry size may also lead to increased productivity, especially for higher technology industries. Work on China, Brazil, and Korea suggests such scale economies are also crucial in developing countries. For example, for Korea, Henderson, Lee, and Lee (2001) find that increases in own industry size generate big productivity gains in heavy industry and the more modern manufacturing sector such as transport equipment and high tech, while magnitudes of scale externalities are smaller in traditional industries such as textiles and foods processing found in smaller towns and cities. With the exception of high tech industries, there is limited evidence that these manufacturing industries benefit from greater city size per se; rather they benefit from being in greater clusters of like, or interrelated activities. These are a powerful incentive to cluster, or agglomerate industrial activity together, as a country industrializes and develops.

In addition to this body of somewhat received knowledge, there is recent work examining particular specific, important issues. First is recent evidence that, at least certain components dissipate rapidly over space, sometimes within a matter of a few blocks (Arzaghi and Henderson, 2008; Bayer, Ross, and Topa, 2008), emphasizing the benefits of close clustering. Second, the traditional literature has struggled to separate out scale economy benefits from benefits of unobserved attributes of the local environment that are correlated with agglomeration and may also affect productivity. But there is also a selection issue: are bigger clusters really more productive or do they just attract more productive firms or more productive workers. Some work has started to look at worker selection (Combes et al., 2009).

The second strand of literature has a more dynamic focus. In economic growth models, cities are viewed as the engines of growth for an economy—dense interactive locations where knowledge is exchanged, innovations spurred, and sophisticated skills developed (Lucas, 1988; Black and Henderson, 1999a). In modeling and empirical work, this is often summarized by the notion of localized knowledge accumulation, where such accumulation is measured by the educational level of the population in a city. In a careful, although controversial study,[1] using panel data on plant productivity, Moretti (2004) suggests that, for manufacturing plants in the U.S., a 1 percentage point increase in the percent college educated in a city leads plant productivity [TFP] to rise by 0.6

[1]Using wage data, Ciccone and Peri, 2006 find little evidence of knowledge accumulation, compared to basic scale externalities.

percent. Work on R&D (Carlino, Chatterjee, and Hunt, 2007) and advertising (Arzaghi and Henderson, 2008) which, as we will note below, are industries which are important components of big city economic bases that suggests high density is a key factor in promoting localized external economies of scale and knowledge spillovers. In developing countries, bigger cities are also typically the point of technology importation and adaptation, the consequences of which have not been rigorously explored.

Theory and empirical evidence suggest scale and knowledge externalities may interact, so that scale benefits are enhanced by knowledge accumulation—information spillovers are more beneficial the more educated the population. There is no work yet that looks at this interaction directly. Unfortunately, researchers tend to look at scale externalities or knowledge effects separately. Some indirect evidence is in Henderson and Wang (2007). There the assumption is that knowledge enhanced scale externalities will lead to increased city sizes as the scale benefits of cities are enhanced relative to urban diseconomies, discussed below. Based on a panel data set of metropolitan areas worldwide, Henderson and Wang find as a causal effect that a one standard deviation increase in the percent high school educated leads to a 9 percent increase in metropolitan population at city sizes of 1 million. They also find that this effect increases with city size, suggesting that knowledge effects are important the bigger the city, perhaps due to differing economic compositions of bigger versus smaller cities and the benefits of knowledge externalities for different types of production. This finding suggests that cities increase in size over time with knowledge accumulation. The authors show that, in the right tail of the size distribution—metro areas over 100,000—metro area sizes over the last 40 years have doubled, most predominantly in countries with rapid human capital accumulation.

This discussion presumes two notions—city sizes are limited and cities have different industrial compositions and related different sizes. These notions arise from our modeling and knowledge of aspects of the urban hierarchy.

The Urban Hierarchy

Limited city sizes and specialization

Scale economies explain why industrial activity agglomerates in cities. On the other side are urban diseconomies and other factors that dissipate the benefits of agglomeration, which are why cities are limited in size. Following traditional systems of cities analysis (Henderson, 1974; see Duranton and Puga, 2004 for a review), real output per worker is postulated to be an inverted U-shape function of local scale, as measured by total urban employment. At low levels of employment, there are the high scale economies from clustering firms together as explained above, as well as exploitation of pecuniary economies in reducing transport costs of shipping intermediate inputs among firms and to consumers. However, in order to have limited city sizes (rather than one giant city in the economy), at some point as cities get big, these economies may

dissipate, costs of commuting and urban diseconomies will escalate, and the extent of a city's market will be stretched, so real income per worker peaks and then declines with further increases in city size. The new economic geography literature stresses the regional size of the market as a limiting factor to a city's size as well.

On the diseconomy side, United Nations data (UNCHS) suggest that, typically around the world, moving from a city of 250,000 to one of 2.5 million is associated empirically with an 80 percent increase in commuting times and housing rental prices (Henderson, 2002). Richardson (1987) argues from data for four developing countries that moving from a small city to a megacity raises per capita investment costs per family (in urban infrastructure) by threefold. But in contrast to the vast literature on scale externalities, good econometric evidence on urban diseconomies is lacking.

The idea of a scale economy-diseconomy trade-off leads to the notion of urban specialization. Small and medium size cities in the U.S., Korea, Japan, and other countries are highly specialized. For example, if one looks at the U.S. in 1950 or 1970, based on work by Alexandersonn (1959), Bergsman, Greenston, and Healy (1972), Henderson (1988), and Black and Henderson (2003), one can classify U.S. metro areas into steel, textile, auto, shipbuilding, aircraft, pulp and paper, petrochemical, and the like cities. This has two features. In a context where 60–70 percent of local employment is in nontraded goods and services (e.g., wholesale, retail and personal services, construction, nonmetallic materials, utilities, etc.), cities have 10–35 percent of their employment in just one subindustry that produces traded goods for export beyond the city. In small cities, much of remaining employment is in supporting or complementary subindustries to the main export one. Second, most cities have zero (or almost zero) employment in most manufacturing activities. Looking back at a more industrialized U.S. to be compared with industrializing developing countries, in 1970 for steel mills, metal stamping, engines and turbines, farm machinery, electronic computing, household appliances, photo equipment, knitting mills and so on, over 80 percent of metro areas had under 250 workers in any of these sectors and many of those had absolutely zero employment (Henderson, 1988). Ellison and Glaeser (1997) have a nice treatment of specialization and industrial concentration in the U.S. today, as do Duranton and Overman (2005) for the U.K. Lee (1998) finds similar patterns for Korea.

