

**Distance, Land, and Proximity,  
Economic analysis and the evolution of cities**

**Gilles Duranton**

**January 1999**

*Research Papers in Environmental and Spatial Analysis* No. 53 (Department of Geography & Environment, London School of Economics).

# **Distance, Land and Proximity, Economic analysis and the evolution of cities**

*Gilles Duranton*

## **Abstract**

This paper attempts to a synthesis on the long-run evolution of cities using an economic perspective. It defends the idea that urban growth for *pre-industrial cities* has been limited by the tyranny of distance. Then, technological progress, by fostering mobility, reinforced economies of agglomeration and thus allowed for larger cities. This led to the development of *industrial cities*. Nowadays however, technological progress in communications and telecommunications seems to challenge the rationale for agglomeration in cities as more and more economic interactions can be realised at arm's length. Increasing mobility may have turned into a threat for cities. Hence predictions about the demise of cities. Nonetheless, it is argued that the 'tyranny of proximity' may provide a fairly strong glue keeping *post-industrial cities* together.

## Introduction

The city of the future is the future of our world. *The European* 6-12 June 1996.

The objective of this paper is to explore the long-run evolution of cities in the light of economic analysis. This may seem ambitious but the idea is not to produce a definitive account of urban history but to attempt to unravel changes and continuities in the evolution of cities from an economic perspective. This work is easier than it seems for two reasons. First, historical work on cities is very often of the highest quality. Second, the economic analysis of cities over the last forty years has made tremendous progress. By insisting on facts as much as on the mechanisms, my aim is to present the issue of urban evolution through the questions and the methods used by economists interested in cities. Consequently this synthesis is about work done by economists and it is purposefully that the work of other social scientists is not discussed here. For instance, on the changing nature of cities, authors like Castells (1989) or Sassen (1991) are conspicuously ignored. My objective is to try to promote exchange between people in different quarters and therefore I present mainly reflections derived from economic analysis and not an integrated multi-disciplinary approach that remains to be undertaken.

This article defends the idea that technological progress over time, through a decrease in the costs of transport and telecommunications, has led to a modification in the nature of cities and relaxed the constraints preventing their growth. Historically, we can distinguish first the pre-industrial cities whose growth was mainly limited by the costs of agricultural supplies. In their case, we will speak of a tyranny of distance. The main institution of the pre-industrial city was the guild, which regulated both the urban population and its fragmented labour market. Then improvements in transports strongly reduced the power of the tyranny of distance and led to very strong urban growth. They also made inter-city trade much cheaper. This increase in the size of market areas of firms enabled them to use more capital intensive methods of production, i.e., industrial technologies. This gave birth to the much larger industrial city whose growth is primarily limited by the cost of commuting. We speak then of a tyranny of land. In the industrial cities, the land market replaced the guild as the major institution regulating the number of inhabitants in the city.

Today, improvements in transport and telecommunication technologies enable ever-larger cities as they make the physical constraints preventing urban growth ever less binding. However, those improvements also threaten the very existence of cities since they increase the mobility of goods and services, which before could only be found in cities. Therefore, cities

could lose their power of attraction and disappear. Nonetheless, the necessity of direct personal interactions (the tyranny of proximity) may keep cities together and allow for large sustainable post-industrial cities. Those arguments are developed in greater detail in what follows. Each of the three types of cities defined above in turn is discussed in the next three sections.

## **Pre-industrial cities: the tyranny of Distance**

### *The city as a magnet*

Popular clichés have for very long identified cities with magnets. We are going to endorse this vision. This 'magnet theory of cities' is more interesting than it at first seems, if we push the analogy with the laws of physics further and identify people with 'magnetic dust'. Let us first assume only this: each individual is a magnetic particle attracting and attracted by other particles. We will later on try to understand what this magnetism is about, but for now let us explore only its implications. Imagine an initial situation where dust is spread all over the existing land. It is obvious that this situation is not stable. A small particle can join another one, forming an embryo of magnet. The superior magnetic power of this pair will attract other particles from its immediate neighbourhood. This will lead to the formation of a rather large magnet. This magnet has many features in common with real cities: (i) it attracts people; (ii) its location is rather stable because of the inertia of its weight; and finally (iii) like real cities, it has some internal order.

As a theory of city formation and city growth, our magnet theory, as it is, is nonetheless flawed in a major way. If we pursue its implications further, we can see that all the existing dust should ultimately be part of ever-larger magnets and soon all the magnets should collapse to form only one very large magnet. It is only then that the process we set in motion earlier on will stop. Clearly, this prediction is supported neither by the present nor by any past reality.

What's wrong then? It is either that our magnetic analogy is completely irrelevant or that we miss something. The fact that human beings have sought to live close to each other seems rather robust and we shall not question this hypothesis before the last section of the paper. The simple magnet theory we have just exposed may just be too 'strong'. It predicts the emergence of only one big city, whereas the real world for the last 5000 years has comprised many cities, some big and some small, and part of the population still live in the countryside. It seems reasonable to keep our hypothesis of mutual attraction but to introduce some 'frictions'.

This is what we shall be doing in this section and the next. The alternative would be to replace our hypothesis of unconditional mutual attraction by one of limited attraction. Apart from its ad-hoc aspect, it seems that the implications of this alternative are not as interesting as those of the two frictions we will successively introduce.

### *The tyranny of distance*

The first friction we should appeal to is the most basic one. Unlike magnetic particles, people need to be fed and cultivating food requires space. This element implies that a fraction of the population will thus remain attached to the land and will not be able to congregate in the city. Still, the argument is not complete since our model keeps predicting the existence of only one city. We need to consider one more key ingredient. Food is costly to transport and the larger the city the higher the marginal cost of transport. This element was historically the first major brake on urban growth. We shall refer to this cost of shipping goods as the *tyranny of distance*. Before elaborating further on this, let us turn first to some empirical evidence. They are given by urban historians (e.g., Pirenne, 1925, Duby 1981, Braudel 1984, Bairoch 1988 among many) who all recognise the tyranny of distance and the existence of agricultural surpluses as the major determinants of the urban landscape until the early 19<sup>th</sup> century. The argument is most clearly articulated by Bairoch (1988) when he explains the absence of cities in Palaeolithic Britain. He argues that the human density at the time was around 0.1 inhabitant per km<sup>2</sup> and that the (very low) agricultural surplus was at most 10%. Consequently, to feed a city of 10,000 inhabitant would have required a hinterland of around 200,000 km<sup>2</sup>, that is more than the area of Britain.

This link between the formation of cities and the existence of an agricultural surplus is robust. Disagreements start when one tries to investigate more precisely the nature of the relationship between cities and agricultural surpluses. Even though it is clear that an agricultural surplus is necessary for cities to exist, it is not obvious that it necessarily leads to the formation of cities. Markets do not require large permanent settlements. The point was made first by Max Weber (1958). Nonetheless, according to Bairoch (1988), there are very few cases (if any) of agricultural surpluses not leading to the creation of cities. Of course, this does not imply that the chief cause of city formation is economic in a narrowly defined way: religious and defence motives have been instrumental. (And this is why our magnet metaphor is quite useful as it allows us to leave temporarily in the shade the exact nature of this attraction.) The direction of the causality is the source of another debate. Considering causality

running from agricultural surpluses to cities seems rather natural. Jane Jacobs (1969) however argues that cities might have been formed in the first place, with rising agricultural productivity to follow. This hypothesis regarding a cumulative and self-reinforcing causality from agricultural surpluses to cities and back seems perfectly valid. Obviously, the activity carried on in the city influences the agricultural sector.<sup>1</sup> Nonetheless it seems that for early cities, we need to think primarily of them resulting from rising agricultural productivity. Otherwise, if we think that urbanisation preceded and caused rising agricultural productivity, this implies that the urban spillovers on the agricultural sector must have been incredibly large and fast for the urban population not to die of starvation in the interval between the formation of the city and the agricultural improvements.

