

# Urbanization, Agglomeration and Economic Development

John M. Quigley  
University of California  
Berkeley, CA  
quigley@econ.berkeley.edu

## Abstract

This paper reviews the linkages between urbanization and economic development. It articulates the relationship between urban density and potential increases in productivity – through specialization, complementarities in production, through the diffusion of knowledge and mimicry, and simply through size and scale. The factors limiting the efficient sizes of cities are analyzed. The paper reviews empirical knowledge – from underdeveloped countries as well as high-income industrial societies – about the importance and magnitudes of these productivity gains. The analysis documents the close link between gains in economic efficiency and the urbanization of populations in most parts of the world.

Key words: external economy, urban density, growth, unpriced externalities

JEL codes: O14, O15, O47

March 2008

This paper was prepared for presentation at the Seminar on Growth Strategies and Urbanization, April 17, 2008. An earlier version was presented at the Commission on Growth and Development workshop. I am grateful for the comments of Patricia Annez and Robert Buckley. Portions of this draft have benefited from conversations with Vernon Henderson and Stephen Malpezzi.

## **I. Introduction**

Less than a year ago, the United Nations Population Fund released a report forecasting greatly increased levels of urbanization during the next two decades, especially in the developing world (United Nations, 2007a). The report declared that for the first time in recorded history, more than half the world's population resides in urban - - not rural -- areas. At roughly the same time, another agency of the United Nations (UN-Habitat) issued a report highlighting the slums and deplorable living conditions in cities in developing countries and estimating that by the end of 2007 there would be more than a billion slum dwellers, largely in developing countries (United Nations, 2007b). This latter report argues (p. 12) that in many cases the economic circumstances of urban migrants are worse than those of rural peasants. Four years earlier, it had been reported, also by the UN (United Nations, 2003), that surveys of member governments eliciting their attitudes towards urbanization found that the "vast majority" of these governments would wish to shift populations back to rural areas and stem the tide of urbanization that has been experienced around the world.

How bad can urbanization be for the development process? Urbanization simply cannot be all that bad for those who live in cities in developing countries – after all, in most instances migrants to cities and urban areas could simply return to rural life if they thought such moves would improve their economic circumstances. Why all the fuss about urbanization and development?

In this essay, we consider the evidence about the mechanisms increasing economic efficiency in cities and the record of cities in facilitating economic output and improving the consumption opportunities available to workers. Much of this evidence is

based on observations on highly-developed countries, but a growing body of evidence is based upon analyses of developing economies in low-income countries. This evidence clearly supports the conclusion that cities are important in facilitating economic growth, increased productivity and rising incomes in poor and rich nations alike. Policies to facilitate, not inhibit, urbanization will improve economic conditions in developing countries.

The analysis suggests a variety of broad policies – predispositions for policy – that would improve resource allocation and increase incomes in poorer countries.

## **II. Why Cities?**

A simple but useful point of departure for investigating the link between cities and economic development is consideration of the threshold question: why do people and firms choose to locate in the close proximity characterized by cities? Clearly autarchy, a uniform distribution of populations over space, would reduce competition for locations, thus the rents paid by households and firms, making them both better off (Starrett, 1974).

The answer to this question starts to provide a rationale for the existence of cities. Clearly there must be compensating benefits of urban location, either in terms of cost reduction, output enhancement, or else utility gains, to make dense location and the payment of location rent sensible choices for households and firms.

The putative utility gains from urbanization have been the subject of much speculation and analysis by non-economists. The once-popular refrain “How’re ya gonna keep them down on the farm...?” trivializes the principal insight. However, the vivid prose of Jane Jacobs (1969) argues seriously that the potential for variety in consumption is itself valuable to consumers. As long as the higher density of cities is associated with

greater variety – in people, in goods, in 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, 2001 explores some of these models.)

In contrast, the productivity gains, cost reductions, or output enhancements associated with collocation have been the subject of extensive analysis by professional 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 an industrial site as well as the costs of delivering outputs – finished products – to local, national and world markets. From this perspective, it is no accident that many of the large cities of the world developed along waterways (Rappaport and Sachs, 2003), 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.

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, and 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 woolen industry dictated large mills near cheap water power, and the residences of workers came to be concentrated nearby. The development of denser settlements – industrial plants and tenements – allowed firms to operate at scales where 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 nineteenth and early twentieth century (e.g., Pittsburgh, Manchester, Detroit) reflects the importance of internal economies of scale.

