Scale Economies and Cities

Indermit S. Gill • Chor-Ching Goh

This paper summarizes the policy-relevant insights of a generation of research on scale economies. Scale economies in production are of three types: internal economies associated with large plants, localization economies that come from sharing of inputs and infrastructure and from greater competition among firms, and urbanization economies that are generated through diversity and knowledge spillovers. The benefits (and costs) of localization and urbanization are together called "external (dis) economies" because they arise due to factors outside any single household, farm or firm. The empirical literature yields some stylized facts. Internal scale economies are low in light industries and high in heavy industries. External scale economies are amplified by economic density and dissipate with distance from places where economic activity is concentrated. Scale economies are most visibly manifest in towns and cities. To simplify somewhat, towns allow firms and farms to exploit internal scale economies, medium-sized cities help firms in an industry exploit localization economies, and large cities and metropolises provide urbanization economies to those who locate within or nearby. Scale economies have implications for policy makers. The first is that because urban settlements rise and thrive because market agents demand their services, they should be seen as creatures of the market, not creations of the state. The second is that because settlements of different sizes provide differing services, towns, cities, and metropolises are more often complements for one another, not substitutes. Third, as a corollary, policymakers should aim to improve the functioning of urban settlements, and not become preoccupied with their size. JEL codes: O1, O4, O3, R3, R11

Even a casual examination of the economic geography of most countries reveals that the topography of production and population is bumpy. People, firms, jobs, services, and opportunities tend to agglomerate in relatively small parts of a country, frequently leaving the remaining areas to agriculture or to nothing. In the United States, half of the nation's GDP is produced on just 4 percent of the

The World Bank Research Observer

doi;10.1093/wbro/lkp022 Advance Access publication November 26, 2009 25:235–262

land area (Easterly and Levine 2001). As one zooms in to a smaller spatial scale, this pattern is replicated. New York is the most economically dense state in America. This in turn is because of the extreme concentration of production in the metropolitan area of New York. And the metro area contains New York County, the densest county in the nation, as well as three of the next ten densest counties (Ciccone and Hall 1996).

Besides being bumpy—even spiky—the spatial distribution of economic activity is sticky. New York has been the largest U.S. city for over a century (Kim and Margo 2004, p. 2994). No new urban agglomeration has emerged in France since 1911, and no city that was relatively big has died (Eaton and Eckstein 1997, pp. 447–8). Once established, a dense agglomeration of economic activity tends to persist. A country's economic geography exhibits inertia.

The spikiness and stickiness of a country's economic geography can be explained by a single factor: scale economies. Scale economies can be internal or external to firms and farms. Internal scale economies are the cost advantages that a firm reaps by producing in large plants. As Adam Smith recognized during his visit to a pin factory more than two centuries ago, they include the productivity advantages from the division of labor which is made possible by large scale production. Internal increasing returns create a tendency for a firm to locate its activities in one or a few locations, thereby contributing to bumpiness.

But internal economies cannot, by themselves, explain the existence of cities, because they cannot explain why many firms might choose to base themselves in a single location and often stay there for long periods. For this, it is necessary to turn to external economies—the benefits of scale that arise outside a firm or industry. These are the advantages that a bookshop on Dadabhai Naoroji Road in Mumbai gets by being close to other book stalls. They are also the advantages which financial firms derive from locating in large and diverse cities such as São Paulo or Shanghai. Agglomeration economies can take the form of either "localization" economies (the benefits from colocation of similar producers) or "urbanization" economies (the benefits from conglomeration of diverse producers). Firms benefit from locating close to other firms in either the same or different industry. And unless all of them move elsewhere together, they would lose these benefits.

However, the logic can only be stretched so far. For people and firms, being in cities comes at a price. Traffic in central London moves at only 11 miles per hour (Shaffer and Santos 2004), the same speed as horse drawn carriages a hundred years ago. Bangkok is notorious for congestion and Beijing for pollution. Housing in Delhi is amongst the most expensive in the world. Crime and slums are prevalent in many developing countries' primary cities. These are the costs of density, the *diseconomies* of agglomeration, which limit the productivity advantages to firms, and the welfare benefits to inhabitants, when they locate in dense urban areas. Combined with the differing propensities of activities to benefit from

agglomeration economies, they help to explain why economic activity within a country is not restricted to a single center, but is instead spread across centers of differing sizes. Good policies can relax the constraints generated by the congestion and overcrowding of land and resources; bad policies can result in the diseconomies quickly stifling the benefits that come from scale economies.

But caveats notwithstanding, the benefits of producing large amounts in a single plant or place have only increased as transport costs have fallen in the two centuries since Adam Smith visited the pin factory. The potential of scale economies and how access to world markets helps to exploit them can be seen in the city of Dongguan in China's Pearl River delta. In 2005 one plant there manufactured 30 percent of the magnetic recording heads used in disk drives worldwide. In several products, Dongguan's factories account for over 40 percent of global production. Of the parts and components used in personal computers, 95 percent can be sourced within Dongguan city (Gill and Kharas 2007).

Places like Dongguan have rushed headlong into the world of scale economies. While economic crises slow down and even temporarily reverse these transformations, millions of people in the developing world enter this new realm every year, and the implications—for them and for policymakers—are enormous. This paper provides a brief synopsis of work by a generation of economists who have sought to understand these scale economies. It summarizes the evidence of scale economies in production, focusing on "agglomeration economies," whose exploitation requires locating in areas that are densely populated by other producers. Finally, we discuss the lessons for policymakers in the developing world. Our main conclusions are:

- The sectoral transformations that countries must undergo as they develop have a spatial "dual." As they transition from agricultural to industrial to service-oriented modes of production firms and workers leave behind not just their villages and their agrarian occupations, but also a world where scale does not matter so much. They enter not just settlements that are larger and denser, but also a world in which producers, distributors, and consumers enjoy agglomeration economies.
- Research over the last generation indicates that different types of human settlements facilitate agglomeration economies for different forms of production. An oversimplified (but not incorrect) generalization is that market towns facilitate plant-level scale economies in marketing and distributing produce, medium-sized cities provide localization economies for manufacturing industries, while the largest cities provide urbanization economies that foster the growth of business, government, and educational services.
- Because of coordination failures associated with spatial transformations, policymakers have often seen cities as constructs of the state. In reality, like firms

and farms, cities and towns are creatures of the market. Just as firms and farms deliver final and intermediate goods and services, towns and cities deliver agglomeration economies to producers and workers. While a longer discussion of policy implications is outside the scope of this survey, two clear lessons emerge by recognizing the importance of agglomeration economies. The first is that settlements of different sizes serve functions that are complementary. Large, medium, and small cities should not be seen as substitutes—the focus should be on function, not size. The second is that it gets harder to keep settlements well-functioning as urbanization advances, but reasonable land markets, policies that address negative externalities like congestion and pollution, and universal provision of basic services remain important at every stage of the spatial transformations that are necessary for development.

Theoretical Advances

Adam Smith introduced scale economies, factor mobility, and transport costs as central to understanding the nature and causes of the wealth of nations. But until the 1980s most economists were happier to anchor their inquiries to another concept introduced in *The Wealth of Nations*: the "invisible hand" of perfect competition.[1] Constant returns led to convenient characterizations of the economy where all firms and workers were identical, so one firm or worker could be considered representative of all. Scale economies were inconvenient—they required acknowledging that specialization differentiated people and products.

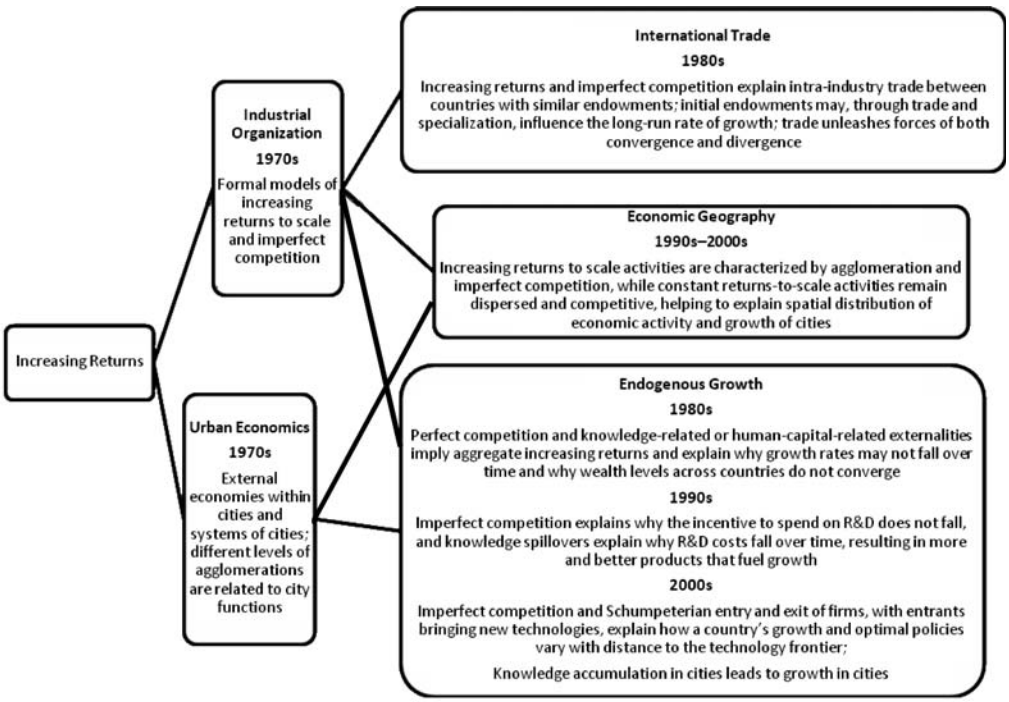
But perfect competition is an artificial construct. It assumes infinitesimally small firms with little or no influence over market prices, even in the immediate vicinity of a firm. The assumption of constant returns to scale also implies the problem of "backyard capitalism"(Krugman, 1991). That is, if small-scale production is as efficient as large-scale, every household would produce a range of goods and services in its backyard. Occasionally, the contradiction between these two phenomena—internal scale economies on the one hand, and perfect competition on the other—would surface. Because of the technical difficulties of modeling increasing returns to scale, though, it would quickly be buried again.[2]