Let us now examine the details of our amended magnet model. The food requirement of the population determines the agricultural employment, whereas the rest of the population is free to congregate in the cities, producing manufactured goods.<sup>2</sup> This will lead to a two-way trade between cities and the countryside as the urban population exports manufacturing goods and imports agricultural products. And given the weight of the tyranny of distance, links between cities and their immediate agricultural hinterland were much more important than links with other cities (Bairoch, 1988). The tyranny of distance is thus a particularly useful tool to think about pre-industrial cities. Nonetheless, as noted earlier (Duranton, 1998), the tyranny of distance does not offer any precise prediction about the size of cities. It just sets bounds on city size. This can be explained in the following way. The maximum feasible size is such that all the agricultural surplus from the fringes of the supply region is completely consumed in the transport process. (Needless to say that this size is not very desirable since the marginal inhabitant increases transports costs much more than production in the city.) Further reasoning can also lead us to set a lower bound for the city size. Below this threshold, the gain for the existing urban population of one extra urban dweller (e.g., more efficient defence of the city, greater specialisation of labour, etc.) more than offsets the extra shipping costs of food. In other words, below this lower bound, all the inhabitants of the city are willing to welcome any newcomer. Thus, as a major conclusion at this point, it must be noted that the tyranny of distance does not fully determine everything and only sets bounds. And since the lower bound is also an optimal size, cities will tend to be too large in absence of further regulation.

### *The necessary institutions*

This naturally implies that to improve our understanding and sharpen our predictions,

institutions and political forces need to be investigated seriously. In European cities between circa 1200 and circa 1750, guilds played a prevalent role.<sup>3</sup> According to several authors (Weber, 1958, Le Roy La Durie in Duby, 1981, and especially Friedrichs, 1995) urban guilds had a regulating role by limiting the entry of workers in their area of supervision. By defending the interests of their members, guilds probably also kept city size at a fairly efficient level (i.e., they achieved a good balance between the benefits of internal labour specialisation and the costs of external supplies). In practice, the guilds let in a regular flow of newcomers to compensate for the excess mortality in cities. These limits on urban size were clearly a positive effect of guilds. However, the same defence of the interests of their members also led guilds to slow down innovation and to delay the adoption of new labour-saving productive processes. Thus urban guilds in pre-industrial cities must be assessed in light of static gains (efficient city size) and dynamic losses (slower technological change).

This is corroborated by the very slow evolution of the urban landscape before the Industrial Revolution. Urbanisation rates in Europe between 1000 and 1800 rose from 9.7% to 12.2%. The size of the largest cities nonetheless rose more significantly. In 1300, there were only 7 cities in Europe with a population between 50,000 and 100,000 and 5 slightly above 100,000 whereas in 1700, 21 cities had a population between 50,000 and 100,000 and 11 above 100,000 including 2 above 500,000.<sup>4</sup>

This analysis of urban guilds goes a long way towards a better understanding of pre-industrial cities but other political features also mattered. For instance, the existence of a powerful monarchy with a sizeable aristocratic class requiring a lot of personal services is probably the most important element to consider if we want to understand the existence of large and parasitic cities in post-Renaissance Europe (see Ades and Glaeser, 1995, or Bairoch, 1988). This type of analysis can also be extended to consider the effects of taxation over the countryside and those of the structure of land ownership on the urban landscape (see, Duranton, 1998).

An interesting prediction of our small model is that, with improvements in agricultural and transport technologies, cities should grow in parallel. This is strongly supported by empirical evidence (Robson 1973, Eaton and Eckstein, 1997, among others). It must also be said that the tyranny of distance also has implications for the location of cities. With the tyranny of distance, it is no real surprise that more prosperous cities tended to be found in fertile areas and in locations with specific advantages in shipping costs such as a confluent of rivers, a port, etc.

At this stage, the main characteristics of the pre-industrial cities can be underlined.

Cities were producing manufactured goods and exchanging them against food with their surroundings so that the main distributive opposition in pre-industrial economies was between landowners and urban craftsmen. The main driving force was the tyranny of distance, which was mediated by powerful local institutions (guilds) regulating the urban population by means of direct control by the means of quantities (i.e., guilds restricted directly the entry in the city instead of imposing an entry fee for residents).

This last point is fairly interesting. Why should an institution regulate by the quantities, whereas it is a well-established tenet of modern economics that price regulations are in most cases better than quantity regulations? It seems that the most plausible answer lies in the production structure of pre-industrial cities. Production was very decentralised at the personal level so that self-employment was a dominant form of employment. For these individuals, labour specialisation was extensive and pre-industrial labour specialisation often involved important occupation-specific investments. Consequently, a new entrant in a city did not have a marginal effect on the local economy (as would be the case with a homogenous workforce), but an important discrete effect on a small fraction of the urban workforce. In short, the new entrant was not one more worker among many but for instance a new butcher among 4 or 5. Hence, given these important occupation-specific fixed costs, each occupation was keen on regulating the entry within its own area of expertise. The aggregation of this at the city level implies a regulation by the quantities. This argument is consistent with the bottom-up structure of the urban guilds (Friedrichs, 1995).

### *The pervasiveness of the tyranny of distance*

Was the tyranny of distance putting constraints only on interactions between cities and their hinterland? Clearly not. The tyranny of distance was pervasive in the pre-industrial world. For instance, it is also a useful tool for a better understanding of the patterns of agricultural production around the city. This was noted long ago by von Thünen (1826) who can be held as a founding father of modern economics. Indeed, his 'Isolated State' is the first real analytical piece known in economic analysis. In his 'model', he assumed only one city located in the middle of a featureless plain. Of course he was aware that cities are not isolated islands but that was not important for his purpose. His assumption of a flat plain has also been much criticised. However, this is the great strength of his approach. He was able to show that heterogeneity could rise out of homogeneous land because of agricultural transport costs. This is a much greater achievement than being able to generate heterogeneity out of heterogeneity (especially



when the pattern is pervasive). His empirical predictions were that more labour intensive crops should be located closer to the centre, that for a given crop, less labour intensive methods should be used as the distance from the city increases, and that more perishable goods should be cultivated in the outskirts of cities. All those predictions are strongly supported by the facts (see Dicken and Lloyd, 1990, chapter 1).

The tyranny of distance was not only about agricultural goods. Even though agriculture was by far the most important victim of the tyranny of distance until the 18-19<sup>th</sup> century, raw materials and manufactured goods were also subjected to the tyranny of distance. Let us examine these two elements in turn. With industrialisation, raw materials like coal and iron became much more important than ever before. This led to a new geography of productive activities and the transformation of formerly small towns into major industrial cities. As mentioned above, the institutional rigidities like the reluctance of guilds to accept new technologies clearly played a role in preventing the location of these new activities in established cities. Nonetheless, it can be argued that this development of new cities was primarily caused by the tyranny of distance applied to raw materials instead of agricultural products.

It must be noted that the second milestone of location theory (if Thünen's monocentric model is assumed to be the first), the classical location theory (Weber, 1909), tackles this issue in a rather simple and elegant way. The location of plants according to classical location theory can be shown to follow a simple program of cost minimisation. Let us take as given the location of the sources of raw materials as well as that of the major markets, assuming the availability of labour everywhere at the same price and a technology using factors in constant proportions. The optimal location is then such that it minimises total transport costs.

Modern economic theorising nonetheless showed some discontent with this type of approach (Krugman, 1991). The chief reason is that the most interesting aspect of the story, that is the concentration of economic activity in some location, was taken as given (the marketplace is an exogenous variable in both classical and neo-classical location theory). In his very influential book, Krugman (1991) considers an economy with two regions and two sectors. For simplicity he assumed the existence of transport costs that affected only differentiated manufacturing goods and not agricultural goods. Agricultural workers are assumed to be tied to their land, whereas manufacturing workers are mobile between regions. Finally, consumers have a preference for variety (which is equivalent to consumers with heterogeneous tastes). Krugman then showed that lower transport costs lead to more agglomeration thus generalising the tyranny of distance to manufactured goods. Let us examine

this model in more details.