### **III. Cities and Growth**

If transport costs and internal scale economies were the only economic rationale for cities, the effects of urbanization upon economic growth more generally would be quite limited. The economic importance of cities would be strictly determined by the technologies available for transport and production. And the secular reductions during the past century in both 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. But the importance of cities to the modern economy hardly emphasizes internal scale economies at all, but rather external effects, spillovers, and external economies of scale. And these factors have become more important with increased industrialization, technical progress, and with economic development.

These external effects can be characterized along a variety of dimensions, and there are many taxonomies. One useful taxonomy distinguishes among productivity gains arising from specialization, those arising from transactions costs and complementarities in production, more generally from education, knowledge and mimicking, and those arising simply from the proximity to large numbers of other economic actors.

#### **A. Specialization**

The gains from specialization arise because denser conurbations with a larger number of firms producing in proximity can support firms which are more specialized in producing intermediate products. Specialization can lead to enhanced opportunities for

cost reduction in goods production – for example, when the components of intermediate goods can be routinized in production, or when the components of final products can be mechanized or automated. But the gains from specialization extend to the production of services as well. For example, specialized legal services may be provided more efficiently by firms which concentrate on taxation, copyright law, secured transactions, and so forth. In both intermediate goods and in services, specialization increases the opportunities for cost reduction.

The potential gains from specialization are further enhanced by the opportunities for sharing inputs among firms, and these are facilitated by larger and more dense urban areas. Specialized services -- for example, repair, printing, advertising, and communications – can be provided to a wide spectrum of goods 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 these gains 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. This specialization, in turn, leads to cost reductions.

## **B. Transactions Costs and Complementarities**

Externalities arising from transactions costs and complementarities in production can result because a 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, the opportunity for better skill matches reduces the search costs of workers with differentiated skills and the search costs 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 greater 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 workers of lower skill will end up working with more physical capital; they will be more productive and earn higher incomes. The return on workers' human capital and on employers' physical capital thus rises with the stock of human capital in the city – even when production at each worksite is undertaken with 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 rendering equipment more useful in resale markets.

### **C. Education, Knowledge Mimicking**

The notion of complementarities in labor market matching can be distinguished from externalities arising from the collocation of workers of similar education and skills in dense urban areas. The effects of aggregate levels of schooling in urban areas upon aggregate output may be distinguished from the effects of an individual's schooling on her 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

technique among firms, the copying and innovation in style, and the genealogy of patents among firms are all examples of local externalities in production which are encouraged by urban density and concentrations of skilled workers near to each other. These economies may arise with spatial concentration by industry (localization economies) or with higher densities of diverse industries (urbanization economies).

#### **D. The Law of Large Numbers**

Finally, considerable cost savings may arise simply from the presence of large numbers of economic actors in close proximity. For example, to the extent that fluctuations in demand are imperfectly correlated across firms in an urban labor market, employment can be stabilized – since some firms will be hiring workers while other firms will be contracting. To the extent that fluctuations in demand for products are uncorrelated across buyers, firms need carry less inventory, since some consumers are buying while others are 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 upon better information. This is true on the buying and selling sides of markets for purchasing inputs, storing intermediate products, and selling outputs.

#### **IV. Limitations on City Sizes**

The external effects of the urban environment on productivity described above all point to larger and more dense accommodations and indicate that there is a strong positive relationship between urbanization and economic development. What are the

limits, if any, to the extent of urbanization and thus the efficient sizes of cities? There are at least three sources limiting the sizes of cities and affecting 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.

### **A. 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.)

Alternatively, the wages offered as cities expand must increase enough to offset the higher costs that 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. The operation of these factors, housing and land prices, together with wages and output prices, limit the efficient sizes of cities.

### **B. Unpriced Externalities of Urban Density**

The increased transport costs and higher densities of cities may bring their own externalities, and if these are large enough they will limit the extent of urbanization. Of course, if these externalities are unpriced, they will not limit urbanization sufficiently. In higher income countries, air pollution from vehicles is typically underpriced, and until recently congestion in cities was hardly ever priced. Externalities from vehicle accidents are seldom compensated. In low-income countries, there may be additional external costs of higher-density living in the form of potential disease, epidemic, or conflagration.

