



CHAPTER 4

Urbanization, Agglomeration, and Economic Development

John M. Quigley

In 2007 the United Nations Population Fund released a report forecasting rapidly rising levels of urbanization over the next two decades, especially in the developing world. It noted that for the first time in history, more than half the world's population resides in urban areas. The same year UN-HABITAT issued a report highlighting the slums and deplorable living conditions in cities in developing countries. That report (UN-HABITAT 2007) estimated that by the end of 2007 there would be more than 1 billion slum dwellers, most of them living in developing countries. It claims that in many cases the economic circumstances of urban migrants are worse than those of rural peasants. In 2003 the United Nations surveyed member governments eliciting their attitudes toward urbanization. It found that the "vast majority" of these governments would have liked to shift populations back to rural areas and to stem the tide of urbanization.

Is urbanization actually bad for development? If life in urban areas were worse for urban dwellers than the life they left behind, presumably they would leave the city. So why all the fuss about urbanization and development?

The author is grateful to Patricia Annez and Robert Buckley for their comments, as well as to Vernon Henderson and Stephen Malpezzi.

This article considers the evidence on the mechanisms increasing economic efficiency in cities and examines the record of cities in facilitating economic output and in improving the consumption opportunities available to urban residents. Much of this evidence is based on observations from highly developed countries, but a growing body of evidence is based on analyses of developing countries. The evidence clearly supports the conclusion that cities are important facilitators of economic growth, increased productivity, and rising incomes in poor and rich nations alike. Policies to facilitate, not inhibit, urbanization are likely to improve economic conditions in developing countries. The analysis suggests a variety of broad policies that would improve resource allocation and increase incomes in such countries.

Why Cities?

Why do people and firms choose to locate in cities? A uniform distribution of populations over space would reduce competition for locations and thus the rents paid by households and firms, making both better off (Starrett 1974). There must therefore be compensating benefits of urban location—cost reduction, output enhancement, utility gains—to make dense location and the payment of location rent rational choices for households and firms.

The putative utility gains from urbanization have been the subject of much speculation and analysis by noneconomists. In vivid prose Jane Jacobs (1969) argues that the potential for variety in consumption is valuable to consumers. As long as the higher density of cities is associated with greater variety—in people, goods, and services—there are some utility gains to those who value diversity. These gains compensate consumers for some or all of the increased location rents in cities. It is not hard to incorporate a taste for variety into economists' models of consumer preferences (Quigley 1998 and 2001 explore some of these models).

The productivity gains, cost reductions, and output enhancements associated with collocation have been the subject of extensive analysis by economists. The historical reasons for city formation and the rationalization for the payment of location rents emphasize transport costs and internal economies of scale to the exclusion of other factors (see, for example, Hoover 1975). Transport costs refer to those incurred in delivering inputs (raw materials and labor) to industrial sites as well as the costs of delivering outputs (finished products) to local, national, and world markets. It is no accident that many of the large cities of the world developed along waterways, where ocean vessels facilitated lower-cost shipment of products to far-flung markets, or along trade routes, at *entrepôts* where the transshipment of products had already been established (Rappaport and Sachs 2003).

After the Industrial Revolution the internal scale economies arising from factories and production facilities provided a new rationale for

urbanization. The factory system replaced cottage industry. The new division of labor required larger facilities and more workers at these facilities for the production of commodities. The economies of scale in the wool industry dictated large mills near cheap water power and nearby workers. The development of denser settlements—industrial plants and tenements—allowed firms to operate at scales at which average costs could be reduced. Aggregate rents and the higher wages paid by firms to workers were more than offset by the value of increased output. The growth of many large cities in the developed world in the 19th and early 20th centuries (Detroit, Manchester, Pittsburgh) reflects the importance of internal economies of scale.

Cities and Growth

If transport costs and internal scale economies were the only economic rationale for cities, the effects of urbanization on economic growth more generally would be limited. The economic importance of cities would be determined strictly by the technologies available for transport and production. Reductions in travel cost and in the scale of the “best practice” manufacturing plant surely would have made cities less important to the health of national economies during the past century.

In fact, the importance of cities to the modern economy hardly emphasizes internal scale economies at all. Instead, the emphasis is on external effects, spillovers, and external economies of scale, factors that have all become more important with increased industrialization, technical progress, and economic development.

These external effects can be characterized along a variety of dimensions. One useful taxonomy distinguishes among productivity gains arising from specialization; from transaction costs and complementarities in production; from education, knowledge, and mimicking; and from proximity to large numbers of other economic actors.

Specialization

The gains from specialization arise because denser aggregations of urban communities (conurbations) with a larger number of firms producing in proximity can support firms that are more specialized in producing intermediate products. Specialization can lead to enhanced opportunities for cost reduction in goods production when the production of components of intermediate goods can be routinized or the components of final products mechanized or automated, for example. The gains from specialization extend to the production of services as well. Specialized legal services, for example, may be provided more efficiently by firms that concentrate in specific areas (taxation, copyright law, secured transactions, and so forth). In both intermediate goods and services, specialization increases the opportunities for cost reduction.