The main point made in Krugman's (1991) core-periphery model is that the agglomeration of economic activity is endogenous and depends in a non-trivial and non-linear way on the transport costs of manufactured goods, thus making the tyranny of distance more subtle in its effects than envisioned until now. With high transport costs (or trading costs more generally) industrial activity is dispersed. To understand this, imagine both high transport costs and a high degree of agglomeration of industrial activities in one region. This is just impossible in equilibrium. Some firms would find it worth leaving the agglomerated region in order to be able to serve profitably the agricultural workforce remaining in the second region. This would enlarge the market in the depressed region, thus reinforcing the movement of firms towards it. This is the famous circular and cumulative causation. Eventually, economic activity should be evenly spread across regions.

When transport costs are lower, multiple equilibria are possible. To understand this, imagine a situation with agglomeration of industry in one region again. A single industrial firm may not have any incentive to move to the depressed region, as the demand from the agricultural workforce in the periphery can now be served from the industrial region since transport costs are not too high. However, if many firms decide to move at the same time, the region with little economic activity will suddenly constitute a much larger market for industrial goods following the arrival of this significant amount of firms. This in turn makes the coordinated movement of those firms self-fulfilling. Thus, for intermediate values of transport costs, both agglomeration and dispersion can be equilibrium outcomes.

For even lower level of transport costs, the equilibrium with dispersion of economic activity is no longer sustainable. A core-periphery structure is the only outcome. The argument is the following. A small advantage in one region makes it more attractive since this region can offer a (slightly) larger variety of goods at a lower price. This leads more firms to move from the other region to this one. In short, this triggers a movement of circular causation, where the advantage of the leading region is reinforced until all industrial activities are agglomerated there.

To summarise, in Krugman's model there is a tension between a market size effect (pulling industrial workers together) and the cost of serving a fixed demand tied to the land (dispersing industrial workers). Two major conclusions emerge. As before, when the tyranny of distance weakens, agglomeration is more likely.<sup>5</sup> However, the transition may not be smooth and sudden changes (in the sense of catastrophic agglomeration or dispersion) occur. This type of idea is not completely new and in this respect Krugman follows authors like Myrdal (1957)

and Hirschman (1958). His value-added lies in the precise description of how economic agglomeration occurs and under which conditions. What should be underlined is that by making relatively simple assumptions (mainly a manufacturing sector with differentiated goods and transport costs), he is able to generate rich and subtle results. Even more importantly, his contribution triggered a large wave of research on this topic. These subsequent contributions contributed to polishing the raw edges of the model just described.

### *The tyranny of distance today*

What we have seen until now suggests that the decline of transport costs should lead to larger cities and to the death of the tyranny of distance. Obviously, the tyranny of distance is not as important as it used to be. Certainly physical transport costs have been declining very significantly, making the tyranny of distance a less decisive factor in determining the nature and size of cities. For instance, according to Jones (1998), freight charges decreased threefold between 1920 and 1990 and air revenue per passenger-mile was divided by six between 1930 and 1990. However the tyranny of distance may still be more prevalent than many tend to think. Even though they do not bear any direct relationship to cities, the following facts remind us that the tyranny of distance is not over

- (i) Trade flows still show a very strong distance bias. In the past gravity-like equations were famous for being good predictors of trade flows. We could expect those gravity equations to become less relevant in a world of much lower transport costs. Not at all: recent empirical work (see Evenett and Keller, 1997, for instance) shows that gravity equations still perform extremely well.
- (ii) In relation to this, it can be remarked that the tyranny of distance is not only about direct monetary costs. Time is crucial as well. As suppliers of US firms, Mexican maquiladoras are competitive despite the much higher costs of unskilled labour in Mexico compared with Southeast Asia even after controlling for productivity differences. The main reason for this does not lie in the difference in the monetary cost of shipping but because it takes at most three days for Mexican maquiladoras to deliver components anywhere in the US against three weeks for Southeast Asian suppliers.

## **Industrial cities: the tyranny of Land**

### *The industrial city*

Before discussing the tyranny of land and the internal organisation of the industrial city we must explain its emergence. Let us first remark that the Industrial Revolution was a time of great changes affecting all aspects of life nearly everywhere in Europe. The urban system was not left untouched and the industrial city replaced its pre-industrial counterpart. As far as cities are concerned, those changes took place gradually over a century, making of course the term *Industrial Revolution*, even more inaccurate than usual. The starting date may be around 1770 in the UK and after 1800 in continental Europe. The first crucial change was the rise of new cities, many enjoying locational advantages with respect to coal and iron. This has been already discussed. The other two important changes are closely linked with each other. Urban size and urbanisation rates increased dramatically. Besides, as noted by prominent historians like Braudel (1984), there was a shift from merchant activities to industrial production in cities. Finally the urban specialisation increased. These three points deserve further elaboration.

The usual identification of pre-industrial cities with purely merchant places or market places is partly misleading even though a lot of trade did take place in pre-industrial cities. However, pre-industrial cities also produced most of the manufactured goods at the time. Even Braudel (1984), a major advocate of the 'merchant pre-industrial city', underlines that nowhere in rural areas could the productivity and income of rural specialised workers be compared to those of urban craftsmen. In pre-industrial cities, the production of manufactures was indeed a major activity. These two viewpoints (merchant vs. manufacturing pre-industrial city) can be reconciled when we remark that production in pre-industrial cities was usually undertaken by individual craftsmen who were highly specialised. This mode organisation of production involves a high volume of trade within the city. Hence the identification of the pre-industrial city with a market place.

Thus, the shift from pre-industrial to industrial cities was a consequence of another shift from 'dis-aggregated' or 'decentralised' economic activity using pre-industrial technologies to a more 'integrated' industrial production. This led to the concentration of the workforce in larger units of production. Let us consider the shoe industry as an example. It is dealt with by Friedrichs (1995) for pre-industrial cities, and Marshall (1890) and Chandler (1977) for industrial cities. Pre-industrial specialisation was typically about shoe-makers specialising in shoes for female or for male customers or in shoes made with different materials. Industrial

specialisation by contrast was about workers specialising in part of the production process. Pre-industrial specialisation was mostly co-ordinated through markets, whereas industrial specialisation gives a prominent role to hierarchies. Hierarchies internalise many transactions. The corollary is the reduced volume of exchange on the markets. Thus, when it became 'industrial', the city lost some of its importance as a market place as more of its activity was carried on within 'hierarchies' (or factories) and less was transacted through markets (Chandler, 1977).

But why did the city become industrial? At the level of each producer, lower transport costs implied that its market area was not limited to one city any more. This may have induced firms to use more capital intensive methods of production with larger fixed costs, i.e., to generate the shift from pre-industrial to industrial technologies. During this transformation towards capital-intensive methods of production, the local trade of manufactures against agricultural products performed by individuals - typical of pre-industrial cities - was overtaken by a movement of industrial goods between cities orchestrated primarily by managers within firms.<sup>6</sup> The idea of lower transport costs leading to the adoption of industrial methods of production is quite seductive but we must remain careful as to how far we can exploit it, since we still lack firm evidence on this. Lower transport costs were it seems a necessary condition for industrialisation since transport costs, trade and urban specialisation are tied with each other. Whether or not they were a causal factor of industrialisation is still an open question.<sup>7</sup>

Furthermore, following the Industrial Revolution and lower transport costs, cities have become more specialised because of economies of localisation (at the sector level) pulling firms in the same industry together in the same city. As a consequence, the bulk of the industrial city's external trade was no longer with its hinterland but with other industrial cities, trading industrial goods against industrial goods. This idea might be somewhat controversial, as the idea of pre-industrial cities being heavily specialised is quite popular. However, this is largely a myth. For instance, according to Duby (1981), less than 20% of the population of the city of Gent (now in Belgium) during the 13<sup>th</sup> century was employed in the textile sector, which was the city's main area of specialisation. Yet, Gent was one of the most specialised cities of its time. More generally, according to Bairoch (1988), inter-city trade before the Industrial Revolution was for most cities below 5% of the production. This increased specialisation of the industrial city was of course made possible by lower transport costs, making inter-city trade easier. At the same time, these improvements in transport made the two-way trade with the hinterland less crucial. In short, lower transport costs relieved cities from the tyranny of distance.<sup>8</sup>

Associated with this was an important increase in the size and population of many cities. We saw earlier that changes in the pre-industrial urban system were very slow. The rate of urbanisation in Europe went up by less than 2.5 points between 1000 and 1800. But it jumped from 12.1% in 1800 to 18.9% in 1850 and to a staggering 37.9% in 1900. Furthermore, there was only one millionaire city in Europe in 1800 against 9 in 1900 and 110 in 1980 (including 9 above 5 million inhabitants). In this respect, this transformation of pre-industrial cities into industrial cities brought about by the Industrial Revolution changed the nature of our problem. We now have to understand objects, the industrial cities, which are much bigger than pre-industrial cities and, as we will argue, internally organised on different principles.

