

Urban Growth: Is It a Never Ending Story?

*Roberta Capello**

Abstract

In the post-industrial society or knowledge economy, which are the determinants that explain urban growth? The aim of the paper is to give an answer to this question by critically analyzing the contributions that exist in the urban economic literature.

Keywords: Efficient Urban System; Urban Growth; Relational Capital; Functional Innovation

1. The Constantly Increasing Size of Cities

In the real world, the number of people living in cities is growing in all countries and continents. The urbanisation process is a phenomenon which, in the last decade, has been increasingly intense in the developing countries. The share of urban population in the more developed continents, such as Europe and North America, is extremely high, and at the world scale is nearly 50%. This percentage, according to official forecasts (BBC News, 2011), is expected to rise yet further in future decade. As a consequence of increasing population, cities physically expand through processes which have been defined as ‘ville éclatée’, ‘ville éparpillée’, ‘ubiquitous city’ and more recently ‘metropolisation’. The population of large cities is continuing to grow, though sometimes more slowly than previously (Camagni, 1996); this continuous trend puts at the fore front of the theoretical reflections Alonso’s challenging questions ‘how big is big enough?’ and ‘how big is too big?’ (Alonso, 1964).

In fact, the constantly increasing size of cities encountered in the real world is in contrast with the famous ‘optimal city size’ theory, which envisages a size above which an increase in physical dimension decreases the advantages of agglomeration. The declining rate of urban population growth recorded in the last decade in most developing countries appears to be common to all cities, independently of physical size, and represents a general slowing down, rather than a specific crisis in the larger cities. Indeed, during the seventies, there were negative population growth rates in the urban system of the Po Valley in Northern Italy not only in the major cities, but also a number of secondary centres of 75,000 to 150,000 inhabitants (8 out of 19) and even some smaller towns of 20,000 to 75,000 inhabitants (27 out of 113) (Camagni et al. 1985; Camagni et al. 1986). According to the theory, however,

* Full Professor of Regional Economics, Politecnico of Milan (roberta.capello@polimi.it)

medium-sized towns are expected to increase their size, since the advantages associated with the physical dimension are still higher than location costs.

The paper aims at reviewing the theoretical approaches developed within urban economics that made an effort in interpreting urban growth. In particular, starting from the criticisms to the optimal city size theory contained in the literature (sec. 2), the paper organises the theoretical steps forward made in the literature organising them into two major theoretical paradigms (sec. 3 and 4). Moreover, the paper aims at providing evidence of the two theoretical paradigms by recalling two empirical analyses developed by the research group the author belongs to (sec. 5).

2. Optimal City Size: an Old and Unsolved Problem

Since the sixties, urban economists and geographers have put at the forefront of their reflections the problem of urban growth and of optimal urban size, trying to reply to the intriguing questions of Alonso ‘how big is big enough?’ and ‘how big is too big?’ (Alonso 1964), and the issue is still on the agenda of both scientists and policy makers. In particular, the main question was to identify whether increasing returns to urban size exist. The reply was for the first time given by the optimal city size theory, which claims that urban location (average) advantages increase when the city size increases (Alonso 1971), due to externalities that stem from: a) consumptions and investments in public services, b) large markets of outputs; c) large and diversified markets of inputs. These externalities are well known as ‘agglomeration economies’ (Glaeser et al. 1992; Rosenthal and Strange 2001). On the other side, average urban costs decrease while the city size increases: expenses for the implementation of fixed capital infrastructure decrease while the number of people using them increases (Richardson 1972).

A general consensus exists within the literature on urban size on the fact that agglomeration economies exist up to a certain threshold, after which urban benefits start to decrease; after that threshold, in fact, opposite mechanisms start to act and change positive into negative elements, ‘agglomeration economies’ into ‘agglomeration diseconomies’, while average urban costs start to increase, diminishing the net agglomeration advantage. Congestion, high urban rents, environmental costs are all elements that explain the decrease in agglomeration advantages, and the city, as each economic resource intensively exploited, shows decreasing returns.

Following this logic, average urban location costs and benefits have both a U shaped form: the former decrease and then increase, the latter increase and then decrease. The size of the city for which the difference between location total costs and advantages is the size that maximizes the net benefits, and for this reason it is identified as the ‘optimal city size’.