Why is there specialization? As noted above, the empirical literature suggests that general standardized manufacturing activity benefits from agglomeration of the own industry, not so much from the general level of local agglomeration. If only the level of own industry activity matters for productivity improvements, standardized manufacturing generally won't want to locate in large metro areas with their high wage and rent costs. They will want to locate in small, specialized cities where own industry economies of scale are maximized, relative to urban size diseconomies. High tech industries where the overall level of local agglomeration also contributes to productivity may gravitate toward larger, more diverse metro areas consistent with our earlier

TABLE 1: New York County (Manhattan): 1997

	All Industries	Head-Quarters	Financial Headquarters	Financial Services	Security Brokers	Business Services	Advertising
Share of nation's private employment	1.8%	3.0	11.7	12	25	7.5	15

Data source: Economic Census of U.S.

discussion of why industries that are undergoing rapid technological progress tend to be found in larger cities. Examples include the aircraft industry in Los Angeles or the R&D portion of electronics in Tokyo (Fujita and Ishii, 1994). This difference in location of standardized versus high tech manufacturing is related to notions of the product cycle, discussed below.

Larger Metro areas and the overall hierarchy

At the other extreme to smaller specialized manufacturing cities, and now in countries like the U.S., smaller cities specialized in specific consumer service activities, like retirement services, health services, and insurance are huge metro areas that are truly global financial and service oriented cities such as New York or London. Such cities produce almost everything including small scale personal order manufactured products, but overall a very small share of their employment is in manufacturing, and an enormous share is in key business and financial activities. That employment also accounts for a significant portion of the nation's employment in key business and financial activities. Table 1 illustrates New York City's dominance nationally in financial and business services, especially security brokers and advertising. Note that New York City has a low concentration of headquarters generally, except for financial headquarters. Banking, investment, security, advertising, legal and accounting services alone make up 25 percent of Manhattan's employment. These activities are in the largest cities because they benefit the most from overall agglomeration economies and the diverse economic environments of large cities.

Cities like New York and London have an economic-legal environment conducive to development of global services: a free standing, transparent legal system to enforce contracts; a transparent, competitive financial sector including open securities markets; transparent accounting practices and credit rating systems; and the like. Without these strong institutions, it is difficult to compete for international business. Such cities also serve as cultural centers, with a large uninhibited culture industry, with thousands of often young people engaged in fashion design, architecture, art, and theater. Some global cities such as Tokyo also have a strong high tech and R&D sector as noted above. All these industries are characterized by a strong degree of creativity and enjoy the type of urbanization economies envisioned by Jane Jacobs (1969).

The presumption is that New York and London have high concentrations of business and financial services because (1) these industries benefit the most in terms of within and cross-industry localized scale externalities, and (2) in the realization of these, scale effects benefit the most from the high density buildings and employment found in these cities. We are only in the early stages of the empirical examination of scale externalities in services. Arzaghi and Henderson (2008) examine the benefits of clustering of advertising agencies in New York City and Carlino et al. (2007) explore the role of city size and density for innovation.

In between global cities and small and medium size specialized cities are an array of large metro areas which tend to also specialize relatively but still have fairly diverse industrial bases. In these cities, the share of service activity tends to increase as metro area sizes increase. We still have a long way to go in understanding these patterns and how they change over time, as we will discuss below. But we have made some progress in overall empirical and theoretical modeling of these general notions.

For example, Au and Henderson (2006a,b) try to put together the following notions: (1) cities are limited in size with inverted U-shape real income per capita curves against city sizes, and (2) in an urban hierarchy the share of manufacturing declines as city size increases. They estimate city production functions for value added, or GDP per worker for prefecture level (or above) cities in China.[2] They structurally try to disentangle scale economies and diseconomies underlying the net scale effects embodied in an inverted U, focusing on per person commuting costs as the source of urban diseconomies, but allowing for limited market potential for cities. The urban hierarchy is represented by the manufacturing to service ratio, which they postulate decreases in bigger cities, as one moves up the urban hierarchy. They find inverted U-shape curves, which shift out, as cities become increasingly service oriented. Relative to the peak of the inverted U, where real income per capita is maximized, they find an asymmetry: for a given loss in population, there is a sharp per capita income loss from having an under-sized city, but an equal size increase in population generates a much smaller loss. We will return to this point in the second section.

Changing Industrial Composition with Development

Our discussion of the largest cities focused on New York and London. While these are now service oriented cities, that was not always the case; and,

[2]Au and Henderson (2006a,b) use 1996–1997 data on 212 prefecture level cities for China to follow this more aggregative approach looking at cities as a whole. They estimate how value-added per worker in the nonagricultural sector of the city varies with total nonagricultural employment. Determinants include the capital stock to labor ratio, share of accumulated FDI in capital stock, distance to the coast, education and scale measures. As motivation, they specify a model in which the manufacturing to service ratio represents a parameter of a structural model, in which greater uses of local services by local manufacturers indicates more "sophisticated" types of production, with greater out-sourcing of local business services of producers.

in developing countries at lower income levels, the largest cities in a country may be heavily manufacturing oriented and are also the places where the country's limited public infrastructure and high skill labor are concentrated. The concentration of manufacturing has in part to do with "learning," and the fact that a development path may emphasize growth through development of export oriented manufacturing. The largest cities, which are the most accessible to foreign investors, are the entry point for foreign technologies, which local firms are learning to adopt. They are incubators for new firms trying to discover the best product lines and production methods as modeled in Duranton and Puga (2001) and studied for Korea and Colombia by Lee (1989), Lee and Choe (1990), and Mohan (1994). The ideas and the immediate discussion to follow are closely related to the product cycle literature in international trade.

As a country develops economically, the largest cities become inefficient locations for standardized manufacturing locations for several reasons. First, firms and industries as a whole, have accomplished much learning and adoption of foreign technologies and no longer benefit so much from the learning environment of the largest cities. These cities become very expensive locations with high rents and labor costs. Infrastructure and skilled labor is in greater relative abundance in other locations. And, finally, the business service sector is expanding, demanding the large city locations, and outbidding manufacturing for central city lands in those cities.

The first step is that manufacturing decentralizes to periurban and suburban locations of the largest cities, where rents and wages are cheaper, but firms still have access to the infrastructure of the largest cities and often government offices to deal with red tape. This suburbanization process has been analyzed extensively for the U.S. historically, as well as Colombia, Indonesia, Japan, and Korea in recent decades. There is a nice discussion of some of the evidence in Mohan (1994). The final step is deindustrialization of the largest metro areas, with manufacturing moving to smaller cities and rural areas.

Table 2 shows the current spatial allocation of manufacturing and business services today versus in 1910 for the U.S. In 1910, manufacturing was more heavily concentrated in the largest cities. But as discussed above, today there is a distinct increase in manufacturing share of local employment as we move down the urban hierarchy with rural areas being the most manufacturing intense. In 1910, business services were a small part of the U.S. economy; today they are an enormous part and there is a distinct increase in their share of local employment as we move up the urban hierarchy.