### *The tyranny of land and the rise of land markets*

The second principle, *the tyranny of land* (in the sense of intra-urban land scarcity leading to high commuting costs) is crucial if we want to understand large industrial cities. The point is again made very clear by Bairoch (1988). He argues that in a typical city of 100,000 persons (assuming a reasonable density of 35,000 inhabitants per km<sup>2</sup>), one can walk from any part of the city to its centre in at most 10-15 minutes, which can be taken as negligible. By contrast, in a city of one million people, the same walking can take up to one hour. In a large city, land is a much scarcer resource. To put this in another way, in a large city, commuting and shopping costs become much stronger brakes on urban growth than shipping costs. With the industrial city, the tyranny of land replaces that of distance.

The first analytical studies of cities by Alonso and Muth in the 1960s (Alonso 1964, Muth, 1969) took this as a starting point. The approach they developed is a real achievement for many reasons. Firstly, they developed a clear and simple explanatory framework to understand industrial cities. This is known as the monocentric model. Secondly, this framework leads to empirical predictions that were tested with success. Thirdly, their simple framework could be and has been extended to understand many other aspects of cities. The monocentric model builds on two very simple ideas. In a first approximation, it assumes that the city has a Central Business District (or CBD) where all economic activity is performed. This CBD is given exogenously (which brings us back to our magnet theory). Then it is assumed that urban dwellers need to commute to the CBD. Commuting is of course costly, which makes land close to the centre a scarce resource. This is why we can speak of a tyranny of land.

As in the previous section, these technological assumptions alone are not sufficient for

us to derive any result about land use and city size. Again, institutions regulating entry into the city and the allocation of land therein need to be considered. The main institution performing those functions in industrial cities is the land market. The guilds in pre-industrial cities were directly preventing the entry of newcomers. By contrast in the industrial city, land prices determined by the market, make it more or less attractive for potential newcomers. In short, the visible hand (i.e., the guild) was replaced by the invisible hand (i.e., the land market) to determine the allocation of people in the cities.

The resolution of Alonso's general model is technically fairly difficult but the underlying principle is simple. Alonso's method of resolution consists, for each inhabitant, or potential inhabitant, in calculating the maximal amount he or she would be willing to pay to live in a given place. Then a bid-rent curve can be derived. As the land market is assumed to be competitive, each parcel of land will be allocated to its highest bidder. In equilibrium, people face a trade-off. Rents are higher closer to the CBD but commuting costs are lower. By living further away from the CBD, consumers pay lower rents but face higher commuting costs. The city stops when the willingness to pay of an additional consumer to reside in the city is below the agricultural land rent.

This very simple model delivers a set of interesting predictions. The rent curve should be decreasing with the distance from the centre; larger cities should have higher rents; and lower commuting costs should lead to larger cities. In other words, the model predicts that lower commuting costs, i.e., technological progress, should increase city size and relieve us from the tyranny of land.<sup>9</sup> At this point it must be said that this model is pretty remarkable if one judges it by its ratio of explanatory (and predictive) power over simplicity.

### *The rationale for new urban institutions*

Before going further into the analysis of the monocentric model, we may try to understand why the land market replaced the guild as the main regulating institution. This is actually rather surprising as this greater importance of the land market comes within a framework of general retreat of the markets at the local level as we saw above. The answer may again have to do with production structures. Labour, after the Industrial Revolution, became a much more homogeneous commodity. Industrial labour was (and still is) specialised at the task level but large occupation-specific investments typical of pre-industrial craftsmanship nearly disappeared with industrialisation. It is more appropriate to speak of fragmented than specialised labour with industrial production. This implies that any newcomer in the industrial

city only had a marginal impact over the other workers, instead of a large discrete impact on a small fraction of the existing urban population as in the pre-industrial city. Thus the guild as a quantity instrument lost some of its usefulness.

Furthermore, the demise of the guild system coincided with the rise of national institutions. Those national institutions favoured the development of new cities from the top and avoided overpopulation in some cities. The important co-ordination role of urban guilds regarding urban size was thus partly transferred to national institutions. Another reason is that urban guilds were dominated by labour (or the most skilled and senior fraction of it), whereas industrialisation coincided with a more important role for capital in production. For instance, large firms were also able to lead the development of new cities, the famous factory-towns, which epitomise the heyday of industrialisation (see Duranton, 1999, for more on this). The corollary of this is of course that the main distributive conflict in the economy shifted from the opposition between landowners and urban labour to that between capital and labour. A final reason for the demise of the guilds lies in the evolution of the physical boundaries of the city. Pre-industrial cities were geographically limited and well-defined, usually by a wall. When they became much larger, cities lost their well-defined boundaries along with their specific legal status and it became much more difficult for guilds to keep controlling cities in the way they did before.

To summarise what has been seen until now we can see that the industrial city differs from the pre-industrial city in at least three dimensions:

- (i) Commuting costs replace shipping costs as the main force limiting urban growth.
- (ii) Productive activities are carried on and co-ordinated more within firms and less through the marketplace. This follows the greater opportunities for long distance trade and the subsequent adoption of industrial production processes.
- (iii) There is a notable exception to the fading importance of the market at the local level. In industrial cities, the land market replaces the guild as the main regulating institution within the city.

### *The development of urban economics building on the tyranny of land*

The monocentric model did not remain this simple and elegant piece of abstract theorising for very long. It has been tested empirically and extended theoretically to account for other aspects of urban life. On the empirical side, the monocentric model is a success since rents in many cities behave according to its predictions. Until recently, cities which are obviously not



monocentric such as Los Angeles, were the exception and not the rule.<sup>10</sup> We shall come back to this below. The empirical literature on the topic is large but fairly consistent in its message: the monocentric model is doing well (Helsley and Rosenthal, 1994, Cheshire and Sheppard, 1995, or DiPasquale and Wheaton, 1996).

A large fraction of the theoretical work on cities since the 1970s has taken the monocentric model as a departure point. A first type of work has consisted in extending the initial Alonso-Muth framework without any major amendment. Let us take a few examples. Instead of a homogeneous population, assume two sorts of workers with high and low wages. Assume also that commuting costs are higher for workers with high wages (their opportunity cost of not working is higher). Then if the demand for land is inelastic (as it is in many European cities), this implies that workers with high wages will prefer to live close to the centre. By contrast, if we assume a very elastic demand for land, rich people may substitute distance for larger land lots, leading to a more US style of land use pattern with the suburbanisation of the rich. As the demand for land can in principle (using panel data) be measured independently from the distance to CBD, this can provide a good hypothesis to explain differences between US and European cities. It can be tested against some other plausible explanations like the more atomised institutional setting in the US leading to larger differences in the provision of public goods. A definitive conclusion on this has not yet been provided (Mieszkowski and Mills, 1993). Another possible extension is to be more precise about the construction sector, assuming for instance that the marginal construction cost of a building is increasing and convex in relation to its height. Besides, demand depends on the location of the building and its size. From these reasonable premises, it is possible to derive the size of buildings as a function of distance from the centre. Here again, the predictions fare well (i.e., height should decrease in the distance from the CBD). Another fruitful direction has been to extend the monocentric model with a single city to a framework with many cities so that the whole urban system could be considered (Henderson, 1974).