The potential gains from specialization are further enhanced by the opportunities for sharing inputs among firms, opportunities that are facilitated by larger and denser urban areas. Specialized services—repair, printing, advertising, communications—can be provided to a wide spectrum of producers if the density of establishments is high enough.

These external gains from specialization may arise because firms producing for final demand are themselves more spatially concentrated by industry or product, giving rise to localization economies. But they may also arise because firms producing diverse goods for final demand are more densely packed in space, giving rise to urbanization economies. In either case the environment permits more specialization among firms producing intermediate goods and services, which leads to cost reductions.

Transaction Costs and Complementarities

Externalities arising from lower transaction costs and better complementarities in production can emerge because larger urban scale can facilitate better matches between worker skills and job requirements or between intermediate goods and the production requirements for final output. In the labor market, for example, better opportunities for skill matches reduce the search costs of workers with differentiated skills and of employers with differentiated demands for labor. Complementarities in production between physical and human capital suggest that when the pool of urban workers has a larger stock of human capital, firms that expect to employ these workers will invest more in physical capital. With costly search and imperfect matching in urban labor markets, some low-skill workers will end up working with more physical capital, making them more productive and raising their incomes. The return on workers' human capital and employers' physical capital thus rises with the stock of human capital in the city, even when production at each worksite is characterized by constant returns to scale.

The same principle—externalities arising from better matches in larger urban environments—applies to specialized machines in production and to entrepreneurs in firms. Better matches can also reduce the potential losses from bankruptcy by making it easier to resell equipment.

Education, Knowledge, and Mimicking

The notion of complementarities in labor market matching can be distinguished from externalities arising from the collocation of workers with similar education and skills in dense urban areas. The effects of aggregate levels of schooling in urban areas on aggregate output may be distinguished from the effects of individuals' schooling on their individual earnings. Productivity spillovers—educated or skilled workers increasing the productivity of other workers—may arise in denser spatial environments regardless of whether the urban industrial structure is diversified or specialized. The diffusion of techniques among firms, the copying and innovation in style, and the genealogy of patents among firms are all examples of local exter-

nalities in production fostered by urban density and the concentration of skilled workers. These economies may arise with spatial concentration by industry (localization economies) or higher densities of diverse industries (urbanization economies).

The Law of Large Numbers

Considerable cost savings may arise simply from the presence of large numbers of economic actors in close proximity. To the extent that fluctuations in demand are imperfectly correlated across firms in an urban labor market, employment can be stabilized, because some firms will be hiring workers while other firms will be laying workers off. To the extent that fluctuations in demand for products are uncorrelated across buyers, firms need to carry less inventory, because some consumers will be buying while others will not. The decisions of large numbers of imperfectly correlated economic actors in close proximity can provide a form of natural insurance.

The basic insight from the law of large numbers is straightforward; it is possible to get a better estimate of the moments of a distribution with a larger sample size. This allows all economic actors to make decisions based on better information. This is true on the buying and selling sides of markets for purchasing inputs, storing intermediate products, and selling outputs.

Limitations on City Sizes

The external effects of the urban environment on productivity described above all point to larger and denser accommodations and indicate that there is a strong positive relation between urbanization and economic development. What are the limits, if any, to the extent of urbanization? What are the efficient sizes of cities? At least three sources limit the size of cities and affect the efficiency of city sizes: land and transport costs, unpriced externalities of urban life and higher densities, and explicit public policies affecting the gains from urbanization.

Land and Transport Costs

Important factors limiting city sizes arise from the same technological considerations that spawn cities in the first place. Increased housing and land prices mean that the attractiveness of larger cities for residents declines, holding the wages offered constant. (Any decline in the attractiveness of cities will, of course, be less pronounced if the consumption externalities of cities are large.) The wages offered as cities expand must increase enough to offset the higher costs workers must bear if they choose to live and work in these locations. The efficiency gains in production from higher densities must be at least as large as the increased wage payments required. Together with wages and output prices, housing and land prices limit the efficient sizes of cities.

Unpriced Externalities of Urban Density

The increased transport costs and higher densities of cities may bring their own externalities. If these are large enough, they will limit the extent of urbanization. Of course, if these externalities are unpriced, they will fail to limit urbanization sufficiently. In developed countries air pollution from vehicles is typically underpriced; until recently congestion in cities was rarely priced. Externalities from vehicle accidents are seldom priced. In developing countries there may be additional external costs of higher-density living in the form of higher risks of disease, epidemics, or fires, all of which are unpriced. To the extent that these factors are underpriced in cities, potential rural migrants do not face the marginal costs of urban life. Hence migration will be excessive and cities larger than their efficient sizes.