These external effects are certainly unpriced. To the extent that these external effects are underpriced in cities, potential rural migrants do not face the marginal costs of urban life. Hence migration will be in excess, and cities will be larger than their efficient sizes.

### **C. Explicit Policies**

Finally, explicit governmental policies, especially in developing countries, may provide strong indirect incentives affecting the extent and distribution of urbanization. For example, it was widely thought that governments in developing countries favored producers and consumers in urban areas at the expense of rural and agricultural workers, for example, by imposing below-market prices for agricultural output and above-market prices for urban products. Structural adjustment policies, widely adopted from the 1980s onwards, have greatly reduced the scope for this urban bias and the distorted migration signals inherent in these subsidies. But a continuing form of central government bias may be a set of policies favoring particular cities, for example, the national capital. Policies favoring the locations which benefit elites and bureaucrats may be enacted as a result of rent-seeking behavior or as a consequence of corruption by elites. Questionable policies may include direct public investments in plant and equipment, or in infrastructure, capital controls on investment across cities, differences in rules imposed on cities for access to capital markets, or for obtaining licenses and permissions. In some countries, these restrictions have included explicit limitations on labor mobility as well.

### **D. Summary**

All of the factors suggest reasons 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 industries and across industries. And they facilitate sharing and risk-pooling by their very size.

Even recognizing the potential negative externalities of larger cities, these factors still suggest that real wages in larger cities in developed and developing countries will exceed those in smaller cities. In the development context, it is hardly necessary to stress that urban productivity will be higher than rural productivity and that this differential will facilitate migration from the labor surplus hinterland to the more productive urban areas.

Early models of rural-urban migration, beginning with Kuznets, recognized the free flow of labor from unproductive agriculture to urban employment, tending to equalize wages, as a vital part of the development process. In the 1970s analysts emphasized the importance of minimum wage rules in cities and the tendency towards equalization of expected wages across sectors. These models, beginning with those of Harris and Todaro in 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 to limit mobility to the productive cities rather than to remove barriers to competition in the labor market. Indeed, comparing high levels of official unemployment in cities with equally “official” statistics from rural areas, which ignore disguised rural unemployment, often seems to be deployed by agrarian romantics revealing an anti-urban bias. (See Lall, Selod and Shalizi, 2006, for a more balanced discussion of these issues – but one which reaches a similar conclusion, e.g., pp. 47-48.)

## **V. Empirical Evidence: Productivity Gains**

### **A. Basic Findings**

Despite the attention paid to agglomeration economies – going back to observations by Alfred Marshall in the 1890s – verification of efficiency gains by direct observation proved difficult at first, even using data from advanced industrial economies. A number of early studies estimating aggregate production functions are suggestive, but most of these efforts lacked critical data (for example, measures of capital stock), making inferences about the importance of external effects problematic. These issues are reviewed in Eberts and McMillen (1999) and in Rosenthal and Strange (2004).

More recent work using micro data sets on firms and establishments for the U.S. has overcome most of these measurement problems. For example, Henderson's (2003) recent analyses of machinery and high tech industries test directly for the presence of localization economies (agglomeration within an industry) and urbanization economies (agglomeration across industries), by estimating plant-level production functions. A panel of plants across counties and metropolitan areas makes it much easier to test for the importance of local conditions upon the productivity of plants and their levels of output. Henderson's results verify 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, then more talented entrepreneurs will certainly seek out these more productive locations. To account for this simultaneity, more sophisticated statistical methods are necessary. Henderson certainly recognizes this in applying more appropriate statistical methods of

estimation in his study of high tech and machinery industries, but the instruments he relies upon (i.e., measures of the local environment) are weak, rendering the statistical results problematic.

A recent unpublished paper by Greenstone, Hornbeck and Moretti (2007) solves this identification problem. The authors study the effects of the opening of large-scale plants (“million dollar plants”) on the productivity of pre-existing plants nearby, using a panel of establishments from the same data source relied upon by Henderson. For each of the million dollar plants, the authors have information on the county chosen for investment and also on the county which was not ultimately selected -- but which was under final consideration by the parent firm for the location of the new plant. This permits a direct side-by-side comparison of the course of productivity in plants located in the counties chosen for exogenous investment with the course of productivity in plants located in the “losing” counties. The authors found clear evidence of a discontinuity in total factor productivity in plants after the opening of a large plant nearby. Total factor productivity was enhanced in those pre-existing plants located in the “winning” counties, but not in the losing counties, confirming the existence of urbanization economies. This is important evidence.

