Settlement Size Dynamics in History

Denise Pumain

SETTLEMENT SIZE DYNAMICS IN HISTORY

Denise Pumain, UFR de Géographie, Université Paris I, INED, 27 rue du Commandeur, 75014 Paris, France

Paper presented at the 7th European Colloquium of Theoretical and Quantitative Geography, Stockholm, 5-8 sept. 1991

INTRODUCTION

Settlement patterns have been for a long time objects of concern for geographers. However, there is still a conceptual link lacking between the comprehensive theories of their evolution, including determinants of rural depopulation and urbanization, and the statistical observations which are made about the distribution of their sizes.

The scattering of population over a territory always leads to aggregates of various sizes: households, hamlets, villages, towns and cities. There is on the long run a trend towards an increase in the size of the aggregates and an increase in the concentration of the population in the aggregates. This evolution has not always been continuous in the past. However, the process of concentration of the population has been almost continuous and of an unprecedented intensity since at least two centuries in many developped countries: more than 80% of the total population is now concentrated in the largest towns and cities, it was only about 20% before. From now on, more than a half of the total population of the world is concentrated in cities.

Despite the many difficulties in defining and delimitating urban areas in a comparative way through space and time, all the studies which were based on a rigorous statistical material have showned the persistence of the same kind of statistical distribution of city sizes over the whole process of spatial concentration of the population, either for historical periods (for instance, De Vries, 1984 and Bairoch, Batou, Chevre, 1988) or for contemporary evolution (Moriconi, 1991). However, several questions remain unsolved. Traditionally, the evolutions of rural and urban settlements have been considered separately. Does the so-called "rank-size" distribution characterize only cities and towns or all settlements? How can the various interpretations of such a statistical shape of the distribution of sizes of settlements be integrated in a unified theoretical construction?

The aim of this paper is to recall briefly the explanations which are already well-known to geographers, and to review related recent work by archeologists and historians, with major empirical and theoretical contributions but partly questionable interpretations. Finally, a more general interpretative framework will be provided, with the help of concepts from the synergetics.

1- CLASSICAL EXPLANATIONS OF THE HIERARCHY OF URBAN SIZES

Usually the cities belonging to the same region or to the same country are markedly differentiated according to their size. There is a strong evidence about a kind of invariance in the statistical distribution of city sizes. The size distribution is highly skewed, with many small towns located close to each other, less and more distant medium-sized cities and only a few very large cities which are located in average far away from each other. The number of cities is roughly in an inverse geometrical progression with their size. This distribution was identified (Le Maître, 1682) and formalized (Auerbach, 1913) a long time ago. It is usually known as the "rank-size rule" as described by Zipf (1949): when plotted on logarithmic coordinates, the population size of a city is a linear function of its rank in the urban hierarchy. The cumulated frequency distribution of the sizes of cities is also similar to a Pareto function (Simon, 1955) or to a lognormal distribution (Gibrat, 1931). The slope of the Pareto distribution, or the dispersion of the lognormal function, are indicators of the degree of concentration of urban populations (in other words, it is a measure of the inequalities in size among cities).

1.1 Spatial equilibrium

Such types of size distributions for cities can be found everywhere in the world (Zipf, 1949, Rosen and Resnick, 1980, Parr, 1985). Several kinds of explanations have been put forward to explain the distribution of city size in urban systems (for a review, see Pumain, 1982 or Sheppard, 1982). This distribution was at first considered as expressing an **equilibrium.** Zipf considered the "rank-size distribution" as an equilibrium between two counterbalancing forces, a "force of concentration" and a "force of diversification". Both forces resulted from the "law of least effort" tending either to concentrate people and activities on the same place of production and consumption in order to avoid transportation costs, or to locate production and consequently the consumption in the vicinity of the natural ressources to be exploited, ressources whose distribution is likely to be rather dispersed. His explanation of why the two forces equilibrate in order to produce the perfect regular size distribution of cities was however not really argued.

Tentative explanations of city size distribution as the most probable state of an urban system are not convincing either. Berry and Garrison (1958) have suggested without demonstration that the Pareto distribution would correspond to a state of maximum entropy within the urban system. Curry (1964) has derived a skewed distribution of city sizes from the principle of entropy maximization (in the sense of information theory) in a process of allocation of individual cities among the various possible city sizes. Many weaknesses have been found however in his interpretation: the theoretical resulting distribution is rather different from the observed ones (Haran, Vining, 1973), the model supposes a closed urban system (Chapman, 1970), and the explanation is only shifted from the size distribution toward the constraints which have to be defined to get it as a result of the process (Pumain, 1982).

Central place theory - whose inventor may not be W. Christaller (1933) but J. Reynaud (1841, as discovered by Robic, 1982) - provides a static explanation for the size, spacing, and hierarchical functional organization of cities. "Sociability" according to J. Reynaud, "centrality", according to W. Christaller, are the key concepts for understanding the grouping of population and of non-agricultural activities in the most accessible, central, places. The regular pattern of nested urban levels is considered as an equilibrium between supply and demand of goods and services. Supply is limited by a minimum threshold in the number of consumers which is required for its profitability, whereas demand is limited in extent by a decreasing accessibility around the centers and a corresponding increase in the relative transportation cost. In this last model, the theoretical distribution of city sizes is supposed to be discontinuous, like a staircase where each level corresponds to a specific bunch of urban functions. Several authors have tried to make it compatible with a rank-size distribution by introducing stochastic variations in the hierarchical levels (Beckmann, 1958, Beguin, 1979). One may indeed argue that several independant and random factors are blurring the size limits of each level and smoothing the size distribution: the location of industrial functions, which does not depend on city size, and the local variations in the demand per capita and in the elasticity of supply according to profit are the main explanatory factors for such a randomization of the model. However, Parr (1970) argues that the rank-size distribution has a larger validity, because it contents all the plausible hierarchical distributions, whereas all hierarchical structures which are derived from central place theory are not compatible with the actual size distributions of cities.

1.2 Genetic explanations

Of particular interest are the models which derive a skewed distribution of sizes from the growth process of the population of cities. Several models of the distribution of growth among cities have been designed in order to demonstrate how they could lead to a Pareto distribution (Steindl, 1965 ; Simon, 1955) or to a lognormal distribution (Gibrat, 1931) of the sizes of cities. Most of them were stochastic models. Their specifications could be different but they were all based on the same kind of main hypothesis. Gibrat demonstrates that any distribution of city sizes will become a lognormal distribution if the following growth process is observed:

- cities are growing at each small time interval of an amount of inhabitants which is proportional to their size (it is equivalent to say that the growth rates of cities are independent of their sizes) ;

- the distribution of growth rates among cities is independent from one time interval to another.

In other words, the "law of proportional effect" assumes that the causes of urban growth are diverse and numerous, that the effect of one of them is small when compared to the total effect of all causes, and that this global effect is multiplicative, and proportional to the size of cities. The model looks at first as a simple statistical interpretation of the distribution of city size in an urban system.

Due to its simplicity, this model has been tested several times (for instance Robson, 1973, Pumain, 1982, Guérin-Pace, 1990). The results are rather convergent:

- a proper description of the urban growth process is actually given by the "law of proportional effect". On the one hand, on average, the variations in the number of inhabitants of a city are roughly proportional to its size, or, in other words, in the whole urban system, urban growth rates are randomly distributed among cities, irrespective of their size. On the other hand, the spatial distributions of the growth rates are rather independent from one period of time to the other. Both rules hold for the largest part of the growth process in any urban system.

- it may be added that, despite non negligible fluctuations of the ratio between the growth rates of urban population and of the number of cities, there is a rather strong parallelism in the evolutions of both rates and then the hypothesis of a not too unstable relationship between those two processes of expansion in urban systems can also be hold for true. Of course, this last observation relies upon the degree of relevance of the geographical definition which is used for the delimitation of towns and cities. In a simpler way, it can be said that the dynamic of the rural settlements whose population is large enough to let them grow above the threshold of urban size is similar to the one of the cities already belonging to the urban system.

As a first approximation, the "law of proportional effect" can be seen as a very general process, which may provide at least a satisfying statistical explanation of the lognormal or Paretian structure of urban systems. However, to become part of an urban theory, this process has to be more closely related to the intrinsic properties of urban systems. At first, one may wonder if the hierarchical organization is a typically urban and recent feature or if it can be hold for true for any settlement system an at any historical time.

2 HIERARCHY OF RURAL SETTLEMENTS : EVIDENCE FROM ARCHEOLOGY

In most interpretations, as in those mentioned above, the strongly differenciated size distribution and the hierarchical organization have been described as a distinctive feature of towns and cities systems, without questioning if they could be a more general property of any settlement system, including rural ones. It would be of the highest importance to know if other types of distributions can be observed: is it possible for instance to imagine a normal distribution of settlements sizes? Such a distribution would suggest that there exists an optimal size for human settlement, depending upon the availability of ressources in a given distance for a group, and/or upon the internal management capacities of the group. It would imply that under those circumstances a permanent settlement is not viable below or above a given threshold of size.

If such patterns of settlements could be found, for instance for agricultural communities, the explanation of the distribution of city sizes would have to be searched in the specificity of the urban functions. If it is impossible to find evidence for any normal distribution of settlement sizes, the explanation of the lognormal distribution can be more general and may be limited to the study of the aggregation of the population over a territory.