A second approach, more critical of the monocentric model, tries to tackle more seriously the assumption of a competitive land market. More realistic is the idea that the land and housing markets are search-types of markets because all the products sold on those markets are unique (not simply in terms of location). This was studied by Wheaton (1990) who shows that housing vacancies are perfectly compatible with equilibrium behaviour and that the perfectly competitive assumption is just a limit case in a search framework. Another angle of attack is to remark that spatial interactions often lead to spatial externalities, i.e., interactions that are not mediated properly by the land market. Spatial externalities make the price

mechanism less efficient in mediating exchange. Pollution, traffic congestion and neighbourhood effects are three typical examples among many. This has led to planning intervention by public authorities. The positive and normative aspects of those planning regulations now constitute a sizeable body of literature. Nonetheless, the tyranny of land remains in the background, as the issue here is about trying to reach a better modelling of the land market and deriving the most suitable rules and institutions to deal with the scarcity of land. (In short, how does the land market work and how well does it perform?) Government intervention cannot suppress the tyranny of land. At best it can alleviate it.

A third strand of literature tries to justify the monocentric assumption more carefully or to amend it in order to allow for richer patterns. This has led to the analysis of multicentric cities (Yinger, 1992) and more recently to that of edge-cities (Henderson and Mitra, 1996, following the work of Garreau, 1991). Despite the breadth of issues covered by contemporary urban economics, let us remark that the monocentric model and its extensions belong to a Thünen tradition, where cities are viewed mostly isolated islands.<sup>11</sup> This neglects the issue of urban specialisation (leading to different types of cities) and the relative location of cities. In short two important pieces of urban geography, the rank-size rule and central place theory, have not yet been successfully incorporated by the monocentric styles of approach.<sup>12</sup>

### *Explaining urban magnetism*

Now we need to open the black-box of the CBD and try to understand the magnetic properties of the city centres instead of assuming them in an ad-hoc fashion. Opening this black-box leads us to consider two possible explanations. Being together in cities has obvious economic and social costs. Above, the focus was put mainly on commuting and shipping costs but it must also be said that crime rates and pollution tend to worsen with the size of cities (Glaeser, 1998). To counterbalance those costs, some benefits are needed. Without them cities would not exist. The benefits of living together can be purely anthropological. We may like to interact with each other since we are 'social animals'. These anthropological reasons seem obvious but do not seem to warrant large cities. They have been incorporated in a formal setting by Beckman (1976). The alternative (and complementary) explanation is that urbanisation has economic gains, be it about production or consumption.

The analysis of economic motives for agglomeration in cities has a long tradition starting with Marshall (1890). He distinguished three different motives for economic agglomeration.<sup>13</sup> He

talked about labour market pooling, the availability of intermediate goods and technological spillovers. The labour market pooling argument reflects the incentive for firms to locate in large cities where they can recruit specialised workers more easily. For workers, cities may also be a place of interest as the presence of more firms may give them an incentive to specialise more. With more potential employers, cities may also offer some form of insurance against unemployment to workers. More generally, cities, by reducing distances, tend to reduce search costs. This was studied formally by Helsley and Strange (1990). The argument about intermediate goods is also fairly straightforward. Suppliers may congregate in cities where the market for their product is larger. In turn, buyers of intermediates may prefer a location with a wider range of suppliers. They are the famous forward and backward linkages first introduced by Hirschman (1958) and formally analysed by Venables (1996).

The last Marshallian motive, technological spillovers, i.e., the better circulation of ideas in cities, is usually fairly popular with students of cities. Nonetheless, this last argument still awaits further research. Proposing a model with urban spillovers is not difficult. But a first problem is that the model should be replicating some known stylised facts. The truth of the matter is that what we do not have many. The second problem is that such a model should also generate some testable implications (and this notoriously difficult given the reliance on externalities which are by definition un-priced). The third problem is to explain why spillovers are often only local and do not diffuse everywhere.<sup>14</sup> These Marshallian categories are very useful tools to think about cities. Their main limitation is their exclusive focus on the production side, whereas cities also attract people because of the variety of (public and private) consumption goods they propose, many of them non-tradable, or tradable only at a cost. This is the underlying motive for agglomeration in Krugman (1991). We may take this as a fourth economic motive for the existence of cities.

Those four economic arguments along with non-economic human interactions allow us to replace our ad-hoc magnet theory by something more operational. In particular, the opening of the black-box of urban attraction enables us to analyse the effect of technological progress on agglomeration motives. In this respect, Fujita and Ogawa (1982) produced a seminal study where they introduced the endogeneity of agglomeration into a traditional urban economics framework using an information theoretic argument. Instead of directly assuming a CBD, they only assume that firms are more productive when they receive information more intensively. The quantity of information that firms receive depends in turn on the number of firms in the region and is subject to distance decay. Using these assumptions, they show that when the decay is very important (bad communication technologies), firms will form a CBD since for

them higher rents in the CBD are more than offset by productivity gains due to the higher information density. But when the circulation of information gradually improves, the urban structure can undergo profound changes with, as possible outcomes, the emergence of edge-cities on the fringes or the complete breakdown of the monocentric structure. This framework was later enriched by Ota and Fujita (1993) who allow firms to spatially separate their activities. They show that lower communication costs can lead information intensive activities (e.g., headquarters) to remain in the CBD, whereas less information intensive activities are relocated to the periphery. The recent evolution of metropolises in developed economies supports the implications of these models (Ingram, 1998).

Using what we have seen until now, we can try to assess the impact of technological progress on cities. The prediction is fundamentally ambiguous. Technological progress improves personal transportation, and thus allows for ever-bigger cities. But it also improves the tradability of specialised goods and the circulation of information over long distance, thus threatening the glue keeping cities together. We have heard this argument again and again. New technologies reduce or destroy the localised aspect of specialised labour or goods, making both distance and land irrelevant. This argument is true but only at the limit. If all mobility costs are zero, the very idea of physical space becomes meaningless. We are however probably centuries away from this limit. Whether it will be reached one day is not even clear, be it only because of the existence of different time zones and physical limits to speed. What is of interest for us is the transition path. There, the ambiguity remains since both the costs of cities and their benefits decline. New concepts need to be introduced to clarify this issue.

## **Post-Industrial City: The tyranny of Proximity**

### *The post-industrial city?*

In the last two sections, we saw that the mobility of primary tasks, that of intermediate goods, and that of tacit information used to be very low and thus could be neglected in explaining pre-industrial and industrial cities. This very low level of mobility gave cities a crucial advantage in the production of manufactures. The corollary of this proposition is that the pre-industrial and industrial cities can be primarily defined by the forces limiting their expansion, the successive tyrannies of distance and land manifesting themselves through the costs of shipping and commuting. We saw that technological progress not only relieved cities from these two

tyrannies (though not fully) but that it has also increased the mobility of formerly localised goods, factors and information, making our assumption of immobility increasingly inaccurate. How can we conceptualise cities in such an environment? Let us define as the post-industrial city, an urban environment where information, factors of production and goods are left immobile by choice and not because of physical constraints as in the pre-industrial and industrial cities. Pre-industrial and industrial cities were viewed as the results of the interactions between nearly immobile objects (e.g., factories, some services, local knowledge) and imperfectly mobile goods and factors. The analysis of the post-industrial city by contrast will be mainly concerned with the interactions between objects and agents, all potentially mobile.