Then, during the 1970s, a simple model of increasing returns to scale (Dixit and Stiglitz 1977) opened the door for researchers to the same realm that so many firms and workers had inhabited since the Industrial Revolution. By the late 1980s scale economies were standard features of the explanations for international trade. By the early 1990s growth theorists had accepted the need to incorporate imperfect competition among firms into aggregate formulations of an

economy. By the mid-1990s theorists were beginning to show how these ideas could be used to understand the spatial distribution of economic activity, including the rise of towns and cities (figure 1).

The literature on the microeconomic foundations of agglomeration economies has flourished in the last two decades by combining models of scale economies and insights about urban economics that highlight tensions between the benefits and costs from concentration (Mills 1967; Dixit 1973; Henderson 1974a). In general, researchers have recognized that economic growth has different impacts on firms and workers depending on their sector and location. The underlying reason is the love for variety in consumption and the economies of scale in

Figure 1. Recognizing the Importance of Scale Economies: 30 Years of Theoretical Advance



	Key Publications
Industrial Organization	Spence (1976); Dixit and Stiglitz (1977)
Urban economics	Mills (1967); Diamond and Mirrlees (1973); Henderson (1974a)
International trade	Krugman (1980, 1981); Ethier (1982); Helpman and Krugman (1987); Grossman and Helpman (1995)
Economic geography	Krugman (1991b); Fujita, Krugman, and Venables (1999); Baldwin and Martin (2004)
Endogenous growth	Romer (1986, 1990); Lucas (1988); Grossman and Helpman (1991); Aghion and Howitt (1992, 2005); Rossi-Hansberg and Wright (2007); Duranton (2007)

Source: Adapted from Gill and Kharas (2007) with suggestions from Vernon Henderson and John McCombie.

production; the proximate causes are product differentiation, monopolistic power, specialization, and location externalities. The formal recognition of scale economies, externalities, and imperfect competition makes economic theory conform more closely to the world which policymakers inhabit. The main insight is that activities with increasing returns due to reasons external to a firm tend to be concentrated in cities, while those displaying constant returns remain more dispersed.

A Guide to Scale Economies

Countries develop through sectoral shifts out of traditional subsistence-based agricultural activities into manufacturing and services. Along the way firms rather than farms become the dominant production unit. The production of differentiated manufactured goods and services increases as a share of the economy's output relative to agricultural produce. Production technology shifts away from constant returns to increasing returns to scale. And over time, scale-augmenting technical change generates scale economies. Imperfect competition becomes the rule.

The benefits of greater scale in production can be either internal or external to a producer. Internal scale economies are the *within-firm* efficiency gains associated with production in sizeable quantities in a single plant. External economies are those that come from producing in large quantities in a single place. They are traditionally classified as localization economies arising from *within-industry* economic interactions, and urbanization economies arising from *between-industry* interactions. External economies are synonymous with "agglomeration economies," which comprise the benefits of proximity to other producers of the same commodity or service, and diversity, or being close to other producers of a wide range of commodities and services. This paper deals mainly with production-related scale economies.[3]

Internal economies arise from the size of a plant's ability to exploit fixed costs better. For example, a larger steel mill can get volume discounts from suppliers—implying fixed costs of transport and trade—and reap the benefits of division of labor within the firm. But the steelmaker's location near other firms is not the main source of gains from internal scale economies. *Localization economies* arise from the presence of a large number of firms in the same industry in the same place. For example, textile firms cluster to share a large pool of suitably skilled and unskilled labor and other intermediate input suppliers. *Urbanization economies* arise from the presence of a large number of different industries in the same place. For example, a hedge fund can benefit from locating near business schools, banking and other financial service providers, management consulting companies, legal and accounting firms, as well as manufacturers. This diversity breeds new ideas and innovation.

As summarized in Duranton and Puga (2004), agglomeration economies depend not just on size (that is, on how big a city or industry is). The reasons for producers to gain from proximity to others lie in the sharing of capital inputs, information, and labor; improving the matches between production requirements and types of land, labor, and intermediate inputs; and learning about new techniques and products. A plant in an isolated location can benefit from internal scale economies, but unless it is situated in an area of density it cannot enjoy the benefits associated with localization or urbanization economies. Cities bring together large pools of skilled labor and suppliers of specialized intermediate inputs and by so doing enhance employer–employee and buyer–seller matches. Input sharing is an important channel for agglomeration economies. Averaging across industries, a firm's relocation from a less dense location (of fewer than 500 neighboring employees in the same industry) to a denser location (of 10,000–25,000) results in a 3 percent increase in purchased input intensity (see, for example, Holmes and Stevens 2002).

This paper does not discuss such scale diseconomies as congestion, crime, noise, flooding, and pollution because the empirical literature on the subject is thin. While there is ample theoretical discussion (see, for example, Tolley 1974; Henderson 1975, 1977; Arnott 1979; Sasaki 1998), there is little empirical evidence of such relationships. In a seminal paper, Mills and Ferranti (1971) rebut the conventional statement that pollution distorts city size. They argue that urban population is not the only factor that affects pollution, and increased recycling, treatment of waste, and improved efficiency of urban transportation systems can significantly reduce pollution. Henderson (1974b) and Tolley and Crihfield (1987) show that policy measures that reduce diseconomies will enhance the welfare of population and profit of production, and, as a result, cities will become even bigger.

A large city provides a large labor pool to firms and a dependable supply of specialized labor, providing more job options to workers and reducing the volatility of labor demand. Smaller specialized cities expose workers to greater industry-specific shocks but provide favorable match-specific advantages. In both cases the concentration of economic activity lowers the search costs between firms and workers, which results in fewer unfilled vacancies, lower risk of job loss, and shorter duration of unemployment. Costa and Kahn (2000) document that because of better matching married couples with university education are increasingly found in large cities, up from 32 percent in 1940 to 50 percent in 1990 in the US. The matching mechanisms depend on information transmission among people (see Bikhchandani, Hirshleifer, and Welch 1998 for a survey of the literature on social learning). Cities make it easier for producers to find inputs and for consumers to experiment and discover new suppliers.

Learning mechanisms are another important explanation for agglomeration in cities. Marshall (1890) emphasizes how cities favor the diffusion of innovations and ideas, while Jacobs (1969) stresses how the environment offered by cities improves the prospects for generating new ideas. Lucas (1988) also suggests that learning encompasses, not just cutting-edge technology, but incremental mundane knowledge (knowing how, knowing who, and so on) through intended and unintended communications, and that cities are the best place for knowledge transmission. Knowledge spillovers are difficult to measure because they can seldom be traced through transactions. With patent citations, however, it is possible to identify a paper trail for some, although by no means all, knowledge spillovers. Jaffe, Trajtenberg, and Henderson (1993) find that patent citations are spatially concentrated, with citations between them 5 to 10 times more likely to come from the same standard metropolitan statistical areas in the United States as originator patents. Jaffe (1986) also finds evidence for local R&D spillovers among U.S. firms where the number of patents per dollar of R&D spending is higher for firms located in areas with above average R&D spending.

Another strand of research focuses on workers as the primary vehicles of knowledge. This implies that economies with substantial labor mobility across industries will exhibit a greater spread of ideas and growth. Rauch (1993) shows that wages are higher where average education levels are high, since workers will be more productive and employers will be willing to pay high wages in competing for them. Moretti (2004a, 2004b) finds a positive effect of the presence of college graduates on a city's wages. Charlot and Duranton (2003) use survey data to show that workplace communication is more extensive in urban areas and that this communication affects wages positively.

Internal Scale Economies

Internal increasing returns to scale are well documented in manufacturing, based on cost and value-added data, engineering estimates, and trade data and markups. Table 1 summarizes the stylized findings in the literature. Internal scale economies range from negligible among light industries to high among heavy and high-tech industries. Based on engineering estimates of cost savings and the minimum efficiency scale, increasing returns are found in transport equipment, chemicals, machinery, engineering, and paper and printing. Internal scale economies are negligible in rubber and plastics, leather goods, footwear, clothing, and textiles. Estimates based on cost and value-added data at different scales of output—from cross-section, time series, and panel data of firms—point to similar findings.

