The evolution of the US urban structure from a long-run perspective (1900-2000)

González-Val, Rafael

30 November 2007

The evolution of the US urban structure from a long-run perspective (1900-2000)

Rafael González-Val

Department of Economic Analysis

Universidad de Zaragoza

Abstract: This paper analyses the evolution of the size distribution of cities in the United States throughout the 20^{th} century. In particular, we are interested in testing the fulfilment of two empirical regularities studied in urban economics: Zipf's law, which postulates that the product between rank and size of a population is constant, and Gibrat's law or the law of parallel growth, according to which the growth rate of a variable is independent of its initial size. For this parametrical and non-parametrical methods have been used. These laws have already been studied for the American case with the most populous cities or with MSAs. The main contribution of this work is the use of a new database with information on all the cities, thus covering the entire distribution. The results show that although if the sample is considered as a whole the fulfilment of Zipf's law is rejected, Gibrat's law is accepted for all the period considered.

Keywords: Zipf's law, Gibrat's law, city size distribution, urban growth.

JEL classification: R00, C14.

Address: Rafael González-Val,

Dpto. de Análisis Económico, Universidad de Zaragoza

Facultad de CC. Económicas y Empresariales

Gran Vía, 2, 50005 Zaragoza (Spain)

E-mail: rafaelg@unizar.es

1. Introduction

One of the stylized facts in urban economics is that the size distribution of cities in many countries can be approximated by a Pareto distribution, and it is an extensively studied empirical regularity that the parameter of this distribution, the Pareto exponent, is close to the unit[1]. This empirical regularity has produced theoretical developments which explain the fulfilment of Zipf's law, justifying it analytically, associating it directly with an equilibrium situation and relating it to parallel growth (Gibrat's law). Much of this theoretical literature takes as reference the case of the US, assuming an exponent equal to 1. The purpose of this work is to test if this supposition is true, studying the evolution of the size distribution of cities in the US during the 20th century.

Gabaix [1999] presents a model based on local random amenity shocks, independent and identically distributed, which through migrations between cities generate Zipf's law. The main contribution of the work is to justify the fulfilment of Zipf in that the cities in the upper tail of the distribution follow similar growth processes, that is, that the fulfilment of Gibrat's law involves Zipf's law. Thus, the explanation for the smaller cities' having a smaller Pareto exponent is that the variance of their growth rate is larger[2] (deviations from Zipf's law appear due to deviations in Gibrat's law). Córdoba [2004] concludes that under certain conditions Gibrat's law is not just *an* explanation of Zipf's law, but is *the (statistical)* explanation.

Rossi-Hansberg and Wright [2007] develop a model of urban growth which generates Zipf's law in two restrictive cases (when there is no physical capital and productivity shocks are permanent, or when city production is linear in physical capital and there is no human capital, depreciation is 100% and productivity shocks are temporary), and identifies the standard deviation of industrial productivity shocks as the key parameter which determines dispersion in the size distribution of cities. Eeckhout [2004] presents a model which also relates the migration of individuals between cities with productive shocks, obtaining as a result a lognormal and non-Paretian distribution of cities, although satisfying Gibrat's law. Duranton [2006] offers a model of urban economics with endogenous growth based on knowledge spillovers which in the stationary state reproduce Zipf's law for cities in the upper tail of the distribution; it also introduces some extensions which give empirically observed results (for example, a concave relationship between the rank and population logarithms).

To sum up, these theoretical models rest on local externalities, whether amenities or shocks in production or tastes, which must be randomly distributed independently of size, and identify deviation from Zipf's law with a distribution of these shocks which is not independent of size. Other works also show the empirical relevance of other variables distributed clearly heterogeneously, such as climate or geographical advantages (access to the sea, bridges, etc).

These theoretical developments arise in response to numerous empirical works which explore the relationship between the growth rate and Zipf's law. For the American case, Beeson et al. [2001] conclude, based on data on the population of counties, that during the period 1840-1990 there is a weak convergence in the population considering the entire sample, but if the counties of the western frontier are excluded, the conclusion is the opposite: divergence in the population.

Both Krugman [1996] and Gabaix [1999] use data from metropolitan areas from the Statistical Abstract of the United States and conclude that for 1991 Pareto's exponent is exactly equal to 1.005. This implies the fulfilment of Zipf's law for this specific year. For a dynamic analysis, Ioannides and Overman [2003] use data from metropolitan areas from 1900 to 1990 and arrive at the conclusion that Gibrat's law is fulfilled in the urban growth processes and that Zipf's law is also fulfilled approximately well for a wide range of city sizes. However, their results suggest that local values of Zipf's exponent can vary considerably with the size of cities. Nevertheless, Black and Henderson [2003] arrive at different conclusions for the same period (because they use different metropolitan areas). Zipf's law will be fulfilled only for cities in the upper third of the distribution, while Gibrat's would be rejected for any sample size. These results highlight the extreme sensitivity of conclusions to the geographical unit chosen and to sample size.

To close the debate, Eeckhout [2004] demonstrates that if we consider all the cities for the period 1990 to 2000 the city size distribution follows a lognormal rather than a Pareto distribution, so that the value of Zipf's parameter is not one, as earlier works concluded, but is slightly above $1/2$, as well as fulfilling Gibrat's law for the entire sample. The shortcoming of this work is that this is a short-term analysis, as only two decades are considered. The explanation is the data of only the last two decade are computerised. The aim of the present study is to generalise this analysis for all of the 20th century and extract long-term conclusions. For this it uses a new database which covers all the cities included in the ten-yearly census carried out by the US Census Bureau. Section 2 presents the database, sections 3 and 4 concern Zipf's and Gibrat's laws respectively, and section 5 concludes.

2. The Database

A very basic description of the territorial configuration of the United States is that the territory is divided into states, these are divided into counties, and finally, cities are the base of the system (although their relationship with counties is not always hierarchical; a city may belong to various counties). However, a slightly more detailed view shows that the reality is a little more complex. The US Census Bureau deals with several different types of geographical units. If our interest centres on analysing the evolution over time of the urban structure of the US, the first decision to make is which should be the unit of study. The literature usually chooses between two possibilities: cities or metropolitan areas.

What is a city?

We identify cities as what the US Census Bureau denominates as places. This generic denomination, since the 2000 census, includes all incorporated and unincorporated places.

The US Census Bureau uses the generic term incorporated place to refer to a type of governmental unit incorporated under state law as a city, town (except the New England states, New York, and Wisconsin), borough (except in Alaska and New York), or village and having legally prescribed limits, powers, and functions. On the other hand there are the unincorporated places (which were renamed Census Designated Places, CDPs, in 1980), which designate a statistical entity, defined for each decennial census according to Census Bureau guidelines, comprising a densely settled concentration of population that is not within an incorporated place, but is locally identified by a name. Evidently, the geographical limitation of unincorporated places may change if settlements move, so that the same unincorporated place may have different boundaries in different census.