Table 3 shows the decentralization process for Korea. From 1970 on in Part (a), Seoul metropolitan area's share of manufacturing in the national capital region (Kyonggi province) declined steadily, with a huge decline in the late 1970s and early 1980s. Since the early 1980s, industry has further decentralized. In Part (b), excluding the satellite cities of the three major metropolitan areas, the share in national manufacturing employment of rural areas and smaller cities rose from 25 percent in 1983 to 42 percent 10 years later, at a time when those areas were actually losing population.

TABLE 2: Manufacturing versus Business Services in the U.S. Urban Hierarchy

Metro Area Population 1995	Share of Manu. in Local Employ. 1995	Share of Business Services in Local Employ. 1995	Metro Area Size 1910	Share: Manu. in Local Employ. 1910	Share: Business Services in Local Employ. 1910
Over 2.5 m	0.14	0.21	4 largest	0.35	0.062
1–2.5 m	0.15	0.19	0.1 m–0.6 m employ.	0.35	0.051
0.5 m–1 m	0.16	0.18	Under 0.1 m employ	0.31	0.046
0.25 m–0.5 m	0.19	0.16			
Under 0.25 m	0.19	0.13			
Nonmetro	0.27	0.09	Nonmetro	0.25	0.044
Nation	0.17	0.18	Nation	0.30	0.050

Source: Kolko (1999).

TABLE 3: Manufacturing Decentralization in Korea

Part (a). Share of Seoul Metropolitan Area in Kyonggi Province*

	1970	1980	1983	1993
Population	62%	63	67	61
Manufacturing employment	76%	61	45	30

Part (b). Share of National Manufacturing Employment

	1983	1993
Seoul	21%	14
Pusan and Taegu	23	14
Satellite metro areas of Seoul, Pusan, and Taegu	30	30
Other cities, rural areas of satellite city provinces	26	42

*Excludes Inchon metro area. Source Lee (1998) and author's calculations from Korean Census data.

Source: Lee (1998).

For Japan, Table 4 shows a similar process. The share of the big three—Tokyo, Nagoya, and Osaka—in national manufacturing peaked in 1970 and then has declined steadily, with manufacturing moving to ex-urban, small city, and rural areas. There is also considerable information on Japan on how the structure of manufacturing has changed in cities. Not only has manufacturing decentralized, the manufacturing activities remaining in large cities are distinctive in terms of function within the firm. Fujita and Ishii (1994) illustrate this using the nine major electronic firms in Japan, indicating whether different types of units are found in central business districts of metro areas (CBDs), suburbs, or nonmetropolitan areas. Headquarter units of firms are found mostly in city CBDs, where business and financial services locate.

TABLE 4: Manufacturing Decentralization in Japan

	1955	1965	1970	1995
Share of Tokyo, Nagoya, and Osaka in national manufacturing GDP	39%	43	39	27
Share of big 3 in total national GDP	26%	33	34	31

Source: Based on Figure 7 in Fujita et al. (2004).

R&D and trial production facilities are found disproportionately in suburban areas of the largest metro areas, where these experimental facilities still benefit from being in large metro areas. But mass production plants are disproportionately in nonmetropolitan areas.

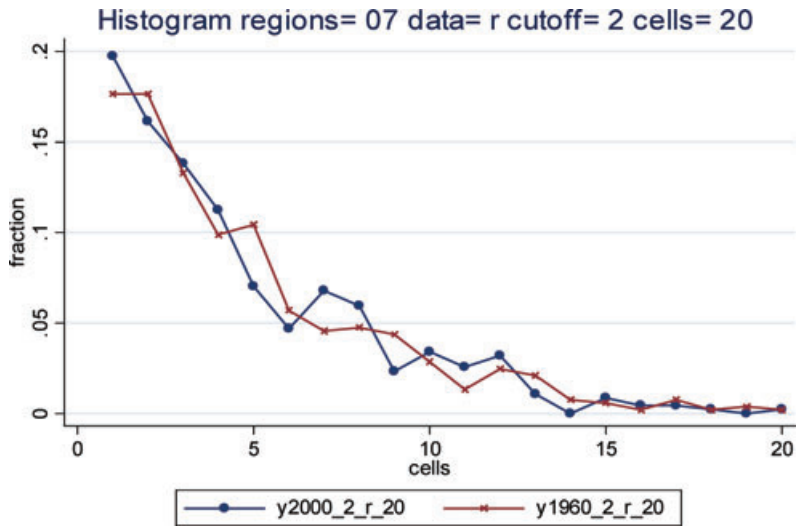
Churning and Stable Size Distributions

For developing countries, this evolution of the location of manufacturing activities raises two related but distinctly different points. First is the notion of churning of individual industries modeled in Duranton (2007) and second is the idea that large cities retain their relative size ranking over long periods of time. In terms of churning, individual cities even in developed countries change industrial activity over time. In Duranton, industries innovate in a Grossman-Helpman quality ladder model, but there can be cross-industry innovation. If industry A in city A discovers an innovation that yields the dominant technology for industry B, then industry B producers need to move to city A to access that technology, because this knowledge availability is assumed to be very localized. So with innovation industries tend to move around.

Black and Henderson (1999b, 2003) show that over the last 100 years in the U.S., although the absolute number of cities in the top 5 percentile by size has grown, cities in that category that were there 100 years ago are there today (see also Eaton and Eckstein (1997) on this point for Japan and France). The relatively big stay big. However, this is not the case for individual industries. Cities that have the highest concentrations of employment in a particular industry are much more likely to have huge relative losses in that industry overtime and to gain high concentrations of some other industry.

Despite both this churning and the deconcentration of manufacturing from the largest cities which occurs with economic development and despite the incredible technological revolutions of the last century, there is a remarkable feature about cities. The relative size distribution of cities—the share of cities which are relatively large versus small—within large countries and across countries has remained rock stable. Figure 2 illustrates the raw data for the world for 1960 versus 2000 (Henderson and Wang, 2007).