Likewise, increasing output brings about disproportionate savings in the European car, truck, and consumer durables industries as well as in the

Table 1. Internal Scale Economies Are Low in Light Industries and High in Heavy Industries

Findings of returns to scale	*Data source*	*Studies*
Low: apparel, leather, footwear, wood products *High*: machinery, pharmaceuticals, instruments, iron and steel, petroleum and coal products	Trade data	Antweiler and Trefler (2002)
Low: leather goods, clothing, wood prods, textiles *High*: Vehicles, transport equipments, chemicals, printing and publishing	Engineering estimates of costs and changes in minimum efficiency scale	Junius (1997) and Prateen (1988)
Low: footwear, apparel, food products, leather *High*: tobacco, pharmaceuticals, computing machinery, railroad equipment	Markups for manufacturing in 14 OECD countries	Junius (1997) and Martins, Scarpetta, and Pilat (1996)
Low: apparel, leather products, textiles *High*: electric, gas, and sanitary services; motor vehicles and equipment, chemicals, tobacco	Markups for two-digit sectors in the US, 1953–84 (24 sectors)	Roeger (1995)
Low: textiles, milk products, lumber mills, fish oil and meal products *High*: metal, transport equipment, cement products	Production function estimates for 1963 in Norway (27 industries)	Griliches and Ringstad (1971)
Low: clothing, knitting, leather, textiles *High*: petroleum, basic and fabricated metal, transport equipment	Cost and profit data in 167 four-digit industries for 1970 in Canada	Baldwin and Gorecki (1986)
Low: textiles, apparel, lumber, wood *High*: paper products, chemicals, fabricated metals, stone, clay, glass	Dynamic cost functions for 450 U.S. industries, 1958–89	Paul and Siegel (1999)
Low: clothing, knitting, leather, textiles *High*: petroleum, basic and fabricated metal, transport equipment	Labor productivity estimates for 1965–70 for 90 four-digit industries in Canada	Gupta (1983) and Fuss and Gupta (1981)
Low: apparel, wood products *High*: other chemicals, food products, printing and publishing	Firm-level production functions for 6,665 plants in Chile, 1979–86	Levinsohn and Petrin (1999)

manufacturing sectors in the United States, Norway, and Chile. Based on trade data, a third of all goods-producing industries have increasing returns to scale. Manufacturing industries with the highest plant-level economies and industry-level externalities are petroleum and coal products (1.40), petroleum refinery (1.19), pharmaceuticals (1.31), machinery (1.11), and iron and steel (1.15). Industries with constant returns are footwear, leather, textiles, apparel, and furniture and fixtures. Similar findings can be found with markups. Because increasing

returns to scale confer some market power on firms, markups of price over marginal cost can be a proxy for plant-level scale economies. Markups for U.S. industries range from 15 percent in apparel to over 200 percent in electricity, gas, and sanitary services.

External Scale Economies

The literature distinguishes between localization and urbanization economies. Urbanization economies come from industrial diversity which foster the exchange of ideas and technology, and produce greater innovation and growth. They tend to be more important for more sophisticated activities requiring learning and new or high-tech activities. Spatial concentration of a diversity of people and products reduces knowledge production costs because information transmission, competition, spying, imitation, learning, innovation, and commercialization of new ideas are easier. Feldman and Audretsch (1999) document that in the United States a staggering 96 percent of innovations are in metropolitan areas.

Localization economies arise from input sharing and competition within the industry and come from geographically concentrated groups of firms linked by the technologies they use, the skills they require, the markets they serve, and the products and services they provide. Competitive pressures that force firms in the same sector to innovate or fail also lead to productivity growth. Conditions tend to be competitive when upstream and downstream firms and associated institutions in a particular industry (say, electronic machinery or petrochemicals) cluster.

The extent of agglomeration economies is influenced by their geographical scope. Density of economic activity and the distance between economic agents influence the productivity gains of scale economies. Table 2 summarizes some stylized findings in the literature. Evidence from Brazil and the United States indicates that doubling the distance to dense metropolitan centers reduces productivity by 15 percent; and doubling the distance from 280 to 550 kilometers reduces profits by 6 percent. The concept of distance can be generalized, in this context, from distance in physical space to distance in industrial space. For example, spillovers between industries are more likely if the industries in question share related scientific spaces and are, therefore, closer in industrial space (Feldman and Audretsch 1999, pp. 409–29). Furthermore, the extent to which distance attenuates agglomeration economies differs for different types of agglomeration. For example, knowledge spillovers that rely on face-to-face communication will decay more quickly with distance than the home market effect (Venables 2006, pp. 61–85).

In contrast to distance, density raises productivity for plants and workers. A large literature on agglomeration economies suggests that doubling city size will

Table 2. External Economies Increase with Density and Decline with Distance

Findings	*Data source*	*Studies*
External economies increase with economic density . . .		
Doubling economic density increases productivity by 6 percent	1988 data on output per worker in U.S. states	Ciccone and Hall (1996)
Doubling employment density increases productivity by 4.5–5.0 percent	Nonagricultural value added per worker in NUTS regions, 1980s	Ciccone (2002)
A one-standard deviation increase in own-industry employment share raises employment growth by 16–31 percent	1970–19: five manufacturing industries in 224 US metro areas	Henderson, Kuncoro, and Turner (1995)
A 10 percent increase in local own-industry employment raises plant output by 0.6 to 0.8 percent for fixed inputs	Korean city-industry data for 1983, 1989, 1991–93	Henderson, Lee, and Lee (2001)
Own-county effect on plant productivity significant but no effect from neighboring county	Plant-level data on productivity, 1972–92 in 742 US counties	Henderson (2003b)
. . . ***and decline with distance***		
Increasing distance from the city center by 1 percent leads to a 0.13 percent decline in productivity	1980 data for 356 new manufacturing firms in Brazil	Hansen (1990)
Doubling the distance to a regional center lowers profits by 6 percent	Firms in auto-part and agricultural machinery (Brazil and the US)	Henderson (1994)
Doubling travel time to a city center reduces productivity by 15 percent	Data for eight industries in Brazil	Sveikauskas, Townroe, and Hansen (1985)
Effects of own-industry employment on new plant births attenuate within the first five 1-mile concentric rings	12 million U.S. establishments from Dun & Bradstreet, Marketplace database	Rosenthal and Strange (2003)

increase productivity by 3–8 percent (see, for example, Shefer 1973; Mera 1973; Segal 1976; Kawashima 1975; Sveikauskas 1975; Moomaw 1981, 1983; Bartlesman, Caballero, and Lyons 1994). Density of activity allows more refined specialization and a wider variety of intermediate inputs. In addition to static agglomeration economies, there are also dynamic agglomeration economies. For example, doubling the density of economic activity in European regions may increase Total Factor Productivity growth by 0.42 percentage points a year (Angeriz, McCombie, and Roberts 2008). While evidence of external economies comes primarily from developed countries (Carlton 1983; Wheeler and Mody 1992; Carlino 1979; Hay 1979), there is also evidence of external economies in developing countries. A survey of 12,400 firms in the manufacturing sector in 120 cities in China points to the higher productivity of firms in more populous cities (World Bank 2006). In the Republic of Korea, Henderson, Shalizi, and Venables (2001) find that a plant in a city with 1,000 workers could, without altering its input mix, increase output by 20–25 percent simply by relocating to

a city that has 15,000 workers in the same industry. Kuncoro (2008), focusing on four broad industry groups in Indonesia, finds that agglomeration economies vary over time and that localization, rather than urbanization, economies tend to dominate.

Despite the availability of land, almost all recent development in the United States has been less than one kilometer from earlier developments (Burchfield and others 2006). Even today, only about 2 percent of the land area of the United States is built-up or paved. This extreme clustering of firms and workers in cities can be explained only by agglomeration economies. As countries develop, their economies become more knowledge-based and service-oriented. Their spatial concentration of activity also rises as knowledge spillovers which require close proximity become more important.

As countries develop and manufacturing and service activities become more important, firms cram in closer to harness agglomeration economies. In the United States (Ellison and Glaeser 1997), the United Kingdom (Devereux, Griffith, and Simpson 2004), and France (Maurel and Sedillot 1999) about 75–95 percent of industry is localized. For instance, in the United States more than a third of the total production of aerospace engines is concentrated in three cities: Hartford with about 18 percent of total employment, and Cincinnati and Phoenix with another 18 percent together (Rosenthal and Strange 2004). Using continuous space without considering administrative boundaries and based on concentration of plants, more than half of the United Kingdom's 122 four-digit industries are localized and only 24 percent are dispersed (Duranton and Overman 2002).