Explicit Policies

Explicit governmental policies, especially in developing countries, may provide strong indirect incentives affecting the extent and distribution of urbanization. Governments in many developing countries favored producers and consumers in urban areas at the expense of rural and agricultural workers (by imposing below-market prices for agricultural output and above-market prices for urban products, for example). The structural adjustment policies widely adopted since the 1980s have greatly reduced the scope for this urban bias and the distorted migration signals inherent in these subsidies, but some policies still favor certain cities, particularly national capitals. Policies favoring the locations that benefit elites and bureaucrats may be adopted as a result of rent-seeking behavior or corruption. Questionable policies may include direct public investments in plant, equipment, or infrastructure simply because certain cities are favored by elites; capital controls on investment across cities; and differences in rules imposed on cities for access to capital markets or obtaining licenses and permissions. In some countries—for example, China—restrictions include explicit limitations on labor mobility as well.

Summary

All of the factors suggest why productivity is higher in larger cities than in smaller cities. Larger cities permit greater specialization and admit more complementarities in production. They facilitate spillovers and learning within and across industries. And they facilitate sharing and risk-pooling by their very size.

Even given the potential negative externalities of larger cities, these factors suggest that real wages in larger cities in developed and developing countries will exceed those in smaller cities. Urban productivity will be higher than rural productivity, and the differential will facilitate migration from the labor-surplus hinterland to more productive urban areas.

Early models of rural–urban migration, beginning with Kuznets, recognized that the free flow of labor from unproductive agriculture to urban employment tended to equalize wages and was a vital part of the develop-

ment process. In the 1970s analysts emphasized the importance of minimum wage rules in cities and the tendency toward equalization of expected wages across sectors. These models, beginning with that of Harris and Todaro (1970), reconciled high levels of wages and worker productivity with unemployment in cities in developing countries. Inexplicably, the reasoning behind these models has been used by some to “justify” actions by governments in developing countries that limit mobility to productive cities rather than remove barriers to competition in the labor market. Indeed, agrarian romantics with an antiurban bias often compare high levels of official unemployment in cities with “official” statistics from rural areas that ignore disguised rural unemployment (see Lall, Selod, and Shalizi 2006, especially pp. 47–48, for a more balanced discussion of the issue of bias that also suggests that the concern is greatly overblown).

[[AU: please add Harris and Todaro to reference list]]

[[AU: Please confirm the cited page numbers are 47–48; your fax wasn't fully legible]]

Empirical Evidence on Productivity Gains

Despite the attention paid to agglomeration economies—going back to observations by Marshall in the 1890s—verification of efficiency gains by direct observation initially proved difficult, even using data from advanced economies. A number of early studies estimating aggregate production functions are suggestive, but most of these efforts lacked critical data (such as measures of capital stock), which made inferences about the importance of external effects problematic (see Eberts and McMillen 1999; Rosenthal and Strange 2004).

More recent work using micro data sets on firms and establishments in the United States has overcome most of these measurement problems. Henderson's (2003a, 2003b) analyses of machinery and high-tech industries, for example, tests for the presence of localization economies (agglomeration within an industry) and urbanization economies (agglomeration across industries) by estimating plant-level production functions. Using a panel of plants across counties and metropolitan areas makes it much easier to test for the effect of local conditions on the productivity of plants and their levels of output. Henderson's results show that productivity in single-establishment firms is higher as a result of localization economies.

Even with appropriate micro data, however, simple statistical models may lead to misleading inferences. If agglomeration economies do enhance firm productivity, more talented entrepreneurs will seek out these more-productive locations. More-sophisticated statistical methods are needed to account for this simultaneity. Henderson handles the problem by applying more-appropriate statistical methods of estimation in his study of high-tech and machinery industries, but the instruments he relies on (measures of the local environment) are weak, rendering the statistical results problematic.

Greenstone, Hornbeck, and Moretti (2007) solve this identification problem. They study the effects of the opening of “million-dollar plants” on the productivity of nearby plants, using a panel of establishments from the

same data source used by Henderson. For each of the million-dollar plants, they collect information on the county chosen for investment and on the county that had been under final consideration by the parent firm but not ultimately selected. The authors find clear evidence of a discontinuity in total factor productivity in plants after the opening of a large plant nearby. Total factor productivity rose in preexisting plants located in the “winning” counties but not the losing counties, confirming the existence of urbanization economies. This finding is important.

A variety of less direct approaches have been employed to make inferences about agglomeration. Rosenthal and Strange (2004), among others, study the location of firm births. To avoid the problems associated with the data on factor inputs (including the legacy of sunk capital), they investigate new establishments. This plausibly allows them to take the existing economic geography of regions as exogenous. Their results suggest that births are substantially more likely to occur where there is a concentration of firms in the same industry (see also Carlton 1983). To the extent that profit-seeking entrepreneurs are drawn to more-productive locations, this result emphasizes the importance of localization economies.