A variety of less direct approaches have been employed to make inferences about agglomeration. Rosenthal and Strange (2003), among others, have studied the locations of firm births. Problematic data on factor inputs (e.g., the legacy of sunk capital) are not necessary for inquiries about new firms, and new establishments may plausibly take the existing economic geography of regions as exogenous. This empirical work suggests that births are substantially more likely to occur where there is an existing concentration of

firms in the same industry. (See also earlier work done by Carlton, 1983.) To the extent that profit-seeking entrepreneurs are drawn to more productive locations, this emphasizes the importance of localization economies.

Alternatively, the study of the spatial distribution of wages and rents may provide indirect evidence about economies of agglomeration. In more productive regions, the marginal product of labor will be higher, and wages will be higher. Analogously, locations where industrial rents are higher are those possessed of compensating differentials in productivity. Investigations using U.S. data on wages are reported by Wheaton and Lewis (2002) and using U.S. data on rents by Gabriel and Rosenthal (2007). An analogous investigation using rent data from Japanese prefectures to make inferences about agglomeration has also been undertaken by Dekle and Eaton (1999). All of these investigations document agglomerative economies.

Finally, patterns of employment growth may provide indirect evidence on the importance of agglomeration. If agglomeration economies enhance productivity, then more productive regions will grow more rapidly. An important paper by Glaeser, Kallal, Scheinkman, and Shleifer (1992) used aggregate employment data from U.S. metropolitan areas to confirm these effects. Henderson, Kuncoro and Turner (1995) conducted a more precise test using employment in manufacturing.

The specific mechanisms transmitting these urbanization and localization efficiencies have been studied by economists. Perhaps the clearest evidence of external effects in local labor markets comes from education and training. Early studies by Rauch (1993), predating the Lucas (1998) hypothesis, identified the external effects of schooling upon wages in cross-sectional models of wage determination, using U.S. cities as units of

observation. Recent work by Moretti (2004) extends this analysis to explain longitudinal as well as cross-sectional variation 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 upon 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 recent study of the service sector. Arzaghi and Henderson (2006) analyzed advertising firms in Manhattan, documenting the substantial increases in productivity attributable to the networking opportunities arising from the nearby location of similar firms.

Beyond these direct inquiries into the production of goods and services, it has been widely reported that incomes have grown more rapidly in U.S. cities with high initial levels of human capital (e.g., Glaeser, Schenkman and Shleifer, 1995), consistent with skill acquisition and diffusion through the interaction of workers in more dense urban areas (Glaeser and Mare, 2001; Duranton and Puga, 2004).

Lacking direct observations on workers' interactions, economists have evaluated one important "paper trail" of these interactions, namely 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 spatial decay of patent citations with distances between firms and between inventors (Jaffee, Trajtenberg, and Henderson, 1993). This work has

provided explicit confirmation of the importance of geographic spillovers in the development of new knowledge.

Of course, there has been direct anthropological study by sociologists and others, observing worker interactions in dense locations, most authoritatively (Saxenian, 1994) among highly educated workers in Santa Clara County, California (“Silicon Valley”) and along the technical corridor (“Route 128”) outside Boston. These investigations are broadly consistent with the results of quantitative investigation by economists.

## **B. Corroboration From Developing Countries**

Many of the models reported in the previous section have been adapted, extended and applied using data from low-income and developing countries. Much of this work has been pioneered by Vernon Henderson and his collaborators. Henderson’s 1998 book includes a detailed chapter estimating the extent and importance of agglomeration economies in Brazil using detailed industrial census data. Henderson found clear evidence of external economies of scale, localization economies for two-digit industries. (Unfortunately, for at least some cities analyzed by Henderson, a single industry was dominant, meaning that factor prices and populations are endogenously determined. This is a major limitation.) This work on Brazil is similar to (but much more primitive than) the work reported by Henderson (1993) and by Greenstone, Moretti, and Hornbeck (2007) using U.S. data. In a more recent analysis of city growth in Brazil, Henderson and his and his collaborators analyzed aggregate data for 123 cities observed during three decades beginning in 1970 (da Mata, Deichmann, Henderson, Lall, and Wang, 2007). The authors laid out an ambitious model of the structure of supply and demand for output at the municipal level, and they estimated relationships 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 upon the growth rates of cities. Improvements in labor force quality and in the initial levels of educational attainment matter significantly and importantly for economic growth. In these respects, recent work on Brazil confirms and extends the conclusions of Glaeser, Scheinkman and Shleifer (1995) in their earlier study of American cities.