Spatial clustering is more pronounced with high-skill and high-technology industries (electronic computing machinery, process control instruments, semiconductors, pharmaceuticals) than light industries (Kim 1995; Henderson, Lee, and Lee 2001). This is consistent with the documented findings of higher scale effects in heavier industries. High-skill and high-tech industries have more capital-intensive production technology. They are also likely to benefit more from the various mechanisms that generate external economies, as discussed earlier.

Services are even more spatially concentrated than manufacturing for two reasons. First, they tend to use less land per employee. Banks, insurance companies, hospitals, and schools can operate comfortably in high-rise buildings that economize on land and allow for very high density. Second, because of external economies, business services have even greater potential for agglomeration, as firms serve one another: every bank needs advertising, and every advertising firm needs a bank account. The potential for codependence and agglomeration is thus intrinsic to services (Feldman 1994; Dekle and Eaton 1999; Feldman and Audretsch 1999).

Services are among the most agglomerated activities in the United States (Fafchamps and Desmet 2000; Kolko 2007). With the rising prominence of services economic activity will become more concentrated. Jobs in the United States became more spatially concentrated between 1972 and 1992, driven primarily by the rising localization of service activities in larger cities as small and medium size counties lost jobs to the more urban areas (Carlino and Chatterjee 2001; Desmet and Fafchamps 2006). As communication costs fall, services become more tradable, allowing providers to take advantage of narrower specialization and agglomeration economies.

An Urban Hierarchy

The relative importance of localization or urbanization economies differs across the urban hierarchy of cities, depending on a product's life cycle, an industry's tenure in the city, and the city's amenities, natural resources, current industrial profile, historical industrial environment, and local culture. An urban hierarchy exhibits some stylized patterns. Larger metropolises tend to be more diversified and service-oriented. They innovate, invent, breed new firms, and expel mature industries (Calem and Carlino 1991; Glaeser and others 1992; Black and Henderson 1999; Fujita and Ishii 1999; Feldman and Audretsch 1999; Combes 2000; Duranton and Puga 2001, 2004; Desmet and Fafchamps 2006). Smaller cities tend to be industrially specialized. They produce or manufacture standardized products and receive relocated industries from diversified cities (Moomaw 1981, 1983; Nakamura 1985; Sveikauskas, Gowdy, and Funk 1988; Henderson, Kuncoro, and Turner 1995; Henderson 1997a, 1997b, 2003a; Kolko 1999; Glaeser and Mare 2001; Rosenthal and Strange 2001, 2003). And the relative city-size distribution and industrial concentration in cities tend to be stable over time. The stylized observation in most countries is an urban hierarchy of few metropolises and many smaller cities with complementary economic functions.[4]

Improved infrastructure and falling transport costs have encouraged standardized manufacturing production to move out of high-rent centers to smaller cities. Production in large cities focuses on services, research and development, and nonstandardized manufacturing (Glaeser and others 1995; Fafchamps and Desmet 2000). Glaeser and Kahn (2001) document that while it is common to find that manufacturing activities de-concentrate from city centers to their suburbs, services do not. The relocation of manufacturing to suburbs has also been documented in developing countries, for example Colombia, Indonesia, Korea, and Thailand (Henderson, Kuncoro, and Nasution 1996).

Even after controlling for natural comparative advantage, externalities are still important in explaining the patterns of specialization and diversity among cities.

Table 3 summarizes the relative importance of localization and urbanization economies. Intraindustry localization economies have the largest impact after 3–4 years, and eventually peter out after 5–6 years. Interindustry urbanization economies persist much longer. In fairly mature metropolises, competition and diversity help industrial growth. This greater importance of across-industry knowledge spillovers in metropolises seems consistent with the urbanization economies of between-sector innovation. The production of nontraditional items is more concentrated in diverse cities, while standardized traditional goods are located in smaller specialized cities. Similar patterns are found in the United States

Table 3. Localization and Urbanization Economies Vary by City Size and Profile, and by Industry Life Cycle

Findings	*Data source*	*Studies*
Localization economies are more important for heavy industries; urbanization economies are more important for light industries	Data for two-digit manufacturing industries in Japan	Nakamura (1985)
Localization economies become less important, giving way to urbanization economies as cities expand in size	Cross-sectional data for the US and Brazil	Henderson (1986)
Labor pooling is more important in newer and expanding markets while knowledge spillovers and specialized asset-sharing are dominant in mature markets	Firm employment data for four U.S. metro areas and two-digit industries	Hammond and Von Hagen (1994)
For mature capital goods industries, localization effects are significant. For new high-tech industries, both localization and urbanization effects are present	Panel data of 742 urban counties for 1970–87	Henderson, Kuncoro, and Turner (1995)
For all industries localization and urbanization effects are important. For traditional industries most effects die out after four years. For high-tech industries, the effects persist longer	Five traditional and three new high-tech industries in 224 metropolitan areas between 1970 and 1987	Henderson (1997a)
Historical industrial environments of cities matter. In fairly mature cities, urbanization economies encourage industrial growth	Data for the largest industries 1956–87 in 170 U.S. cities	Glaeser and others (1992)
In high-tech industries, a one-standard-deviation increase in industrial diversity raises productivity by 60 percent, but diversity has no effect on light industries	Korean city-industry data for years 1983, 1989, 1991–93	Henderson, Lee, and Lee (2001)

(Henderson 1997b), Japan (Fujita and Ishii 1999), and Korea (Kolko 1999; Henderson, Lee, and Lee 2001).

While new firms often start in diverse cities, they move to specialized ones after they mature (Feldman and Audretsch 1999). Of all new plants in France, for example, 84 percent were created in cities with above-median diversity. Some 72 percent of firm relocations are from a diverse city to a specialized one. In the United States almost all product innovations take place in metropolitan areas. Own-industry specialization in a city has a negative effect on innovative output. However, industrial diversity and city size both have a positive effect on innovative output. Trial plants are based in large cities in Japan, but mass production plants are in small cities or rural areas. Young firms need a period of experimentation. In the early learning phase, diversified cities act as "nurseries" for firms to try out a variety of processes. Once a firm identifies its ideal process, it can begin mass production in specialized cities, where all firms share similar processes or specializations (Fujita and Ishii 1999; Duranton and Puga 2001).

The different economic functions served by cities can also be seen in the clustering of headquarters from different sectors and concentrations of business services in a few large cities, while production plants from each sector congregate in smaller specialized cities. Duranton and Puga (2005) document that in 1950 there was little difference across U.S. cities in their proportions of managers and production workers. Although the largest cities already housed more managers, there was no clear ranking by city size. By 1980, however, the differences across cities had increased substantially, and a clear ranking by size had emerged. Larger cities had become specialized in management and information-intensive activity, which benefit from face-to-face contacts, and smaller cities had become specialized in production. This pattern became even more marked over the following decade.

As discussed above, at the top of this hierarchy are a few large cities serving as incubators for new industries which benefit most from *urbanization economies*. Below these metropolises are medium-sized cities, serving as regional hubs of transportation, finance, and commerce, as well as regional centers of advanced public health, higher education, and culture. They are typically more specialized, focusing on manufacturing. In these cities, *localization economies* tend to dominate. Finally, towns are linked to cities and connected to a mass of rural areas at the bottom of the hierarchy. They are facilitators of *internal scale economies* for mills and factories and market centers for farm products and rural output. Towns draw sustenance from surrounding rural areas and their prosperity also spills over to villages through nonfarm employment opportunities.[5]

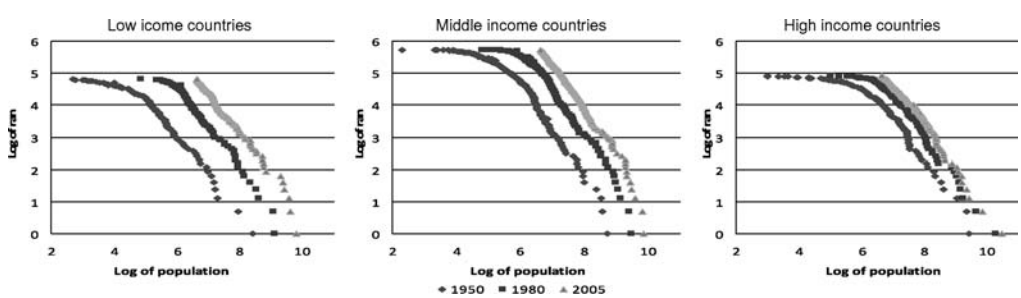
Externalities mean that history matters. Two otherwise identical enterprises in the same city could benefit differently from the local agglomeration depending on how long each has been in the city. Similarly, two otherwise identical cities would

offer different types of external economies depending on their history (Glaeser and others 1992; Henderson, Kuncoro, and Turner 1995). Such intangibles include the local stock of knowledge relevant for an industry or a labor force with specific acquired skills. The influences of history and specialization are consistent with the remarkably stable relative city-size distribution and the industrial concentration in specific cities over time even as incomes and populations grow. This pattern of stable relative size distribution is robust across income levels (figure 2).