The study of the spatial distribution of wages and rents may provide indirect evidence on economies of agglomeration. The marginal product of labor and wages will be higher in more-productive regions. Analogously, locations where industrial rents are higher are those with offsetting differentials in productivity. Wheaton and Lewis (2002) use U.S. data on wages, Gabriel and Rosenthal (2004) use U.S. data on rents, and Dekle and Eaton (1999) use data from Japanese prefectures to document agglomerative economies.

Patterns of employment growth may provide indirect evidence of the importance of agglomeration. If agglomeration economies enhance productivity, more-productive regions will grow more rapidly than less-productive regions. Glaeser and others (1992) use aggregate employment data from U.S. metropolitan areas to confirm these effects. Henderson, Kuncoro, and Turner (1995) conduct a more precise test using employment in manufacturing.

Economists have studied the mechanisms transmitting these urbanization and localization efficiencies. Perhaps the clearest evidence of external effects in local labor markets comes from education and training. Early studies by Rauch (1993), testing the Lucas (1988) hypothesis, identified the external effects of schooling on wages in cross-sectional models of wage determination, using U.S. cities as units of observation. Moretti (2004) extends this analysis to explain longitudinal as well as cross-sectional variations in wages across labor markets.

Perhaps the most persuasive evidence of the importance of educational externalities comes from Moretti’s (2004) analysis of educational spillovers and productivity in the United States. This research is based on the estimation of total factor productivity and the effects of education at the level of the individual plant or establishment.

These productivity findings are confirmed, at least roughly, in a study of the service sector by Arzaghi and Henderson (2006). They analyze advertising firms in Manhattan, documenting the substantial increases in productivity attributable to the networking opportunities arising from the proximity of similar firms.

It has been widely reported that incomes have grown more rapidly in U.S. cities with high initial levels of human capital (see, for example, Glaeser and others 1992). This finding is consistent with skill acquisition and diffusion through the interaction of workers in denser urban areas (Duranton and Puga 2001; Glaeser and Maré 2001).

Lacking direct observations on workers' interactions, economists have examined one important paper trail of these interactions: data on patent applications and awards. Patent applications list the addresses of the holders of antecedent patents as well as the addresses of patent applicants. This makes possible the study of the localization of patents and the analysis of the decay of patent citations as a function of the distances between firms and between inventors (Jaffe, Trajtenberg, and Henderson 1993). This work provides explicit confirmation of the importance of geographic spillovers in the development of new knowledge.

Anthropological studies by sociologists and others have observed worker interactions in dense locations. The results of Saxenian's (1994) study of highly educated workers in Santa Clara County, California (Silicon Valley) and along Route 128 (the technical corridor outside Boston) are broadly consistent with those of quantitative investigation by economists.

Corroboration from Developing Countries

Many of the models reported in the previous section have been adapted, extended, and applied using data from developing countries. Much of this work has been pioneered by Vernon Henderson and his collaborators. Using detailed industrial census data, Henderson (1988) estimates the extent and importance of agglomeration economies in Brazil. He finds clear evidence of external economies of scale at the two-digit industry level. (The fact that in some cities a single industry is dominant, meaning that factor prices and populations are endogenously determined, is a major limitation.) This work is similar to (but much more primitive than) the work of Greenstone, Hornbeck, and Moretti (2007) using U.S. data. In a more recent analysis of city growth in Brazil, Henderson and his collaborators analyze aggregate data for 123 cities over three decades beginning in 1970 (da Mata and others 2007). Using an ambitious model of the structure of supply and demand for output at the municipal level, the authors estimate relations describing the evolution of city sizes in Brazil and their decennial growth. The empirical results indicate that increases in the sizes of local markets and their access to domestic markets have very strong effects on the growth rates of cities. Improvements in labor force quality and in the initial levels of educational attainment significantly affect economic growth, extending the conclusions of Glaeser and others (1992).

Other direct investigations of agglomeration and productivity have been undertaken in China, Indonesia, India, and the Republic of Korea. Henderson, Lee, and Lee (2001) report evidence of localization economies for Korean industry, including transport and traditional industry. They analyze metropolitan-level data for 23 Korean industries in 5 major groups between 1983 and 1993, a time of rapid deconcentration of economic activity from Seoul to smaller metropolitan areas. They estimate aggregate production functions by using census estimates of capital stock and labor and testing for the importance of the potential urbanization and localization economies provided in Korean cities. Their results confirm the importance of localization economies in Korean industry, especially in heavy industry and transportation. They also find significant localization economies in machinery and high-tech industries and, to a lesser extent, in traditional manufacturing. Lee and Zang (1988) find similar results by applying somewhat different statistical models to the same basic source of data.