2.1 Hierarchy of rural settlements

To our knowledge, the number of studies dealing with the distribution of rural settlement sizes is very limited. A pioneer work by Baker (1969) identified a "reversal of the rank-size rule" - that is, actually, a lognormal distribution - for rural settlements of the nineteenth century in a French region (figure 1). Unwin (1981) also found a similar distribution for medieval settlements in Nottinghamshire. Burtchett (1969) insisted on the "size continuum" of small settlements in the Central Loire Valley, observed between 1846 and 1946. On the contrary, Grossman (1982) and Sonis and Grossman (1984) suggested the existence of a "concentration plateau" around a middle size of rural settlements, for Samaria, a region west of the Jordan river. This settlement size distribution is better adjusted by a platykurtic right-skewed size frequency distribution than by the rank-size model (those last authors did not try however to adjust a lognormal distribution; figure 2).

Therefore, available studies do not allow to decide clearly if the general model for the statistical distribution of rural settlements sizes differs in a systematic way from the urban one. More empirical material and analysis are necessary. However, it may be hold for true that the distribution of rural settlement sizes is always highly skewed, and that it is closer to a lognormal distribution than to a normal one.

In any case, the specificity of the urban growth of the last two centuries has led to individualize urban systems in describing them as distinct from the other parts of the settlement systems. But they derive from the same model of distribution. This can be understood from figure 3, which represents the size distribution of the French settlement system on a gausso-logarithmic paper. Clearly, it is a combination of two lognormal distributions differing by their slope, at a threshold of size which is close to the rural-urban threshold (2 000 inhabitants). This slope is a measure of inequalities in settlement size (it corresponds to the standard deviation of the distribution of logarithms of sizes). As towns and cities have been growing at a rate in average twice or three times the rate of growth of rural populations, size inequalities are much larger among urban than among rural settlements.

2.2 Evidences from archeology

As archaeologists mainly deal with rural -and even pre-agricultural- settlement systems, their results may also be considered in order to compensate our lack of more recent information. Obviously, the difficulties in delimitating settlements and in measuring their sizes are still larger in this case. But systematic recordings and comparisons lead to very interesting results. Reviewing recent archeological litterature about the recorded size of settlements, and using superficies as a measure of settlement size, R. Fletscher (1986) assessed the generality of what he called a "hollow curve" shape for such distributions, at least since sedentary communities appeared (figure 4). One may conclude from those observations that there is a strong evidence for a very early and permanent trend to a lognormal distribution of human settlements, even if the large number of settlements recorded in the smallest size class is partly an artefact.

The evolution of those lognormal distribution of settlement sizes over time is however not a coutinuous one. According to R. Fletscher, "long periods of stability in average and maximum site and sellement size have been punctuated by brief periods of marked size increase" (p.65). He distinguishes three major periods: the development of sedentary communities, the initial formation of urban settlements, and the Industrial revolution, for which "maximum settlement/site sizes successively approached and exceeded approximately 1.5 ha, 100 to 150 ha and about 150 sq km". Such area thresholds "seem to have some significance as indicators of constraints on cultural processes in compact occupation areas" (p.69). Can it be suggested that, within the context of new communication and information technologies, we are entering a fourth stage where large conurbations or more or less continuously built-up metropolitan areas would become the most common pattern for settlements, with maximal sizes around 15 000 sq.km? This prediction would be in agreement with the theoretical suggestions made for instance by L. Suarez-Villa (1988).

3 DIFFERENT URBAN HIERARCHIES THROUGH HISTORY ?

Whereas assessing the early validity and universality of a lognormal type of distribution as a model of reference for settlement size distribution, archaeology also confirms another important idea about urban settlement. It seems that from the very beginnings the city never existed as an isolated entity, but always as a node in a network of cities (Renfrew, Cherry, 1987, van der Leeuw, 1990). Whatever the theory about the origins of urban functions, either political and religious control over a society and its territory, or handicraft production (Duby, 1980), or economic trade and exchanges (Bairoch, 1985), it seems very difficult to imagine an isolated city but on the contrary, following Berry (1964), it is necessary to conceive it as "a system within a system of cities".

The size distributions of towns and cities belonging to such systems seem to maintain the same statistical shape since a very long time, over centuries. For instance in Europe, and despite the numerous difficulties in evaluating the number of inhabitants, it is hold as highly probable by several historians that such an hierarchical organization of cities exists at least since the Roman Empire. The existence of the hierarchy is not restricted to periods of demographic growth, since it has maintained itself even during the regressive periods of the Middle Ages.

3.1 From "immature" to "mature" urban systems?

Quite a few recent publications by archeaologists and historians elaborate on Zipf's idea that a regular rank-size distribution can be better observed in territorial entities with a long and stable history within the same boundaries (which he called "national unity"), and on Berry's hypothesis (1961) that lognormal city size distributions reveal the best achieved and economically well-integrated territorial organizations and that all observed historical distributions should tend toward this model. G.A. Johnson (1980) identifies several examples of "convex" rank-size distributions in various archaeological and historical contexts, where the largest settlement is smaller than the rank-size rule would predict, and which he relates to a low level of system integration (figure 5). The anthropologist C.A. Smith (1982 and 1990) refers to "immature" urban systems when the top of the rank-size curve is flat, that is when a group of the largest cities have similar sizes. According to her, such systems are "characteristic of the premodern world as well as the underdeveloped economic systems of the modern world". She quotes examples from Russell (1972) in the case of European medieval cities as well as Vapnarski's study of Argentinian urban network (1975) and her own findings for central latin America. She argues that modernization brings a steeper slope of the distribution, which comes closer to a lognormal one. Primacy, which is tied to local conditions, may appear in both cases.

3.2 Methodological uncertainties

Whatever appealing, the theory of urban systems evolving from immature to mature rank-size distributions in the course of a territorial integration process still needs more convincing demonstrations. Most of the empirical material and the statistical methods which were presented until now can be criticised. First, the reference which is taken is the simplest Zipf's model, the rank-size distribution with a slope of -1. Empirical adjustments show however that the actual adjusted slope may vary, from around 0.6 to 1.3: for instance De Vries (1984) found a value of 0.63 for the European urban system in 1500. Judgments about degrees of primacy do depend on this parameter since in the case of a slope of 0.6, the size of the first city would be expected to be 1.5 time the size of the second one, and almost 2.5 with a slope of 1.3.

Moreover, many of those studies are relying upon very small samples of cities. It is well known that most of the distorsions in city size distributions are observed in their upper part, and that the type of statistical adjustment which can be made depends heavily upon the number of cities under consideration, within the same urban system (Guérin-Pace, 1990).

Third, many counter-examples to an historical evolution from primate or oligarchic distributions to more lognormal ones can be found. Cuervo (1989) after a detailed investigation of the evolution of primacy indexes for latino-american cities, showed that neither a unique evolutionary model could be derived, nor a straightforward connexion between the shape of the urban systems and the trends of economic integration. It is well known that France is an example of a well-integrated urban system which functions with a primate distribution of city sizes since a very long time. On the contrary, the city size distribution of the world was of a rank size type, well before it could be considered as an integrated system.

3.3 Conceptual limitations

J. De Vries (1990) also questioned those interpretations because of "the use of a largely arbitrary norm (the rank-size rule) and the confusion over measurement techniques". He adds "the problem of arbitrariness in the delimitation of regions". He argues that "a distribution with a shallow slope can happen not only when many cities are of similar size and presumably perform similar functions in relatively autarchic hinterlands", but also when "two or more regions are wrongly combined, that is when the system is misspecified" (p.51).

Historical studies of urban systems raise the difficult problem of circular reasoning in delimitating a suitable geographical framework. In many studies of contemporaneous urban systems, the national boundaries are used as limits, because they usually strongly reduce the number of interactions between cities located inside and outside. The delimitation of an urban system remains however a difficult problem, which has not been solved until now in a satisfying way. We know from empirical observation and from urban theory that the sphere of relations of any city is highly dependent upon its size. Small towns may well have relationships of interdependence only at a regional scale, whereas medium-size ones will be more strongly connected at a national level, and large cities at a supranational level. The delimitation of the relevant urban system should therefore vary according to the size of the cities being considered, as well as in the course of time, when interconnections are progressively settled. This kind of variable geometry for the space of reference has not yet been satisfactorily integrated into any evolutionary model.

Whereas the lognormal distribution can be taken as a very general model for the distribution of settlement size, one should avoid considering it as a norm and be very careful about the interpretation of particular deviations from this model. Since a long time statisticians have repeated warnings about the risks of "overinterpretations" of statistical distributions. Especially in the case of long-tail distributions, large similarities may exists between observations and many theoretical models (Quandt, 1964) and a slightly better statistical adjustment is not a reason for believing more in a particular statistical model. In any case, such models are probably too simple for a dynamic description of an urban system.

3.4 Stages in the development of urban systems

There is until now no sufficient proof of any regular trend in the evolution of the qualitative form of such distributions over historical time and more research should be carried out in this direction. However, during the course of time, a quantitative transformation does occur: a general increase in settlement size concentration is observed. This trend was neither regular nor continuous in the past. In Europe's history, De Vries (1984) identified two types of stages for urban development (figure 6): periods of "urban concentration", like 1500-1750 and 1850-1970, where urban population is increasing within the existing cities, and in most of those cases the largest cities are growing much faster than the small ones; and periods of "city creation", like 800-1300 and 1750-1850, where on the contrary urban population is growing because of the creation of new, and of course small, towns. The alternance of city growth and city creation does alter the shape of the city size distribution: when more cities are created, the slope is reduced, whereas the rapid growth of existing urban population and its correlation with city size tends to increase the concentration of the distribution of sizes.