This relationship is known as Zipf 's law. As Eeckhout (2004) notes, as early as 1682 Alexandre Le Maître observed a systematic pattern in the size of cities in France. Black and Henderson (2003) show that the relative sizes of U.S. cities have been stable over the last century. Similarly, Eaton and Eckstein (1997) and Dobkins and Ioannides (2001) find a pattern of stability in France, Japan, and the United States. They observe that the relative populations of the top 40 urban areas in France (1876–1990) and Japan (1925–85) remained essentially unchanged.

There is also persistence in the industrial concentration in specific cities (Henderson 1997b). Kim (1995) shows a high (0.64) correlation coefficient of regional localization for two-digit industries in the United States between 1860 and 1987 at the state level. Dumais, Ellison, and Glaeser (2002) also find that, for most industries, agglomeration patterns were strikingly stable over the period 1972–92. Likewise, Henderson (2003a) finds stable specialization patterns over 30 years of nine three-digit industries. Among mature industries the persistence in employment patterns across cities is high and the convergence in individual industry employment across cities is slow (Dunne, Roberts, and Samuelson 1989a, 1989b; Herzog and Schlottmann 1991; Davis and Haltiwanger 1992). Brezis and Krugman (1997) document that cities which underwent major sectoral overhauls were the exceptions. The norm has been the "lock-in" of industrial

Figure 2. Urban Hierarchies Remain Stable over Time and across Incomes



Source: United Nations (2006).

structure in specific cities, which can be explained by localization economies. These cities can better compete and, over time, retain plants and employment in that industry. A larger scale of own-industry activity historically means that firms in that locality today will operate more productively with greater accumulated knowledge about technology, sources of supply of different quality inputs, and local culture and its effect on the legal, business, and institutional climate. Henderson, Kuncoro, and Turner (1995) find that these localization advantages are relevant for traditional manufacturing industries which explain the longevity of particular clusters in certain locations, for example the world-class cutlery cluster in Solingen, Germany, since 1348 (Van der Linde 2003).

The evolving urban hierarchy is an enduring feature of economic development. Settlements of different sizes are linked through complementary economic functions.

Seen another way, places at different positions within the urban hierarchy are at different stages of urbanization. The types of scale economies that dominate in these places will depend on their stage of urbanization. Areas of *incipient* urbanization, with urban shares of about 25 percent, are predominantly agricultural or resource based and need policies to encourage internal economies of scale. Areas of *intermediate* urbanization, where urban population shares are about 50 percent, have seen economic alliances strengthen within and between urban areas. Many firms and plants in the same sector conglomerate to take advantage of shared inputs and knowledge spillovers. In such areas, the promotion of localization economies and congestion control is the highest priority. In areas of *advanced* urbanization, with urban shares of about 75 percent, productivity and consumption benefits arise from urbanization economies associated with the diversity and intensity of economic activity.

Implications for Policy

Towns, cities, and metropolises provide farms, firms, and families with the benefits of proximity, but the geographical compactness of activity also produces congestion, pollution, and social tension that can offset these benefits. Millions in Mumbai live in slums, with little or no access to basic amenities. High levels of crime are an accepted feature of urban living in Latin America. Workers in Jakarta, Kinshasa, Lagos, and Manila spend on average 75 minutes commuting to work. Even in better-managed cities like Seoul and Shanghai, downtown traffic averages 8 kilometers an hour (World Bank 2002). Crime, squalor, and congestion are the diseconomies of agglomeration which can erode the benefits of rising density.

Restricting the growth of cities is not the answer. While there is little evidence that agglomeration economies of megacities have been exhausted (Alonso 1971;

Sheshinski 1973; Kanemoto, Ohkawara, and Suzuki 1996; Au and Henderson 2006), the appropriate policies involve a balancing act between controlling negative externalities and steps to exploit the positive externalities from agglomeration. Most countries undergoing rapid urbanization have not matched urban growth with a commensurate expansion of transport infrastructure. While traffic in developing countries is increasing with per capita income along a path similar to that of the richer countries (World Bank 2002; McCrae 2006), road capacity has not increased enough. Successful cities react to growing congestion with spatially connective infrastructure. But preceding most such infrastructure investments in all successful cities are robust and versatile land market institutions. Land use and building regulations must adapt to shifts in population and production as urbanization advances. They must enable the integration of private and public use of land to provide space for transport infrastructure.

The congestion, grime, and crime-related costs that accompany and undermine rising concentration will, however, differ through the various stages of urbanization. Policies should be calibrated and sequenced accordingly. Towns require mainly the application of common institutions. Cities require the addition of spatially connective policies. And metropolises require all this and spatially targeted efforts such as slum development programs. The success of a policy initiative is predicated on the implementation of the ones introduced before it.

For areas of incipient urbanization, the policy goal in towns and villages includes such *spatially blind institutions* as flexible and efficient working of land markets, especially enforcement of property rights and provision of security, sanitation, streets, schools, and other basic services. For areas of intermediate urbanization, such spatially neutral efforts must be augmented by steps to ease congestion so that the benefits of rising density are more widely shared. *Spatially connective infrastructure* includes mass transit systems, roads and railways, and traffic demand management. Connections among rapidly urbanizing areas facilitate the relocation of production and people to places which supply most efficiently the desired internal or external economies. Finally, for metropolises or areas of advanced urbanization, policy must also address division within urban settlements, as reflected in the residential, as well as the economic and social, segregation of the poor in slums. *Spatially targeted interventions* such as slum upgrading and subsidized housing often become necessary for helping to facilitate urbanization economies.

Consistent with the implied prioritization, economic history and analysis of contemporary experience signal the need for a bedrock of spatially blind policies for equipping every place with the foundation of urbanization, while allowing markets to select the locations at which agglomeration happens. The United Kingdom in the nineteenth century is illustrative. The Reform Act of 1832 and the Municipal Corporations Act of 1835 allowed municipal governments to take

over privately owned sewerage, water, and gas systems to ensure universal provision of such basic amenities.[6] By the 1880s British land market institutions were mature enough to begin unifying patchy systems, separating sewage and drainage systems from the water systems, and extending the reach of basic services to poor areas (Land Enquiry Commission 1914; Hargan 2007). This facilitated rapid urban growth in the following decades (Offer 1981, p. 291).[7]

New York City provides another example of how versatile land market institutions allow a rapidly urbanizing city to accommodate complex public transit networks. Building a new or expanded transport network requires the purchase of contiguous plots of land, and holdouts can extract huge rents or thwart the project. The United States had a well-defined system of property rights by the mid-nineteenth century. As New York's transport networks expanded, its institutions could adapt. The 1916 Zoning Resolution was amended numerous times over the following decades as market needs changed (Dunlap 1992). Successful cities have flexible zoning laws to allow higher value users to bid for the valuable land. The interaction between such institutions and connective infrastructure enabled the density of Manhattan, the Bronx, Brooklyn, and Queens to increase from 230 people per square kilometer in 1820 to more than 5,000 in 1900 and about 12,000 today.[8]

Flexible land market institutions, the universal provision of basic social services, and investments in connective infrastructure help to make spatially targeted interventions such as slum clearance programs more likely to succeed. Hong Kong is an example (Bristow 1984; Adams and Hastings 2001; Cullinane 2002). Consistent with a tradition of minimal government intervention, Hong Kong's government contracted urban redevelopment to an organization led by private sector interests. Over five decades, slums were gradually eliminated. Similar success in a sustainable elimination of slums because of strong aspatial institutions coupled with connective infrastructure can be found in the United States (Chandler 1992; von Hoffman 1996; Hall 2002; Abreu 2008), Sweden (Swedish Council for Building Research 1990; Nesslein 2003; Borgegård and Kemeny 2004; Hall and Vidén 2005; Abreu 2008), and Costa Rica (Hall 1984; Trackman, Fisher, and Salas 1999; Ruster and Imparato 2003; World Bank 2007; Abreu 2008).

While an area's stage of urbanization (for example town, city, or metropolis) determines policy priorities, a nation's urban share can be a good indicator of the overall complexity of its urbanization challenges. Different parts of a country urbanize at different speeds. In the same way, more urbanized countries have more of their people in high-density areas and less urbanized countries have most people in rural areas. In the simplest case, one area may characterize an entire country, as in Singapore. For larger countries, more refined disaggregation can help determine priorities at different levels of government.

In a famous article Tinbergen (1952) proposed that one policy instrument is needed to address each policy objective. Urbanization's challenges increase with the stage of urbanization, so the required number of policy instruments increases as well. Fortunately for developing nations, the capacity of markets and governments grows as they urbanize. But these policies must be introduced in the right sequence. The institutional foundations of an inclusive urbanization have to be instituted early in the development process. Investments in infrastructure should come next. When implemented in settings where the institutions related to land markets and social services function reasonably well, and where transport infrastructure is adequate, targeted interventions aimed at reducing slums and squalor can be successful.

Notes

Indermit S. Gill and Chor-Ching Goh are, respectively, the Chief Economist and Senior Economist in Europe and Central Asia Region of the World Bank; email address: cgoh@worldbank.org. The authors would like to thank Sarah Boeckmann, Jean-Jacques Dethier, Vernon Henderson, John McCombie, Diego Puga, Mark Roberts, three anonymous referees, and the editor of this journal for helpful suggestions.