In related empirical work on Indonesia, Henderson and Kuncoro (1996) report substantial localization economies for many industries and less-pronounced urbanization economies. They estimate models of the choice of location for plants and the establishment of small and medium-size firms in Java. Their results indicate that manufacturing plants are much more likely to choose locations that include mature establishments and plants in the same or related industries. These results are consistent with the work on firm births in the United States by Rosenthal and Strange (2001), who report that entrepreneurs actively seek out localization and agglomeration to improve productivity and profits.

Deichmann and others (2005) extend Henderson and Kuncoro's work by analyzing a large sample of plant locations for the entire country. Their statistical analysis documents the importance of localization economies and the influence of existing firms in the same industry in affecting location choice. The econometric results suggest the importance of existing backward linkages to suppliers in determining location choice. Urbanization economies per se are much less important.

Simulations based on these statistical results illustrate the difficulties faced by lagging regions in attracting new economic activity. Au and Henderson (2006) use aggregate data on some 285 Chinese cities to estimate the effects of urban agglomeration on productivity, using detailed data on GDP by metropolitan area in three categories. The aggregate productivity relation exhibits an inverted *U* shape in metropolitan size and scale, as expected. The estimated urban agglomeration benefits are high, and it appears that a large fraction of cities in China are undersized, as a result of migration controls imposed at the national level. These results are consistent with earlier and less complete work by Chen (1996). Some of the policy implications of this line of research are discussed in CERAP (2007).

The evidence from India includes an analysis of the relation between urban populations and total factor productivity by state and industry over

a 16-year period (Mitra 2000). Of more significance, perhaps, is the analysis of plant-level data by Lall, Koo, and Chakrovorty (2003), who use micro data on establishments from the 1998 Indian Survey of Industries to estimate the parameters of a translog cost function. They provide direct separate estimates of the elasticity of costs with respect to four different measures of agglomeration for eight industrial groupings and three size classes of plants. The results provide strong support for the importance of urbanization economies in reducing costs per unit of output. The fact that this result holds across all industries and plant size classes suggests that urbanization economies may apply to other developing economies as well.

Summary

The quality of the evidence from developing countries cited above is probably lower than that obtained from developed countries, if only because more-reliable data on economic activity are available for a longer period for developed countries. Nevertheless, the quantitative results obtained in developing countries in Asia and Latin America are remarkably consistent with those obtained in developed economies. Comparable evidence from developing countries in Africa is conspicuously absent (Collier 2007).

Urbanization and localization do support increases in productivity. Of course, it may be that the economic returns to mimicking successful ideas or investments are especially high in developing countries, where mimicking could result in too little entrepreneurial activity, a Hausmann and Rodrik (2002, 2006) make. But there is no systematic evidence that the potential returns to mimicking are greater in poorer countries than in richer ones. As the evidence on patent citations suggests, denser and more-specialized local economies may simply generate a larger stock of entrepreneurial capital to be copied.

Of course, none of this evidence establishes a tight causal link between urbanization and economic development (see Henderson 2003a, 2003b for a balanced discussion). Moreover, evidence from elsewhere suggests that urbanization is not a sufficient condition for economic development (Fay and Opal 2000). Nevertheless, there is overwhelming evidence that in both developing and industrial countries, enhanced by the localization and urbanization features of cities increases productivity.

Efficient City Sizes

Given the productivity advantages of larger cities documented in the previous section, one would expect urbanization to be a natural concomitant of increased output and well-being in developing countries. City sizes are determined by the trade-off between the increased productivity and incomes in larger cities on the one hand and the increased rent and transport costs consumers confront on the other. To the extent that rural

workers contemplating moving to cities do not adequately account for congestion, pollution, and the risk of epidemics in making their decision, cities will be “too large”—but not by much. Some of these externalities can be eliminated by improved technology or investments in public health.

It is surprising that a cohesive body of literature—or much economic literature at all—relating these externalities to levels of urbanization in developing countries does not exist. Case studies on the linkage between traffic fatalities and economic growth have been conducted (Kopits and Cropper 2005), but no case studies have examined the link between traffic fatalities and urbanization or investigated the linkage between externalities from traffic fatalities and levels of urbanization. It is relatively straightforward to estimate the correlation between the incidence of health problems and communicable diseases (such as diarrhea and tuberculosis) and urbanization at the country level or to estimate the correlation between access to water and sanitation on the one hand and urbanization on the other. Indeed, Evans (2007) reports that the rate of infant mortality in developing countries is higher in urban slums; many correlations can be investigated on line (using, for example, the World Bank’s WDI Online). Although these correlations barely hint at the causal mechanisms at work, the results are widely interpreted as if there were a causal mechanism. Absent definitive analysis, at this point one can conclude only that unpriced externalities are probably a bit more important in distorting migration flow to cities in developing countries than in developed countries. These distortions can be reduced in all countries by direct pricing or the imposition of indirect levies, such as urban property taxes.

