changes in spatially delineated public goods. *International Economic Review* 45, 1047–77.
Tolley, G. 1974. The welfare economics of city bigness. *Journal of Urban Economics* 1, 324–45.
United Nations Population Division. 2004. *World Population Prospects: The 2004 Revision Population Database.* Online. Available at http://esa.un.org/unpp, accessed 28 June 2005.

urban growth

Urban growth – the growth and decline of urban areas – as an economic phenomenon is inextricably linked with the process of urbanization. Urbanization itself has punctuated economic development. The spatial distribution of economic activity, measured in terms of population, output and income, is concentrated. The patterns of such concentrations and their relationship to measured economic and demographic variables constitute some of the most intriguing phenomena in urban economics. They have important implications for the economic role and size distribution of cities, the efficiency of production in an economy, and overall economic growth. As Paul Bairoch's magisterial work (1988) has established, increasingly concentrated population densities have been closely linked since the dawn of history with the development of agriculture and transportation. Yet, as economies move from those of traditional societies to their modern stage, the role of the urban sector changes from merely providing services to leading in innovation and serving as engines of growth.

Measurement of urban growth rests on the definition of 'urban area', which is not standard throughout the world and differs even within the same country depending upon the nature of local jurisdictions and how they might have changed over time (this is true even for the United States). Legal boundaries might not indicate the areas covered by urban service-providers. Economic variables commonly used include population, area, employment, density or output measures, and occasionally several of them at once, not all of which are consistently available for all countries. Commuting patterns and density measures may be used to define metropolitan statistical areas in the USA as economic entities, but major urban agglomerations may involve a multitude of definitions.

The study of urban growth has proceeded in a number of different directions. One direction has emphasized historical aspects of urbanization. Massive population movements from rural to urban areas have fuelled urban growth throughout the world. Yet it is fair to say that economics has yet to achieve a thorough understanding of the intricate relationships between demographic transition, agricultural development and the forces underlying the Industrial Revolution. Innovations were clearly facilitated by urban concentrations and associated technological improvements. A related direction focuses on the physical structure of cities and how it may change as cities grow. It also focuses on how changes in commuting costs, as well as the industrial composition of national output and other technological changes, have affected the growth of cities. Another direction has focused on understanding the evolution of systems of cities – that is, how cities of different sizes interact, accommodate and share different functions as the economy develops and what the properties of the size distribution of urban areas are for economies at different stages of development. Do the properties of the system of cities and of city size distribution persist while national population is growing? Finally, there is a literature that studies the link between urban growth and economic growth. What restrictions does urban growth impose on economic growth? What economic functions are allocated to cities of different sizes in a growing economy? Of course, all of these lines of inquiry are closely related and none of them may be fully understood, theoretically and empirically, on its own. In what follows we address each in turn.

Urbanization and the size distribution of cities

The concentration of population and economic activity in urban areas may increase either because agents migrate from rural to urban areas (urbanization) or because economies grow in term of both population and output, which results in urban as well as rural growth. Urban centres may not be sustained unless agricultural productivity has increased sufficiently to allow people to move away from the land and devote themselves to non-food producing activities. Such 'symmetry breaking' in the uniform distribution of economic activity is an important factor in understanding urban development (Papageorgiou and Smith, 1983). Research on the process of urbanization spans the early modern era (the case of Europe having been most thoroughly studied; De Vries, 1984) to recent studies that have applied modern tools to study urbanization in East Asia (Fujita et al., 2004). The 'New Economic Geography' literature has emphasized how an economy can become 'differentiated' into an industrialized core (urban sector) and an agricultural 'periphery' (Krugman, 1991). That is, urban concentration is beneficial because the population benefits from the greater variety of goods produced (forward linkages) and may be sustained because a larger population in turn generates greater demand for those goods (backward linkages). This process exploits the increasing returns to scale that characterize goods production but does not always lead to concentration of economic activity. The range of different possibilities is explored extensively in Fujita, Krugman and Venables (1999). These ideas have generated new lines of research; see several related papers in Henderson and Thisse (2004).

The process of urban growth is closely related to the size distribution of cities. As the urban population grows, will it be accommodated in a large number of small cities, or in a

small number of large cities, or in a variety of city sizes? While cities have performed different functions in the course of economic development, a puzzling fact persists for a wide cross-section of countries and different time periods. The size distribution of cities is Pareto-distributed, is 'scale-free'. Gabaix (1999) established this relationship formally. He showed that, if city growth is scale independent (the mean and variance of city growth rates do not depend on city size: Gibrat's Law) and the growth process has a reflective barrier at some level arbitrarily close to zero, the invariant distribution of city sizes is a Pareto distribution with coefficient arbitrarily close to 1: Zipf's Law. (Empirical evidence on the urban growth process as well as Zipf's Law is surveyed by Gabaix and Ioannides, 2004.)