It is therefore the way the growth is distributed inside the settlement system which is responsible for the transformations in the shape of city size distribution. In Great Britain as in France, during the last two centuries at least, the largest cities have taken a larger advantage from the urbanization process than the smallest ones (Robson, 1973, Guérin, Pumain, 1990). There is a slight positive correlation between growth rates and the size of cities (figure 7), which is higher with the logarithm of size than with the size itself. It is responsible for a larger increase in the absolute value of the slope of the rank-size distribution over time than it would be expected from Gibrat's model (figure 8). On the contrary, the correlation between city size and city growth was negative in a developing urban system like the American one during the XIXth century (Madden 1955 and Moore 1958), it also has become negative during the last fifteen years in the most urbanized countries, leading to a slight decrease in the concentration of city sizes (figure 7).

This trend was interpreted by several scholars as a complete reversal, in the framework of a "counterurbanization" process (Champion, 1989): for instance J.B. Parr (1985), (who uses a different representation and computes a slope whose value is the inverse of the one shown in fig. 8) considers the "U-shaped" curve that he gets as characteristic of the urbanization process (figure 9). However, it can be shown that this effect is partly an artefact due to the data that were used (Frankhauser, Guerin-Pace, Pumain, forthcoming), and this interpretation may be at variance with the observations of the eighties and the previsions made for the nineties, which show again, at least in Europe, a trend toward an increasing concentration of urban population.

Such long-term studies of urban growth show that instead of trying to interpret city size distributions as the product of a momentaneous equilibrium of forces of of any hierarchical order, it is necessary to consider them as integrating a long history of urban growth and development of the settlement system. The development of urban systems can be modelled and simulated. According to modelling capacities or to the size of the system under consideration (the whole settlement system or only its urban part), either the lognormal distribution or a Pareto distribution can be chosen. One could search to relate the general model of settlement size distribution to a generic process for the aggregation of populations. This model should be a dynamic one: there is no reason to think of urban systems as being in any equilibrium, and genetic models should be used for explaining more specific historical sequences which can be conceived as particular realizations of a more general dynamic model. The aim is not to add a further possible explanation to a model which is already "overidentified", but to establish closer connections between its parameters and specific observations about the evolution of urban systems. The universality of the lognormal configuration leads to think of it as a very strong and stable attractor in the dynamics of the urbanization process.

4 TOWARD A DYNAMIC INTERPRETATION

In the evolution of urban systems, major regularities can be observed. The most important feature, at the macroscopic scale of the whole system of cities, is the persistency of the spatial and hierarchical configuration of the system: despite of the growth of urban population and of the increases in the number and sizes of cities, the general spatial pattern and the statistical form of the size distribution of cities is preserved over time. The dynamics of a system of cities can then be formalized in using analogies with other open systems situated "far of equilibrium" and whose spatio-temporal dynamic properties have been investigated in physical sciences.

In a most formal and parsimonious way, urban systems can be conceptualized as hierarchical open systems (Prigogine, Stengers, 1979). They are hierarchical, in the sense that they can be described at three significant levels: the elements are the individuals, the urban residents (urban households could also be used to define that elementary level); the subsystems are the individual cities, which are defined only by the number of elements that they contain, and this measure of their size strongly differentiate them (usually from a few thousands to several millions of people); the whole system is the total set of cities, where exchanges of elements between subsystems are taking place. Urban systems are also open systems, since they exchange individuals with their environment, for instance migrants to and from rural areas or other countries.

Such a theory is not as poor a description, as it could seem at first sight, for an urban system. In a most general sense, a city can be defined as a permanent agglomeration of many people over a limited space. Of course, this definition is too simple and a minimum size of the aggregate and of its density, which may not always be the same, has also to be considered. On the one hand, when describing a city simply by its number of inhabitants, it becomes less unreasonable to compare cities through centuries and from one civilization to another. On the other hand, the population size indeed is a very synthetic description of the relative importance of a city as compared to the others, since it is very often rather well correlated with many other social, economic and cultural indicators. It is very often considered as the first "dimension", that is the most important differentiation factor among a set of cities (Parr, Reiner, 1980). A dynamic theory of urban size in urban systems can then induce a progress in the understanding of the whole logic of the urbanization process.

A dynamic model could consider the shape of the urban hierarchy as an **attractor** for the state of any urban system. The model would show by simulation how is this **structural stability** compatible with the permanent **fluctuations** in the size of the individual cities at the meso-level, and with the even more rapid changes among the individuals who compose the urban population at a micro-level. A dynamic theory of the urban systems should then connect the **slow dynamic** of their **city-size distribution** with the more **rapid dynamic** of the **individual cities competing for attracting the**

populations. But it should also integrate a few specific properties of the urban systems, which are both geographical and historical objects.

4.1 A stochastic model

An evolutionary model of urban systems has to be a stochastic one, and should incorporate at first the Gibrat's hypothesis. Why does urban growth follow the two fundamental hypothesis of the process described by Gibrat? It is possible to understand why cities grow in average at the same speed: as they are all connected in the same system, they are all submitted to the same general conditions (demographic and economic growth, socio-economic and technical innovations...) which are responsible for the urbanization process. The variety of local conditions and the slight differences in the timing of the diffusion of innovations may explain the quasi random fluctuations in the growth rates of individual cities.

What remains to be explained is why urban growth is in average proportional to urban size. Urban growth has two components, natural growth and net migration. The natural growth of populations is always multiplicative, proportional to their initial size: when there is no limitations of resources, it is exponential, otherwise it follows a logistic curve. It would require a very peculiar segregation of the age groups among cities, according to their size (for instance all young people in small towns and old ones in the largest cities) before this rule is changed. Net migration is also roughly proportional to city size, as long as the inmigration and outmigration flows are also proportional to that size. This is normally the case, as it is described by the analogic models of gravitation which are used most of the time to adjust the observations of migratory flows or to predict them. According to spatial migration theory (Courgeau, 1984), interaction flows between various places are proportional to the probability of contacts between their residents, and this probability is a function of the size of the places.

It is then the multiplicative aspect of the spatial process of the agglomeration of population which has to be explained, in order to find out a proper explanation of the size distribution of cities. The main characteristic of the urban systems is their hierachic structure, which reveals a principle of ordering in the system at the macro-level; it can be interpreted as resulting from a specific process of interactions between the individuals at the micro-level, which define the spatial variations in their reproductive and migratory behaviour. There is no explicit intention in the behaviour of the individuals regarding the city-size distribution. It is in this sense that urban systems can be described as self-organized systems. If one or both of the behavioral features of urban populations would change from what they have used to be during the centuries of slow and then rapid urbanization, or if they cease to be rather homogeneous over space, the evolution and the structure of the urban systems could become very different.

Simulations with a dynamic model including explicitly such individual behaviour parameters should allow, among other results, to estimate the time scale which would be needed to alter in a significant manner the configuration of urban systems.

4.2 Non-stochastic aspects of urban growth

However, systematic deviations from the stochastic process described by Gibrat were observed: first is the trend to increasing concentration of city sizes due to a correlation between growth rates and size of cities; a second type of deviations from a random process is less frequent and seems less continuous over time but may have stronger effects on the geographical sturcture of the system: city growth rates are not independent from one time interval to the other, they exhibit a slight temporal autocorrelation.

Both features seem related to specific properties of urban systems. They can be interpreted in the following way: if urban growth was a perfect competition process in a general context of spatial diffusion of innovations, it would lead only to random fluctuations and in average the growth rate of cities would be the same. This explains the large relevance of a stochastic process like the one defined by Gibrat. But the process of urban change is not perfectly ubiquitous as one would expect from a simple theory of the diffusion of innovations in space. The hierarchical diffusion process gives an initial advantage to the largest cities. Two effects can explain the advantage of the largest cities. First their large size increases the probability of appearance and use of innovations, and especially of linked innovations. Second, because of their size, they have more interactions in the urban network and can capture the innovations which come from elsewhere.

This explains the systematic advantage of the largest cities for urban growth, at least at the beginning of the cycles of the diffusion of innovations. The stages during which smaller cities grow faster may be regarded as periods of the end of such a process (Robson, 1973).

There is a third reason for large cities having an advantage. Despite of the general growing trend, cities whose share in the total urban population is decreasing are in average smaller than the growing ones. Their geographical distribution is rather regular, dispersed all over the country, in all regions, most of the times in the intervals between largest and more rapidly growing cities. This can be interpreted as an effect of the shrinking of space (Janelle,1969, Juillard, 1970). Within the course of time, the speed of the communications is increasing, the range of the largest cities is extended in such a way that the smaller cities are short-circuited, even for low level functions. As interurban competition is taking place in a shrinking spatio-temporal framework, the urban hierarchy tends to become always more and more contrasted (Guérin-Pace, 1990). This process is not continuous either. By using a dynamic theory of Pareto distribution, P. Frankhauser (1991) was able to show that, in the case of France, a reversal could be detected in the general trend towards a higher urban concentration, well before the turn-around appeared in the slope of the rank-size distribution (figure 10).