1. There were, of course, some exceptions, most notably in the work of Young (1928) and Kaldor (1972). These economists did not provide the technical solutions associated with the modeling of increasing returns to scale.

2. Warsh (2006) provides an entertaining and informative account of intellectual progress in dealing with increasing returns to scale. The advances are based on the special features of ideas, highlighted elegantly in Romer (1986, 1989). By adding knowledge explicitly to formulations of economic growth, economists are able to recognize the centrality of ideas and the importance of increasing returns, but this also requires recognizing the proliferation of imperfect competition.

3. Discussion of agglomeration economies dates as far back as Smith's consideration of specialization and the division of labor (Smith 1979); Marshall's information spillovers, searching and matching processes, and input sharing (Marshall 1890); and more recently interindustry supply linkages (Chinitz 1961); learning-by-doing (Arrow 1962); and cross-fertilization of ideas and innovation (Jacobs 1969). There are also consumption externalities from agglomeration, which are understudied. They are not covered here (but see Glaeser, Kolko, and Saiz 2001; Sinai and Waldfogel 2004).

4. This was recognized in geography as long ago as and.

5. Capello and Camagni (2000) argue that the real issue is not the size but the functional characteristics of a city, which depends on its position within the urban system. While negative externalities start to erode scale economies after a certain city size, urban development also generates conditions leading to structural readjustments which will create new economic advantages and innovative measures which will address diseconomies. These structural adjustments may be sectoral transformations toward higher-order functions or increases in external linkages with other cities. They provide econometric evidence of these processes among 58 Italian cities that confirm the integrative nature of an urban network.

6. Based on the historical collection of transport pamphlets the London School of Economics library has converted into digital format and made available for public use. See Roberts (2008).

7. More policy examples can be found in Chapter 7 of the *World Development Report 2009: Reshaping Economic Geography* (World Bank 2008).

8. Population and area figures are available at the US Census Bureau website.

References

The word *processed* describes informally reproduced works that may not be commonly available through libraries.

Abreu, Maria. 2008. "Effectively Dealing with Slums." Unpublished manuscript as background paper for the *World Development Report 2009 Reshaping Economic Geography.*

Adams, David, and E.M. Hastings. 2001. "Urban Renewal in Hong Kong: Transition From Development Corporation to Renewal Authority." *Land Use Policy* 18(3):245–58.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 60(2):323–51.

_____. 2005. "Growth with Quality-Improving Innovations: An Integrated Framework." In Philippe Aghion, and Steven Durlauf, eds., *Handbook of Economic Growth*, Volume 1A. Amsterdam, Netherlands: North-Holland.

Alonso, William. 1971. "The Economics of Urban Size." *Regional Science Association Papers* 26: 67–83.

Angeriz, Alberto, John McCombie, and Mark Roberts. 2008. "Estimates of Returns to Scale and Spatial Spillovers for EU Regional Manufacturing." *International Regional Science Review* 31(62): 62–87.

Antweiler, Werner, and Daniel Trefler. 2002. "Increasing Returns and All That: A View from Trade." *American Economic Review* 92(1):93–119.

Arnott, Richard J. 1979. "Optimal City Size in a Spatial Economy." *Journal of Urban Economics* 6: 65–89.

Arrow, Kenneth J. 1962. "Economic Welfare and the Allocation of Resources for Invention." In Richard R. Nelson, ed., *The Rate and Direction of Inventive Activity.* Princeton, NJ: Princeton University Press.

Au, C.C., and Henderson, V. 2006. "Are Chinese Cities Too Small?" *Review of Economic Studies* 73: 549–76.

Baldwin, John R., and Paul K. Gorecki 1986. *Role of Scale in Canada–U.S. Productivity Differences in the Manufacturing Sector 1970–1979*. Toronto, Canada: University of Toronto Press.

Baldwin, Richard E., and Philippe Martin. 2004. "Agglomeration and Regional Growth". CEPR Discussion Paper 3960.

Bartlesman, Eric J., Ricardo J. Caballero, and Richard K. Lyons 1994. "Customer and Supplier-Driven Externalities." *American Economic Review* 84(4):1075–84.

Bikhchandani, Suchil, David Hirshleifer, and Ivo Welch. 1998. "Learning from the Behavior of Others: Conformity, Fads, and Information Cascades." *Journal of Economic Perspectives* 12(3): 151–70.

Black, Duncan, and J. Vernon Henderson 1999. "A Theory of Urban Growth." *Journal of Political Economy* 107 (2): 252–84.

_____ 2003. "Urban Evolution in the USA." *Journal of Economic Geography* 3(4):343–73.

Borgegård, Lars-Erik, and Jim Kemeny. 2004. "Sweden: High-Rise Housing for a Low-Density Country." In R. Turkington, R. van Kempen, and E. Wassenberg, eds., *High-Rise Housing in Europe: Current Trends and Future Prospects*. Delft, Netherlands: Delft University Press.

Brezis, Elise S., and Paul Krugman. 1997. "Technology and the Life Cycle of Cities." *Journal of Economic Growth* 2(4):369–83.

Bristow, R. 1984. *Land-Use Planning in Hong Kong: History, Policies, and Procedures*. Hong Kong: Oxford University Press.

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew Turner. 2006. "Causes of Sprawl: A Portrait from Space." *Quarterly Journal of Economics* 121(2):587–633.

Calem, Paul S., and Gerald A. Carlino 1991. "Urban Agglomeration Economies in the Presence of Technical Change." *Journal of Urban Economics* 29(1):82–95.

Capello, Roberta, and Roberto Camagni. 2000. "Beyond Optimal City Size: An Evaluation of Alternative Urban Growth Patterns." *Urban Studies* 37(9):1479–96.

Carlino, Gerald A. 1979. "Increasing Returns to Scale in Metropolitan Manufacturing." *Journal of Regional Science* 19(3): 369–72.

Carlino, Gerald, and Satyajit Chatterjee. 2001. "Employment Deconcentration: A New Perspective on America's Postwar Urban Evolution." Philadelphia, PA: Federal Reserve Bank of Philadelphia Working Paper Series 01-4.

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):440–9.

Chandler, Mittie Olion. 1992. "Public Housing Desegregation: What Are the Options?" *Housing Policy Debate* 3(2):509–34.

Charlot, Sylvie, and Gilles Duranton. 2003. "Communication Externalities in Cities." London, UK: Centre for Economic Policy Research Working Paper Series 4048.

Chinitz, Benjamin J. 1961. "Contrasts in Agglomeration: New York and Pittsburgh." *American Economic Review* 51(2):279–89.

Christaller, Walter. 1933. *Central Places in Southern Germany.* London, UK: Prentice Hall.

Ciccone, Antonio. 2002. "Agglomeration Effects in Europe." *European Economic Review* 46(2): 213–27.

Ciccone, Antonio, and Robert E. Hall 1996. "Productivity and the Density of Economic Activity." *American Economic Review* 86(1):54–70.

Combes, Pierre-Philippe. 2000. "Economic Structure and Local Growth: France 1984–1993." *Journal of Urban Economics* 47(3):329–55.

Costa, Dora L., and Matthew E. Kahn 2000. "Power Couples." *Quarterly Journal of Economics* 115(4): 1287–315.

Cullinane, Sharon. 2002. "The Relationship between Car Ownership and Public Transport Provision: A Case Study of Hong Kong." *Transport Policy* 9(1):29–39.

Davis, Steven J., and John C. Haltiwanger 1991. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation." Chicago, IL: Federal Reserve Bank of Chicago Working Papers 91–5.

_____. 1992. "Gross Job Creation, Gross Job Destruction, and Employment Reallocation." *Quarterly Journal of Economics* 107(3): 819–63.

Dekle, Robert, and Jonathan Eaton. 1999. "Agglomeration and Land Rents: Evidence from the Prefectures." *Journal of Urban Economics* 46(2):200–14.

Desmet, Klaus, and Marcel Fafchamps. 2006. "Employment Concentration Across U.S. Counties." *Regional Science and Urban Economics* 36(4):482–509

Devereux, Michael P., Rachel Griffith, and Helen Simpson. 2004. "The Geographic Distribution of Production Activity in the U.K." *Regional Science and Urban Economics* 34(5):533–64.

Diamond, P.A., and J.A. Mirrlees 1971. "Optimal Taxation and Public Production; I: Production Efficiency" and "II: Tax Rules." *American Economic Review* 61:8–27 and 261–78.

_____. 1973. "Aggregate Productions and Consumption Externalities," *Quarterly Journal of Economics*, LXXXVI: 1–24.

Dixit, Avinash. 1973. "The Optimum Factory Town." *Bell Journal of Economics and Management Science* 4(2):637–51.

Dixit, Avinash, and Joseph E. Stiglitz 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67(3):297–308.

Dobkins, Linda H., and Yannis M. Ioannides 2001. "Spatial Interactions Among US Cities, 1900–1990." *Regional Science and Urban Economics* 31(6):701–32.

Dumais, Guy, Glenn Ellison, and Edward L. Glaeser 2002. "Geographic Concentration as a Dynamic Process." *Review of Economics and Statistics* 84(2): 193–204.