They are the statistical counterpart of the incorporated places. The difference between them in most cases is merely political and/or administrative. Thus for example, due to a state law of Hawaii there are no incorporated places there; they are all unincorporated.

Metropolitan Areas

We begin by examining briefly the US Census Bureau's definition of metropolitan areas, those which could be considered "official" MAs, and then go on to see which MAs are most commonly used in the literature, which will not be the "official" ones due to the periodical changes in methodology.

The definition of a metropolitan area is from the Office of Management and Budget (OMB), based on data provided by the US Census Bureau. Standard definitions of metropolitan areas were first issued in 1949 by the then Bureau of the Budget (predecessor of OMB), under the designation "standard metropolitan area" (SMA). The term was changed to "standard metropolitan statistical area" (SMSA) in 1959, and to "metropolitan statistical area" (MSA) in 1983. The term "metropolitan area" (MA) was adopted in 1990 and referred collectively to metropolitan statistical areas (MSAs), consolidated metropolitan statistical areas (CMSAs), and primary metropolitan statistical areas (PMSAs). Finally, the term "core based statistical area" (CBSA) became effective in 2000 and refers collectively to metropolitan and micropolitan statistical areas.

Without entering into each definition[3], what interests us is the basic criterion used to define a MSA, as CMSAs and PMSAs are still MSAs which fulfil certain conditions. Thus, according to the OMB definition, qualification of an MSA requires the presence of a city with 50,000 or more inhabitants, or the presence of an urbanized area[4] and a total population of at least 100,000 (75,000 in New England). However, this criterion has changed over the course of the 20th century. Thus, the original criterion of 1950 only required the city of 50,000 inhabitants.

But leaving aside changes of denomination, what interests us is that there have also been changes in standards to define what is a metropolitan area in 1958, 1971, 1975, 1980, 1990 and 2000. Changes in the definitions of these statistical areas since the 1950 census have consisted chiefly of:

i. The recognition of new areas as they reached the minimum required city or urbanized area population.

ii. The addition of counties (or cities and towns in New England) to existing areas as new decennial census data showed them to qualify.

In some instances, formerly separate areas have been merged, components of an area have been transferred from one area to another, or components have been dropped from an area.

Because of these historical changes in geographic definitions, one has to be very cautious when comparing data from these statistical areas from different dates. The Census Bureau itself suggests that is some cases it may be preferable to maintain a consistent area definition through time, and provides historical metropolitan area definitions for 1999, 1993, 1990, 1983, 1981, 1973, 1970, 1963, 1960 and 1950.

Given the above, when carrying out an empirical study which requires temporal analysis there are three options:

1) Use a consistent yardstick, i.e., take the MSA definition from one year and apply it to all of them. The work of Bogue [1953] consists of this. He took the definitions of MSAs for 1950 and reconstructed the population of these areas for the period 1900-1940.

2) Use the areas as defined in each period; i.e., use the 1960 definitions for 1960, the 1970 definitions for 1970, and so on. This option presents problems, as a metropolitan area changes a good deal from decade to decade.

   This criterion is used in the works of Dobkins and Iaonnides [2000, 2001] and Iaonnides and Overman [2001, 2003]. For the period 1900-1950 they use the MSAs of Bogue [1953], and for the rest, the contemporary definitions of each period, with some adaptations.

3) Choose the geographical area which defines the metropolitan area and use it consistently. Thus, Black and Henderson [2003] built their database of metropolitan areas based on the situation in 1990, and applied the definition they created based on the county limits (approximately half the metropolitan areas consist of a single county) to earlier periods. That is, they took the geographical area which defined the MSA in 1990, examining the counties which made it up. To avoid inconsistencies, they eliminated from the sample the counties which had changed, building up what they defined as a set of "common denominator counties". But the number of MSAs was not fixed; they used a cut-off point, determined by a variable criterion which depended on the mean and minimum size.

In any case, any definition of metropolitan areas will be arbitrary to some degree. Ehrlich and Gyourko [2000] investigated a variety of possible definitions (considering up to four ways of measuring metropolitan areas) and tested their results for robustness between the different definitions and over time, using descriptive statistics (mainly distribution by deciles). This work shows that there are different definitions of MSA which are not mutually compatible (unlike places), which often include non-urban areas, and sometimes include various counties, and thus mix different administrative units. While consistency is obtained in the qualitative results – prior to the Second World War there is a clear skewing of the population share into the largest metropolitan areas, and following the Second World War the top decile of areas by population loses share (almost all of which is picked up by the next-largest decile) - perceptible quantitative differences are seen according to the method chosen.

Which geographical unit to choose?

The above has shown some of the possible characteristics of the two possible options, cities and MSAs. City boundaries respond to a legal and political limit, but may not reflect what a city is really in economic terms. Thus it is argued that MSAs represent better the job market and are more appropriate for analyzing the impact of large infrastructure projects. On the other hand, cities capture better the effect of local externalities (especially spillovers of human capital, which operate on a very local level), or the impact of taxes or education policies.

However, city boundaries change over time as their population grows, and some divide and others are annexed. But this drawback is also seen in MSAs, whose definitions change over time. Also, taking into account that the Census Bureau introduced them in 1950, if we want to carry out a long term analysis we will be obliged to reconstruct the areas for earlier years, and if we also want that analysis to be consistent, we will have to correct somehow the change of definitions for each period.

But there is a powerful reason for choosing cities: as Eeckhout [2004] demonstrates the use of MSAs bias the analysis from the start. A MSA may include several large cities; also, due to the minimum size restriction of at least 50,000 inhabitants, smaller cities are excluded from the analysis, which thus does not cover the entire size distribution of cities.

However, in practice most research done has been on data from MSAs. Few researchers choose cities, and those that do choose only the biggest. For example, Kim [2002, 2007] takes cities with populations over 25,000 and Dobkins and Iaonnides [2001] take cities with populations over 50,000. This is due to a practical reason, the availability of data. Also, the data of only the last two decades are computerized (1990 and 2000). To quote Eeckhout himself: *"Ideally…one would like to analyze the entire size distribution over time… Unfortunately, due to the lack of available data covering the entire size distribution, those further analyses are at this point in time not possible"*.

The data

Our base is the available data of all incorporated places for each decade of the 20th century[5]. Two details should be noted. First, that all the places corresponding to Alaska, Hawaii, and Puerto Rico for each decade are excluded, as these states were annexed during the 20th century (Alaska and Hawaii in 1959, and the special case of Puerto Rico, which was annexed in 1952 as a Free Associated State) and data do not exist for all periods. Their inclusion would produce geographical inconsistency in the samples, which would not be homogenous in geographical terms and thus could not be compared.