Other direct investigations of agglomeration and productivity have been undertaken in Korea, Indonesia, China, and in India as well.

Evidence of localization economies was reported for Korean industry, including transport and traditional industry, by Henderson, Lee, and Lee (2001). The authors analyze metropolitan level data for 23 Korean industries in five major groups at the metropolitan level during the 1983-1993 period, a time of rapid deconcentration of economic activity from the capital Seoul to smaller metropolitan areas. The authors estimate aggregate production functions, relying upon census estimates of capital stock and labor, testing for the importance of the potential urbanization and localization economies provided in the sample of Korean cities. The results confirm the clear importance of localization economies in Korean industry, most importantly in heavy industry and transportation. Significant localization economies were also found in machinery and high-tech industries, and to a lesser extent in “traditional manufacturing.”

These results are confirmed by the contemporaneous work of Lee and Zang (1988) applying somewhat different statistical models to the same basic source of data from the Korean Census of Manufacturing.

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

A more recent and quite ambitious paper by Deichmann, Kaiser, Lall, and Shalizi (2007) extends these results for Indonesia, 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 also suggest the importance of existing backward linkages to suppliers in affecting location choice. Urbanization economies, *per se*, seemed much less important.

Simulations based upon these statistical results illustrate the difficulties faced by lagging regions in attracting new economic activity.

Au and Henderson (2006) used aggregate data on some 285 Chinese cities to estimate the effects of urban agglomeration on productivity. This analysis was facilitated by detailed data reporting GDP by Chinese metropolitan area in three categories. The aggregate productivity relationship exhibits an inverted U shape in metropolitan size and scale, as expected. The estimated urban agglomeration benefits are quite high, and it

appears that a large fraction of cities in China are under-sized due to 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 relationship between urban populations and total factor productivity by state and industry over a sixteen-year period (Mitra, 2000). Of more significance, perhaps, is the analysis of plant-level data undertaken by Lall, Koo, and Chakrovorty (2003). Lall and his collaborators use these micro data on establishments, from the Indian Survey of Industries in 1998, to estimate the parameters of a translog cost function. The authors provide direct estimates of the elasticity of costs with respect to four different measures of agglomeration, separately for eight industrial groupings and three size classes of plants. The results provide rather strong support for the importance of urbanization economies in reducing costs per unit of output. This finding is consistent across all industries and size classes of Indian plants. This paper provides important evidence on urbanization economies in India, and perhaps in other developing economies as well.

### **C. Summary**

The scientific quality of the evidence from developing countries cited above is probably lower than that obtained from more advanced countries – if only because more reliable data on economic activity are available for a larger period for developed countries. Nevertheless, the quantitative results obtained from low-income and developing countries in Asia and Latin America is remarkably consistent with that

obtained from more advanced economies. Comparable evidence from low-income countries in Africa is conspicuously absent (Collier, 2007).

Urbanization and localization do support increases in productivity. Of course, it may well be that the economic return to mimicking successful ideas or investments is especially high in developing countries, and mimicking could result in too little entrepreneurial activity in low-income countries. (This allegation is made by Hausmann and Rodrik, 2002, 2006.) But there is no systematic evidence that the potential returns to mimicking are greater in poorer countries than in richer ones. And, 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 really proves a tight causal link between urbanization and economic development. (See Henderson, 2003, for a balanced discussion.) And there is evidence from elsewhere that urbanization is certainly not a sufficient condition for economic development to occur (Fay and Opal, 2000). Nevertheless, it seems quite clear that productivity is enhanced by the localization and urbanization features of cities, in developing economies as well as industrialized countries. The cumulative evidence is overwhelming.

## **VI. Efficient City Sizes**

Given the productivity advantages of larger cities in developing countries documented in the previous section, we should expect urbanization to be a natural concomitant of increased output and well-being in low-income countries. City sizes are determined by the tradeoff between the increased productivity and incomes in larger conurbations and the increased rent and transport costs consumers confront in larger

cities. To the extent that congestion, pollution, and the risk of epidemic are not considered appropriately by rural workers contemplating moves to cities, the cities will be “too large,” but not by much. Some of these externalities can be eliminated by improved technology; others by investments in public health.