Dunlap, David W. 1992. "Some Land-Use History Highlights." *New York Times*, April 12.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. 1989a. "Plant Turnover and Gross Employment Flows in the US Manufacturing Sector." *Journal of Labor Economics* 7(1):48–71.

Dunne, Timothy, Mark J Roberts, and Larry Samuelson. 1989b. "The Growth and Failure of US Manufacturing Plants." *Quarterly Journal of Economics* 104(4):671–98.

Duranton, Gilles. 2007. "Urban Evolutions: The Fast, the Slow, and the Still." *American Economic Review* 97(1):197–221.

Duranton, Gilles, and Henry G. Overman 2002. "Testing for Localisation Using Micro-Geographic Data." London, UK: Centre for Economic Performance Discussion Paper 540.

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.

_____. 2004. "Micro-Foundations of Urban Agglomeration Economies." In J. Vernon Henderson, and Jacque Thisse, eds., *Handbook of Urban and Regional Economies*, Volume 4. Amsterdam: North Holland.

_____. 2005 . "From Sectoral to Functional Urban Specialization." *Journal of Urban Economics* 57(2): 343–70.

Easterly, William, and Ross Levine. 2001. "What Have We Learned from a Decade of Empirical Research on Growth? It's Not Factor Accumulation: Stylized Facts and Growth Models." *World Bank Economic Review* 15(2):177–219.

Eaton, Jonathan, and Zvi Eckstein. 1997. "Cities and Growth: Evidence from France and Japan." *Regional Science and Urban Economics* 27(4–5):443–74.

Eeckhout, Jan. 2004. "Gibrat's Law for (All) Cities." *American Economic Review* 94(5):1429–51.

Ellison, Glenn, and Edward L. Glaeser 1997. "Geographic Concentration in US Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy* 105(5):889–927.

Ethier, Wilfred J. 1982. "National and International Returns to Scale in the Theory of International Trade." *American Economic Review* 72(3):389–405.

Fafchamps, Marcel, and Klaus Desmet. 2000. "The Changing Spatial Distribution of Economic Activity Across US Counties." Oxford, UK: Economic Series Working Paper 43.

Feldman, Maryann P. 1994. *The Geography of Innovation*. Boston, MA: Kluwer.

Feldman, Maryann P., and David B. Audretsch 1999. "Innovation in Cities: Science-Based Diversity, Specialization and Localized Competition." *European Economic Review* 43(2):409–29.

Fujita, Masahisa, and Ryoichi Ishii. 1999. "Global Location Behavior and Organizational Dynamics of Japanese Electronics Firms and Their Impact on Regional Economies." In Alfred D. Chandler Jr., Peter Hagstrom, and Orjan Solvell, eds., *The Dynamic Firm: The Role of Technology, Strategy, Organization, and Regions*. Oxford, U.K.: Oxford University Press.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables 1999. *The Spatial Economy. Cities, Regions, and International Trade*. Cambridge, MA: The MIT Press.

Fuss, Melvyn A., and Vinod K. Gupta 1981. "A Cost Function Approach to the Estimation of Minimum Efficient Scale, Returns to Scale, and Suboptimal Capacity : With an Application to Canadian Manufacturing." *European Economic Review* 15(2):123–35.

Gill, Indermit, and Homi Kharas. 2007. *An East Asia Renaissance: Ideas for Economic Growth*. Washington, DC: The World Bank.

Glaeser, Edward L., José A. Scheinkman, and Andrei Schleifer. 1995. "Economic Growth in a Cross-Section of Cities." *Journal of Monetary Economics* 36(1): 117–43.

Glaeser, Edward L., and Matthew E. Kahn 2001. "Decentralized Employment and the Transformation of the American City." Boston, MA: Harvard Institute of Economic Research Working Paper Series 1912.

Glaeser, Edward L., and David C. Mare 2001. "Cities and Skills." *Journal of Labor Economics* 19(2): 316–42.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. "Consumer City." *Journal of Economic Geography* 1(1):27–50.

Glaeser, Edward L., Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100(6):1126–52.

Griliches, Zvi, and Vidar Ringstad. 1971. *Economies of Scale and the Form of the Production Function: An Econometric Study of Norwegian Manufacturing Establishment Data*. Amsterdam, Netherlands: North-Holland.

Grossman, Gene M., and Elhanan Helpman. 1991. "Quality Ladders in the Theory of Growth." *Review of Economic Studies* 58(1):43–61.

_____. 1995. "Comparative Advantage and Long Run Growth." In J.P. Neary, ed., *International Trade*. Aldershot, UK: Edward Elgar.

Gupta, Vinod K. 1983. "Labor Productivity, Establishment Size, and Scale Economies." *Southern Economic Journal* 49(3):853–9.

Hall, Carolyn. 1984. "Regional Inequalities in Well-Being in Costa Rica." *Geographical Review* 74(1): 48–62.

Hall, Peter. 2002. *Cities of Tomorrow*. Oxford, UK: Blackwell Publishing.

Hall, Thomas, and Sonja Vidén. 2005. "The Million Homes Programme: A Review of the Great Swedish Planning Model." *Planning Perspectives* 20(3):301–28.

Hammond, George William, and Jürgen von Hagen. 1994. "Industrial Localization: An Empirical Test for Marshallian Localization Economies." Working Paper 917, Centre for Economic Policy Research, London.

Hansen, Eric R. 1990. "Agglomeration Economics and Industrial Decentralization: The Wage-Productivity Trade-Offs." *Journal of Urban Economics* 28(2):140–59.

Hargan, Jim. 2007. "Prince Charles's Poundbury." *British Heritage*. May.

Hay, Donald Andrew. 1979. "The Location of Industry in a Developing Country: The Case of Brazil." *Oxford Economic Papers* 31(1):93–120.

Helpman, Elhanan, and Paul Krugman. 1987. *Market Structure and Foreign Trade*. Cambridge, MA: MIT Press.

Henderson, J. Vernon. 1974a. "The Sizes and Types of Cities." *American Economic Review* 64(4): 640–56.

_____. 1974b. "Optimum City Size: The External Diseconomy Question." *Journal of Political Economy* 82(2):373–88.

_____. 1975. "Congestion and Optimum City Size." *Journal of Urban Economics* 2:48–62.

_____, 1977. *Economic Theories and Cities*, New York: Academic Press.

_____. 1986. "Efficiency of Resources Usage and City Size." *Journal of Urban Economics* 19(1): 47–70.

_____. 1994. "Where Does an Industry Locate?" *Journal of Urban Economics* 35(1):83–104.

_____. 1997a. "Externalities and Industrial Development." *Journal of Urban Economics* 42(3): 449–70.

_____. 1997b. "Medium Size Cities." *Regional Science and Urban Economics* 27(6):583–612.

_____. 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.

Henderson, J., Vernon Ari Kuncoro, and Damhuri Nasution. 1996. "The Dynamics of Jabotabek Development." *Bulletin of Indonesian Economic Studies* 32(1):71–95.

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):479–501.

Henderson, J. Vernon, Zmarak Shalizi, and Anthony J. Venables 2001. "Geography and Development." *Journal of Economic Geography* 1(1):81–105.

Herzog, Henry W., and Alan M. Schlottmann 1991. *Industry Location and Public Policy*. Knoxville, TN: University of Tennessee.

Hoffman A., von. 1996. "High Ambitions: The Past and Future of American Low-Income Housing Policy." *Housing Policy Debate* 7(3):423–46.

Holmes, Thomas J., and John J. Stevens 2002. "Geographic Concentration and Establishment Scale." *Review of Economics and Statistics* 84(4):682–91.

Jacobs, Jane. 1969. *The Economy of Cities*. New York: Vintage.

Jaffe, Adam B. 1986. "Technological Opportunity and Spillover of R&D: Evidence from Firms' Patents, Profits, and Market Value." *American Economic Review* 76(5): 984–1001

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidence by Patent Citations." *Quarterly Journal of Economics* 108(3): 577–98.

Junius, Karsten. 1997. "Economies of Scale: A Survey of Empirical Literature." Kiel, Germany: Kiel Institute of World Economics Working Paper Series 813.

Kaldor, Nicholas. 1972. "The Irrelevance of Equilibrium Economics." *Economic Journal* 82(328): 1237–55.

Kanemoto, Yoshitsugu, Toru Ohkawara, and Tsutomu Suzuki. 1996. "Agglomeration Economies and a Test for Optimal City Sizes in Japan." *Journal of the Japanese and International Economies*, 10(4):379–98.

Kawashima, Tatsuhiko. 1975. "Urban Agglomeration Economies in Manufacturing Industries." *Papers in Regional Science* 34(1):157–75.

Kim, Sukkoo. 1995. "Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in US Regional Manufacturing Structure, 1860–1987." *Quarterly Journal of Economics* 110(4):881–908.

Kim, Sukkoo, and Robert A. Margo 2004. "Historical Perspectives on U.S. Economic Geography." In J. Vernon Henderson, and Jacque-Francois Thisse, eds., *Handbook of Regional and Urban Economics, Volume 4: Cities and Geography*. Amsterdam: North-Holland.