And second, the unincorporated places, which began to be accounted for from 1950. The US Census Bureau established size restrictions for their inclusion (except in 2000, when they were all counted). Although the overall criterion is usually that they have over a thousand inhabitants, there are differences in each decade. However, these settlements existed earlier and so their inclusion again faces us with a problem of sample inconsistency, as our proposal is to carry out a dynamic analysis. Also, only in the last period, in 2000, are they all included, so that even if we entered all the unincorporated places for which we have data each year, a bias would be introduced. The alternative, excluding them, is the only way to obtain homogenous samples for comparisons between decades and for robust conclusions[6]. As a result we decided to exclude unincorporated places from the sample in order to carry out a long term analysis of the 20th century with a homogenous sample. The results obtained will thus be conditioned by this choice of sample size. But bear in mind that the unincorporated places are known to be less populated settlements, so that the loss of representativeness of the sample is minimal.

Table 1 presents the number of cities for each decade and the percentage that the cities in the database represent of the total population of the US. Table 2 shows the descriptive statistics. The sample reflects the urbanization process which took place throughout the 20^{th} century. Thus the population of cities goes from less than half the total population of the US in 1900 (46.99%) to 61.49% in 2000. The number of cities increased by 82.11%, from 10596 in 1900 to 19296 in 2000. From the beginning of the century to 1930 there was a rapid increase both in the number of cities and in the percentage of the total population that they represent. This informs us of an urbanization process which has two manifestations: on one hand, already existing cities which are capable of attracting new population (the mean value of inhabitants per city grows over time, as can be seen in Table 2) and on the other, the growth in the number of cities. After this decade growth slows and stabilises at around 64% until the last decades (from 1970 to 2000) when it falls to 61.49%.

Also, a glance at the minimum values of each decade enables us to state that absolutely all incorporated places for which data exist are included, without size restrictions.

Table 3 shows both the mean growth rates for the whole period (g_p), calculated from gross growth rates, defined as $g_{it} = \frac{S_{it} - S_{it-1}}{S_{it-1}}$, where S_{it} is the population of the city i in the year t, and the annual mean growth rates (g_a), which are calculated from the mean growth rates for the whole period applying that $(1+g_a)^{10} = g_p$. It can be observed that indeed, the first decades of the century saw strong growth rates for city sizes. Between 1940 and 1980 the high growth rates seem to recover, and then fall in the last two decades. The two periods of lowest growth, 1930-1940 and 1980-1990, coincide with the two periods of lowest growth of the total population in United States history, 7.3% and 9.8% respectively[7].

3. Zipf's law

The aim of this work is to study the temporal evolution of the American city size distribution during the 20th century. For this we will use Pareto's distribution [1896] as a statistical approximation, also known as power law, originally used to study the distribution of incomes. If we use s to denote the relative size of the city[8] and R for its rank, a power law links relative size of the city and rank as follows:

$$R(s) = As^{-a}, \qquad (1)$$

where A and a are parameters. This expression is applied to the study of very varied phenomena, such as the distribution of the number of times different words appear in a book, the intensity of earthquakes or the flow of rivers. It has been used extensively in urban economics to study the city size distribution (see, for example, Eeckhout [2004] and Ioannides and Overman [2003] for the US case). It has also been used recently to study the country size distribution (Rose [2006]).

Zipf's law is an empirical regularity which appears when Pareto's exponent of the distribution is equal to the unit ($a = 1$). The term was coined after a work by Zipf [1949], which observed that the frequency of the words of any language is clearly defined in statistical terms by constant values. Or, applied to our variable, that ordered from largest to smallest, the relative size of the second city is half that of the first, the relative size of the third is a third of the first, and so on.

Parametric approach

The expression (1) of Pareto's distribution is usually estimated in its logarithmic version:

$$\ln R = K - a \ln s, \qquad (2)$$

where K is a constant.

It is useful to test if Pareto's parameter is more or less than the unit and the evolution of this coefficient in time, that is, for the different cross-sections available. Indeed, if the coefficient is greater (smaller) than the unit this would indicate that the relative sizes would tend to be more (less) homogenous regarding size. Also, if an evolution is growing (decreasing) in time we would have a process of convergence (divergence) as regards the size of cities.

Equation (2) can be represented as a graph. Figure 1 shows the Zipf plots which relate the logarithm of rank (y axis) with the logarithm of relative size (x axis) for the initial and final periods. The behaviour of other decades, which is not shown, is identical. If Zipf's law were fulfilled, the points would represent a decreasing straight line with a slope equal to minus one. However, a non-linear and clearly concave behaviour is observed.

Table 4 shows the results of the OLS estimation[9] of Pareto's exponent. The residues resulting from this regression usually present problems of heteroskedasticity, so to analyse the significance of the parameters the corrected standard error proposed by Gabaix and Ioannides [2004] is used: $\text{GI s.e.} = \hat{a} \cdot \left(2/N\right)^{1/2}$, where N is the sample size.

The results indicate that when the entire sample is taken, Pareto's exponent is always less than the unit and thus Zipf's law is not verified. Also, the estimations decrease over time, which would indicate that for the entire sample (including all the cities for each year) a divergent behaviour was produced. This divergence would be explained not so much by differences in the growth rate of cities but by the appearance of new cities which enter with very small relative sizes.

However, if we consider different cross-sections of the sample we can observe different behaviours. Thus, for the 1,000 biggest cities the exponent grows over time, so that we can state that for the biggest cities the trend has been convergence: they have become closer in relative size. For the 5,000 biggest cities the exponent remains stable, and from there the exponents decrease in time for different size samples.

It is also corroborated that as Eeckhout [2004] showed in theory, if the underlying distribution is lognormal the estimated value of Pareto's exponent depends negatively on the cut-off point, so that as we increase the sample size and include ever smaller cities, the estimated coefficient decreases (but not always; in principle, starting with a small sample and going on to a slightly larger one, as for example from 100 cities to 500, the coefficient can grow).

Non-parametrical approach

Once proven that Zipf's law is not fulfilled for the entire sample, we wonder what distribution best fits the data. For this, we estimate the empirical distribution of the data using an adaptive kernel[10].

Figure 2 shows the results for four representative decades. The graphs show the estimated density or probability in the y axis, and the logarithm of relative size in the x axis. The scale is the same in all graphs to enable comparison. It is observed that starting in 1900 from a very leptokurtic distribution with much density concentrated in the mean value of the distribution, its pointing increases until 1930. Starting from this decade, in which as we have seen in Table 3 there is slowdown in the growth of the urban population, the distribution loses kurtosis and concentration decreases until in 1970 it reaches a distribution very similar to lognormal, which it maintains until 2000. This evolution can also be seen in Figure 3, which shows the empirical cumulative density functions estimated for 1900 and for 2000. The y axis shows the accumulated probability, while the x axis shows the logarithm of relative size. It can be observed that in the year 2000 (the solid line) probability accumulates much more slowly than in 1900 (the dotted line), which indicates a change to a less concentrated distribution. Unfortunately, we must limit ourselves to a graphic examination of the distributions, as the normality tests are not robust for such high sample sizes. The most usual, Kolmogorov-Smirnov and Jarque-Bera, systematically reject the null hypothesis of normality for such large samples.