Figure 2 divides world cities into 20 cells, where across each cell there is approximately an equal percentage increase in mean city size of the cell. The

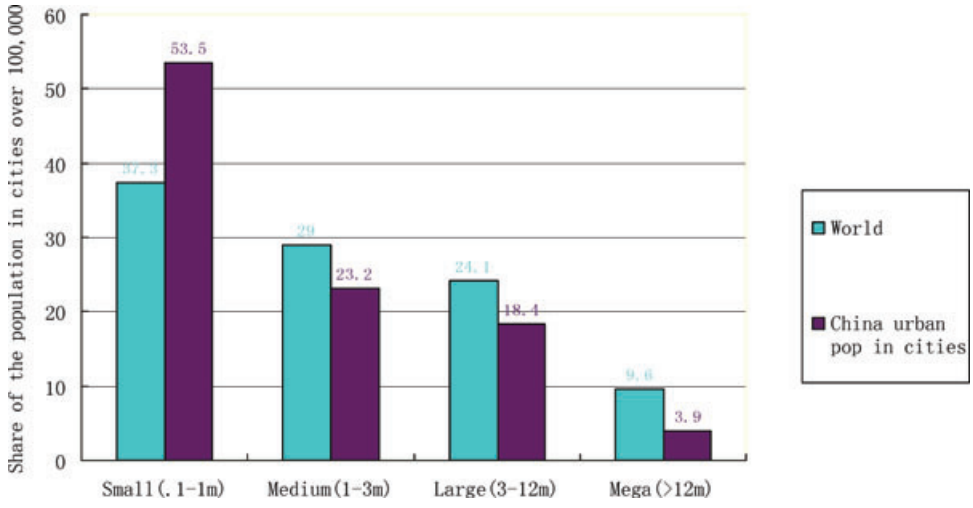


Source: Henderson and Wang, 2007.

FIGURE 2: Stable Size Distribution of Cities.

graph shows the fraction of metro areas worldwide in each cell. The graph is normalized in two ways: (1) city sizes for each year are divided by average city size in the decade, recognizing that sizes of all city cities are growing absolutely over time with knowledge accumulation; and (2) the graph represents the right tail of the size distribution of human settlements, with the 1960 minimum size being 100,000 people, and the sample chosen to represent the same size cut of the overall distribution in any decade. Note the complete overlap of these size distributions in 1960 and 2000, despite massive urbanization and technological change over the time period. A similar exercise for the U.S. for 1900 versus 1990 shows the same thing (Black and Henderson, 2003). Why might there be such stability?

This stability in part has sparked the Zipf's Law literature (Gabaix, 1999a,b) which argues that again the right tail of this overall distribution, overtime, and across countries, always follows a Pareto distribution, arguably with a constant exponent of minus one. Gabaix's argument is simple. Cities experience shocks to individual production or consumption amenities which induce population inflows or outflows. If the upshot of these stochastic events is that individual city growth rates follow Gibrat's Law (the growth rate at any instant is independent of size), then Zipf's Law will emerge. Rossi-Hansberg and Wright (2007) adapt the Black and Henderson (1999a) model to generate Zipf's Law and Duranton's (2007) churning paper generates an approximation to Zipf's Law. All these papers have one catch: to generate the result there must be an exogenous lower bound to which city sizes can fall in the event of



Source: For the world, data are based on Henderson and Wang, 2007. For urban population in cities in China, data are from the 2000 Census and kindly supplied by Yang Du of CASS.

FIGURE 3: Share in Each Category of Metropolitan Areas of the Urban Population in 2000 (in cities over 100,000).

a series of bad shocks; otherwise, typically a lognormal distribution emerges. No one has yet to satisfactorily model this lower bound—why cities can't disappear, rather like ghost towns in the American West. Durability of housing capital and reductions in housing prices in the face of negative shocks is one mechanism that would retard the disappearance of cities (Glaeser and Gyourko, 2005; Henderson and Venables, 2009).

Finally, there is another aspect to city size distributions implied in Figure 2 and illustrated in Figure 3. Most of the world's urban population is housed in metro areas that are not megacities. Even restricting the comparison to metro areas over 100,000 population (housing about 60 percent of the world's urban population), megacities account for under 10 percent of the population among these larger cities. Figure 3 also shows China compared to the world, for reference in the next section. Note China's urban population tends to live in cities that are smaller than in other countries.

Summary of Some Facts

We started this section by presenting evidence that urbanization and development go hand-in-hand, that urban scale economies in production are the basis for large cities to exist while diseconomies in living limit their sizes, and that there is an urban hierarchy where smaller cities tend to specialize in standardized production of different manufacturing and service activities,

while bigger cities have more diverse economic bases, that are focused on high tech manufacturing and certain business services. But there are dynamic patterns in urban development as well. In early stages of economic development, large cities tend to be manufacturing oriented, but as development proceeds manufacturing decentralizes to the hinterlands and the largest cities become service oriented. Finally, despite the massive urbanization of the last 50 years in much of the world, the incredible technological changes, and an enormously expanded role for government, the relative size distribution of cities has remain rock stable. And most of the world's urban population lives in metro areas and cities well under 3 million population. Armed with these facts and the notions underlying them, we now turn to a research agenda, driven by a set of very basic unanswered questions about today's urbanization in developing counties, some of it driven by the strong role of government in of today's countries that are rapidly urbanizing.

2. AN AGENDA FOR AN URBANIZING WORLD

We discuss research related to two issues. Most of this research is new and there is a long way to go in to establish solid findings to properly inform policy debate about urbanization in developing countries. The first concerns how spatially concentrated the urban population should be in a country, which has strong policy implications in terms of location of and investment in public facilities and managing urban populations.

Spatial Concentration

The literature

There is a literature which examines the determinants of spatial concentration across countries. Researchers have utilized Gini measures of concentration, Pareto parameters, and Hirschman-Herfindahl measures to describe spatial concentration. However in the past such measures were not available for a large sample of countries, since they require population data for all metropolitan areas over a certain size in a country. Instead researchers until recently relied on a measure of urban primacy, typically the share of the largest metropolitan area in the national urban population.

Central to Ades and Glaeser (1995) and Davis and Henderson (2003) is the idea that national governments tend to favor certain regions or cities of a country, typically the national capital region with a variety of advantages—better access to capital markets, better access to import and export licenses, better fiscal conditions, and disproportionately better provision of public goods. Aspects of this favoritism for Brazil, Indonesia, Korea, and China are discussed in Henderson (1988), Henderson and Kuncoro (1996) and Lee and Choe (1990), Jefferson and Singhe (1999) for example. Foremost is some type of rent-seeking behavior: bureaucrats who are the only suppliers of licenses earn more rents if they don't decentralize provision of licenses, which would

create competitors in the rent seeking-allocation business. And government officials like to improve living conditions in the places where they live. But among informed policymakers, there can also be a belief that the national capital region is more efficient place to locate production. Policymakers may tend to see the scale economies benefits of cities reflected in rising nominal wages, and less the diseconomies and rising costs that limit real income increases.

Favoritism induces firms to locate in favored locations, drawing in workers seeking jobs. As a result, in Ades and Glaeser and Davis and Henderson, favored cities tend to be much larger than nonfavored ones. Henderson and Wang (2007) also show that institutions matter—the greater fiscal decentralization the more decentralized the population. Fiscal decentralization whether through greater federalism or as part of democratization allows hinterland regions to be more fiscally independent and to set more of their own regulations, allowing them to better compete with, say, the national capital region.