All these observations should lead to a reformulation of a dynamic model including hypothesis which would take into account the specific nature of urban systems. Whereas the "law of proportional effect" can be seen as a very general process for the evolution of any hierachic open system, trivial since just explaining the Paretian structure, the **selectivity** in the growth process, according to the size or to other properties of the cities (for instance, when new specializations are created in the system of cities), appears now as a specific feature of the evolution of the urban systems. This can no longer be considered as a residual in a stochastic theory of the growth of cities, but that it has to be integrated in a dynamic model.

CONCLUSION

Either formalized by a lognormal or by a Pareto distribution, the rank-size rule is an over-identified model for which many non contradictory explanations can be found. This model is valid not only for city sizes but for the whole settlement system. Therefore a larger theory has to be elaborated for describing the process of spatial aggregation of population. The hypothesis of Gibrat's model are a good starting point but for a macroscopic description only.

If a dynamic model of city sizes in urban systems is to be conceived within the framework of self-organizing systems theory, it will derive the macro-level hierarchic structure of the system from a micro-level process of interactions between the individuals. Even if this interactive process is reduced to hypothesis concerning the natural growth and the migratory behaviour of the individuals, this process cannot be defined independently of the previous state of the system: the behaviour of individuals is influenced by the size of the settlements and by the relative location of cities. The advantage which is given to the largest cities is not only a matter of size but also of the shrinking-space trend in the spatial organization of the system.

This may give to the urban systems a large inertia, which is also a characteristic of their dynamics and can be explained by the fact that a city is not only a grouping of inhabitants, but also has a material, built-up structure. This "concrete" inscription of the past processes of accumulation of wealth and values, slow down the momentaneous trends which would tend to change the ordering of cities in the urban hierarchy. Few major bifurcations may be expected in the historic trajectory of urban systems. It may be that any modelling using only a conceptualization of cities in terms of aggregation of people will proove irrelevant.

This kind of formalization could however provide measures of the degree of freedom of local and national governments when they pretend to intervene in the distribution of population over a territory. We should answer such questions as: is an hierarchical organization of city sizes the only possible state for any urban system? Is the trend towards a higher concentration of urban population unavoidable? Dynamic modelling of urban systems would help us to explore and to measure the "costs" of other possible solutions. Experiments with computer simulations are (almost) painless for the inhabitants!

References

Allen P. M. 1978: Dynamique des centres urbains. Sciences et Techniques, 50, 15-19

Auerbach F. 1913 : Das Gesetz der Bevölkerungskonzentration. Petermans Mitteilungen, 59, 1, 74-76.

Bairoch P. 1985 : De Jéricho à Mexico. Paris, Gallimard, 708 p.

Bairoch P., Batou J., Chèvre P. 1989 :La population des villes européennes de 800 à 1850. Genève, Droz, 303 p.

Baker A.R.H. 1969: Reversal of the rank-size rule: some nineteenth century rural settlement sizes in France. RThe Professional Geographer, 386-392.

Beckmann M.J. 1958: City hierarchies and the distribution of city size. Economic Development and Cultural Change, 6, 243-248.

Beguin H. 1979: Urban hierarchy and the rank-size distribution. Geographical Analysis, 11, 2, 149-163.

Berry B.J.L. 1961 : City Size distribution and Economic Development. Economic Development and Cultural Change, 9, 571-587.

Berry B.J.L. Garrison W.L. 1958: Alternate explanations of urban rank-size relationships. Annals of the Association of American Geographers, 1, 83-91.

Champion A.G. 1989: Counterurbanization. London, Arnold, 266 p.

Chapman G.P. 1970: The application of information theory to the analysis of population distribution in space. economic Geography, 2, 317-333.

Christaller W. 1933: Die Zentralen Orte in Süddeutschland. Jena, G. Fisher.

Courgeau D. 1980: Analyse quantitative des migrations humaines. Paris, Masson.

Cuervo M. 1989: Primatie et développement économique en Amérique latine. Univeristé Paris-Val de Marne, thèse de doctorat.

Curry L. 1964: The random spatial economy: an exploration in settlement theory. Annals of the Association of American Geographers, 138-146.

De Vries J. 1984: European Urbanization, 1500-1800. London and Cambridge, (Mass.).

De Vries J. 1990: Problems in the Measurement, Description, and Analysis of Historical Urbanzation, in Van der Woude A.D. De Vries J., Hayami A. eds., Urbanization History, Oxford, Clarendon Press, 43-60.

Duby G. ed. 1986 : Histoire de la France urbaine. Paris, Seuil.

Fletcher R. 1986: Settlement archaeology: world-wide comparisons. World Archaeology, 18, 1, 59-83.

Frankhauser P. 1990: Beschreibung der Evolution urbaner Systeme mit der Mastergleichung. Universität Stuttgart, Theoretische Physik, Doktorthesis.

Gibrat R. 1931 : Les inégalités économiques. Paris, Sirey.

Grossman D. 1982: Northern Samaria: a process-pattern analysis of rural settlement. Canadian Geographer, 110-127.

Guérin-Pace F. 1990: Analyse et modélisation du développement spatial d'un système de villes. Université Paris VII, thèse de doctorat.

Guérin-Pace F., Pumain D. 1990: 150 ans de croissance des villes françaises. Economie et Statistique, n°.230.

Haran E.G.P. Vining D.R. 1973: On the implications of a stationary urban population for the size distribution of cities. Geographical Analysis, 5, 4, 296-308.

Janelle D.G., 1969 : Spatial reorganization : a model and concept. Annals of the Association of American Geographers, 348-368.

Juillard E., 1970 : L'armature urbaine de la France préindustrielle. Bulletin de la Faculté des Lettres de Strasbourg, n° 6.

Le Maître A. 1682: La Métropolitée, ou de l'Etablissement des villes capitales, de leur utilité passive ou active, de l'union de leurs parties et de leur anatomie, de leur commerce etc... Amsterdam, B. Boekholt.

Lepetit B., 1988 : Les villes de la France moderne. Paris, Albin Michel.

Moore F.T. 1958: A note on city size distributions. Economic Development and Cultural Change, vol.7, n°1, 17-37.

Morrill R. 1965: Migration and the Spread and Growth of Urban Settlement. Lund Studies in Geography, serie B, n°26.

Parr J.B. 1970: Models of city size in an urban system. Papers and Proceedings of the Regional Science Association, 25, 221-253.

Parr J.B. 1985: A note of the size distribution of cities over time. Journal of Regional Economics, vol.18, 99-212.

Parr J.B., Reiner T.A., 1980 : A note on the dimensions of a national urban settlement system. Urban Studies, 2, 223-230.

Pred A.R., 1966 : The spatial dynamics of US Urban-Industrial Growth, 1800-1914: Interpretive and Theoretical Essays. Cambridge, M.I.T. Press.

Pred A.R., 1977 : City Systems in Advanced Economies. London, Hutchinson, 256 p.

Prigogine I. Stengers I. 1979: La nouvelle alliance. Paris, Gallimard.

Pumain D., 1982 : La dynamique des villes. Paris, Economica, 231 p.

Pumain D., 1984 : L'évolution séculaire de la trame urbaine française, in Quant T. ed., Géoscopie de la France, Paris, Minard, 153-174.

Pumain D. Saint-Julien T. 1978: Les dimensions du changement urbain. Paris, CNRS.

Pumain D., Sanders L., Saint-Julien Th., 1989 : Villes et auto-organisation. Paris, Economica, 191 p.

Quandt R.E. 1964: Statistical discrimination among alternate hypothesis and some economic regularities. Journal of Regional Science, 5, 2, 1-23.

Renfrew A.C. Cherry J. 1987: Peer-polity Interaction. Cambridge University Press.

Reynaud J. 1841: Villes, article de l'Encyclopédie Nouvelle, repris dans Pinchemel Ph., Robic M.C., Tissier J.L. ed.: Deux siècles de géographie française, Paris, CTHS, 1984, 29-35.

Robic M.C. 1982: Cent ans avant Christaller, une théorie des lieux centraux. L'Espace Géographique, 1, 5-12.

Robson B. 1973: Urban Growth, an approach. London, Methuen.

Rosen K.T. Resnick M. 1980: The size distribution of cities: an examination of Pareto Law and Primacy. Journal or Urban Economics, 8, 165-186.

Russell J.C. 1972: Medieval regions and their cities. Bloomington, Indiana.

Sheppard E. 1982: City size distribution and spatial economic change. International Regional Science Review, vol. 7, 127-151.

Simon H. 1955: On a class of skew distributions. Biometrika, 42, 425-440.

Smith C.A. 1982: Modern and Premodern Urban Primacy. Comparative Urban Research, 11, 80.

Smith C.A. 1990: Types of city-size distributions: A comparative analysis, in Van der Woude A.D. De Vries J., Hayami A. eds., Urbanization History, Oxford, Clarendon Press, 20-42.

Sonis M. Grossman D. 1984: Rank-size rule for rural settlements. Socio-Econ. Plan. Sci., 18, 6, 373-380.

Spence N. Gillepsie A. Goddard J. et al., 1981: British cities, an analysis of urban change. Urban and Regional Planning Series, vol. 26.

Steindl J. 1965: Random processes and the growth of firms. New York, Hofnur, 249 p.