4. Gibrat's law

The above section has shown what we may consider to be a snapshot of the distribution of American cities during the 20th century. For each decade we obtained the graphic representation of the distribution and the estimated coefficients of Pareto's exponent for different sample sizes, which enabled us to conclude if there had been important variations in the distribution, or if concentration had increased or decreased. However, a more rigorous dynamic analysis demands that we work with growth rates. We are particularly interested in seeing if there is fulfilment of Gibrat's law or the law of parallel growth, which postulates that the growth of a variable is independent of its initial size; Gibrat [1931] observed that the size distribution (measured by sales or the number of employees) of companies tends to be lognormal, and his explanation was that the growth process of companies could be multiplicative and independent of the size of the company. It is interesting to test this over the entire 20th century from a long term perspective.

Parametric approach

The parametric approach consists of estimating growth regressions, which relate the growth rate with initial size (the ever popular β-convergence in economic growth). We take two specifications; in one growth depends on the initial relative size, while in the other the exogenous variable is a mean of the relative size of the two periods:

Specification I: $$\frac{s_{t+1}}{s_t} = K + a \cdot \frac{s_t + s_{t+1}}{2}, \qquad (3)$$

Specification II: $$\frac{s_{t+1}}{s_t} = K + a \cdot s_t. \qquad (4)$$

Note that the variable is the relative size, so that we are monitoring relative or effective growth, not gross growth. This means that the population of a city may have grown, but if others have grown more, the average rises and thus in terms of relative size it has shrunk. This can be seen from the decomposition of the ratio s_{it+1}/s_{it} :

$$\frac{s_{it+1}}{s_{it}} = \frac{\dfrac{S_{it+1}}{\bar{S}_{it+1}}}{\dfrac{S_{it}}{\bar{S}_{it}}} = \frac{\dfrac{S_{it+1}}{\sum_{i=1}^{N_{t+1}} S_{it+1} \Big/ N_{t+1}}}{\dfrac{S_{it}}{\sum_{i=1}^{N_t} S_{it} \Big/ N_t}} = \frac{S_{it+1}}{S_{it}} \cdot \frac{\sum_{i=1}^{N_t} S_{it}}{\sum_{i=1}^{N_{t+1}} S_{it+1}} \cdot \frac{N_{t+1}}{N_t}.$$

This means that relative growth can be produced not only by the increase in population of the city; it also happens if the number of cities rises or the total population of all the cities decreases.

Table 5 shows the results of the OLS estimations decade by decade and for a pool of the observations of the whole century. The conclusion is that the parameter $\hat{a}$ is not significant for any period with either of the specifications, which adds evidence in favour of Gibrat's law and the independence of growth in relationship to relative size. The only exception is the period 1980-1990, where the estimated coefficients are significant and positive (although very close to zero), which would indicate that a positive relationship existed between growth and size, with the largest cities gaining the most population. Remember that this is the period of least growth in urban population of the entire 20th century, about 1.69 % (Table 3), and the second lowest period of growth of the total population in the history of the United States, at 9.8%.

Non-parametrical approach

The earlier results confirm the fulfilment of Gibrat's law. However, Quah [1993] points out the problems of regressions towards the mean, so current in studies of economic growth, and proposes using non-parametric methods, specifically, transition matrices. We will use the methodology followed by Eeckhout [2004] and Ioannides and Overman [2003]. It consists of taking the following specification:

$$g_i = m(s_i) + \varepsilon_i, \qquad (5)$$

where g_i is the normalized growth rate[11] (subtracting the mean and dividing by the standard deviation) and s_i is the logarithm of relative size, and instead of making suppositions about the functional relationship of m and supposing a linear relationship, as in equations (3) and (4), $\hat{m}(s)$ is estimated as a local average around point s and is smoothed using a kernel, which is a symmetrical, weighted and continuous function around s.

In order to analyse the entire period 1890-2000 all the growth rates are taken between consecutive periods. And the Nadaraya-Watson method is used, exactly as it appears in Härdle [1990], based on the following expression[12]:

$$\hat{m}(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s - s_i) g_i}{n^{-1}\sum_{i=1}^{n} K_h(s - s_i)}, \qquad (6)$$

where K_h denotes the dependence of the kernel K (in this case an Epanechnikov) on the bandwidth h (0.5). Starting from this calculated mean $\hat{m}(s)$, the variance of the growth rate g_i is also estimated, again applying the Nadaraya-Watson estimator starting from:

$$\hat{\sigma}^2(s) = \frac{n^{-1}\sum_{i=1}^{n} K_h(s - s_i)(g_i - \hat{m}(s))^2}{n^{-1}\sum_{i=1}^{n} K_h(s - s_i)}. \qquad (7)$$

The estimator is very sensitive, both in mean and in variance, to atypical values. Thus, the growth rate, both in mean and in variance, of the smallest cities usually has much higher values than for the rest. If we examine the smallest 5% of cities the differences are even greater[13]. This is logical; we are considering cities of under 200 inhabitants, where the smallest increase in population is very large in percentage terms. For example, the value which distorts the mean and the variance in 1940-1950 is Pine Lake (De Kalb, Georgia), which goes from 2 inhabitants in 1940 to 566 in 1950. But we need not consider such extreme changes; any city with fewer than 50 inhabitants which sees some population growth increases a great deal in percentage terms. Thus, we decided to eliminate this 5% of the smallest distribution observations, as they are characterised by very high dispersion in mean and in variance, and they distort the results. This is not a great loss in terms of representativeness of the sample, as the size of the last city excluded is under 180 inhabitants.

Gibrat's law implies that growth is independent of size in mean and in variance. As growth rates are normalized, if Gibrat's law were strictly fulfilled and growth were independent of size, the estimated kernel would be a straight line on the zero value. Values different to zero involve deviations from the mean. And variance would also be a straight line, supposing that variance does not depend on the size of the variable analysed.

Figures 4 and 5 show the estimated kernels of growth and the variance of growth, respectively, for all the 20th century (a pool of 162,403 observations). It is noticeable that the estimation of growth is nearly a straight line around zero, meaning that as a mean we can accept that during the whole period growth was independent of size, and Gibrat's law was fulfilled. Regarding variance, Figure 5 shows that even if the smallest 5% of observations are eliminated, the smallest cities show greater variance than the rest of the sample. But it should be noted that starting from the zero value (being in a logarithmic scale, this corresponds to a size relative to 1, i.e., cities whose size is equal to the mean) variance begins to decrease, becoming much more homogenous, indicating that the variance of growth is independent of size for cities with a population equal to or greater than the mean (a little over 3,000 inhabitants at the beginning of the century and almost 9,000 at the end).

5. Conclusions

This work analyses the evolution of the urban structure of the United States during the entire 20th century, analysing whether two empirical regularities profusely studied in urban economics are fulfilled: Zipf's and Gibrat's laws. The main contribution consists of using a database of cities (understood as incorporated places), created from the census of what is now the US Census Bureau, which permits us to cover almost the entire city size distribution.

Regarding Zipf's law, the results indicate that when the entire sample is taken, Pareto's exponent is always less than the unit and thus Zipf's law is not verified. Also, the estimations decrease over time, which would indicate that including all the cities for each year a divergent behaviour was produced. This divergence would be explained not so much by differences in the growth rate of cities but by the appearance of new cities which enter with small relative sizes.