What about the explicit policies of governments? Explicit policies of developing countries have inappropriately favored cities at the expense of agriculture, interfering with economic development. The most direct accusation of urban bias was made three decades ago, by Lipton (1976; see also Lipton 1993). Distorting price signals through macroeconomic and national trade policies that raise value added in the urban sector when value added is computed using local prices provide incentives for inefficiently high levels of urbanization; valuing urban products at inflated prices and rural products at deflated prices can make the productivity advantages attributed to cities illusory.

It is not clear how these price distortions can be measured (see Becker and Morrison 1999) or how the implications of this bias could be tested directly. But after two decades of structural adjustment policies advocated by international organizations, it is clear that in most developing countries price liberalization has caused local relative prices to converge closer to world prices, reflecting economic scarcity. Indeed, the World Bank’s 1991 treatise on urban policy documents the contemporaneous effects of structural adjustment policies in removing artificial price advantages of cities and reducing the economic circumstances of the poor in cities in developing countries.

The controversy over policies to undo distortions in relative prices seems dated. But certain limited aspects of “bias” in development policies—such as government policies that favor particular cities or regions for political or ideological reasons—may be of continuing concern.

A remarkable regularity observed across systems of cities is the rank-size rule, according to which the product of the city rank in the size distribution and the city population is roughly constant. This means that the second-largest city in a country is half the size of the first and so forth. This relation (more generally, a power relation) has proved robust over time in the United States (Dobkins and Ioannides 1998) and other countries as well as across countries (Rosen and Resnick 1980; Soo 2005). Many explanations for the general findings are purely mechanical. Fujita, Krugman, and Venables (1999) describe “nihilistic and simplistic” models that generate this pattern. Gabaix (1999) shows that if, over some range of city sizes, the expected growth rate of population and its variance are independent of size, the distribution of city sizes follows a simple power relation. Puga (1998) hypothesizes that the higher costs of spatial interaction and the less elastic labor supply in the 19th century help explain why a smaller share of the national population lives in large old European cities than in large cities in developing countries. As Puga stresses, the nature of increasing and decreasing returns to city size govern the size distribution of cities in the long run. Where, for example, there are stronger external economies of scale in cities, the distribution of city sizes will be more uneven. The exact relation between economies of scale in production and the distribution of city sizes remains elusive, however.

Considerable evidence suggests that political variables affect the distribution of city sizes. Soo’s (2005) analysis of the size distribution of cities in 73 countries suggests that political measures—dictatorial government, measures of political rights and liberties, and the length of time a nation has been independent—are more important than economic variables in explaining deviations from a common exponential relation relating city rank and size.

These results generalize the more primitive analysis by Ades and Glaeser (1995) of the primacy of a single city in national economic life. Ades and Glaeser examine variations in the national population residing in the largest city in a sample of 85 cities over 15 years. Their empirical analysis suggests that countries currently governed by dictatorships have principal cities that are about 45 percent larger than the principal cities found in democracies; democratically governed countries that were governed by dictators in the past have principal cities that are about 40 percent than those in countries with no history of dictatorship. These and similar results survive a variety of tests for causality.

Most of the discussion of “excessive” concentration in cities by economists is framed in terms of the extreme primacy of one or a few cities in many developing countries (see Henderson 1999 and the references he

cites). Surprisingly, little or none of the criticism is based on the empirical evaluation of externalities in developing countries.

Excessive concentration may be abetted by government policy. The mechanisms by which authoritative governments are able to favor particular cities or regions may be difficult to document, however. These mechanisms include the imposition of weaker benefit–cost tests on infrastructure investment or the relaxation of licensing rules in favored cities, explicit allocation of credit to favored regions, and the adoption of decisions that favor investments by public officials and cronies in national capitals. In this sense there may be an urban bias in government policy, which may adversely affect not only rural areas but also most small and medium-sized cities in developing countries.

Some Conclusions

This review and analysis of the literature reveals the strong relation between urbanization on the one hand and economic productivity and development on the other. Based on extensive analyses of data from the United States and other high-income countries and less-extensive analyses of data from developing countries, it suggests that specific mechanisms fostered by urbanization and localization of industry can affect productivity. The evidence does not conclusively show that urbanization is necessary for development or sufficient to increase output and well-being in developing countries, but the case is strong and the causal relation clear.

Urbanization and economic development are intimately related, and the concentration of resources—labor and capital—in cities is a part of this process. To the extent that movements of these factors represent a rational response to market signals about scarcity, there is no reason for concern about the size of any city or the size distribution of cities in general. To the extent that external effects, such as pollution and congestion, are unpriced in cities, conurbations will be too large, but not by much. Public concerns about pricing congested roadways and about water supplies and public health investments to decrease the chances of epidemic are well placed.

From this perspective, the concern with urban slums and low-quality housing, which impose no externalities per se, is less important. Urban poverty in developing countries is not an excuse for adopting policies that limit the extent of urbanization.