These results imply that the size distribution of cities and the process of urban growth are closely related. Eeckhout (2004) extends the empirical investigation by examining in depth all *urban places* in the United States and finds that the inclusion of the lower end of the sample leads to a log-normal size distribution. Duranton (2004) refines the theory by means of a quality-ladder model of economic growth that allows him to model the growth and decline of cities as cities win or lose industries following technological innovations. Ultimately, the movements of cities up and down the hierarchy balance out so as to produce a stable, skewed size distribution. This theory is sufficiently rich to accommodate subtle differences across countries (in particular the United States and France) that constitute systematic differences from Zipf's Law. Rossi-Hansberg and Wright (2004) use a neoclassical growth model that is also consistent with observed systematic deviations from Zipf's Law: in particular, the actual size distribution of cities shows fewer smaller and larger cities than the Pareto distribution, and the coefficient of the Pareto distribution has been found to be different from 1 although centred on it. They identify the standard deviation of the industry productivity shocks as the key factor behind these deviations from Zipf's Law. The evident similarity of the conclusions of those two papers clearly suggests that the literature is closer than ever before to resolving the Zipf's Law 'puzzle.'

Urban growth and city structure

Understanding urbanization and economic growth requires understanding the variety of factors that can affect city size and therefore its short-term dynamics. All of them lead to the basic forces that generate the real and pecuniary externalities that are exploited by urban agglomeration, on one hand, and congestion, which follows from agglomeration, on the other. Three basic types of agglomeration forces have been used, in different varieties, to explain the existence of urban agglomerations (all of them were initially proposed in Marshall, 1920): (*a*) knowledge spillovers, that is, the more biomedical research there is in an urban area, the more productive a new research laboratory will be; (*b*) thick markets for specialized inputs: the more firms that hire specialized programmers, the larger the pool from which an additional firm can hire when the others may be laying off workers; and (*c*) backward and forward linkages. Local amenities and public goods can themselves be relevant agglomeration forces.

The size of urban agglomerations is the result of a trade-off between the relevant agglomeration and congestion forces. Urban growth can therefore be the result of any city- specific or economy-wide change that augments the strength or scope of agglomeration forces or reduces the importance of congestion forces. One example that has been widely used in the literature is reductions in commuting costs that lead to larger cities in terms of area, population, and in most models also output (Chatterjee and Carlino, 1999). Another example is the adoption of information and communication technologies that may increase the geographical scope of production externalities, therefore increasing the size of cities.

Changes of underlying economic factors cause cities to grow or decline as they adjust to their new equilibrium sizes. Another more subtle factor is changes in the patterns of specialization that are associated with equilibrium city sizes. That is, the coexistence of dirty industry with high-tech industry generates too much congestion, and therefore cities specialize in one or the other industry. Adjustments in city sizes and patterns of specialization in turn may be slow, since urban infrastructure, as well as business structures and housing are durable, and new construction takes time (Glaeser and Gyourko, 2005). However, this type of change lead only to transitional urban growth, as city growth or decline eventually dies out in the absence of other city-specific or economy-wide shocks. Even when any of the economy-wide variables, such as population, grows continuously, the growth rate of a specific city will dwindle because of new city creation (Ioannides, 1994; Rossi-Hansberg and Wright, 2004).

Much attention has also been devoted to the effect that this type of urban growth has on urban structure. Lower commuting costs may eliminate the link between housing location choices and workplace location. This results in more concentration of business areas, increased productivity because of, say, knowledge spillovers, and lower housing costs in the periphery of the city. Urban growth can therefore lead to suburbanization as well as multiple business centres, as in Fujita and Ogawa (1982) or Lucas and Rossi-Hansberg (2002). Those phenomena become increasingly important because of the decline in transport and commuting costs brought about by the automobile along with public infrastructure investments. In other words, urban growth is associated with sprawl (Anas, Arnott and Small, 1998).