However, if we consider different sections of the sample we can observe different behaviours. Thus, for the 1,000 biggest cities the exponent grows over time, so that we can state that for them the trend has been convergence: they have become closer in relative size. For the 5,000 biggest cities the exponent remains stable, and from there the exponents decrease in time for different size samples.

It can also be observed that starting in 1900 with a leptokurtic distribution with a great deal of density concentrated in the mean value of the distribution, kurtosis is progressively lost and concentration decreases until in 1970 a distribution very similar to lognormal is reached, which is maintained until the year 2000.

Regarding Gibrat's law, the estimation of growth for the whole period (all the 20th century in the long term) shows that as a mean we can accept that during all the period growth was independent of size and Gibrat's law was fulfilled. And regarding variance, the smallest cities present greater variance than the rest of the sample, although it appears that for cities with a population equal to or more than the mean, variance is more homogenous and thus independent of size.

Appendix A: Normalized growth rates and the different measurements of city size

When growth rates are normalized, subtracting the mean and dividing by the standard deviation, the choice of measurement of size (size, relative size or share of the total) makes no difference, as it means only a change of scale.

If we take size $\left(S_{it}\right)$:

$$\frac{g_i - \overline{X}_g}{\sigma_g} = \frac{\left(\dfrac{S_{it+1} - S_{it}}{S_{it}}\right) - \dfrac{1}{N}\sum_i^N\left(\dfrac{S_{it+1} - S_{it}}{S_{it}}\right)}{\left[\dfrac{1}{N}\sum_i^N\left(\left(\dfrac{S_{it+1} - S_{it}}{S_{it}}\right) - \dfrac{1}{N}\sum_i^N\left(\dfrac{S_{it+1} - S_{it}}{S_{it}}\right)\right)^2\right]^{1/2}} = \frac{\left(\dfrac{S_{it+1}}{S_{it}}\right) - \dfrac{1}{N}\sum_i^N\left(\dfrac{S_{it+1}}{S_{it}}\right)}{\left[\dfrac{1}{N}\sum_i^N\left(\left(\dfrac{S_{it+1}}{S_{it}}\right) - \dfrac{1}{N}\sum_i^N\left(\dfrac{S_{it+1}}{S_{it}}\right)\right)^2\right]^{1/2}}$$

We arrive at the same expression taking the relative size, $s_{it} = \dfrac{S_{it}}{\overline{S}_t} = \dfrac{S_{it}}{\sum_{i=1}^{N_t} S_{it} \Big/ N_t}$:

$$\frac{g_{relative\_size_i} - \overline{X}_{g_{relative\_size}}}{\sigma_{g_{relative\_size}}} = \frac{\left(\dfrac{\dfrac{S_{it+1}}{\frac{1}{N_{t+1}}\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\frac{1}{N_t}\sum_i^{N_t} S_{it}}} - 1\right) - \dfrac{1}{N}\sum_i^N\left(\dfrac{\dfrac{S_{it+1}}{\frac{1}{N_{t+1}}\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\frac{1}{N_t}\sum_i^{N_t} S_{it}}} - 1\right)}{\left[\dfrac{1}{N}\sum_i^N\left(\left(\dfrac{\dfrac{S_{it+1}}{\frac{1}{N_{t+1}}\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\frac{1}{N_t}\sum_i^{N_t} S_{it}}} - 1\right) - \dfrac{1}{N}\sum_i^N\left(\dfrac{\dfrac{S_{it+1}}{\frac{1}{N_{t+1}}\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\frac{1}{N_t}\sum_i^{N_t} S_{it}}} - 1\right)\right)^2\right]^{1/2}} =$$

$$= \frac{\left(\dfrac{N_{t+1}}{N_t} \cdot \dfrac{\sum_i^{N_t} S_{it}}{\sum_i^{N_{t+1}} S_{it+1}} \right) \left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)}{\left[\dfrac{1}{N} \left(\dfrac{N_{t+1}}{N_t} \cdot \dfrac{\sum_i^{N_t} S_{it}}{\sum_i^{N_{t+1}} S_{it+1}} \right)^2 \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)^2 \right]^{1/2}} = \frac{\left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)}{\left[\dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)^2 \right]^{1/2}}$$

And also taking the share of the total, $S_{it} \Big/ \sum_{i=1}^{N_t} S_{it}$:

$$\frac{g_{Share_i} - \bar{X}_{g\,Share}}{\sigma_{g\,Share}} = \frac{\left(\dfrac{\dfrac{S_{it+1}}{\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\sum_i^{N_t} S_{it}}} - 1 \right) - \dfrac{1}{N} \sum_i^N \left(\dfrac{\dfrac{S_{it+1}}{\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\sum_i^{N_t} S_{it}}} - 1 \right)}{\left[\dfrac{1}{N} \sum_i^N \left(\left(\dfrac{\dfrac{S_{it+1}}{\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\sum_i^{N_t} S_{it}}} - 1 \right) - \dfrac{1}{N} \sum_i^N \left(\dfrac{\dfrac{S_{it+1}}{\sum_i^{N_{t+1}} S_{it+1}}}{\dfrac{S_{it}}{\sum_i^{N_t} S_{it}}} - 1 \right) \right)^2 \right]^{1/2}} =$$

$$= \frac{\left(\dfrac{\sum_i^{N_t} S_{it}}{\sum_i^{N_{t+1}} S_{it+1}} \right) \left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)}{\left[\dfrac{1}{N} \left(\dfrac{\sum_i^{N_t} S_{it}}{\sum_i^{N_{t+1}} S_{it+1}} \right)^2 \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)^2 \right]^{1/2}} = \frac{\left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)}{\left[\dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} - \dfrac{1}{N} \sum_i^N \left(\dfrac{S_{it+1}}{S_{it}} \right) \right)^2 \right]^{1/2}}.$$