Suarez-Villa L. 1988: Metropolitan Evolution, Sectoral Economic Change, and the City-Size Distribution. Urban Studies, 25, 1, 1-20.

Unwin P.T.H. 1981: The rank-size distribution of medieval English taxation hierarchies in Nottinghamshire. The Professional Geographer, 350-360.

Van der Leeuw S.E. 1990: Rythmes temporels, espaces naturels et espaces vécus, in Archéologie et Espace, J.L. Fiches et S.E. van der Leeuw eds., Antibes, A.P.D.C.A., 299-346.

Vapnarski C. 1975: The Argentine System of Cities, in Hardoy J. ed., Urbanization in Latin America.

Vlora N.R. 1979: Città e territorio. Bologna, Pàtron, 183 p.

Zipf G.K. 1949: Human behaviour and the principle of last effort. Cambridge (Mass.), Addison Wesley.

Abstract:

Collections of archaeological as well as of historical empirical data lead to held for true that all settlements systems are characterized by their hierarchical organization. The traditional explanations of this property are reviewed. The results of many empirical testings suggest however that an evolutionary theory of the historical process of spatial aggregation of population should be more meaningfully elaborated. It could be a starting point for developing dynamic modelling and simulations of the evolution of urban hierarchies.

Captions for the figures

Fig. 1 Lognormal distribution of rural settlements

Source: Baker A.R.H., 1969 p.390

Baker's data on gausso-logarithmic plot

Fig. 2 Distribution of rural settlements in Samaria after Sonis and Grossman (1984)

Figure 3 Two lognormal distributions for all French settlements (1990)

Figure 4 Settlement size distribution in archaeological sites after R. Fletscher (1986)

Figure 5 "Convex" archaeological and historical rank-size distributions after G.A. Johnson (1980)

Figure 6 Evolution of European city size distribution

Source: De Vries, 1984, p.256 and 265

Figure 7 Urban growth and city size

Source: Robson, 1973

Correlation between log of population size and growth rates

Correlation between successive growth rates

Source: Guérin-Pace, 1990

Figure 8 Evolution of urban concentration over time

Slope of rank-size distribution of French towns (agglomerations) with 2000 inhabitants or more 1831-1982

Source: Guérin-Pace, 1990

Figure 10 A dynamic interpretation of the Pareto exponent

Evolution of the residual variance of the adjustment of 850 French cities size distribution from 1831 to 1982 to a Pareto model

Evolution of an entropy index per inhabitant

$(S(R)/N = \ln N - 1/N \; (n_i \ln n_i))$

Source: P. Frankhauser, 1991, p.100, 103

Figure 1 Lognormal distribution of rural settlements

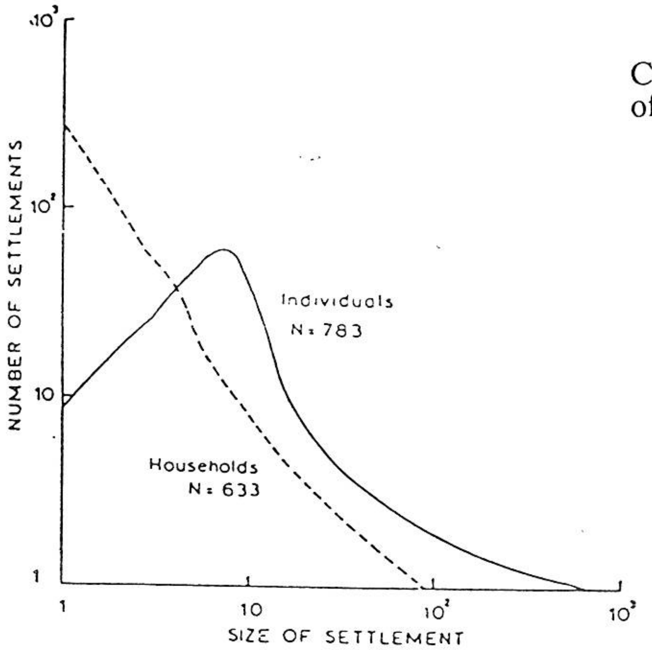
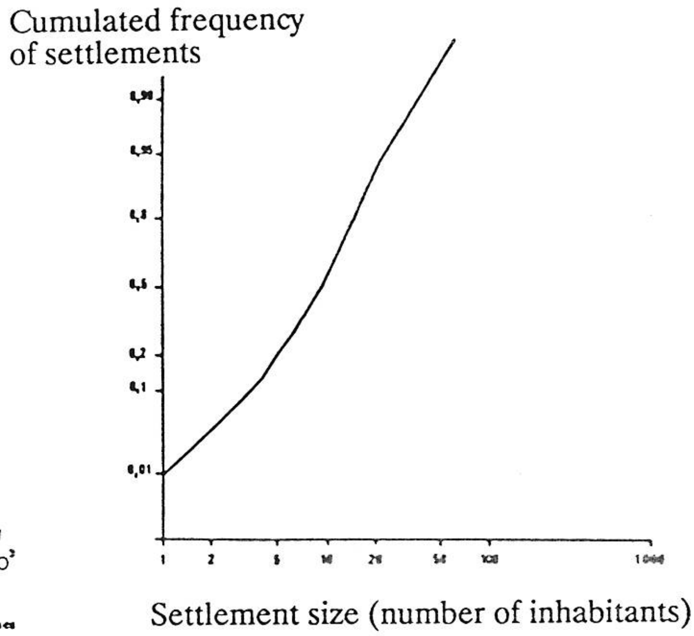


Fig. 3. Frequency distributions for settlement sizes of a random sample of communes on the plateau between the rivers Loire and Loir in 1846.

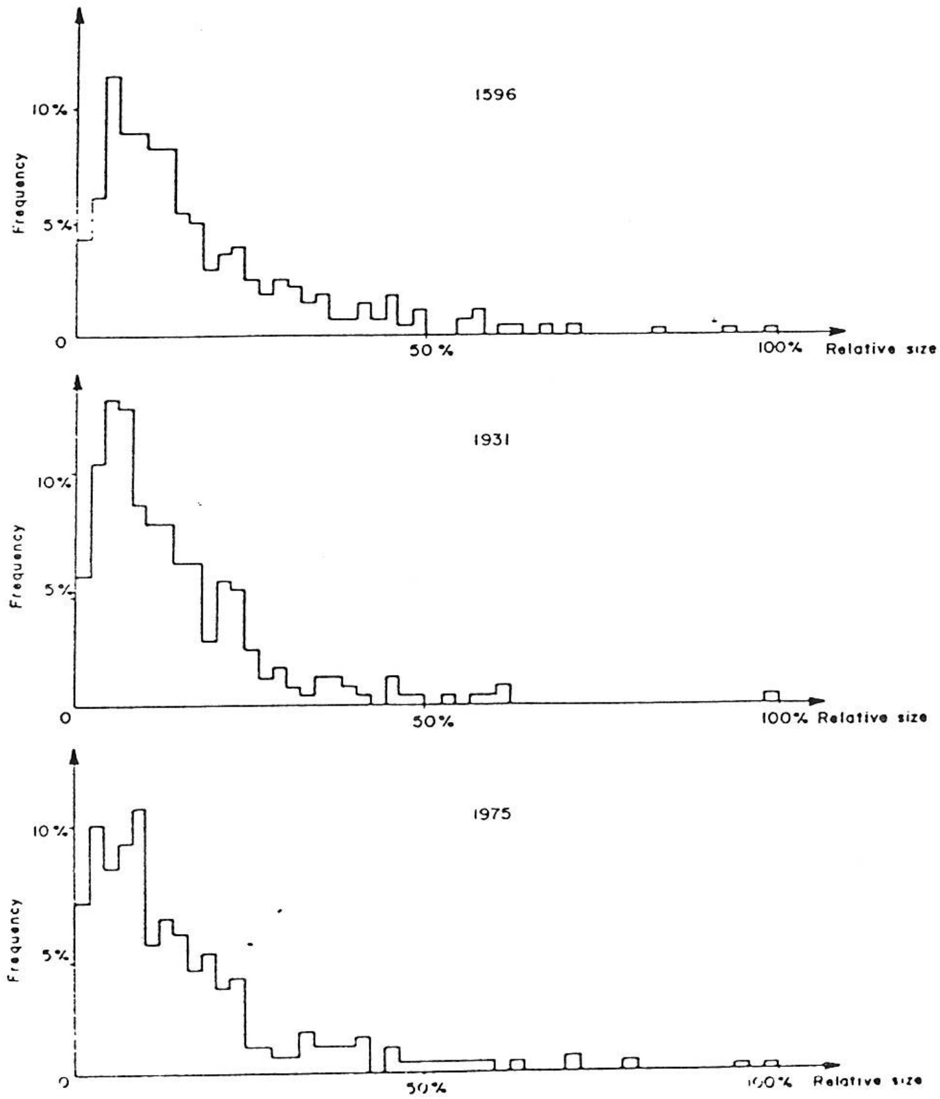


Source: Baker A.R.H., 1969, p. 390

Baker's data on gausso-logarithmic plot

Figure 2 Distribution of rural settlements in Samaria after Sonis and Grossman (1984)