It is quite surprising that there does not seem to be a coherent literature – or much economic literature at all – relating these externalities to levels of urbanization in low-income countries. There are case studies of the linkage between traffic fatalities and economic growth (e.g., Kopits and Cropper, 2005), but not between traffic fatalities and urbanization, much less studies of the linkage between externalities from traffic fatalities and levels of urbanization. It is relatively straightforward to estimate the correlation between the incidence of some communicable diseases (e.g., diarrhea, tuberculosis) and urbanization at the country level and to estimate the correlation between access to water and sanitation, on the one hand, and urbanization, on the other hand. Indeed, many of these correlations may be investigated on line (for example, by relying upon “WDI Online”), and it is reported that infant mortality is higher in urban slums in developing countries. (See Evans, 2007.) But these correlations are barely hints about the causal mechanisms at work. Nevertheless, these results are widely interpreted as if there were some causal mechanism.

At this point, we can only conclude that unpriced externalities are probably a bit more important in distorting migration flow to cities in developing countries than in developed countries. But in both cases, these distortions can be reduced by direct pricing or by indirect levies such as urban property taxes.

What about the explicit policies of governments?

As noted above, there has been widespread belief that explicit policies of developing countries have inappropriately favored cities at the expense of agriculture, and this has interfered with economic development. The most direct accusation of an “urban bias” is due to Lipton three decades ago (1976), but see also Lipton, 1993. If macroeconomic and national trade policies distort price signals, by raising the value added in the urban sector, when value added is computed using local prices, this provides incentives for inefficiently high levels of urbanization. And if urban products are valued at inflated prices while rural products are valued at deflated prices, productivity advantages attributed to cities may be simply illusory.

It is not clear how these price distortions can be measured (See Becker and Morrison, 1989), or how the implications of this bias could be tested directly. But after two decades of structural adjustment policies advocated by the international organizations, it is quite 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 fifteen-year old treatise on urban policy (1991), “an agenda for the 1990s,” documents the contemporaneous effects of structural adjustment policies in removing any artificial price advantages of cities (and in reducing the economic circumstances of the poor in cities in low-income countries). The controversy over policies to undo distortions in relative prices seems somehow dated.

But certain limited aspects of “bias” in development policies may be well taken, and they may be a continuing concern. Of particular concern are government policies that favor particular cities or regions for political or ideological reasons.

A remarkable regularity observed across systems of cities is the rank-size rule, namely that the product of the city rank in the size distribution and the city population is roughly constant (and thus the second ranking city in a country is half the size of the first, and so forth). This relationship (more generally, a power relation) has proved robust across time periods for the U.S. (Dobkins and Ioannides, 1998) and other countries, and across countries as well (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) has hypothesized that the higher costs of spatial interaction and the less elastic labor supply in the nineteenth century help explain why a smaller share of 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. For example, there will be a more uneven distribution of city sizes if there are stronger external economies of scale in cities. But the exact relationships are elusive (Fujita, Krugman and Venables, 1999).

Despite this uncertainty about the economic process governing the rank-size relationship, considerable evidence suggests that political variables affect the distribution of city sizes. For example, Soo’s analysis (2005) of the size distribution of cities across 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 relationship relating city rank and city sizes.

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 analyzed variations in the national population residing in the largest city for a sample of 85 cities over fifteen 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, and democratically-governed countries have principal cities about forty percent larger if they were governed by dictators in the past. 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 therein.) Little or none of the criticism of “excessive” concentration is based upon the empirical evaluation of externalities in developing countries. This is surprising.

Clearly excessive concentration may be abetted by government policy. The mechanisms by which authoritative governments are able to favor particular cities or regions are not hard to visualize, but they may be quite hard to document. These mechanisms may be quite indirect, ranging from weaker benefit-cost tests imposed on infrastructure investment to a relaxation of licensing rules in favored cities, to explicit allocation of credit to favored regions, to decisions favoring investments by public officials and cronies in national capitals.

In this sense, there may be an “urban bias” in government policy, and it may result in discrimination against rural development. But this bias also discriminates against most of the small and medium-sized cities in developing countries as well as rural areas.