Footnotes

[1] However, the Pareto exponent varies greatly between countries (Rosen and Resnick [1980], Soo [2005]). And recent works, Eeckhout [2004], demonstrate its sensitivity to the geographical unit chosen and the sample size.
[2] For the US, Glaeser and Shapiro [2001] study what factors influence the growth rate of American cities (cities of over 100,000 inhabitants and MSAs) using a very wide range of explicative variables (per capita income, average age of the residents, variables in the education level of individuals, temperature, distribution of employment among sectors, public spending per capita, etc.), attempting to capture the influence of local externalities. According to this work, the three most relevant variables would be human capital, climate and transport systems for individuals (public or private).
[3] These can be consulted at http://factfinder.census.gov - *American FactFinder Help*.
[4] An urbanized area (UA), according to the Census Bureau, consists of a central place(s) and adjacent territory with a general population density of at least 1,000 people per square mile of land area that together have a minimum residential population of at least 50,000 people.
[5] While the data of only the last two decades are computerized (US Bureau of the Census, County and City Data Book, Washington DC), the data corresponding to other decades is available in the original documents (US Bureau of the Census, Census of Population, Washington DC). We have created our database from these.
Source for 1900-1980: http://www.census.gov/prod/www/abs/decennial/
Source for 1990: http://www.census.gov/main/www/cen1990.html
Source for 2000: http://www.census.gov/main/www/cen2000.html
[6] Although we have carried out the exercise of using a sample of places which includes both the incorporated places and all the unincorporated places available for each decade. The results obtained would not vary significantly from those obtained with the sample of incorporated places.
[7] Source: http://www.census.gov/population/censusdata/table-4.pdf.
[8] In a long term temporal perspective of stationary equilibrium it is necessary to use a relative measure of size. The chosen measurement is the relative size, defined as: $s_{it} = \frac{S_{it}}{\bar{S}_t} = \frac{S_{it}}{\sum_{i=1}^{N_t} S_{it} \Big/ N_t}$. The other option most used in the literature is to take the share which represents the size of the city over the total population, $S_{it} \Big/ \sum_{i=1}^{N_t} S_{it}$. The results of this section are robust for the three options, size, relative size and share over the total, as the ratios involve only a change of scale.
[9] Gabaix and Ioannides [2004] show that the Hill (Maximum Likelihood) Estimator is more efficient if the underlying stochastic process is really a Pareto distribution. As we will see below, this is not the distribution that the data follow, and so we use the OLS estimator.
[10] The adaptive kernel density estimate is given by $\hat{f}(x) = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} \frac{w_i}{h_i} K\left(\frac{x - x_i}{h_i}\right)$, where the x_i's are the data points (associated with weights w_i), K is a kernel function, and $h_i = h \times \lambda_i$. The local bandwidth factors are proportional to the square root of the underlying density functions at the sample points: $\lambda_i = \lambda(x_i) = \left(G \big/ \tilde{f}(x_i)\right)^{0,5}$, where G is the geometrical mean over all i of the pilot density estimated $\tilde{f}(x)$. The pilot density estimate is a standard fixed bandwidth kernel density estimate obtained with h as bandwidth.
[11] Taking normalized growth rates will mean that the choice of the unit of measurement, size, size relative to the average or share of the total, is indifferent, as it means only a change of scale; the results regarding growth are robust. See Appendix A.
[12] The calculation was done with the KERNREG2 Stata module, developed by Nicholas J. Cox , Isaias H. Salgado-Ugarte, Makoto Shimizu and Toru Taniuchi, and available online at:
http://ideas.repec.org/c/boc/bocode/s372601.html.
This programme is based on the algorithm described by Härdle [1990] in Chapter 5.
[13] The specific values are available from the author upon request.

References

[1] Beeson, P.E., D.N. DeJong and W. Troesken, [2001]. Population Growth in US Counties, 1840-1990. Regional Science and Urban Economics, 31: 669-699.

[2] Black, D. and V. Henderson, [2003]. Urban evolution in the USA. Journal of Economic Geography 3, pp. 343-372.

[3] Bogue, D. [1953]. Population growth in Standard Metropolitan Areas: 1900-1950. Washington, DC: Housing and Home Finance Agency.

[4] Córdoba, J. C. [2004]. A generalized Gibrat's law for cities. Processed, Rice University.

[5] Dobkins L. and Y. M. Iaonnides, [2000]. Dynamic evolution of the US city size distribution. In: Huriot, J. M. and J. F. Thisse (Eds.), The economics of cities. Cambridge University Press, Cambridge, pp. 217-260.

[6] Dobkins L. and Y. M. Iaonnides, [2001]. Spatial interactions among US cities: 1900-1990. Regional Science and Urban Economics 31, 701-731.

[7] Duranton, G. [2006]. Some Foundations for Zipf´s Law: Product Proliferation and Local Spillovers, Regional Science and Urban Economics, vol. 36, pages 542-563.

[8] Eeckhout, J. [2004]. Gibrat's Law for (All) Cities, American Economic Review, American Economic Association, vol. 94(5), pages 1429-1451.

[9] Ehrlich, S. and J. Gyourko, [2000]. Changes in the scale and size distribution of US metropolitan areas during the twentieth century. Urban Studies, Vol. 37, No 7, 1063-1077

[10] Gabaix, X. [1999]. Zipf's law for cities: An explanation. Quarterly Journal of Economics, 114(3):739-767.

[11] Gabaix, X. and Y. M. Iaonnides, [2004]. The evolution of city size distributions. Handbook of urban and regional economics, Vol. 4, J. V. Henderson and J. F. Thisse, eds. Amsterdam: Elsevier Science, North-Holland, 2341-2378.

[12] Glaeser, E. L. and J. Shapiro [2001]. Is there a new urbanism? The growth of US cities in the 1990s. Harvard Institute of Economic Research, Discussion Paper number 1925.

[13] Gibrat, R. [1931]. Les inégalités économiques, Paris: Librairie du Recueil Sirey.

[14] Härdle, W. [1990]. Applied nonparametric regression. Econometric Society Monographs. Cambridge, New York and Melbourne: Cambridge University Press.

[15] Ioannides Y. M. and H. G. Overman [2001]. Cross-sectional evolution of the US city size distribution. Journal of Urban Economics 49, 543-566.

[16] Ioannides Y. M. and H. G. Overman [2003]. Zipf's law for cities: an empirical examination. Regional Science and Urban Economics 33, 127-137.

[17] Kim, S. [2002]. The Reconstruction of the American Urban Landscape in the Twentieth Century. NBER Working Papers 8857, National Bureau of Economic Research. http://www.nber.org/papers/w8857

[18] Kim, S. [2007]. Changes in the nature of urban spatial structure in the United States, 1890-2000. Journal of Regional Science, vol. 47, No. 2, pp.273-287.

[19] Krugman, P. [1996]. The Self-organizing economy. Cambridge: Blackwell, 1996.

[20] Pareto, V. [1896]. Cours d'Economie Politique. Geneva: Droz.

[21] Quah, D. T. [1993]. Galton's Fallacy and Tests of the Convergence Hypothesis, The Scandinavian Journal of Economics, Vol. 95, No. 4, Endogenous Growth (Dec., 1993), pp. 427-443.

[22] Rose, A. K. [2006]. Cities and countries, Journal of Money, Credit, and Banking, 38(8), 2225-2246.

[23] Rosen, K. T. and M. Resnick [1980]. The Size Distribution of Cities: An Examination of the Pareto Law and Primacy, Journal of Urban Economics, vol. 8, pages 165-186.

[24] Rossi-Hansberg, E. and M. L. J. Wright [2007]. Urban structure and growth. Review of Economic Studies 74, 597–624.

[25] Soo, K. T. [2005]. Zipf's Law for cities: a cross-country investigation. Regional Science and Urban Economics, 35: 239-263.

[26] Zipf, G. [1949]. Human Behaviour and the Principle of Least Effort, Cambridge, MA: Addison-Wesley.