This calls for four remarks. Firstly this definition of the post-industrial city does not guarantee its existence (nor its sustainability). We already demonstrated that in the limit case of perfect mobility for information, goods and people, the idea of geographical space loses its meaning. The second important point is that this attempt to analyse post-industrial cities is not a futurology exercise. The transition between the industrial and the post-industrial city has already started. Just like the transition between the pre-industrial and the industrial city took more than a century, the movement towards post-industrial cities is also slow and gradual. The analysis of post-industrial cities is thus warranted by the current situation. Contemporary cities are still very much industrial cities. For instance, commuting and urban congestion are still very important issues. They will probably remain so for very long but the emerging patterns are those of post-industrial cities. Beyond this analysis of contemporary cities, the following pages can be read as an anti-futurology exercise. We shall demonstrate that most futurists have not done their homework properly as many key issues impacting directly on the future of cities remain unanswered (and ignored in their analyses). Thirdly, it must be noted that the analytical perspective on cities is now reversed. In the last two sections we were mainly looking at the forces limiting urban growth taking for granted the propensity for cities to grow. By contrast, this section is primarily concerned with the remaining glue keeping cities together. Finally, note that economic analysis has provided important insights regarding pre-industrial and industrial cities. Thus, the last two sections could provide well articulated summaries building on the work done by economists interested in cities. Unfortunately the tools to analyse post-industrial cities have not been fully developed yet. This section can only be a tentative exploration. In what remains of the paper, we shall first question some current thinking about post-industrial cities. Then, building on what we did for the pre-industrial and industrial cities, we shall propose a candidate organising principle, *the tyranny of proximity*, which captures the

idea that inter-personal physical proximity will remain necessary and emerge as a central feature of post-industrial cities. We shall then discuss the institutions likely to deal with this new constraint.

### *Key questions about technological progress and cities*

The first issue we need to worry about is the precise nature of the decline in transport costs. Until now, we have divided the frictions related to geography into three different categories: shipping costs, commuting costs and telecommunication costs. It seems that the most important decrease nowadays is in telecommunication costs followed by shipping costs, whereas commuting costs seem to decline only very slowly, if at all (Jones, 1998). This observation seems to reinforce the demise-of-cities theory since it implies that dispersion forces remain roughly stable whereas agglomeration forces decrease. The issue is nonetheless more complex than that as we shall see.

Regarding shipping costs and telecommunication costs, we need to know whether the decrease is primarily a decrease in marginal costs or a decline in fixed costs. A decline in marginal transport costs is likely to have implications radically different from those of a decline in the fixed costs associated with transport or telecommunication networks. The issue of increased tradability led by technological progress is even more complex. Technological progress modifies transportation technologies but it also changes the nature of goods as well as the composition of consumer demand following the resultant rise in incomes. Let us consider only dematerialization (i.e., the tendency for production to become lighter keeping its value constant), which seems to be a pervasive implication of current technological progress (Quah, 1996, Coyle, 1997). Dematerialization and the shift towards weightlessness have very ambiguous effects. On the one hand, they allow many goods to be transferred easily over space but on the other hand, they also contribute to shifting demand towards less tradable goods such as personal services. Regarding the magnitude of those changes and their precise nature, it seems that we do not have any clear idea.

Another key question regarding consumer demand is that of the substitutability between tradable and non-tradable goods. It seems unlikely that technology will make all non-tradable good perfectly tradable in any foreseeable future. But in the same time, technological change will enable very soon the supply of tradable substitutes. For instance the video or pay-per-view television is a substitute for the act of going to the movie. The question is not whether pay-per-view television can replace the act of going to see a movie. The answer is it seems a clear no.

The question is more whether pay-per-view television is a sufficiently close substitute to the cinema screen so that given the cheaper price of wireless diffusion, on the one hand, and the costs associated with urban life on the other, it will lead to the demise of movie theatres in central cities. This did not happen with the video but it might occur with pay-per-view television. This question arises again and again for each 'non-tradable attribute' of the large cities. Then, more subtly, depending on the proportion of attributes that will be substituted like this, the remaining attributes still highly advantaged by agglomeration may or may not remain numerous enough to warrant the survival of large cities. On these crucial points, predictions seem very difficult to make as they involve making wild guesses about the evolution of the whole economy.

On the production side, similar concerns may also be raised. Modern telecommunication technologies are no perfect substitutes for face-to-face communication. At the moment, it even seems that both forms of communication complement each other, so that lower telecommunication costs increase the demand for face-to-face communication (Gaspar and Glaeser, 1998). Hence the current decline in telecommunication costs reinforces large cities. But this does not constitute a stable structural relation. Future telecommunication technologies may become better substitutes for direct interactions. Furthermore, in the 'information society', information is set to become very abundant and thus very cheap, so that it may lose of its importance altogether. The future of cities may then depend not on information and other trendy issues but maybe on more mundane questions like the importance of scale economies in the distribution of water. Again, it is not clear whether predictions can be made on this with a reasonable degree of certainty.

A last important black hole in our knowledge barely ever gets mentioned despite its importance. Cities simply display a lot of inertia. Even if very large cities do not make much economic sense in the future, they may remain because of the existing urban infrastructure and the built environment already there. The comparison between teleworking from home and working at the office should not be made by comparing the operating costs in both cases but by comparing the operating costs of office work (knowing that building costs are sunk) and the costs of setting teleworking operations and outer-city accommodation. The difference is not negligible as the replacement value of the existing built environment is of the magnitude of annual GDP. How is this going to affect urban evolution? We do not know, since unfortunately we know much more about dynamics where capital is 'putty-clay' (i.e., where we assume it can be reallocated at no cost) than about dynamics involving vintage capital (as with the built environment).

### *Which organising principle?*

Ending here our discussion on post-industrial cities would not be fair. Emphasising our areas of ignorance may be healthy but it is clearly not constructive. Let us take some risks and propose some ideas about post-industrial cities. Just as the pre-industrial and industrial cities were led by one organising principle, i.e., the successive tyrannies of distance and land, the post-industrial city may have to comply with another tyranny, *the tyranny of proximity*. By proximity I mean very short-distance, direct personal interactions where market mediation is indirect if any. The justification usually given to account for the importance of 'proximity' is that people like to be close to each other regardless of economic aspects. This postulate leads observers to argue that post-industrial cities will be mainly about collective consumption and very little about production. This argument is I think very incomplete.

Personal physical proximity is and will remain crucial for production purposes. The argument builds strongly on the distinction between codified and tacit knowledge as developed by Foray and Lundvall (1996). Existing means of telecommunication are heavily biased towards codified rather than tacit knowledge. However, it may be true that current progresses in telecommunication technologies lead to a better transmission of less fully codified knowledge. Nonetheless it seems impossible that telecommunication technologies can ever replace direct interactions. The reason is that telecommunication is usually purposeful unlike a lot of face-to-face interactions, which are more often incidental. Let us take academic publications as an example to illustrate this. Assume that the quality of one's research depends on his or her 'frontier knowledge' in a discipline. The latest issue of a journal (a typical example of codified knowledge) only gives a picture of the frontier some years ago. Going to conferences and reading recent working papers gives a better access to current thinking. But the real frontier of research is in people's minds. This frontier is constituted of the current imprecise and uncoded ideas that people have in their mind and that will be later codified in the form of papers and books. This type of knowledge is important and is usually gained during lunches, at the end of seminars, and in the coffee room in an accidental way. It seems difficult to replace this serendipity by contacting all the researchers involved in a given area of investigation every week and asking for their latest thoughts. Accidental direct communication cannot be easily replicated electronically. Discussion lists and electronic fora are fine but they are time-consuming, whereas lunch and coffee breaks are dead periods anyway. Indeed, if proximity was not important, researchers would not fight to be in prestigious institutions and instead would always go for the highest present wage. This minimal conception of proximity



(i.e., serendipitous direct personal interactions) clearly has to do with concepts like human capital externalities and informational externalities. However, the analysis of 'proximity' in relation to local externalities is still very much an open question. More work on this is needed but this is clearly far beyond the scope of this paper.

This argument about proximity is just an 'existence theorem' which does not tell us anything about the size of post-industrial cities. However it tells us that the post-industrial city will remain both a consumption community and a working community unlike what is often asserted.

### *Which regulating institution?*

As we saw earlier, both the pre-industrial and the industrial city both had a major regulating institution, respectively the guild and the land market. What will be the major institution of the post-industrial city? The land market is still here and likely to matter a lot for a long time. But my candidate for the first role is nonetheless the network or more precisely 'the personal networks'. It can be rightly argued that there is nothing really new here. Social networks have played an important role in cities for a very long time. The novelty is in the renewed importance of social and personal connections in economic life and in the rise of high-capacity telecommunication networks (in the sense of infrastructure) mixing voice, image and data, along with the decline of the physical constraints of shipping and commuting associated with pre-industrial and industrial cities.