Although demonstrated by a large number of empirical estimations, many criticisms have been made of the neoclassical approach to optimal city size theory. These include the observations that (Capello 2004):

- cities are different from one another. They are characterised by different functions and perform different specialisations (Henderson 1985, 1996). The use of the same urban production function for all cities in econometric analyses estimating optimal city size is extremely restrictive. In the words of Richardson: ‘we may expect the efficient range of city sizes to vary, possibly

dramatically, according to the functions and the structure of the cities in question' (1972 pp. 30).

- If cities are different from one another, the optimal city size may be different, depending on the specific characteristics. Richardson elegantly compares the 'optimal city size' theory with the theory of the behaviour of firms. We would never expect the optimal position for each and every firm to occur at the same level of output, so why should we expect the optimal point in each city to be located at the same population level?.
- Cities exist in an interurban environment. The optimal city size theory, on the contrary, does not consider the spatial context in which cities operate.
- Cities generate a large variety of externalities as a result of the qualitative characteristics of the urban production environment. Already in 1961, Chinitz expressed some doubts about the fact that urban factor productivity depends mainly on the physical size of cities. He emphasised, on the contrary, the importance of a diversified and competitive urban production system as a source of urban productivity. Such a system is able to provide a far larger variety of externalities for small firms than an oligopolistic and specialised urban structure. Chinitz supported his thesis with an empirical analysis of New York, a large and diversified urban area, and Pittsburgh, a highly specialised city (Carlino 1980).

The necessity to overcome the limits of the theory on the optimal city size has increased in recent years, when the urbanisation process has drastically been affected by a rapid growth.

3. Innovating the Urban Structure: Upgrading Urban Functions

The first theoretical step has envisaged the possibility to deal with the fact that cities are not all the same, but produce different goods according to their size. In a post-industrial society, large cities have the undisputable role of knowledge creator, distributing information to medium and small urban centres.

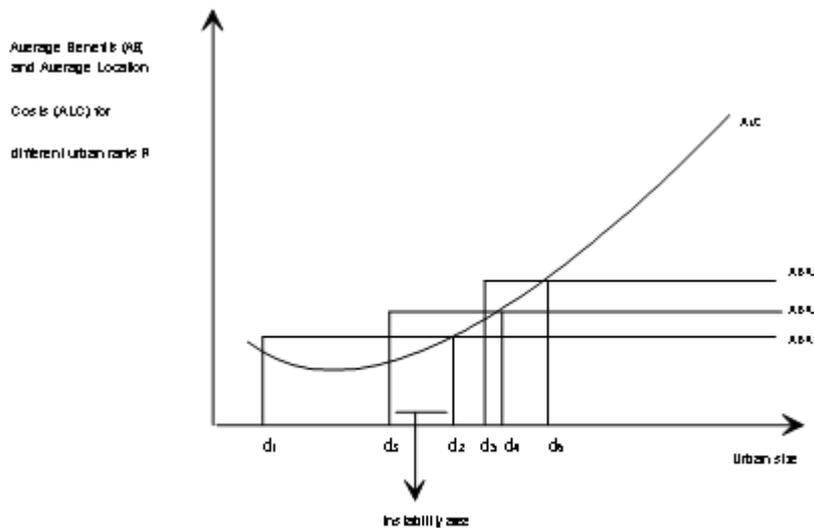
An interesting conceptualization of the role of urban functions (intended as the activities that on average are developed in the city) in explaining urban growth is presented in the SOUDY (Supply Oriented Dynamic Approach) model (Camagni et al. 1986) not yet overcome by other more recent models, which argues that (Figure 1):

- higher order functions are characterised by higher thresholds for the level of appearance in the city (in terms of urban population) (d_1, d_2, d_3, \dots);
- the average (aggregate) location benefit curve (measuring the benefits that a city is able to provide to people and activities located in a city when the city size increases) increases for higher order functions, due to a) growing entry barriers, b) decreasing elasticity of demand which allows extra profits to be gained in all market conditions, and c) increasing possibility of obtaining monopolistic revenues due to the use of scarce, qualified factors;
- the average location cost curve (measuring the costs that a city is able to provide to people and activities located in a city when the city size increases) has the traditional form suggested by Alonso.

When the average location benefit function is compared with Alonso's traditional average location cost curve, the following results are achieved. For each economic function and each associated urban rank, it is possible to define a minimum and a maximum efficient city size, which would increase with the level of the urban function and rank (Figure 1).

An interesting message from this model is that cities can continue to grow from the physical point of view without incurring in diseconomies of congestion, if local public and private sectors innovate by upgrading their functions.