Tables

Table 1. - Size of the database

Year	Cities	% of the total US population
1900	10596	46.99%
1910	14135	54.90%
1920	15481	58.62%
1930	16475	62.69%
1940	16729	63.75%
1950	17113	63.48%
1960	18051	64.51%
1970	18488	64.51%
1980	18923	61.78%
1990	19120	61.33%
2000	19296	61.49%

Excluding Alaska, Hawaii and Puerto Rico

Table 2. - Descriptive statistics of the sample

	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
N	10596	14135	15481	16475	16729	17113	18051	18488	18923	19120	19296
Mean	3376.04	3560.92	4014.81	4642.02	4975.67	5613.42	6408.75	7094.29	7395.64	7977.63	8968.44
Median	640	630	643	638	699	758	828	907	1019	1014	1093
Mode	200	276	300	279	226	217	153	235	139	196	86
Standard deviation	42323.896	49351.239	56781.645	67853.648	71299.371	76064.402	74737.618	75319.588	69167.914	71873.911	78014.749
Minimum	7	4	3	1	1	1	1	3	2	2	1
Maximum	3437202	4766883	5620048	6930446	7454995	7891957	7781984	7894862	7071639	7322564	8008278

Table 3. - Average growth rates of the sample

Period	N	Period mean	Annual mean
1890-1900	7531	31.43%	2.77%
1900-1910	10502	30.56%	2.70%
1910-1920	13578	18.78%	1.74%
1920-1930	15310	15.15%	1.42%
1930-1940	16211	10.39%	0.99%
1940-1950	16420	16.21%	1.51%
1950-1960	17075	20.79%	1.99%
1960-1970	17832	16.25%	1.52%
1970-1980	18321	19.13%	1.77%
1980-1990	18991	1.69%	0.17%
1990-2000	19179	11.80%	1.12%

Table 4. - Pareto coefficients estimated by decade

Truncation point N	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
100	1.050	1.086	1.095	1.101	1.095	1.096	1.167	1.201	1.271	1.313	1.320
	(0.148)	(0.153)	(0.154)	(0.155)	(0.154)	(0.154)	(0.165)	(0.169)	(0.179)	(0.185)	(0.186)
500	1.063	1.060	1.047	1.062	1.088	1.101	1.198	1.233	1.278	1.315	1.341
	(0.067)	(0.067)	(0.066)	(0.067)	(0.068)	(0.069)	(0.075)	(0.078)	(0.08)	(0.083)	(0.085)
1000	1.034	1.060	1.022	1.030	1.065	1.078	1.190	1.211	1.265	1.293	1.319
	(0.046)	(0.047)	(0.045)	(0.046)	(0.047)	(0.048)	(0.053)	(0.054)	(0.056)	(0.058)	(0.059)
5000	0.967	0.978	0.954	0.924	0.941	0.939	0.947	0.949	0.975	0.962	0.963
	(0.019)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)
10000	0.831	0.889	0.884	0.845	0.839	0.828	0.797	0.793	0.806	0.784	0.773
	(0.011)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
14000		0.770	0.810	0.785	0.773	0.752	0.716	0.709	0.719	0.695	0.683
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
15000			0.776	0.763	0.752	0.729	0.695	0.687	0.697	0.673	0.661
			(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)
16000				0.732	0.724	0.702	0.673	0.665	0.675	0.651	0.639
				(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
16700					0.683	0.676	0.656	0.647	0.658	0.634	0.623
					(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
17100						0.642	0.644	0.636	0.648	0.624	0.613
						(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
18100							0.599	0.600	0.617	0.595	0.585
							(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
18400								0.580	0.605	0.584	0.575
								(0.006)	(0.006)	(0.006)	(0.006)
18900									0.572	0.560	0.555
									(0.005)	(0.005)	(0.006)
19100										0.542	0.545
										(0.005)	(0.005)
19200											0.537
											(0.005)

(GI s.e.) Gabaix-Ioannides (2004) corrected standard error. All coefficients are significantly different from zero at the 0.05 level.

Table 5. - Estimated coefficients of parametric growth regressions

			Specification I		Specification II	
Initial year	Final year	N	$\hat{a}$	(s.e.)	$\hat{a}$	(s.e.)
1890	1900	7531	8.28E-04	7.17E-04	3.68E-04	9.72E-04
1900	1910	10502	4.43E-04	4.51E-04	6.62E-04	4.02E-04
1910	1920	13578	4.83E-04	3.50E-04	2.40E-04	3.64E-04
1920	1930	15310	3.14E-04	3.75E-04	5.86E-04	3.63E-04
1930	1940	16211	-1.03E-04	2.42E-04	-1.54E-04	2.40E-04
1940	1950	16420	1.73E-04	1.13E-03	1.37E-05	1.13E-03
1950	1960	17075	6.26E-04	6.66E-04	-1.38E-04	6.16E-04
1960	1970	17832	2.17E-04	7.29E-04	-3.46E-04	7.01E-04
1970	1980	18321	-7.22E-04	6.94E-04	-1.11E-03	6.49E-04
1980	1990	18991	1.07E-03*	3.38E-04	7.05E-04*	3.35E-04
1990	2000	19179	3.78E-04	4.13E-04	2.58E-05	4.20E-04
Pool	Pool	170950	3.43E-04	1.91E-04	3.81E-05	1.88E-04

* Significant coefficients for a confidence level of 95%

Figures

Figure 1. – Zipf plots (relationship between Rank (ln scale) and Relative Size (ln scale))

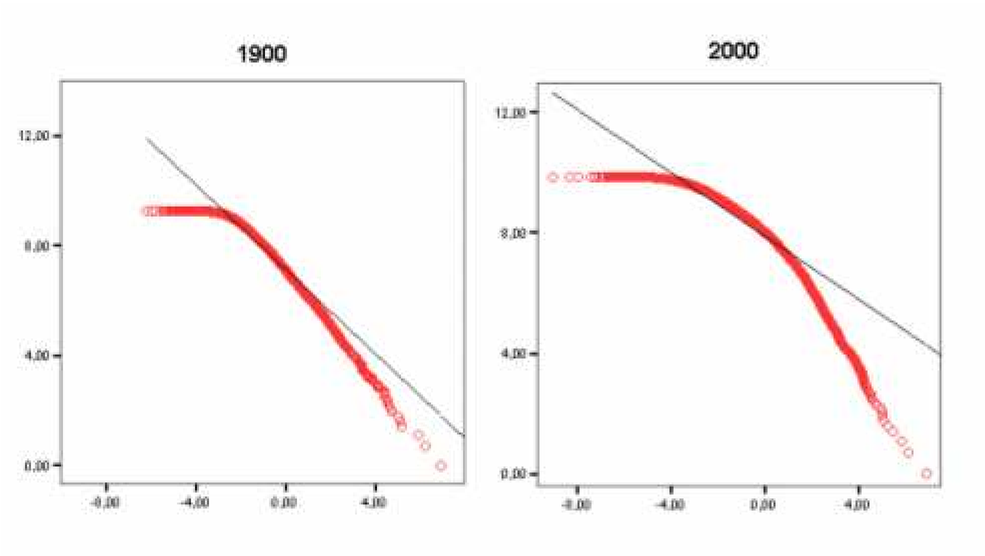

Figure 2. - Adaptive kernels of Relative Size (ln scale)