## **VII. Some Conclusions**

This review documents the strong relationship between urbanization, on the one hand, and economic productivity and development, on the other. This documentation is based upon extensive analyses of data from the U.S. and high-income countries and less-extensive analyses of data from developing countries. It suggests that the effects upon productivity can arise from specific mechanisms fostered by the urbanization and localization of industry. The available evidence does not conclude that urbanization is necessary for the development process, nor is urbanization sufficient to increase output and well-being in low-income countries. But the case is strong, and the causal relationship is nevertheless clear.

Despite equivocation, it is clear that 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 these movements are the sensible 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 – pollution and congestion, for example – are unpriced in cities, conurbations will be too large, but not by a lot. Public concerns about pricing scarce roadways and about water supply and public health investments to decrease the chances of epidemic are well-placed; from this perspective the concern with urban slums *per se* is less important. Urban poverty is not an excuse for policies limiting the extent of urbanization in low-income countries.

It is hard to know how important corruption and anti-democratic policies are in inhibiting or directing flows of factors to and among cities. Their existence in developed as well as underdeveloped countries provides strong argument to allow natural market forces to determine the spatial distribution of labor and capital. Urbanization and Economic development will be increased, and the level of urbanization will also increase.

Increased urbanization certainly facilitates the development process, and explicit policies to discourage urbanization are 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," unpublished paper.
- Au, Chun-Chung and J.V. Henderson (2006). "Are Chinese Cities Too Small?" *Review of Economic Studies*, 73(3): 549-576.
- Becker, Charles M., and Andrew R. Morrison (1999). "Urbanization in Transforming Economies," in Mills, E.S., and P. Cheshire (eds.), *North-Holland Handbook of Regional and Urban Economics*, volume 3. Amsterdam: Elsevier.
- 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.
- Carlton, D.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): 440-449.
- CERAP (China Economic Research and Advisory Programme) (2007). "Urbanization in China," Professional Report #P07-001, Institute of Business and Economic Research. Berkeley: University of California.
- Chen, Y. (1996). "Impact of Regional Factors on Productivity in China," *Journal of Regional Science*, 36(3): 417-436.
- Collier, Paul (2007). *The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It*, London: Oxford University Press.
- Da Mata, D., U. Deichmann, J.V. Henderson, S.V. Lall, and H.G Wang (2007). "Determinants of City Growth in Brazil," *Journal of Urban Economics*, 62(2): 252-272.
- Deichmann, Uwe, Kai Kaiser, Somik V. Lall, and Zmarak Shalizi (2005). "Agglomeration, Transport, and Regional Development in Indonesia," World Bank Policy Research Working Paper #3477.
- Dekle, R. and J. Eaton (1999). "Agglomeration and Land Rents: Evidence From the Prefectures," *Journal of Urban Economics*, 46(2): 200-214.
- Dobkins, L., and Y. Ioannides (1998). "Dynamic Evolution of the U.S. City Size Distribution," in J.-M. Huriot and J.-F. Thisse (eds.), *The Economics of Cities*. New York: Cambridge University Press.

- Evans, Timothy (2007). "Research for Urban Health: Towards a Global Agenda," World Health Organization, Global Summit, Bellagio, Italy.
- Fay, Marianne, and Charlotte Opal. "Urbanization without Growth: A Not-So-Uncommon Phenomenon," World Bank Policy Research Working Paper #2412. Washington, DC: World Bank.
- Fujita, M., P. Krugman, and A. Venables (1999). *The Spatial Economy*. Cambridge, MA: MIT Press.
- Gabaix, Xavier (1999). "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, 114(3): 739-767.
- Gabriel, S.A. and S.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): 438-444.
- Glaeser, E., H. Kallal, J. Scheinkman, and A. Schleifer (1992). "Growth in Cities," *Journal of Political Economy*, 100(6): 1126-1152.
- Greenstone, Michael, Enrico Moretti and R. Hornbeck (2007). "Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants", unpublished paper.
- Hausmann, Ricardo, and Dani Rodrik (2002). "Economic Development as Self Discovery," NBER Working Paper #8952, May 2002.
- Hausmann, Ricardo, and Dani Rodrik (2006). "Doomed to Choose: Industrial Policy as Predicament," unpublished paper, John F. Kennedy School of Government.
- Henderson, J. Vernon (1999). "How Urban Concentration Affects Economic Growth," World Bank Policy Research Working Paper #2326, Washington, DC: World Bank.
- Henderson, J. Vernon, (2003a). "The Urbanization Process and Economic Growth: The So-What Question," *Journal of Economic Growth*, 8(1): 47-71.
- Henderson, J. Vernon (2003b). "Marshall's Scale Economies," *Journal of Urban Economics*, 53(1): 1-28.
- Henderson, J. Vernon (2005). "Urbanization and Growth," in Aghion, P., and S. Durlauf (eds.), *Handbook of Economic Growth*. Amsterdam: North Holland Publishing Company.
- Henderson, J. Vernon, and Ari Kuncoro (1996). "Industrial Centralization in Indonesia," *The World Bank Economic Review*, 10(3): 513-540.