Figure 1: *Efficient Urban Size for Different Urban Functions*



Source: Camagni et al. 1986

The interest of this model is that it overcomes some of the limits of the 'optimal city size' theory, by suggesting:

- the need to replace 'optimal city size' by an interval within which the city size is efficient, i.e. where average benefits exceed average location costs;
- that the interval of efficient city size corresponds to greater urban size, the higher the economic functions developed in the city;
- that, as a result of the previous statement, the economic functions characterising the city are an important determinant of the efficient city size.

With a theoretical approach like the Soudy model (Camagni et al. 1986) the role of functions has been integrated into a theoretically sound model. However, this theoretical step forward is not enough: it is in fact impossible to explain why a city like Zurich, with only 300,000 inhabitants, is specialised in international finance in the same way as the city of New York or Tokyo. In the real world, the urban size is not always characteristic of function. It is on this limit that the second theoretical step forward works on.

4. Relational Capital: the City Network Paradigm

The city network paradigm provides a successfully theoretical framework to overcome the limiting interpretative power of the traditional central-place model. In fact, real city-systems in advanced countries have deeply departed from the abstract Christallerian pattern of a nested hierarchy of centres and markets and reveal (Camagni 1993):

- processes of city specialisation, especially in industry but also in services, which contrast with the prediction of Christaller's model about the hierarchical despecialisation of each centre (Cappellin and Grillenzoni 1983);
- incomplete presence of the whole range of functions in each city;
- presence of higher order functions in centres of lower order;
- horizontal linkages between similar functions and cities, e.g. the financial network among top cities in the worldwide hierarchy;
- synergy linkages among similar centres performing headquarter and advanced production-services functions, as well as among sub-regional industrial districts.

These observations support the idea that new and increasing relationships among centres follow a network logic, where similar specialisation patterns are the main reasons to establish economic relationships.

Born in the field of industrial economics, the concept of network behaviour has efficiently been transferred to urban economics theory, and is extremely useful to interpret the growth of cities in a global world. Cities of intermediate size are being increasingly looked upon as the places that could well host the growth of the years to come - provided that the resulting development does not destroy their present competitive advantage. Non excessive city sizes in fact facilitate environmental equilibrium, efficiency of the mobility system, and the possibility of conserving a sense of belonging as far as the population is concerned.

The limit that the medium sized cities come up against, and which often makes them succumb vis-à-vis the great metropolis, is that of critical mass and centrality. On the one hand, many of the advanced or rare high-level functions call for a considerable minimum market size with respect to both demand for the offered services and, above all, to the supply of the necessary human capital; on the other hand, however, these selfsame functions need great centrality and ready access to the world communication networks. At least, a partial but nevertheless concrete answer can be given to these two needs, an answer that is both conceptual and operative. The need for offering an adequate critical mass can be satisfied thanks to good interconnections between the city and its hinterland, and above all, between the city and other nearby cities of similar size: the so-called 'city-networks' (Camagni 1993).

The city-network model not only makes it possible to upgrade the functions established in the individual centres without necessarily increasing their individual size, but also brings in its wake some advantages of considerable potential as far as mobility management is concerned (Brondoni 2009; Day et al. 2004). In fact, whenever the activities that constitute the productive 'vocation' are integrated into

such a pattern, one can easily imagine that the greater part of the transport interactions will occur within the city itself (or within its productive district).

5. Empirical Evidence from Italian and European Cities

The previous theoretical considerations were empirically tested in two recent empirical analyses run by the research group the author is part of. The first one, conducted by Capello and Camagni (2000), was an econometric estimate of two average urban location costs and benefits functions made dependent on size, high level functions and the degree of a-spatial linkages, on a sample of 58 Italian cities of different size. The results are summarised in Figure 2. The first results regard the variable traditionally interpreted in the literature as the most important source of urban average benefits and costs: urban size.