Why should personal connections become more important in the post-industrial city? They will become relatively more important with the decline of other motives for agglomeration. But they may also become more important in absolute terms. The reason probably lies in the shift towards weightlessness alluded to before. A ton of steel, typical of the industrial world, is easy to buy since its quality can be observed and verified easily. By contrast, the quality of a piece of software is much more difficult to assess and to verify. Hence, trade may have to rely more on 'reputations', repeated interactions and so on, which all imply personal connections. In the same time however, the rise of multi-media communication networks makes anonymous interactions easier. What the result of these opposite forces will be, is unknown.

The concept of personal network allows us to capture the fact that 'proximity' is synonymous with A wanting to be close to B willing to be close to C and D. For instance A is a close colleague of B. B and C play tennis together, whereas C and D are married to each other,

etc. This string of personal interactions all requiring physical proximity is what may keep them together in the post-industrial city. The post-industrial city can then be viewed as the complex web of all those proximity links. The concept of personal networks is nonetheless rather ambiguous since it is not restricted to short-distance relations. Networks may regulate proximity relations but also long-distance interactions. Furthermore, a relationship developed through direct contact may be sustained over long distance and conversely. The complete analysis of the inter-relations between networks and distance is a complex issue, beyond the scope of this paper. Note that this (residual) agglomeration force in post-industrial cities works differently from the way it pulled people together in industrial and pre-industrial cities. In the post industrial city A wants to be close to B who is linked to C who is attached to D. By contrast, in the pre-industrial city, A, B, C and D may have been together primarily because of the benefits in terms of security which accrued symmetrically to all. More generally, attraction in pre-industrial and industrial cities was more 'universal' pulling A, B, C and D together.

It needs to be stressed that our analytic knowledge of networks is quite thin.<sup>15</sup> The concept of network is more often used as a fashionable buzzword than seriously examined. Let us review some important characteristics of networks and see how they may impact on the post-industrial city. The first important characteristic of a social or physical network is its shape. In short, the architecture of a network matters a lot. There is a cost issue. Typically, hub-and-spoke architectures minimise the cost of setting up a network (as they minimise the number of links). Hub-and-spoke networks tend to favour core-periphery types of structure as the hub enjoys privileged access to information. The second important characteristic is the bandwidth or capacity of a network. However weightless, post-industrial production relies on a physical infrastructure whose capacity is not infinite. An important (and ill-understood) part of the development of networks is the race between the demand and the supply of bandwidth. Furthermore even if physical connections can be multiplied, personal connections are limited in number by our finite human capacities. Those two issues, network shape and capacity, if they really matter, should work in favour of the city. The city can indeed be viewed as a port of entry into the world of networks. There could emerge a new geography led by the physical and personal networks, whose nodes would be nothing else but cities.

Other issues are also of importance: how are these networks going to regulate interactions and be regulated? Will network externalities be important enough for everybody to be plugged in or will new divisions between plugged and un-plugged segments of the population arise? Unfortunately, the existing trends are too sketchy for us to be able to say something meaningful on this. For instance, the Internet is nowadays regulated through

quantities (i.e., congestion), which makes proximity to high-bandwidth wires very important. If in the future, Internet access evolves more towards a price regulation, this may change dramatically.

## **Conclusion: Whither cities?**

The first two sections of this paper showed that the two concepts of pre-industrial and industrial cities could be clearly separated. Each type of city had its growth limited by one important physical friction: distance for pre-industrial cities and land scarcity for industrial cities. Each type of city was regulated by a major institution, respectively the guild and the land market and manufactured goods were produced according to a specific mode of production, respectively individual specialised and localised craftsmanship and industrial firms. We also demonstrated that for each type of city each element of the triplet (physical friction, institution and mode of production) was consistent with the other two elements.

Using this framework, we showed that the concept of the post-industrial city kept together by the tyranny of proximity and regulated by networks deserved further investigation. Unlike the previous types of cities, which were about global attraction forces limited by strong physical constraints, the post-industrial cities are about very short distance interactions with less binding physical constraints. Will it be enough to warrant the survival of cities? We think so but we showed that many questions remain unanswered. Furthermore, our investigation was very partial. Other considerations may lead governments to implement very strict land-use rules outside cities and thus keep people in cities to minimise environmental problems for instance.

Ultimately, this analysis mainly focused on two issues: the existence and size of cities. Clearly many aspects were left aside. The evolution of the built environment was completely neglected. The articulation between urbanisation and social change was also ignored along with many other aspects of the internal organisation of cities. Needless to say this deserves further investigation.

## Acknowledgement

I am grateful to Olivier Cottray, Rachel Guillain, Jean-Marie Huriot, Diego Puga, Jenny Robinson and Andr  s Rodr  guez-Pose for their remarks and comments on an earlier version of this paper.

## Notes

1. How important this feedback was and is today remains to be estimated.
2. This is supported by the fact that the urban population in pre-industrial cities of Asia was more important than in pre-industrial Europe due to the higher nutritive return of rice.
3. This was also true, albeit to a lesser extent, for Asian cities (Weber, 1958). Central governments in Asia were more powerful than in Europe. This is possibly due to topographical differences, which implied a greater need for central co-ordination (e.g., control of rivers).
4. All numbers are from Bairoch (1988), unless specified otherwise.
5. Recent extensions of Krugman's work showed that the mobility assumption for manufacturing labour was very important. If we assume immobile labour as well as linkages between firm, it can be shown that the relation between agglomeration and transport costs might be inversely U-shaped. The reason is that for low enough transport costs, firms may choose to move back to the depressed areas where labour is cheaper in order to serve the core. See Puga (1999) for more on this.
6. See Kim (1998) for evidence about the evolution of the US corporate geography.
7. There is paradox here. Industrial cities emerged following innovations leading to industrialisation, whereas we stated that pre-industrial cities were not conducive to innovation. Besides, note that industrialisation was primarily a European phenomenon. The interpretation we propose here rests on the relatively high degree of interactions between cities and the relatively low level of interference from central powers in late pre-industrial Europe. Inter-city competition was thus much more important in late pre-industrial Europe than in the Chinese or Indian empires and it probably limited the effects of guilds attempting to maintain the status quo. By contrast in China or in India, powerful central authorities prevented any form of territorial competition, which reinforced the conservative tendencies of the local powers.
8. Of course, technological change enabling the use of industrial methods of production was also instrumental.

9. This implication holds provided the income elasticity of the demand for land is not implausibly high. Otherwise, the wealth effect induced by technological progress could lead to a much greater consumption of land dominating the effects of cheaper commuting.
10. People should remain cautious on this issue, as the monocentric or non-monocentric properties of a city must not be assessed through armchair empiricism but in the light of measured rent gradients after controlling (at least) for housing quality and for the provision of local public goods.
11. Before moving deeper into this issue, let us add that this very short overview does not do justice to urban economics. The interested reader can refer to a recent textbook. DiPasquale and Wheaton (1996) focus on land and housing markets. McDonald (1997) offers a thorough presentation of the theory sketched above, whereas Bogart (1998) takes a broader and more accessible view on modern urban economics.
12. The literature on the topic is nonetheless not insignificant. One can refer to Abdel-Rahman and Wang (1996) or Fujita *et al.* (1995) for recent tentative explanations on those issues.
13. Nonetheless the very idea of economic agglomeration made economic theorists uncomfortable for many years. The dominant competitive paradigm rests on the assumption of convex preferences, which should lead each consumer to consume many goods, including many accommodations. Typically, with convex preferences two half-houses should be preferred to one house. Most of us by contrast tend to live in only one house because of obvious personal indivisibilities (i.e., non-convexities in the preferences). Incorporating this in mainstream general equilibrium models did not prove easy nor very fruitful unfortunately (Schweizer and Varaiya, 1976, Grimaud and Laffont, 1989). The second major building block of the standard paradigm of economic analysis is that of convex production sets (i.e., constant or decreasing returns to scale in production). However, as shown by Starrett (1978), this assumption should lead to the complete dispersion of economic activity over space and 'backyard capitalism'. The only reason for agglomeration would be a concentration of natural resources. Thus the analysis of cities calls for the use of increasing returns and indivisibilities. It is only in the last 20 years or so that the adequate tools to study increasing returns have been developed.
14. An in-depth discussion of cities and information can be found in Guillin and Huriot (1998).
15. The formal analysis of networks in economic theory started only recently and the tools we have are still rudimentary. We can only hope to see more developments in this area of investigation.