Observations



Model

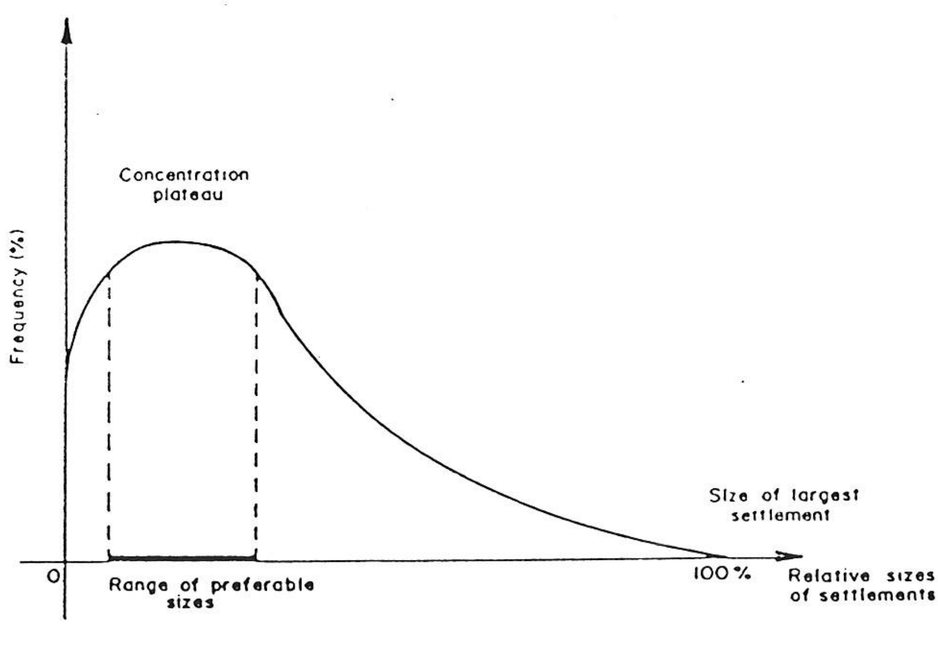


Figure 3 **Two lognormal distributions for all French settlements (1990)**

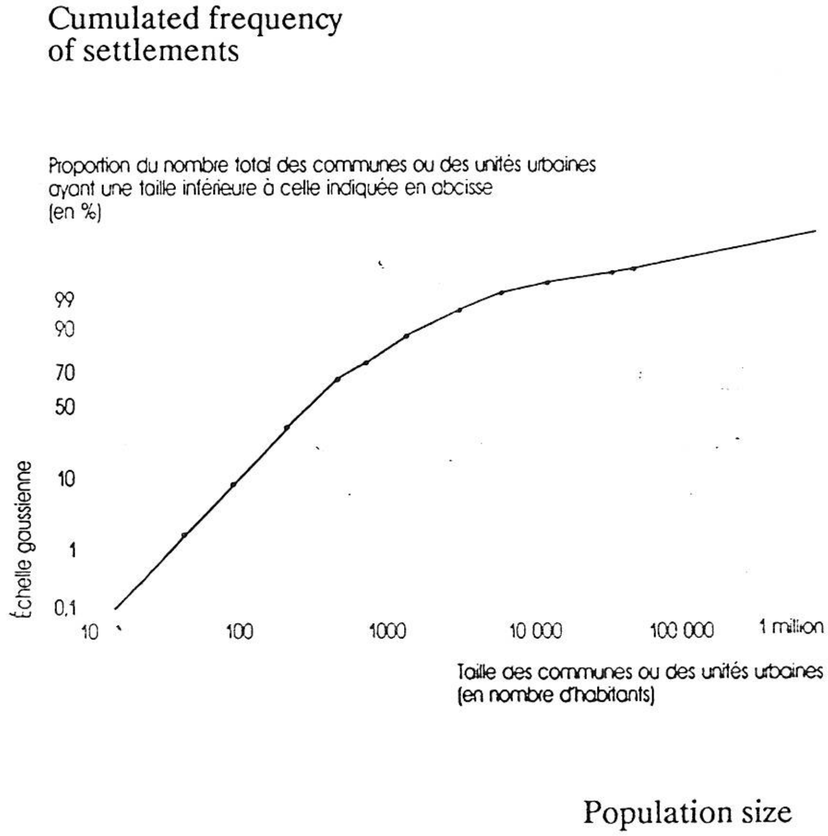
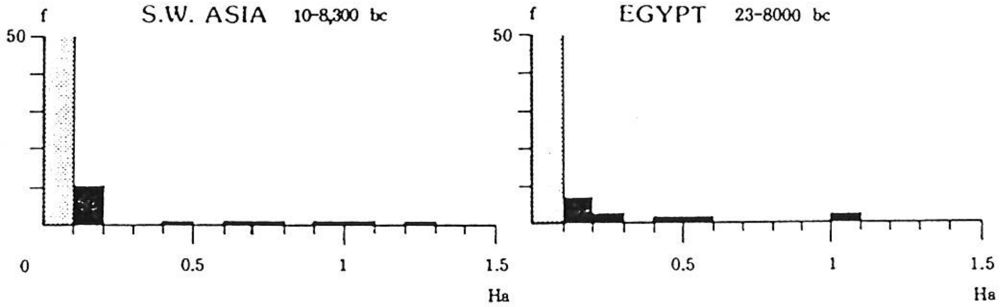
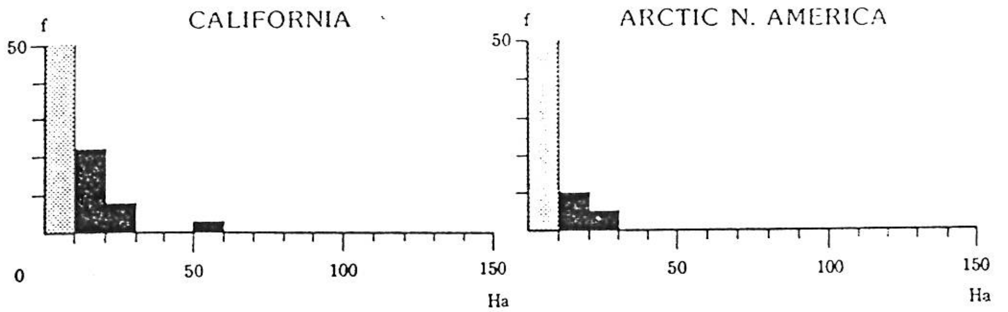


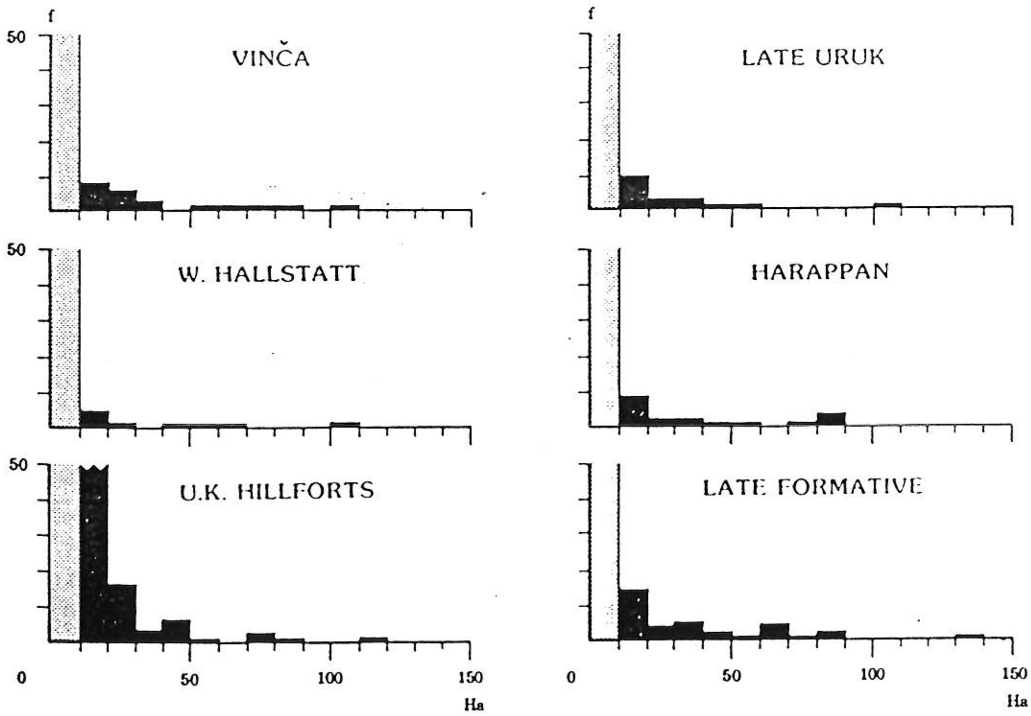
Figure 4 Settlement size distribution in archaeological sites after R. Fletcher (1986)



1 Mobile and semi-mobile communities. Archaeological sites, 0.1 to 1.5 ha. Medium <0.5 ha. Large < 1 ha. Very Large < 1.5 ha.