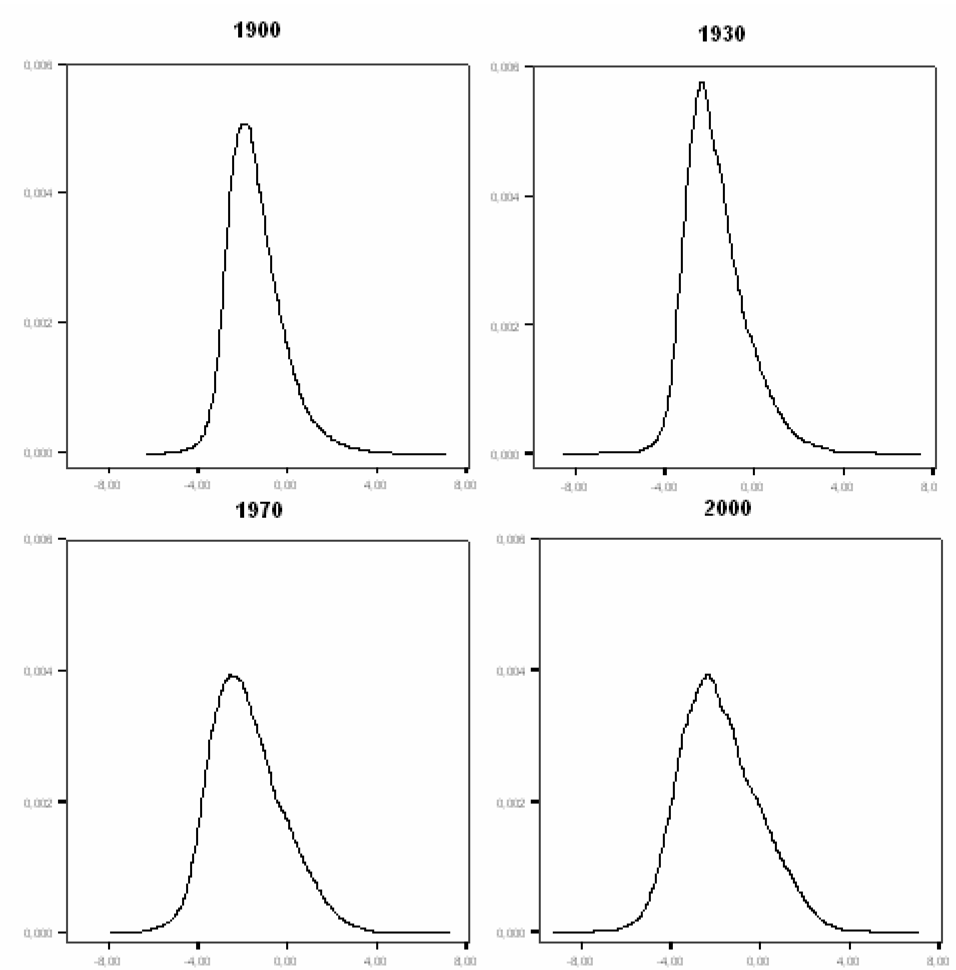

Figure 3. - Empirical cumulative density functions in 1900 and 2000

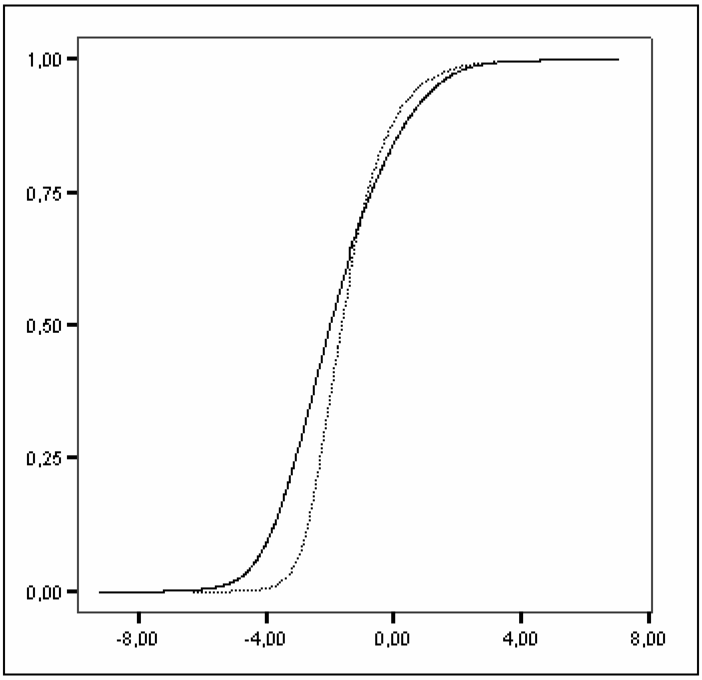

Figure 4.- Kernel estimate (*Bandwidth* 0.5) of Population Growth (1900-2000)

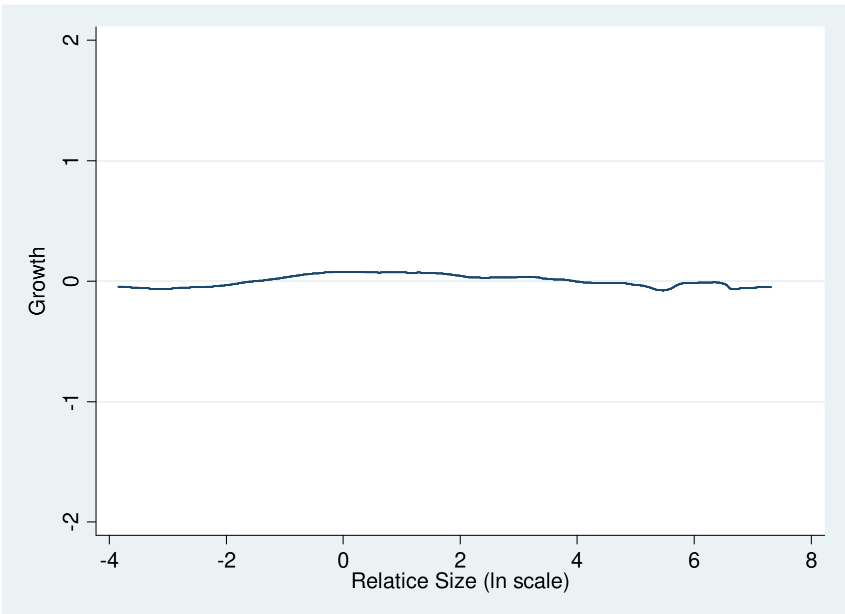

Figure 5.- Kernel estimate (*Bandwidth* 0.5) of the Variance of Population Growth (1900-2000)