Kolko, Jed. 1999. "Can I get Some Service Here? Information Technology, Service Industries, and the Future of Cities." PhD thesis. Harvard University.

_____. 2007. "Agglomeration and Co-Agglomeration of Services Industries." Public Policy Institute of California, San Francisco. Processed.

Krugman, Paul R. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review* 70(5):950–9.

_____ 1981. "Intraindustry Specialization and the Gains from Trade." *Journal of Political Economy* 89(5):959–73.

_____ 1991. "Increasing Returns to Scale and Economic Geography." *Journal of Political Economy* 99(3):483–99.

Kuncoro, Ari. 2008. "Spatial Agglomeration, Site Productivity and Local Governance: Indonesian Experience from 1980 to 2003." In Yukon Huang, and Alessandro Magnoli Bocchi, eds., *Reshaping Economic Geography in East Asia*. Washington, DC: World Bank; EAP Companion Volume to the WDR 2009.

Land Enquiry Commission. 1914. *The Land*, Vol. II. London: Hodder and Stoughton.

Levinsohn, James, and Amil Petrin. 1999. "When Industries Become More Productive, Do Firms?" NBER Working Paper 6893, National Bureau of Economic Research, Cambridge, MA.

Lösch, Auguste. 1940. *The Economics of Location*. New Haven, CT: Yale University Press.

Lucas Jr. Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22(1):3–42.

Marshall, Alfred. 1890. *Principles of Economics*. London, UK: Macmillan.

Martins, Joaquim Oliveira, Stefano Scarpetta, and Dirk Pilat. 1996. "Mark-up Ratios in Manufacturing Industries: Estimates for 14 OECD Countries." Working Paper 162, Organisation for Economic Co-operation and Development, Paris.

Maurel, Françoise, and Beatrice Sedillot. 1999. "A Measure of the Geographic Concentration in French Manufacturing Industries." *Regional Science and Urban Economics* 29(5):575–604.

McCrae, Ian. 2006. "Uncertainties in Road Transport Emissions in Developing Countries." HM Treasury, London.

Mera, Koichi. 1973. "Regional Production Function and Social Overhead Capital: An Analysis of Japanese Case." *Regional Science and Urban Economics* 3(2):157–86.

Mills, Edwin S. 1967. "An Aggregative Model of Resource Allocation in a Metropolitan Area." *American Economic Review*, 57(2):197–210.

Mills, Edwin S., and David M. Ferranti 1971. "Market Choices and Optimal City Size." *American Economic Review* 61(2):340–345.

Moomaw, Ronald L. 1981. "Productivity and City Size: A Critique of the Evidence." *Quarterly Journal of Economics* 96(4):675–88.

_____. 1983. "Is Population Scale a Worthless Surrogate for Business Agglomeration Economies?" *Regional Science and Urban Economics* 13(4):525–45.

Moretti, Enrico. 2004a. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics*121(1–2):175–212.

_____. 2004b. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94(3):656–90.

Nakamura, Ryohei. 1985. "Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities." *Journal of Urban Economics* 17(1):108–24.

Nesslein, Thomas S. 2003. "Markets versus Planning: An Assessment of the Swedish Housing Model in the Post-War Period." *Urban Studies* 40(7):1259–82.

Offer, A. 1981. *Property and Politics, 1870–1914*. Cambridge, UK: Cambridge University Press.

Paul, Catherine, J. Morrison, and Donald S. Siegel 1999. "Scale Economies and Industry Agglomeration Externalities: A Dynamic Cost Function Approach." *American Economic Review* 89(1):272–90.

Prateen, Cliff. 1988. "Survey of the Economies of Scale." Economic Papers 67, Commission of the European Communities, Brussels.

Rauch, James E. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics* 34(3):380–400.

Roberts, Mark. 2008. "Congestion and Spatially Connective Infrastructure: The Case of London in the 19th and early 20th Century." Unpublished manuscript as background paper for the *World Development Report 2009 Reshaping Economic Geography*.

Roeger, Werner. 1995. "Can Imperfect Competition Explain the Difference between Primal and Dual Productivity Measures? Estimates for U.S. Manufacturing." *Journal of Political Economy* 103(2): 316–30.

Romer, Paul M. 1986. "Increasing Returns on Long-Run Growth." *Journal of Political Economy* 94(5):1002–37.

_____. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5): S71–102.

Romer, Paul M., Rosenthal, Stuart S., and William C. Strange 2001. "The Determinants of Agglomeration." *Journal of Urban Economics* 50(2):191–229.

_____. 2003. "Geography, Industrial Organization, and Agglomeration." *Review of Economics and Statistics* 85(2):377–93.

_____. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In J. Vernon Henderson, and Jacque Thisse, eds., *Handbook of Regional and Urban Economics*, Volume 4. Amsterdam:North-Holland: Elsevier.

Rossi-Hansberg, Esteban, and Mark L.J. Wright 2007. "Urban Structure and Growth." *Review of Economic Studies* 74(2):597–624.

Ruster, Jeff, and Ivo Imperato. 2003. *Slum Upgrading and Participation: Lessons from Latin America*. Washington, DC: World Bank.

Santos, Georgina, and Blake Shaffer. 2004. "Preliminary Results of the London Congestion Charging Scheme." *Public Works Management* 9:164–81.

Sasaki, Komei. 1998. "Optimal Urban Growth Controls." *Regional Science and Urban Economics* 28: 475–96.

Segal, David. 1976. "Are there Returns to Scale in City Size?" *Review of Economics and Statistics* 58(3):339–50.

Shefer, Daniel. 1973. "Localization Economies in SMSAs: A Production Function Analysis." *Journal of Regional Science* 13(1):55–64.

Sheshinski, Eytan. 1973. "Congestion and the Optimum City Size." *American Economic Review* 63(2):61–6.

Sinai, Todd, and Joel Waldfogel. 2004. "Geography and the Internet: Is the Internet a Substitute or Complement for Cities?" *Journal of Urban Economics* 56(1):1–24.

Smith, Adam. 1979. *An Inquiry into the Nature and Causes of the Wealth of Nations, 1939 Canaan Edition*. London: Penguin Books.

Spence, Michael, 1976. "Product Selection, Fixed Costs, and Monopolistic Competition." *Review of Economic Studies* 43(2):217–35.

Sveikauskas, Leo A. 1975. "The Productivity of Cities." *Quarterly Journal of Economics* 89(3): 393–413.

Sveikauskas, Leo A., John Gowdy, and Michael Funk. 1988. "Urban Productivity: City Size or Industry Size." *Journal of Regional Science* 28(2):185–202.

Sveikauskas, Leo A., Peter M. Townroe, and Eric R. Hansen 1985. "Intraregional Productivity Difference in Sao Paulo State Manufacturing Plants." *Weltwirtschaftliches Archiv* 121(4):722–40.

Swedish Council for Building Research. 1990. *Housing Research and Design in Sweden*. Stockholm: Swedish Council for Building Research.

Tinbergen, J. 1952. *On the Theory of Economic Policy*. Amsterdam: North-Holland.

Trackman, B., W. Fisher, and L. Salas 1999. *The Reform of Property Registration Systems in Costa Rica: A Status Report*. Cambridge, MA: Law School Harvard University.

Tolley, George. 1974. "The Welfare Economics of City Bigness." *Journal of Urban Economics* 1: 324–45.

Tolley, George, and John Crihfield. 1987. "City Size and Place as Policy Issues." In E.S. Mills, ed., *Handbook of Regional and Urban Economics* Amsterdam: Elsevier North Holland :1285–311.

United Nations. 2006. "World Urbanization Prospects: The 2005 Revision". New York, United Nations, Department of Economic and Social Affairs, Population Division.

Van der Linde, Claas. 2003. "The Demography of Clusters: Findings from the Cluster Meta-Study." In J. Brocker, D. Dohse, and R. Soltwedel, eds., *Innovation Clusters and Interregional Competition*. Berlin, Germany: Springer.

Venables, Anthony J. 2006. "Shifts in Economic Geography and Their Causes." *Federal Reserve Bank of Kansas City Economic Review* Q IV:61–85.

Warsh, David. 2006. *Knowledge and the Wealth of Nations: A Story of Economic Discovery*. New York: W.W. Norton.

Wheeler, David, and Ashoka Mody. 1992. "International Investment Location Decisions: The Case for U.S. Firms." *Journal of International Economics* 33(1–2):57–76.

World Bank. 2002. *Cities on the Move: A World Bank Urban Transport Strategy Review*. Washington, DC: World Bank.

_____. 2006. *Governance, Investment Climate, and Harmonious Society: Competitiveness Enhancements for 120 Cities in China*. Washington, DC: World Bank.

_____. 2007. *World Development Indicators 2007*. Washington, DC: World Bank.

_____. 2008. *World Development Report 2009 Reshaping Economic Geography*. Washington, DC: World Bank.

Young, Allwyn. 1928. "Increasing Returns and Economic Progress." *Economic Journal* 38(152): 527–42.