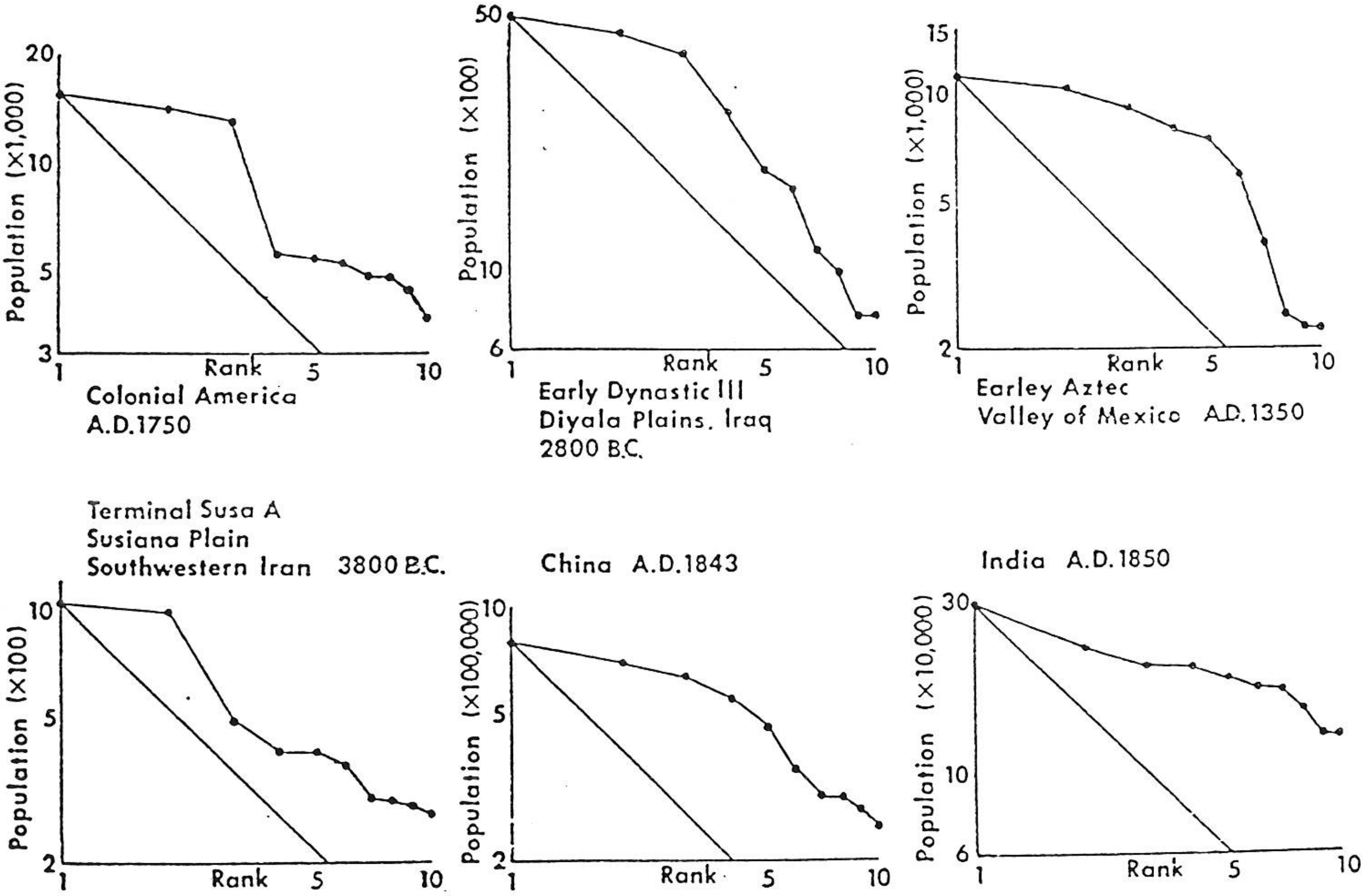


2 Mobile and semi-mobile communities. Archaeological sites, 10 to 150 ha. Medium <50 ha. Large < 100 ha. Very Large < 150 ha.



3 Sedentary communities. Archaeological sites, 10 to 150 ha. Medium <50 ha. Large <100 ha. Very Large <150 ha.

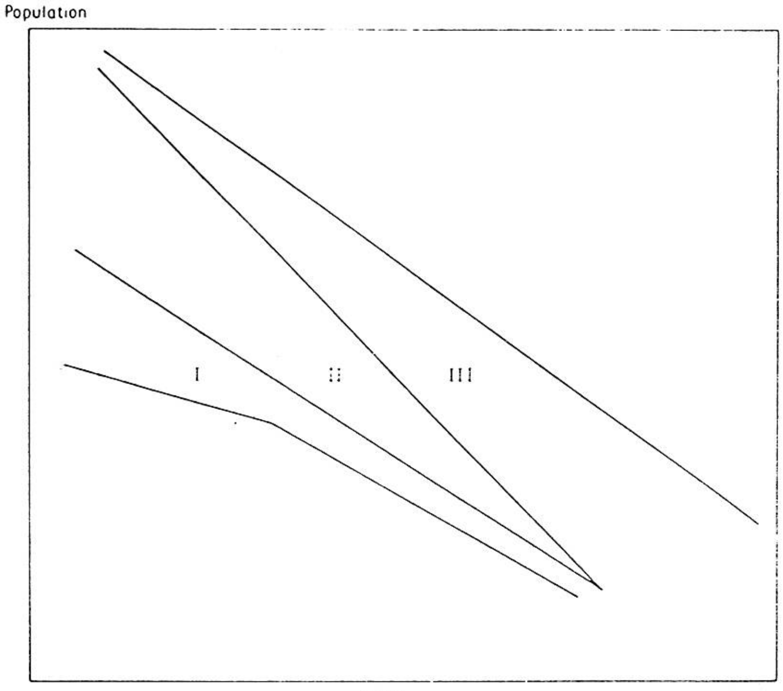
Figure 5 "Convex" archaeological and historical rank-size distributions after G.A. Johnson (1980)



Archaeological and historical convex rank-size distributions.
Data Sources: Colonial America [10]; Early Dynastic III [1]; Early Aztec [3]; Terminal Susa A [14]; China [28]; India [5].

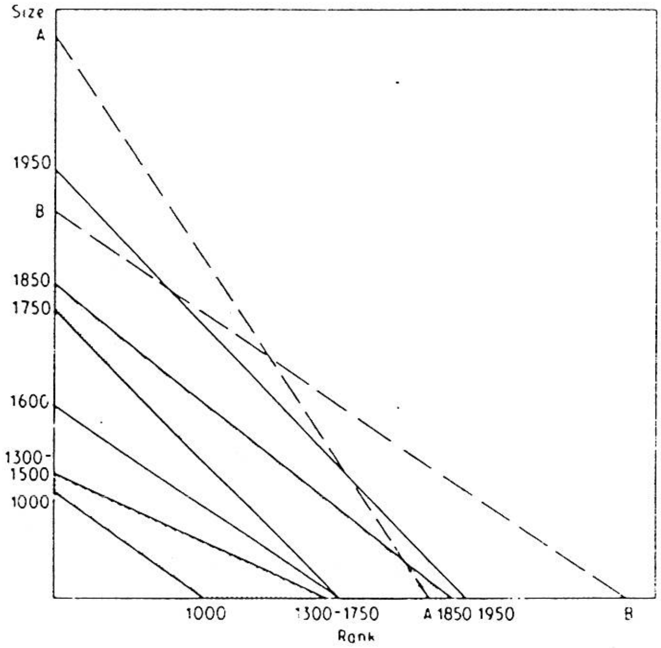
Figure 6 Evolution of European city size distribution

observation: longitunal study of European cities with 10 000 inhabitants or more (1500-1800)



Three phases of early modern urbanization

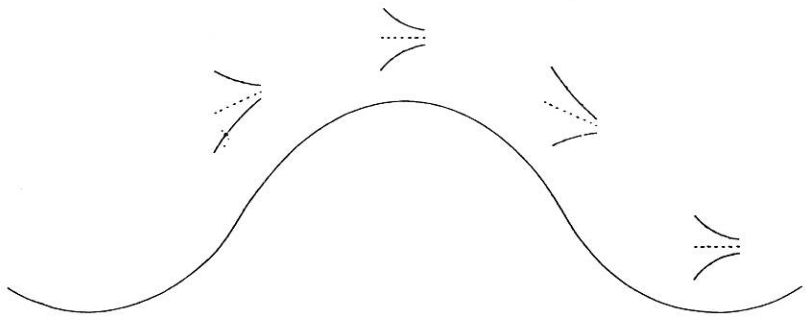
model of stages in urban development



Two modes of European urbanization, 1000–the future

Source: De Vries, 1984, p. 256 and 265

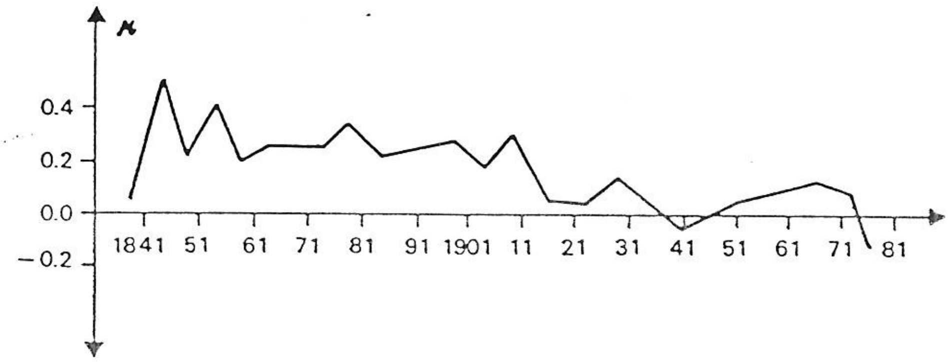
Figure 7 **Urban growth and city sizes**