Urban and national economic growth

Most economic activity occurs in cities. This fact links national and urban growth. An economy can grow only if

cities, or the number of cities, grow. In fact, Jacobs (1969) and Lucas (1988) underscore knowledge spillovers at the city level as a main engine of growth. The growth literature has also argued that, in order for an economy to exhibit permanent growth, the aggregate technology has to exhibit asymptotically constant returns to scale (Jones, 1999). If not, the growth rate in an economy will either explode or converge to zero. How is this consistent with the presence of scale effects at the city level? Eaton and Eckstein (1997), motivated by empirical evidence on the French and Japanese urban systems, study the possibility of parallel city growth, which is assumed to depend critically on intercity knowledge flows together with the accumulation of partly city-specific human capital across a given number of cities. Rossi-Hansberg and Wright (2004) propose a theory where scale effects and congestion forces at the city level balance out in equilibrium to determine the size of cities. Thus, the economy exhibits constant returns to scale through the number of cities increasing along with the scale of the economy. Hence, economic growth is the result of growth in the size and the number of cities. If balanced growth is the result of the interplay between urban scale effects and congestion costs, these theories have important implications for the size distribution of cities and the urban growth process. These implications turn out to be consistent with the empirical size distribution of cities, that is, Zipf's Law, and with observed systematic deviations from Zipf's Law.

To summarize: urban growth affects the efficiency of production and economic growth, and the way agents interact and live in cities. Understanding its implications and causes has captured the interest of economists in the past and deserves to continue doing so in the future.

YANNIS M. IOANNIDES AND ESTEBAN ROSSI-HANSBERG

See also **city and economic development; endogenous growth theory; location theory; new economic geography; power laws; spatial economics; symmetry breaking; systems of cities; urban agglomeration; urban economics; urban environment and quality of life; urban production externalities; urbanization.**

Bibliography

Anas, A., Arnott, R. and Small, K. 1998. Urban spatial structure. *Journal of Economic Literature* 36, 1426–64.

Bairoch, P. 1988. *Cities and Economic Development*. Chicago: University of Chicago Press.

Chatterjee, S. and Carlino, G. 1999. Aggregate metropolitan employment growth and the deconcentration of metropolitan employment. Working paper. Philadelphia: Federal Reserve Bank of Philadelphia.

De Vries, J. 1984. *European Urbanization: 1500–1800*. Cambridge, MA: Harvard University Press.

Duranton, G. 2004. Urban evolutions: the still, the fast, and the slow. Working paper. London: Department of Geography and Environment, London School of Economics.

Eaton, J. and Eckstein, Z. 1997. Cities and growth: theory and evidence from France and Japan. *Regional Science and Urban Economics* 27, 443–74.

Eeckhout, J. 2004. Gibrat's Law for (all) cities. *American Economic Review* 94, 1429–51.

Fujita, M. and Ogawa, H. 1982. Multiple equilibria and structural transition of nonmonocentric urban configurations. *Regional Science and Urban Economics* 12, 161–96.

Fujita, M., Krugman, P. and Venables, A. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

Fujita, M., Mori, T., Henderson, J. and Kanemoto, Y. 2004. Spatial distribution of economic activities in Japan and China. In *Handbook of Regional and Urban Economics*, vol. 4, ed. J. Henderson and J.-F. Thisse. Amsterdam: North-Holland.

Gabaix, X. 1999. Zipf's Law for cities: an explanation. *Quarterly Journal of Economics* 114, 739–67.

Gabaix, X. and Ioannides, Y. 2004. The evolution of city size distributions. In *Handbook of Regional and Urban Economics*, vol. 4, ed. J. Henderson and J.-F. Thisse. Amsterdam: North-Holland.

Glaeser, E. and Gyourko, J. 2005. Urban decline and durable housing. *Journal of Political Economy* 113, 345–75.

Henderson, J. and Thisse, J.-F., eds. 2004. *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: North-Holland.

Ioannides, Y. 1994. Product differentiation and economic growth in a system of cities. *Regional Science and Urban Economics* 24, 461–84.

Jacobs, J. 1969. *The Economy of Cities*. New York: Random House.

Jones, C. 1999. Growth: with and without scale effects. *American Economic Review P&P* 89(2), 139–44.

Krugman, P. 1991. Increasing returns and economic geography. *Journal of Political Economy* 99, 483–99.

Lucas, R., Jr. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22(1), 3–42.

Lucas, R. and Rossi-Hansberg, E. 2002. On the internal structure of cities. *Econometrica* 70, 1445–76.

Marshall, A. 1920. *Principles of Economics*, 8th edn. London: Macmillan.

Papageorgiou, Y. and Smith, T. 1983. Agglomeration as local instability of spatially uniform steady-states. *Econometrica* 51, 1109–19.

Rossi-Hansberg, E. and Wright, M. 2004. Urban structure and growth. Working paper. Stanford: Department of Economics, Stanford University.

urban housing demand

At its core, the demand for urban housing is just the manifestation of the demand for living in urban areas. On net, residence patterns suggest that most people want to live in or near cities, and that desire is increasing over