All this leads to two questions. First, if favoritism leads to excessive concentration, what are the costs of that? Second, what is at the root of such costs and how does that play out in terms of urban living conditions and income inequality across regions. We turn to the second question in the next section. For the first question, in an initial attempt to tackle the problem, Henderson (2003) examines the effect of urban primacy on national economic growth. The paper finds that, for each national size and income level, there is an "optimal degree" of urban primacy, reflecting the agglomeration benefit-urban diseconomy trade-off as it plays out in terms of economic growth. Deviations from that optimal degree—up or down—are costly. A one standard deviation increase above the optimal level leads to a drop in annual growth rate by 1.4 percent, in a standard econometric cross-country-panel-data growth approach. This analysis is limited by the fact that most data points are small countries, where in fact urban primacy is a reasonable measure of concentration. For large countries with potentially a number of major metro areas, the approach is more limited, both because primacy really describes the role of only the biggest city, in a context with many regions and many important cities, and because the sample of large countries is very limited.

Policy debates

Unfortunately this issue of concentration, especially in large countries is very pressing, with countries like China and India contemplating pursuing a strategy of massive increases in urban concentration nationally. Before turning to proposals for these countries, we note what we do see today in large countries, or large regions. We take the U.S. as the example but note the same patterns play out in Brazil, Indonesia, and the EU for example. The U.S. has one huge "supercity," the New York metropolitan region which if one extends it over roughly contiguous urbanized areas into New Jersey, Pennsylvania and Connecticut has a population of 23 million (the combined statistical area [CSA], which adds together various metropolitan areas linked by some degree

cross-commuting). The next region, Los Angeles which might be described as the west coast supercity is 18 million, but after that things drop off with the next 13 regions ranging in population from 3.5 to 9.7 million, with most in the 5–6 million range. While the EU follows a similar pattern, its supercity, the London region, is maybe 15 million. And as we saw in Figure 3, most of the world's urban population does not live in cities of such size.

China and India are contemplating very different urbanization patterns where urbanization would occur mostly in a limited set of giant metropolitan regions. There is a report by the McKinsey Global Institute (2008) [http://www.mckinsey.com/mgi/publications/china_urban_summary_of_findings.asp] on China, which is receiving consideration in China. The report supports one of two scenarios, arguing they would better maximize real income per capita, relative to the current dispersed urbanization patterns where China supports just two supercities, the region for the largest of which supports 17 million people. And as we already saw in Figure 3, China has a low degree of concentration of population into large metro areas compared to the rest of the world. If current policies in China are pursued, the report sees China developing in terms on urban concentration along the lines of the U.S. and the rest of the world.

Instead, the report argues for one of the two alternative scenarios, with similar outcomes. One is a widespread supercity approach where China would develop about 15 supercities, each with an average population of around 25 million. Another is more of a "hub and spoke" system of 11 giant urban network regions averaging 60 million each. The latter is a more in the vein of looking at the northeast corridor within 150 miles of New York City. For India, a report which is considering a similar approach is in the early preparation stages. One proposal under consideration is that India would have over 60 percent of the projected urban population of 500–600 million living in 12 supercity regions, each averaging 25–30 million, with the largest ranging up to 70–80 million.

Whether constructed as supercities or giant urban network regions, we know little about the efficacy of such size urban areas. The Tokyo region at 35 million is the largest such region in the world and it is in a highly developed country with a very specific geography. A New York, London, or Tokyo supports itself by specializing in financial and business services, as well as specialized activities such as fashion apparel, high profile publishing, and theater and the arts. There is a limited demand for such specialized activities nationally; and the U.S., the EU, and Japan each support only one such size region (although maybe Los Angeles is a second for the U.S.). As we saw in Section 1, most other manufacturing and services activities seem to thrive in much smaller environments, with degrees of scale externalities that can only support limited agglomeration. So the proposed scenarios for India and China seem to advocate something that is counter to what we believe may be sustainable. They involve favoritism of not just one region, but a set of regions, at the expense of the rest of the country.

In principle, there are three ways to approach the overall issue and to try to evaluate such proposals. One is to try to compare national economic growth rates under one regime versus another—more spatially dispersed development versus development that emphasizes huge urban regions. But we have no sample of the latter. Second, as in Au and Henderson, we can try to estimate the inverted U that marks real income per capita against supercity or cluster size. For China, Au and Henderson (2006a,b) find that such inverted U's peak well before supercity sizes, but then we don't have estimates for cities which are highly business, IT, and financial service oriented. They don't yet exist in China; and as we already saw for countries like the U.S., the sample is one or at most two. That leaves the third way—trying to pull pieces together so as to assess the impacts. What are the pieces?

For different types of industries, what are the scale benefits of different clusters and when do scale economies peter out? It seems from existing evidence reviewed in Section 1 that if growth is manufacturing driven, manufacturing has historically not thrived in huge clusters. That's where special types of service activity thrive. But we know little about magnitudes of scale externalities for a whole range of business and IT services and high tech manufacturing activities, some of which a country like India could gravitate toward. Knowing more about scale externalities for such industries and how they interact with local knowledge accumulation would help inform policy formulation.

Another piece concerns urban diseconomies, about which we know even less. How do infrastructure costs and commuting or environmental degradation vary with supercity and urban cluster size? As noted in Section 1, Richardson's work based on four case studies suggests very high infrastructure investment costs of servicing large urban areas. But such an evaluation is mixed with a long standing issue with next to no economic research on it. What transport infrastructure works best in huge regions: roads versus mass transit, circumferential versus radial highways, toll-roads versus freeways, and highways versus feeder roads?

More generally, evaluation of urban diseconomies is made more complicated by the high public policy component. Tokyo is very expensive but arguably works well, so that its increase in size from 25 million to 35 million may have had little impact on the cost side. Large, but still smaller urban regions in developing countries (similar to the Mexico City region of over 20 million) have a poor quality of life, which is presumably due to having poor institutions and less human capital in urban administration. We could conceive of trying to estimate the rise in time commuting costs, the rise in per capita infrastructure costs, or the increase in environmental degradation at different levels of development, and under different institutional regimes. But apart from the conceptual issue of cause and effect (institutions and development are affected by urban form) there is a data issue. While the UN and World Bank have worked to try to develop some databases, we simply do not have good measures of relevant variables for a large sample of large urban regions (let alone the sizes which have

been proposed for India and China). But then data availability is also driven by the demand for use of such data, in this case very limited to date.