It is hard to know how important corruption and antidemocratic policies are in inhibiting or directing flows of factors to and across cities. Their existence in developed as well as developing countries provides a strong argument for allowing natural market forces to determine the spatial distribution of labor and capital. Doing so would cause both the level of urbanization and the level of economic development to increase. Increased urbanization unquestionably facilitates the development process. Explicit policies to discourage urbanization are therefore surely misguided.

References

- Ades, A., and E. Glaeser. 1995. "Trade and Circuses: Explaining Urban Giants." *Quarterly Journal of Economics* 110 (1): 195–228.
- Arzaghi, Mohammad, and J. Vernon Henderson. 2006. "Networking off Madison Avenue." Working Paper, Department of Economics, Brown University Providence, RI.
- Au, Chun-Chung, and J. Vernon Henderson. 2006. "Are Chinese Cities Too Small?" *Review of Economic Studies* 73 (3): 549–76.
- Becker, Charles M., Andrew M. Hamer, and Andrew R. Morrison. 1994. *Beyond Urban Bias in Africa: Urbanization in an Era of Structural Adjustment*. Portsmouth, NH: Heinemann.
- Becker, Charles M., and Andrew R. Morrison. 1999. "Urbanization in Transforming Economies." In *Handbook of Regional and Urban Economics*, vol. 3, ed. Edwin S. Mills and Paul Cheshire. Amsterdam: North-Holland.
- Carlton, Dennis W. 1983. "The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables." *Review of Economics and Statistics* 65 (3): 440–49.
- CERAP (China Economic Research and Advisory Programme). 2007. "Urbanization in China." Professional Report P07-001, Institute of Business and Economic Research, University of California, Berkeley.
- Chen, Y. 1996. "Impact of Regional Factors on Productivity in China." *Journal of Regional Science* 36 (3): 417–36.
- Collier, Paul. 2007. *The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done about It*. Oxford: Oxford University Press.
- Da Mata, Daniel, Uwe Deichmann, J. Vernon Henderson, Somik Vinay Lall, and Hyuong Guanghua Wang. 2007. "Determinants of City Growth in Brazil." *Journal of Urban Economics* 62 (2): 252–72.
- Deichmann, Uwe, Kai Kaiser, Somik V. Lall, and Zmarak Shalizi. 2005. "Agglomeration, Transport, and Regional Development in Indonesia." Policy Research Working Paper 3477, World Bank, Washington, DC.
- Dekle, Robert, and Jonathan Eaton. 1999. "Agglomeration and Land Rents: Evidence from the Prefectures." *Journal of Urban Economics* 46 (2): 200–14.
- Dobkins, Linda, and Yannis Ioannides. 1998. "Dynamic Evolution of the U.S. City Size Distribution." In *The Economics of Cities*, ed. Jean-Marie Huriot and Jacques-François Thisse. New York: Cambridge University Press.
- Duranton, Gilles, and Diego Puga. 2001. "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products." *American Economic Review* 91 (5): 1454–77.
- Eberts, Randall W., and Daniel P. McMillen. 1999. "Agglomeration Economics and Urban Public Infrastructure." In *Handbook of Regional and Urban Economics*, vol. 3, ed. Paul C. Cheshire and Edwin S. Mills. Amsterdam: North-Holland.
- Evans, Timothy. 2007. "Research for Urban Health: Towards a Global Agenda." Paper presented at the "Innovation for an Urban World: A Global Urban Summit," Bellagio, Italy, July 3.

- Fay, Marianne, and Charlotte Opal. 2000. "Urbanization without Growth: A Not-So-Uncommon Phenomenon." Policy Research Working Paper 2412, World Bank, Washington, DC.
- Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.
- Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114 (3): 739–67.
- Gabriel, Stuart A., and Stuart S. Rosenthal. 2004. "Quality of the Business Environment versus Quality of Life: Do Firms and Households Like the Same Cities?" *Review of Economics and Statistics* 86 (1): 438–44.
- Glaeser, Edward L., Heidi Kallal, José Scheinkman, and Andrei Schleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100 (6): 1126–52.
- Glaeser, Edward L., and David C. Maré. 2001. "Cities and Skills." *Journal of Labor Economics* 19 (2): 316–42.
- Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2007. "Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants." NBER Working Paper w13833, National Bureau of Economic Research, Cambridge, MA.
- Hausmann, Ricardo, and Dani Rodrik. 2002. "Economic Development as Self-Discovery." NBER Working Paper 8952, National Bureau of Economic Research, Cambridge, MA.
- . 2006. "Doomed to Choose: Industrial Policy as Predicament." John F. Kennedy School of Government, Harvard University, Cambridge, MA.
- Henderson, J. Vernon. 1988. *Urban Development: Theory, Fact, and Illusion*. Oxford: Oxford University Press.
- . 1999. "How Urban Concentration Affects Economic Growth." Policy Research Working Paper 2326, World Bank, Washington, D.C.
- . 2003a. "Marshall's Scale Economies." *Journal of Urban Economics* 53 (1): 1–28.
- . 2003b. "The Urbanization Process and Economic Growth: The So-What Question." *Journal of Economic Growth* 8 (1): 47–71.
- . 2005. "Urbanization and Growth." In *Handbook of Economic Growth*, ed. Philippe Aghion and Steven Durlauf. Amsterdam: North Holland Publishing Company.
- Henderson, J. Vernon, and Ari Kuncoro. 1996. "Industrial Centralization in Indonesia." *World Bank Economic Review* 10 (3): 513–40.
- Henderson, J. Vernon, Ari Kuncoro, and Matthew Turner. 1995. "Industrial Development in Cities." *Journal of Political Economy* 103 (5): 1067–85.
- Henderson, J. Vernon, Todd Lee, and Yung Joon Lee. 2001. "Scale Externalities in Korea." *Journal of Urban Economics* 49 (3): 479–504.
- Hoover, Edgar M. 1975. *An Introduction to Regional Economics*, 2nd ed. New York: Alfred A. Knopf.
- Jacobs, Jane. 1969. *The Economy of Cities*. New York: Random House.

- Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (3): 577–98.
- Kopits, Elizabeth, and Maureen Cropper. 2005. "Traffic Fatalities and Economic Growth." *Accident Analysis and Prevention* 37 (1): 169–78.
- Lall, Somik V., Jun Koo, and Sanjoy Chakravorty. 2003. "Diversity Matters: The Economic Geography of Industry Location in India." Policy Research Working Paper 3072, World Bank, Washington, D.C.
- Lall, Somik V., Harris Selod, and Zmarak Shalizi. 2006. "Rural Urban Migration in Developing Countries: A Survey of Theoretical Predictions and Empirical Findings." Policy Research Working Paper 3915, World Bank, Washington, DC.
- Lee, Y. Jung Joon, and Hyoung-soo Zang. 1998. "Urbanization and Regional Productivity in Korean Manufacturing." *Urban Studies* 35 (11): 2085–99.
- Lipton, Michael. 1976. *Why Poor People Stay Poor: Urban Bias in World Development*. Cambridge, MA: Harvard University Press.
- . 1993. "Urban Bias: Of Consequences, Classes and Causality." *Journal of Development Studies* 29 (4): 229–58.
- Lucas, Robert E., Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3–42.
- Mitra, Arup. 2000. "Total Factor Productivity Growth and Urbanization Economies: A Case of Indian Industries." *Review of Urban and Regional Development Studies* 12 (2): 97–108.
- Moretti, Enrico. 2004. "Workers' Education, Spillovers and Productivity." *American Economic Review* 94 (3): 656–90.
- Quigley, John M. 1998. "Urban Diversity and Economic Growth." *Journal of Economic Perspectives* 12 (2): 127–38.
- . 2001. "The Renaissance in Regional Research." *Annals of Regional Science* 35 (2): 167–78.
- Puga, Diego. 1998. "Urbanization Patterns: European vs. Less Developed Countries." *Journal of Regional Science* 38 (2): 231–52.
- Rappaport, Jordan, and Jeffrey Sachs. 2003. "The United States as a Coastal Nation." *Journal of Economic Growth* 8 (1): 5–46.
- Rauch, James. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics* 34 (3): 380–400.
- Rosen, Kenneth, and Mitchel Resnick. 1980. "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy." *Journal of Urban Economics* 8 (2): 165–86.
- Rosenthal, Stuart S., and William C. Strange. 2001. "The Determinants of Agglomeration." *Journal of Urban Economics* 50 (2): 191–229.
- . 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In *Handbook of Regional and Urban Economics*, vol. 4, ed. J. Vernon Henderson and Jacques-François Thisse. Amsterdam: North-Holland.

[[AU: Please verify first names in the Lee reference; your fax wasn't fully legible]]

- Saxenian, AnnaLee. 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.
- Soo, Kwok Tong. 2005. "Zipf's Law for Cities: A Cross-Country Investigation." *Regional Science and Urban Economics* 35 (3): 239–63.
- Starrett, David A. 1974. "Principles of Optimal Location in a Large Homogeneous Area." *Journal of Economic Theory* 9 (4): 418–48.
- UN (United Nations). 2003. *World Population Policies*. New York: United Nations.
- UNFPA (United Nations Population Fund). 2007. *State of World Population 2007: Unleashing the Potential of Urban Growth*. New York: United Nations.
- UN-HABITAT. 2007. *The State of the World's Cities Report 2006/2007*. London: Earthscan Publications, Ltd.
- Wheaton, William C., and Mark J. Lewis. 2002. "Urban Wages and Labor Market Agglomeration." *Journal of Urban Economics* 51 (3): 542–62.
- World Bank. 1991. *Urban Policy and Economic Development: An Agenda for the 1990s*. Washington, DC: World Bank.