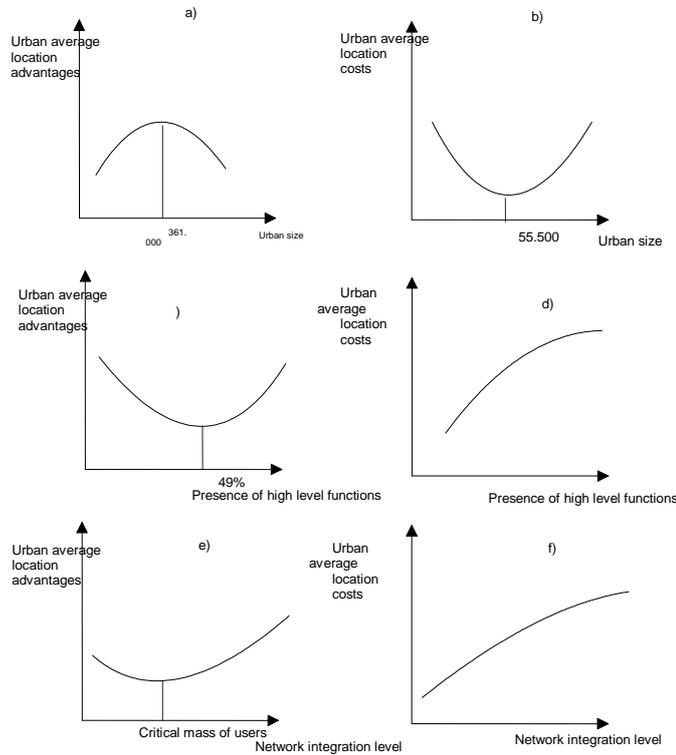
Figure 2 (graphs a and b) shows the estimated location benefits and costs functions for different levels of urban size. In economic terms, the calculated parameters reflect the elasticity of the urban benefits and costs with respect to size, i.e. how the city benefits and urban costs change with an increase in size of 1%, for different urban sizes. The results obtained are in line with the abstract interpretation of the optimal urban size theory. In fact, the curves are 'well-behaved', showing a benefits which increase with urban size up to a certain point (approximately 361,000 inhabitants) and then decrease.

Urban location costs show a decreasing trend up to a certain urban size (approximately 55,500 inhabitants) and an increasing trend afterwards, once again in line with the traditional expectations. Medium-sized cities appear to have a greater endogenous capacity to keep social, economic and environmental costs under control.

The picture changes when the analysis is made on the basis of the different types of economic functions which can characterise a city. The results are quite interesting. The results on the urban benefits are in line with the conclusions suggested by the SOUDY model. The estimated curve confirms the theoretical hypotheses of the SOUDY model (Figure 2c and d): higher order functions guarantee a greater average benefit, due to the positive returns generated, up to a certain share of tertiary activities (49% of its total activities).

The urban costs increase at a decreasing rate when there is a strong presence of high level functions. This means that the increase in value added functions tends to entail congestion and location costs, but that this negative aspect does not occur in a disruptive and uncontrollable way, as in the case of increasing urban size (Figure 2d). The urban average costs increase at a decreasing rate, which indicates that higher order functions produce economic development and also local congestion costs, but with a decreasing marginal productivity, and thus in a more controlled way.

Figure 2: Urban Average Location Advantages and Cost for Different Urban Sizes, Higher Urban Functions and Degree of Networking



Source: Capello and Camagni, 2000

The results of the size elasticity of the urban benefits and costs for different levels of relational capital produce an interesting picture (Figure 2e and f). As far as the benefits are concerned, they decrease up to a certain level of network integration, when they start to increase. These results are stimulating, since they suggest that after a certain threshold level, the city is able to exploit the advantages associated with the interconnected economy and network externality advantages are in full operation. Through the network, the city is able to exploit more dispersed information collection, the acquisition of more know-how and more qualified input factors, as well as a wider market for final goods.

For the urban costs, the picture which emerges is similar to that for different levels of high order functions (Figure 2f). When the level of network integration increases, urban costs increase, too. This is what would be expected: higher levels of network integration stimulate more economic activities and generate higher urban benefits, but with the negative counterpart of an increasing costs. What is rather interesting is that urban costs exhibit decreasing growth rates. Again, this result is different from the exploding situation which occurs when the city size is taken into consideration.

A more recent empirical analysis has been run by Camagni et al. (2012), looking at the same empirical issue from a different angle. In particular, the equilibrium city size is made dependent both on size, and on traditional aspects like urban rent, un-efficient urban structure (sprawl) (Glaeser and Kahn 2003) and social malaise,

the former, and amenities and sectoral diversity the latter, in a traditional Cobb-Douglas form.

What is of particular interest is that *ceteris paribus* cities with higher order functions or higher networking achieve a higher equilibrium size. As Figures 3 and 4 show, *ceteris paribus*, the equilibrium city size increases when higher order functions or higher relational capital increase, suggesting that the SOUDY model and the city network paradigm interpret the reality in the right way.