Spatial Income Inequality

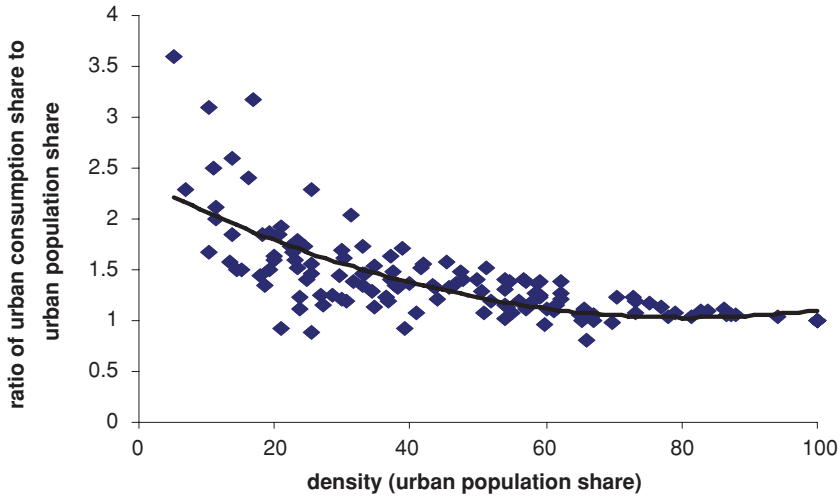
The literature

There is a vast literature documenting changes in spatial inequality for different countries in the development process (e.g., Kanbur, Venables, and Wan, 2006). Underlying this literature is the Williamson (1965) hypothesis that as a country starts to develop interregional income inequality starts to rise, peaks and then declines, with the 2009 WDR reviewing historical evidence for today's developed countries. The notion is that this is part of the development transition. The technological and trade or other policy shocks that may start a country off on a growth path emphasizing manufacturing growth and development lead to rural–urban migration and urban agglomeration in particular regions. But that is not an instantaneous process. Urban labor market opportunities are enhanced relative to rural but the migration response which closes real wage gaps between the urban and rural sector takes time. The more the urbanization process is completed, the smaller the gap. It is not just because population movements have closed the gap per se. The rural sector itself develops, and farming technology is transformed from labor to capital and high skill intensive. Barro and Sala-i-Martin (1992) argue that regional income gaps in Japan were closed by backward regions modernizing. This is also consistent with the earlier discussion which noted the deconcentration of manufacturing (but not population) into hinterland areas which occurs with development.

This idea is also supported by data from the 2009 WDR. Following a common approach today, the WDR focuses on the urban–rural real income gap. Figure 4 shows how per capita urban to rural consumption varies with the degree of urbanization in a country. Note the sharp decline to almost equality at high levels of urbanization. Unfortunately in terms of interpretation, this decline likely mixes two items: declines in wage gaps for the same skill people and changes in skill mix where rural regions relatively upgrade. Figure 5 repeats this analysis looking at the regions within three countries. Each point is a region (province/state) of a country where the vertical axis is the ratio of urban to rural income for the region and the horizontal is the degree of urbanization of the region. Again there is a sharp decline in rural–urban gaps. For India and China, two time periods are shown, where for India inequality also declined over time. What stands out in the graph is China. First, China's urban–rural gaps are huge in comparison to other countries (in a country of almost universal literacy and completion today of at least middle school). Second, those gaps have increased over time.

The rural–urban divide

The concern of the literature and of policy is that, for some countries, the high rural–urban income gaps are not just a part of some transition process, as rural labor moves to cities and the rural sector is upgraded. There is the



Source: Figure 1.11 in Grey Cover Draft of 2009 WDR, March 2008.

FIGURE 4: Urban–Rural Consumption Gaps and Urbanization Levels.

direct evidence noted earlier of policies that are biased in favor in one or two cities or regions of a country relative to others. Second are a set of *local* policies in favored regions which try to counteract the negative aspects of national favoritism, by inhibiting the flow of people into the favored area. The latter is accomplished by making living conditions for migrants into favored cities very unpleasant, as a deterrent to in-migration with a consequence of creating a divide within cities between long-term residents and migrants. Third are national policies that directly try to regulate labor flows and conditions of rural versus urban residents. The last apply mostly to some former "planned" economies, most notably China; but their efficacy has diminished under market reforms.

We use China as an extreme example of the issues and then turn to a discussion of the pressing research agenda. China used its household registration (hukou) system to control initial rural–urban migration through much of the last 30 years. Control involved two facets: first is "leave the land, not the village," meaning allow rural industrialization (town and village enterprises) but hold people in villages by having nonagricultural jobs there, rather than having migration to urban industrial agglomerations. Second, for the urbanization that inevitably occurred especially in a context where capital was disproportionately allocated to cities (Jefferson and Singhe, 1999), control was used to ensure that urbanization was localized and diffuse, spread across many cities. Hence in Figure 3 earlier, there is the high proportion of Chinese living in smaller size cities, compared to the world. These plans were enforced or implemented to some extent through the hukou system

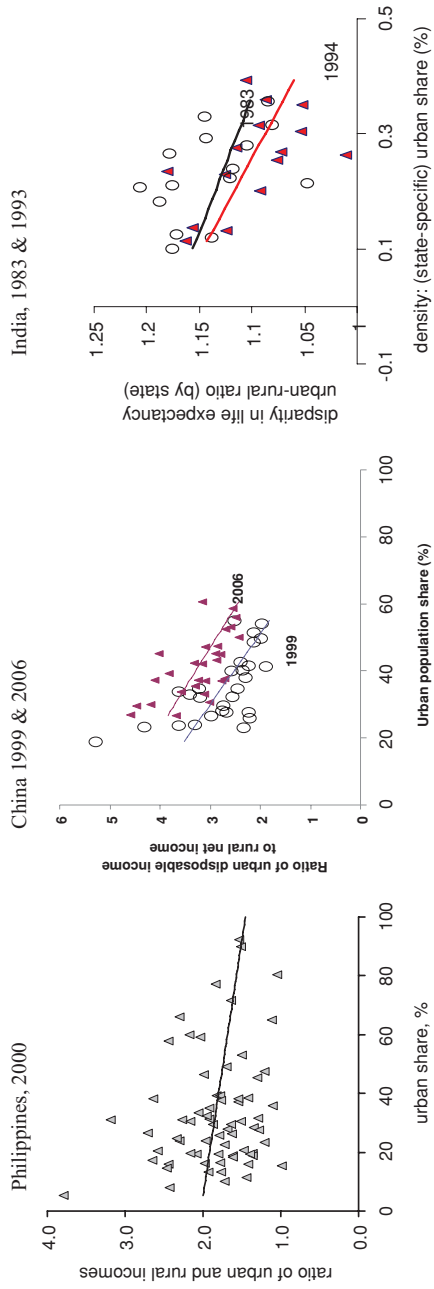


FIGURE 5: Sub-National Urban–Rural Income Disparities.