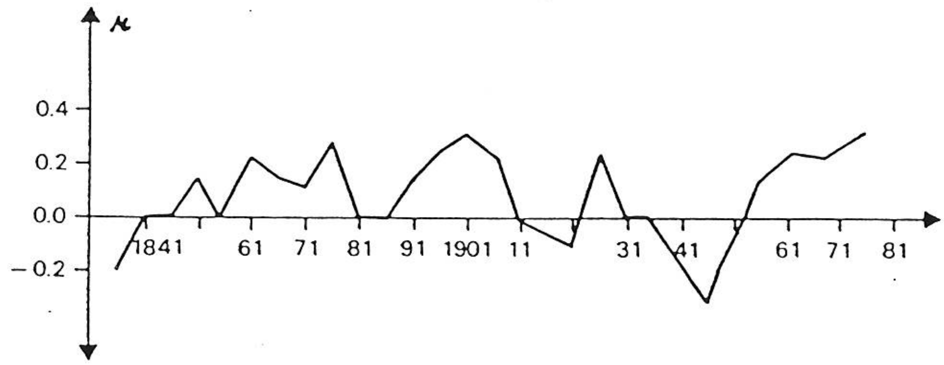
The pattern of urban growth and the stream of innovations over time. The solid line suggests an assumed curve of innovations over time in an economy

Source: Robson, 1973

CORRELATION BETWEEN LOG OF POPULATION SIZE AND GROWTH RATES



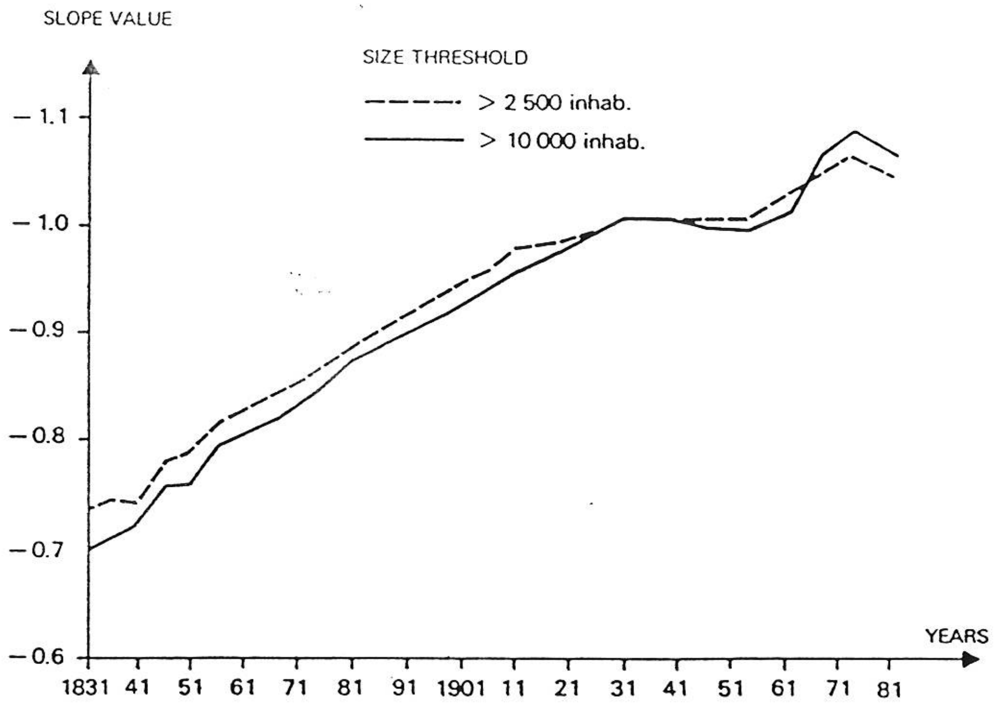
CORRELATION BETWEEN SUCCESSIVE GROWTH RATES



Source: Guérin-Pace, 1990

Figure 8 **Evolution of urban concentration over time**

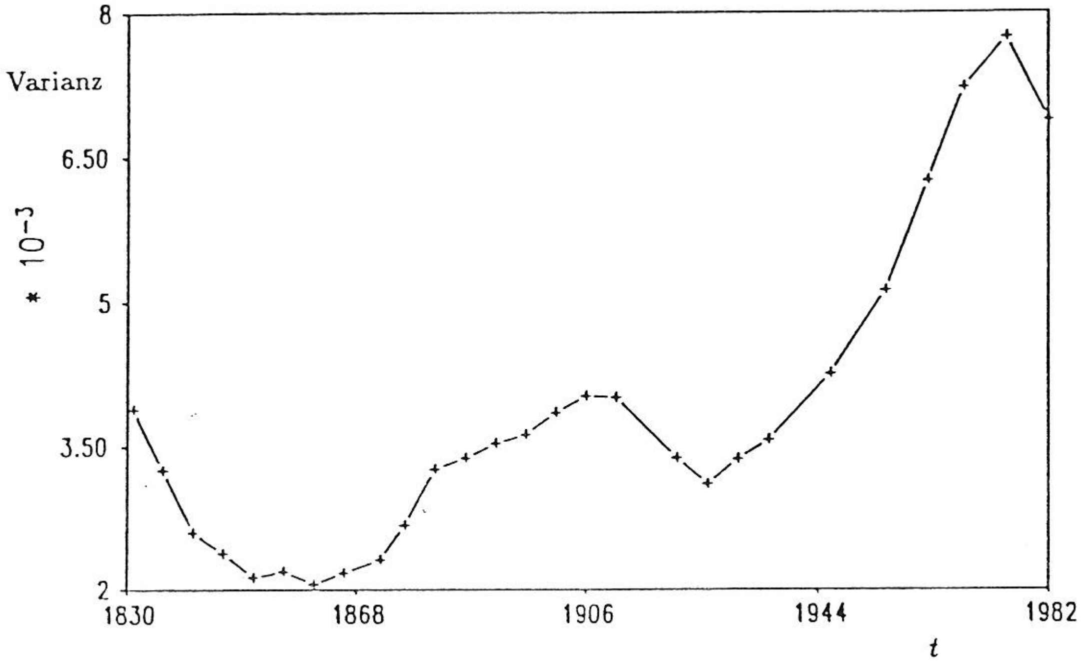
Slope of rank-size distribution of French cities with 2000 inhabitants or more 1831-1982 (a value in equation $P_i = k / r_i^a$)



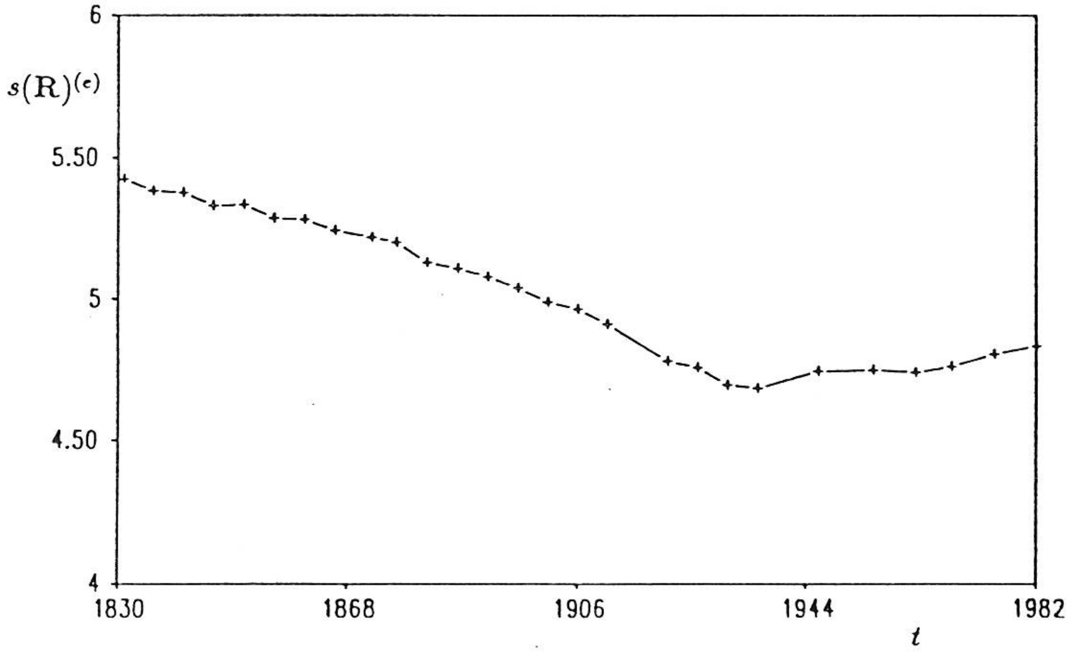
Source: Guérin-Pace, 1990

Figure 10 **A dynamic interpretation of the Pareto exponent**

Evolution of the residual variance of the adjustment of 850 French cities size distribution from 1831 to 1982 to a Pareto model



Evolution of an entropy index per inhabitant
$(S(R)/N = \ln N - 1/N \sum (n_i \ln n_i)$



Source: P. Frankhauser, 1991, p. 100, 103.