Figure 3: Equilibrium City Size for Different Degrees of Urban Networking

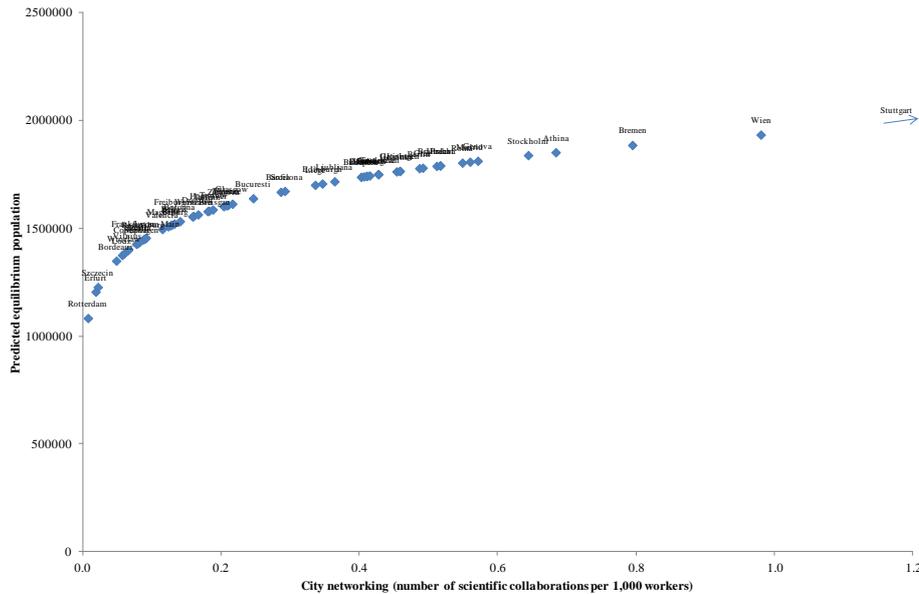
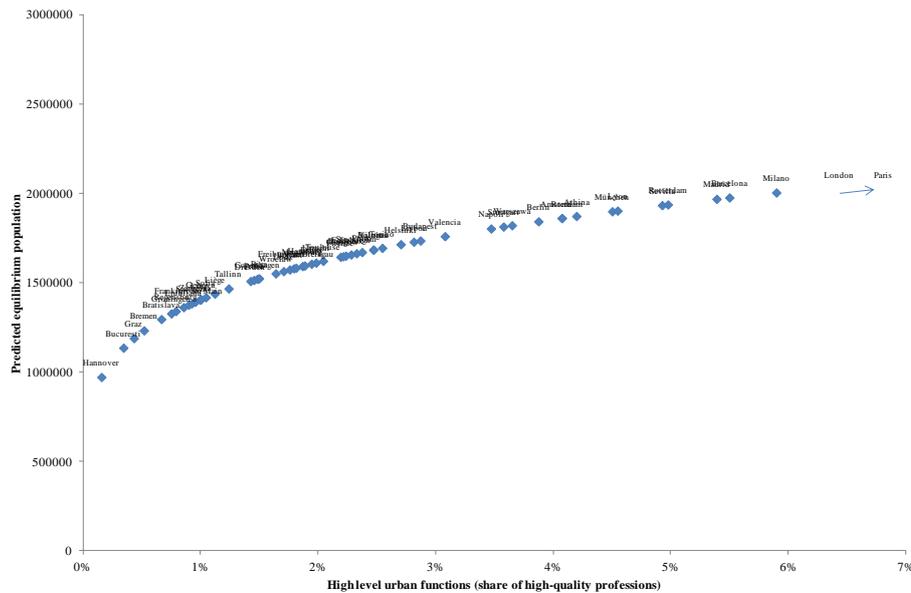


Figure 4: Equilibrium City Size for Different Levels of High Order Functions



Source: Camagni et al., 2011

6. Conclusions

The influence of urban size on urban costs and advantages, as explained by the ‘optimal city size’, theory exists and is important, but cannot be efficiently assessed without overcoming some of the limitations imposed by the theory. What the recent theoretical paradigms suggest is that it is not a problem of optimal city size, but of efficient size, which largely depends on what the city produces, how it produces and the way in which it co-operates within the urban system. Urban size inevitably influences location costs and benefits, however the same also holds for the function the city has and the integration it is able to develop with the urban system, which is not necessarily organised around a ‘territorial logic’, emphasising a gravity-type control over market areas, but around long distance competition and co-operation regardless of the distance barrier.

Moreover, the paper presents two recent empirical works that validate empirically the assumptions of the most recent theoretical paradigms. In particular, the type of economic function and the degree to which the city is integrated in an urban system appear to be strategic elements for the definition of location benefits and costs, analysed in relation to all aspects constituting the city. i.e the social, environmental and economic aspects.