Source: Figure 1.12 in Grey Cover Draft of 2009 WDR, March 2008.

which regulated what locations people could move to, either temporarily or permanently.

The consequences of such policies have been written about extensively (e.g., Chan, 2003; Fujita et al. 2004). First is that China is probably under-urbanized; note in Figure 1 China is well below the "trend" line. Second, many Chinese cities are undersized, as analyzed in Au and Henderson (2006a,b). Third, there are the huge urban–rural income gaps illustrated in Figure 5. Such income gaps for China correspond to both consumption gaps (Knight, Shi, and Song, 2004) and gaps in marginal productivity of labor in the urban versus rural sector, controlling for skill differences (Au and Henderson, 2006a).

Today in China, the ability of the hukou system to directly limit migration has been weakened. Instead, China has adopted an explicit policy similar to what may be implicit policies else, to limit migration to certain key cities, especially the very largest and most favored cities, by making living conditions for migrants there unpleasant. This results in what is called the "double divide"—urban–rural and within cities. Migrants to the largest cities (those who are registered as citizen of a rural area) in China generally cannot obtain housing in the formal sector; they can't rent in the formal sector; and can't obtain a mortgage to purchase. They are forced to rent in "urban villages" which are pockets of crowded housing in slum-like conditions, where land is still under rural governance. Usually such land is at the city fringes, although in cites like Beijing such villages are scattered throughout the city. Second, their children have limited, expensive or no access to state schools and are forced into quasi-legal under-ground schools, with poorly educated teachers, with the schools subject to closures (Kwong, 2004). Finally, such migrants are generally excluded from health insurance, social security, job-training programs, and the like. They are rather like illegal aliens in the U.S., except at least such migrants in the U.S. do not have proscribed areas where they must live and their children can go to public schools (Wu and Rosenbaum, 2007). While we know all this about China, we know less about the long-term consequences. Will urban villages morph fully into favela-style communities not under city governance, which become havens for illegal activities and social unrest?

When we turn to the enormous slum developments in India and countries in sub-Saharan Africa, we ask the more general question of whether slum development is a natural part of development. Does the rush of migrants into cities overburden existing and not fully developed land market institutions and urban management and capabilities, so that formal sector housing and land markets can't respond with adequate supply in the intermediate run. The result is the development and acceptance of a large informal sector, with lack of public facility servicing and public services. Then as the country develops, gradually the informal sector starts to be cleaned up and serviced, and converted to the formal sector.

However, this paper stresses the idea that lack of servicing may be in part intentional, or a strategic policy choice, as in China. Residents of favored

cities do not want to see the benefits of favoritism dissipated through migrants crowding into the city. For Brazil, Feler and Henderson (2009) argue that under-servicing of housing and neighborhoods in which migrants were likely to live was a policy adopted by local districts within metro areas in the 1980's before democratization to try to deflect migrants away from their districts. Such policies were successful in retarding the rate of locality population growth; and localities in a metro area chose such policies strategically in reaction to policies of other districts nearby.

This interplay, where cities are favored by national governments in terms of capital market allocations and licensing but try to deter in-migration of those seeking the resulting job opportunities, seems a key aspect of urban development, but something we have only started to study. This interplay also seems related to the sub-Saharan paradox of the 1970–2000 time period (WDR, 2000)—rapid urbanization in the face of low or no national per capita income growth, along with the development of huge urban slums. In this case, urbanization may have been driven by bias—under-investment in rural infrastructure and spending of government resource revenues in national capital regions. And while slums may be partly the result of overburdened institutions and urban management capabilities, they could also have a strategic component, to try to discourage in-migration. How much does this double-divide contribute to national income inequality and what are the consequences for economic growth? We simply have no answers to these questions.

3. AN END NOTE

The urbanization process in today's developing countries poses critical challenges and questions, especially in the face of the huge roles that governments in some countries have assumed in this process in the 21st century. While urban populations in most countries are widely dispersed across a huge range of cities by size and industrial concentration, consideration is being given to the idea of housing the urbanizing population mostly in huge urban regions, beyond the size of anything in existence. While based on our knowledge of the urbanization process, this may seem like a bad idea, we simply know little about the urban diseconomies of living in such large cities, and our knowledge about the urban scale economies in production has significant gaps.

This enhanced role of government in the urbanization process over the years has resulted in a corresponding bias, where certain regions and cities are heavily favored in terms of capital and fiscal allocations, giving favored regions a cost advantage. This advantage could draw in an enormous number of migrants. To date in many countries favored cities seem to have resisted this in-migration by making living conditions unpleasant for migrants. This raises the issue of increasing inequality, both between people in favored versus other regions and within cities between long term residents and migrants. Again

we don't know the social consequences and the extent to which inequality is heightened by these processes.

REFERENCES

Alexandersonn, G. 1959. *The Industrial Structure of American Cities*. Lincoln: University of Nebraska Press.

Ades, A.E. and E. Glaeser. 1995. "Trade and Circuses: Explaining Urban Giants," *Quarterly Journal of Economics*, 110, 195–227.

Arzaghi, M. and J.V. Henderson. 2008. "Networking off Madison Avenue," *Review of Economic Studies*, 75, 1011–1038.

Au, C.C. and J.V. Henderson. 2006a. "How Migration Restrictions Limit Agglomeration and Productivity in China," *Journal of Development Economics*, 80, 350–388.

———. 2006b. "Are Chinese Cities Too Small," *Review of Economic Studies*, 73, 549–576.

Barro, R. and X. Sala-i-Martin. 1992. "Regional Growth and Migration: A Japan-USA Comparison," *Journal of Japanese and International Economics*, 6, 312–346.

Bayer, P., S. Ross, and G. Topa. 2008. "Place of Work, Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 116, 1150–1196.

Bergsman, J., P. Greenston, and R. Healy. 1972. "The Agglomeration Process in Urban Growth," *Urban Studies*, 9(3), 263–288.

Black, D. and J.V. Henderson. 1999a. "A Theory of Urban Growth," *Journal of Political Economy*, 107, 252–284.

———. 1999b. "Spatial Evolution of Population and Industry in the USA," *Papers and Proceedings of the American Economic Association*, 89, 321–327.

———. 2003. "Urban Evolution in the USA," *Journal of Economic Geography*, 3, 343–373.

Carlino, G., S. Chatterjee, and R. Hunt. 2007. "Urban Density and the Rate of Innovation," *Journal of Urban Economics*, 61, 389–419.

Chan, K.W. 2003. "Urbanization in China in the 1990s," *The China Review*, 3, 37–59.

Ciccone, A. and G. Peri. 2006. "Identifying Human Capital Externalities: Theory with Applications," *Review of Economic Studies*, 73, 381–412.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux. 2009. "Estimating Agglomeration Economies with History, Geology and Worker Effects," in E. Glaeser (ed.), *The Economics of Agglomeration*. Chicago: University of Chicago Press.