- Henderson, J. Vernon, A. Kuncoro and M. Turner (1995). "Industrial Development in Cities," *Journal of Political Economy*, 103(5): 1067-1085.
- Henderson, J. Vernon, T. Lee and J.Y. Lee (2001). "Scale Externalities in Korea," *Journal of Urban Economics*, 49(3): 479-504.
- Hilber, Christian A.L., and Ioan Voicu (2005). "Agglomeration Economies and the Location of Foreign Direct Investment: Empirical Evidence from Romania," unpublished paper.
- Hoover, Edgar M. (1975). *An Introduction to Regional Economics*, second edition, New York: Alfred A. Knopf
- Jacobs, Jane (1969). *The Economy of Cities*. New York: Random House.
- Kopits, Elizabeth, and Maureen Cropper (2005). "Traffic Fatalities and Economic Growth," *Accident Analysis and Prevention*, 37(1): 169-178.
- Lall, Somik V., Jun Koo, and Sanjoy Chakravorty (2003). "Diversity Matters: The Economic Geography of Industry Location in India," World Bank Policy Research Working Paper #3072. Washington, DC: World Bank.
- Lall, Somik V., Harris Selod, and Zmarak Shalizi (2006). "Rural Urban Migration in Developing Countries: A Survey of Theoretical Predictions and Empirical Findings," World Bank Policy Research Working Paper #3915. Washington, DC: World Bank.
- Lee, Y.J. and Zang, H. (1998). "Urbanization and Regional Productivity in Korean Manufacturing," *Urban Studies*, 35(11): 2085-2099.
- Lipton, Michael (1976). *Why Poor People Stay Poor: Urban Bias in World Development*. Cambridge, MA: Harvard University Press.
- Lipton, Michael (1993). "Urban Bias: Of Consequences, Classes and Causality," *Journal of Development Studies*, 29(4): 229-258.
- 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.
- Quigley, John M. (1998). "Urban Diversity and Economic Growth," *Journal of Economic Perspectives*, 12(2): 127-138.
- Quigley, John M. (2001). "The Renaissance in Regional Research," *Annals of Regional Science*, 35(2): 167-178.
- Puga, Diego (1998). "Urbanization Patterns: European Versus Less Developed Countries," *Journal of Regional Science*, 38(2): 231-252.

- Rappaport, Jordan and Jeffrey Sacks (2003). "The United States as a Coastal Nation," *Journal of Economic Growth*, 8(1): 5-46.
- Rosen, Kenneth, and M. Resnick (1990). "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy," *Journal of Urban Economics*, 8(2): 165-186.
- Rosenthal, Stuart S., and William C. Strange (2001). "The Determinants of Agglomeration," *Journal of Urban Economics*, 50(2): 191-229.
- Rosenthal, Stuart and William C. 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 of the series *Cities and Geography*, Amsterdam: North Holland Publishing Company.
- Soo, Kwok Tong (2004). "Zipf's Law for Cities: A Cross-Country Investigation," *Regional Science and Urban Economics*, 35(3): 239-263.
- Starrett, David A. (1974). "Principles of Optimal Location in a Large Homogeneous Area," *Journal of Economic Theory*, 9(4): 418-448.
- United Nations (2003). *World Population Policies*, New York: United Nations Publishing.
- United Nations, UN Population Fund (2007a). *State of World Population 2007: Unleashing the Potential of Urban Growth*. New York : United Nations Publishing.
- United Nations, UN-Habitat (2007b). *The State of the World's Cities Report 2006/2007*. London: Earthscan Publications, Ltd.
- Wheaton, William C. and M.J. Lewis (2002). "Urban Wages and Labor Market Agglomeration," *Journal of Urban Economics*, 51(3): 542-562.
- World Bank (1991). *Urban Policy and Economic Development: An Agenda for the 1990s*. Washington, DC: The World Bank.