Urban growth appears to be a never ending story; the constant positive trend of urban size of world cities has important normative consequences. Cities need a constant monitoring of urban dynamics, of its determinants, also through serious research analyses. Moreover, our cities need efficient urban policies with the aim to upgrade the economic functions within the city, as well as the development of linkages outside the city, such as alliances, co-operation agreements, advanced international transport and telecommunications infrastructure. All of these elements and strategic policies are undoubtedly important for guaranteeing the survival of a modern city.

Bibliography

- Alonso W., *Location and Land Use: towards a General Theory of Land Rents*, Harvard University Press, Cambridge, 1964.
- Alonso W., The Economics of Urban Size, *Papers and Proceedings of the Regional Science Association*, 1971, pp. 67-83.
- BBC News, <http://news.bbc.co.uk/2/shared/spl/hi/world/06/urbanisation/html/urbanisation.stm>, 2011.
- Brondoni S. M., Market-Driven Management, Competitive Space and Global Network, *Symphonya. Emerging Issues in Management (symphonya.unimib.it)*, n. 1, 2008, pp. 14-27. <http://dx.doi.org/10.4468/2008.1.02brondoni>
- Camagni R., From City Hierarchy to City Network: Reflections about an Emerging Paradigm, in Laschmanan T. and Nijkamp P. (eds.), *Structure and Change in the Space Economy*, Springer Verlag, Berlin, 1993, pp. 66-90.
- Camagni R., Diappi L., Leonardi G., Urban Growth and Decline in a Hierarchical System: a Supply-oriented Dynamic Approach, *Regional Science and Urban Economics*, vol. 16, 1986, pp. 145-160. <http://dx.doi.org/10.1016%2F0166-0462%2886%2990017-7>

- Capello R., Beyond Optimal City Size: Theory and Evidence Reconsidered, in Capello R. and Nijkamp P. (eds.), *Urban Dynamics and Growth: Advances in Urban Economics*, Elsevier, Amsterdam, 2004, pp. 57-86.
[http://dx.doi.org/10.1108/S0573-8555\(2005\)0000266004](http://dx.doi.org/10.1108/S0573-8555(2005)0000266004)
- Capello R., Camagni R., Beyond Optimal City Size: An Evaluation of Alternative Urban Growth Patterns, *Urban Studies*, vol. 37, n. 9, 2000, pp. 1479-1497.
<http://dx.doi.org/10.1080/00420980020080221>
- Cappellin R., Grillenzoni R., Diffusion and Specialisation in the Location of Service Activities in Italy, *Sistemi Urbani*, n. 1, 1983, pp. 249-282.
- Carlino G., Contrast in Agglomeration: New York and Pittsburgh Reconsidered, *Urban Studies*, vol. 17, 1980, pp. 343-351.
<http://dx.doi.org/10.1080/00420988020080651>
- Chinitz B., Contrast in Agglomeration: New York and Pittsburgh, *American Economic Review*, vol. 51, 1961, pp. 279-289.
- Day G. S., Reibstein D. J., Managing Brands in Global Markets, in Gatignon H., Kimberly J.R., Gunther R.E. (eds.), *Alliances on Globalizing*, University Press, Cambridge, 2004.
<http://dx.doi.org/10.1017/CBO9780511522093.009>
- Glaeser E., Kahn M. E., *Sprawl and Urban Growth*, Harvard Institute of Economic Research, Discussion Paper n. 2004, Harvard University, 2003.
- Glaeser E., Kallal H., Scheinman J., Shleifer A., Growth in Cities, *Journal of Political Economy*, vol. 100, n. 6, 1962, pp. 1126-52.
- Henderson J., *Economic Theory and the Cities*, Academic Press, Orlando, 1985.
- Henderson J., Ways to Think about Urban Concentration: Neoclassical Urban Systems vs. the New Economic Geography, *International Regional Science Review*, vol. 19, n. 1 & 2, 1996, pp. 31-36.
- Richardson H. W., Optimality in City Size, Systems of Cities and Urban Policy: a Sceptics View, *Urban Studies*, 1972, pp. 29-47.
<http://dx.doi.org/10.1080/00420987220080021>
- Rosenthal S., Strange W., The Determinants of Agglomeration, *Journal of Urban Economics*, vol. 50, 2001, pp. 191- 229.
<http://dx.doi.org/10.1006/juec.2001.2230>