Davis, J. and J.V. Henderson. 2003. "Evidence on the Political Economy of the Urbanization Process," *Journal of Urban Economics*, 53, 98–125.

Duranton, G. 2007. "Urban Evolutions: the Still, the Fact and the Slow," *American Economic Review*, 97, 197–221.

Duranton, G. and H. Overman. 2005. "Testing for Localization Using Micro-Geographic Data," *Review of Economic Studies*, 72, 1077–1106.

Duranton, G. and D. Puga. 2001. "Nursery Cities," *American Economic Review*, 91, 1454–1463.

———. 2004. "Micro-Foundations of Urban Agglomeration Economies," in J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, *Vol. 4*, Amsterdam: Elsevier.

Eaton, J. and Z. Eckstein. 1997. "Cities and Growth: Evidence from France and Japan," *Regional Science and Urban Economics*, 27, 443–474.

Ellison G. and E. Glaeser. 1997. "Geographic Concentration in US Manufacturing Industries: a Dartboard Approach," *Journal of Political Economy*, 105, 889–927.

Feler, L. and J.V. Henderson. 2009. "Exclusionary Policies in Urban Development: Under-Servicing Migrant Households in Brazilian Cities," http://www.econ.brown.edu/students/leo_feler/home.htm

Fujita, M., J.V. Henderson, Y. Kanemoto, and T. Mori. 2004. "Spatial Distribution of Economic Activities in China and Japan," in J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, *Vol. 4*, Amsterdam: Elsevier.

Fujita, M. and T. Ishii. 1994. "Global Location Behavior and Organization Dynamics of Japanese Electronic Firms and Their Impact on Regional Economies," Paper presented for Prince Bertil Symposium on the Dynanmic Firm, Stockholm.
Fujita, M. and H. Ogawa. 1982. "Multiple Equilibria and Structural Transition of Non-Monocentric Configurations," *Regional Science and Urban Economics*, 12, 161–196.
Fujita, J., P. Krugman, and A.J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge: MIT Press.
Gabaix, X. 1999a. "Zipf's Law and the Growth of Cities," *American Economic Association Papers and Proceedings*, 89, 129–132.
———. 1999b. "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114, 739–767.
Glaeser, E. and J. Gyourko. 2005. "Urban Decline and Durable Housing," *Journal of Political Economy*, 114, 345–375.
Harris, J. and M. Todaro. 1970. "Migration, Unemployment and Development: A Two Sector Analysis," *American Economic Review*, 40, 126–142.
Henderson, J.V. 1974. "The Sizes and Types of Cities," *American Economic Review*, 64, 640–656.
———. 1988. *Urban Development: Theory, Fact and Illusion*, Oxford: Oxford University Press.
———. 2002. "Urban Primacy, External Costs, and Quality of Life," *Resource and Energy Economics*, 24, 95–106.
———. 2003. "The Urbanization Process and Economic Growth: The So-What Question," *Journal of Economic Growth*, 8, 47–71.
———. 2005. "Urbanization and Growth," in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Amsterdam: Elsevier.
Henderson, J.V. and A. Kuncoro. 1996. "Industrial Centralization in Indonesia," *World Bank Economic Review*, 10, 513–540.
Henderson, J.V., T. Lee, and Y.-J. Lee. 2001. "Scale Externalities in a Developing Country," *Journal of Urban Economics*, 49, 479–504.
Henderson, J.V. and A. Venables. 2009. "The Dynamics of City Formation," *Review of Economic Dynamics*, 12, 233–254.
Henderson, J.V. and H.G. Wang. 2005. "Aspects of the Rural-Urban Transformation of Countries," *Journal of Economic Geography*, 5, 23–42.
Henderson, J.V. and H.G. Wang. 2007. "Urbanization and City Growth: The Role of Institutions," *Regional Science and Urban Economics*, 37, 283–313.
Jacobs, J. 1969. *The Economy of Cities*. New York: Random House.
Jefferson, G. and I. Singhe. 1999. *Enterprise Reform in China: Ownership Transition and Performance*, New York: Oxford University Press.
Kanbur, R., A.J. Venables, and G. Wan. 2006. *Spatial Disparities in Human Development*, Helsinki: UN University Press.
Knight, J., L. Shi, and L. Song. 2004. "The Urban-Rural Divide and the Evolution of Political Economy in China," Oxford University mimeo.
Kolko, J. 1999. "Can I Get Some Service Here: Information Technology, Service Industries, and the Future of Cities," Harvard University mimeo.
Krugman, P. 1991. "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99, 483–499.
Kwong, J. 2004. "Educating Migrant Children: Negotiations between the State and Civil Society," *The China Quarterly*, 1073–1088.
Lee, K.-S. 1989. *The Location of Jobs in a Developing Metropolis: Patterns of Growth in Bogota and Cali, Colombia*. Oxford: Oxford University Press.
Lee, S.-K. and S.-C. Choe. 1990. "Changing Location Patterns of Industry and Urban Decentralization Policies in Korea," in J.K. Kwon (ed.), *Korean Economic Development*. Santa Barbara, California: Greenwood Press.
Lee, T. 1998. "Essays on Urban Industrial Development: The Case of Korea," Brown University PhD thesis.

Lucas, R.E. 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 12, 3–42.

Lucas, R.E. and E. Rossi-Hansberg. 2002. "On the Internal Structure of Cities," *Econometrica*, 70(4), 1445–1476.

Marshall, A. 1890. *The Principles of Economics*. London: McMillan.

Mohan, R. 1994. *Understanding the Developing Metropolis*, Oxford: Oxford University Press.

Moretti, E. 2004. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *American Economic Review*, 94(3), 656–690.

Richardson, H. 1987. "The Costs of Urbanization: A Four Country Comparison," *Economic Development and Cultural Change*, 33, 561–580.

Rosenthal, S. and W. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies," in J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Urban and Regional Economics*, *Vol. 4*, Amsterdam: Elsevier.

Rossi-Hansberg, E. and E.M. Wright. 2007. "Urban Structure and Growth," *Review of Economic Studies*, 74, 597–624.

Williamson, J. 1965. "Regional Inequality and the Process of National Development," *Economic Development and Cultural Change*, June, 3–45.

World Bank. 2000. *Entering the 21st Century: World Development Report 1999/2000*. Oxford: Oxford University Press.

———. 2009. *Reshaping Economic Geography*. Oxford: Oxford University Press.

Wu, W. and E. Rosenbaum. 2007. "Migration and Housing: Comparing China with the USA," in J. Logan (ed.), *Urban China in Transition*, New York: Wiley-VCH.