## **Bibliography**

- Abdel – Rahman H., and Wang P., 1995, 'Towards a General Equilibrium Theory of a Core-Periphery System of Cities', *Regional Science and Urban Economics* 25, 529-546.
- Ades A., and Glaeser E., 1995, 'Trade and Circuses: Explaining Urban Giants', *Quarterly Journal of Economics* 110, 195-228.
- Alonso WA., 1964, *Location and Land-Use*, Harvard University Press, Cambridge (Mass.).
- Bairoch P., 1988, *Cities and Economic Development*, Chicago University Press.
- Beckman M., 1976, 'Spatial Equilibrium in the Dispersed City', in Y.Y. Papageorgiou (ed.), *Mathematical Land Use Theory*, Lexington Books, Lexington (Mass.), 117-125.
- Bogart W. T., 1998, *The Economies of Cities and Suburbs*, Prentice Hall, Englewood Cliffs (N.J.).
- Braudel F., 1984, *Capitalism and Material Life, 1400-1800*, Fontana Collins (London).
- Castells M., 1989, *The Informational City*, Blackwell, Oxford.
- Chandler A., 1977, *The Visible Hand*, Harvard University Press.
- Cheshire P., and Sheppard S., 1995, 'On the price of land and the value of amenities', *Economica* 62, 247-67.
- Coyle D., 1997, *The Weightless World*, Capston Publishing, Oxford (UK).
- Dicken P. and Lloyd P., 1990, *Location in Space: Theoretical Perspectives in Economic Geography*, Harper Collins, NY.
- DiPasquale D., and Wheaton W., 1996, *Urban Economics and Real Estate Markets*, Prentice Hall, Englewood Cliffs (N.J.).

Duby G., ed, 1981-83, *Histoire de la France Urbaine*, Le Seuil, Paris (5 volumes, 1/ *La Ville Antique*, 2/ *La Ville Mèdièvale* 3/ *La Ville Classique* 4/ *La Ville Industrielle* 5/ *La Ville Contemporaine*).

Duranton G., 1998, 'Labour Specialisation, Transport Costs and City Size', *Journal of Regional Science* 38, 553-573.

Duranton G., 1999, 'Urban Structure, Urbanisation and Growth', forthcoming in J.-M. Huriot and J. Thisse (eds), *Economics of Cities*, Cambridge University Press.

Eaton J., and Eckstein Z., 1997, 'Cities and Growth: Theory and Evidence from France and Japan', *Regional Science and Urban Economics* 27, 443-474.

Evenett S., and Keller W., 1997, 'On Theories Explaining the Success of the Gravity Equation', mimeo.

Foray D., and Lundvall B., 1996, 'The Knowledge-Based Economy - From the Economics of Knowledge to the Learning Economy', in *Employment and Growth in the Knowledge-Based Economy*, 11-34, OECD, Paris.

Friedrichs C., 1995, *The Early Modern City, 1450-1750*, Longman, New York.

Fujita M., Krugman P., and Mori T., 1995, 'On the Evolution of Hierarchical Urban Systems', mimeo.

Fujita M., and Ogawa H., 1982, 'Multiple Equilibria and Structural Transition of Non-monocentric Urban Configurations', *Regional Science and Urban Economics* 12, 161-196.

Garreau J., 1991, *Edge City*, Doubleday, New York.

Gaspar J. and Glaeser E., 1998, 'Information Technology and the Future of Cities', *Journal of*

- Urban Economics* 43, 136-156.
- Glaeser E., 1998, 'Are Cities Dying?', *Journal of Economic Perspectives* 12, 139-160.
- Grimaud A., Laffont J. J., 1989, 'Existence of a Spatial Equilibrium', *Journal of Urban Economics* 25, 213-18.
- Guillain R., and Huriot J. M., 1998, 'Cities, City Systems and Information', mimeo LATEC.
- Helsley R., and Rosenthal S., 1994, 'Redevelopment and the Urban Land Price Gradient', *Journal of Urban Economics* 35, 182-200.
- Helsley R., and Strange W., 1990, 'Matching and Agglomeration Economies in a System of Cities', *Regional Science and Urban Economics* 20, 189-212.
- Henderson J.V., 1974, 'The Sizes and Types of Cities', *American Economic Review* 64, 640-656.
- Henderson J.V., and Mitra A., 1996, 'The new urban landscape: Developers and Edge cities', *Regional Science and Urban Economics* 26, 613-643.
- Hirshman A., 1958, *The Strategy of Economic Development*, Yale University Press, New Haven (Conn.).
- Ingram G., 1998, 'Patterns of Metropolitan Development: What Have We Learned', *Urban Studies* 35, 1019-1035.
- Jacobs J., 1969, *The Economy of Cities*, Harmondworth, NY.
- Jones C., 1998, 'The Upcoming Slowdown in U.S. Economic Growth', National Bureau of Economic Research Working Paper 6284.
- Kim S., 1998, 'The Rise of Multiunit Firms in US Manufacturing', NBER Working Paper 6425.



- Krugman P., 1991, *Geography and Trade*, MIT Press, Cambridge (Mass.).
- Marshall A., 1890, *The Principles of Economics*, McMillan (N.Y.), re-edition 1925.
- McDonald J., 1997, *Fundamentals of Urban Economics*, Prentice Hall, Englewood Cliffs (N.J.).
- Mieszkowski P., and Mills E., 1993, 'The Causes of Metropolitan Suburbanization', *Journal of Economic Perspectives* 7, 135-147.
- Muth R., 1969, *Cities and Housing*, University of Chicago Press.
- Myrdal G., 1957, *Economic Theory and the Under-developed Regions*, Duckworth, London.
- Ota M., and Fujita M., 1993, 'Communication Technologies and Spatial Organisation of Multi-unit Firms in Metropolitan Areas', *Regional Science and Urban Economics* 23, 695-729.
- Pirenne H., 1925, *Medieval Cities, Their Origins and the Revival of Trade*, Princeton University Press, Princeton, New Jersey, Reedited by Double Day Anchor, New York.
- Puga D., 1999, 'The Rise and Fall of Regional Inequalities', forthcoming *European Economic Review*.
- Quah D., 1996, 'The Invisible Hand and Weightless Economy', mimeo.
- Robson B., 1973, *Urban growth: an approach*, Methuen and co, London.
- Sassen S., 1991, *The Global City*, Princeton University Press.
- Schweizer U. and Variaya P., 1976, 'The Spatial Structure of Production with a Leontieff Technology', *Regional Science and Urban Economics* 6, 231-251.

- Starrett D., 1978, 'Market Allocations of Location Choice in a Model with Free Mobility', *Journal of Economic Theory* 17, 21-37.
- Venebles A., 1996, 'Equilibrium Locations of Vertically Linked Industries', *International Economic Review* 37, 341-360.
- Von Thünen H., 1826, *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationaloekonomie*, translated by Wartenberg and Hall, *von Thünen's isolated state*, 1966.
- Weber A., 1909, *Theory of the Location of Industries*, University of Chicago Press.
- Weber M., 1958, *The City*, The Free Press, NY.
- Wheaton W., 1990, 'Vacancy, Search, and Prices in a Housing Market Matching Model', *Journal of Political Economy* 98, 1270-92.
- Yinger J., 1992, 'City and Suburb: Urban Models with more than one Employment Centre', *Journal of Urban Economics* 34, 181-205.