

MPRA

Munich Personal RePEc Archive

New Evidence on Gibrat's Law for Cities

González-Val, Rafael and Lanaspa, Luis and Sanz, Fernando

5 September 2008

Online at <https://mpra.ub.uni-muenchen.de/10411/>
MPRA Paper No. 10411, posted 11 Sep 2008 04:58 UTC

New Evidence on Gibrat's Law for Cities

Rafael González-Val

Luis Lanaspa

Fernando Sanz

Department of Economic Analysis

University of Zaragoza

Abstract: The aim of this work is to test empirically the validity of Gibrat's Law in the growth of cities, using data for all the twentieth century of the complete distribution of cities (without any size restrictions) in three countries: the US, Spain and Italy. On considering the distribution of cities, we find a tendency to divergence. However, this fact does not impede, whether from an empirical or a theoretical point of view, that city size distribution can be adequately approximated with a lognormal distribution. Also, the conclusions which can be obtained as to fulfilment or not of Gibrat's Law depend, first, on the sample size and, second, on the size of the cities being considered (large or small), which means that the results of any study which does not use all the distribution will be relative.

Keywords: Gibrat's Law, city size distribution, urban growth

JEL: R00, C14.

Address:

Dpto. de Análisis Económico, Universidad de Zaragoza

Facultad de CC. Económicas y Empresariales

Gran Vía, 2, 50005 Zaragoza (Spain)

E-mail: llanas@unizar.es

1. Introduction

The relationship between the growth rate of a quantifiable phenomenon and initial size is a question with a long history in statistics: do larger entities grow more quickly, or the other way around? On the other hand, perhaps no relationship exists and the rate is independent of size. A fundamental contribution to this debate is that of Robert Gibrat [1931], who observed that the distribution of size (measured by sales or the number of employees) of firms could be approximated well with a lognormal, and that the explanation lay in the growth process of firms tending to be multiplicative and independent of their size. This proposition became known as Gibrat's Law and was the beginning of a deluge of work exploring the validity of this law for the distribution of firms (see the surveys of John Sutton [1997] and Enrico Santarelli, Luuk Klomp and Roy Thurik [2006]). Gibrat's Law establishes that no regular behavior of any kind can be deduced between growth rate and initial size.

The fulfilment of this empirical proposition also has consequences on the distribution which follows the variable; in the words of Gibrat [1931] himself "the law of proportionate effect will therefore imply that the logarithms of the variable will be distributed following the [normal distribution]". Some years later Michael Kalecki [1945], in a classic article, tested this statistical relationship between lognormality and proportional growth under certain conditions, consolidating the conceptual binomial Gibrat's Law – lognormal distribution.

In the field of urban economics, Gibrat's Law, especially since the 1990s, has given rise to numerous empirical studies contrasting its validity for city size distributions, arriving at a consensus, not absolute but of a majority, that it holds in the long term. Gibrat's Law presents the added advantage that as well as explaining relatively well the growth of cities, it can be related to another empirical regularity well known in urban economics, Zipf's Law, which appears when the so-called Pareto distribution exponent is equal to the unit¹. The term was coined after a work by George Zipf [1949], which observed that the frequency of the words of any language is clearly defined in statistical terms by constant values. This has given rise to theoretical works explaining the fulfilment of Gibrat's Law in the context of external urban local effects and productive shocks, relating them with Zipf's Law and associating them directly to an equilibrium situation. These theoretical works include Xavier Gabaix [1999], Gilles Duranton [2006, 2007] and Juan C. Córdoba [2008].

Returning to the empirical side, there is an apparent contradiction in these studies, as they normally accept the fulfilment of Gibrat's Law but at the same time affirm that the distribution followed by city size is a Pareto distribution, very different to the lognormal. Recently, Jan Eeckhout [2004] was able to reconcile both results, by demonstrating (as Jhon B. Parr and Keisuke Suzuki [1973] affirm in a pioneering work) that if size restrictions are imposed on the cities, taking only the upper tail, this skews the analysis. Thus, if all cities are taken, it can be found that the true distribution is lognormal, and that the growth of these cities is independent of size. However, to date, Eeckhout [2004] is the only study to consider the entire city size distribution. But this is

¹ If city size distribution follows a Pareto distribution the following expression can be deduced: $\ln r = a - b \ln S$, where r is rank (1 for the biggest city, 2 for the second biggest and so on), S is the size or population and a and b are parameters, this latter being known as the Pareto exponent. Zipf's Law is fulfilled when b equals the unit.

a short term analysis², when the phenomenon under study (Gibrat's Law) is, by definition, a long term result.

The aim of this work is to test empirically the validity of Gibrat's Law in the growth of cities, using data for all the twentieth century of the complete distribution of cities (without any size restrictions or with no truncation point) in three countries: the US, Spain and Italy. The following section offers a brief overview of the literature on Gibrat's Law and cities, mainly focusing on the sample sizes used in other studies, and the results obtained. Section 3 presents the databases, with special attention to the US census. From the results we deduce that when we consider the complete distribution of cities (section 4), a tendency to divergence is seen. However, the empirical evidence (section 5.1) shows that this does not impede city size distribution being adequately approximated as a lognormal distribution. In section 5.2 we will try to resolve this apparent contradiction within the theoretical framework developed by Kalecki [1945], proposing a minor modification of his model. Finally, we will relativise the results obtained in this study (and all the earlier ones) showing that the conclusions obtained on the fulfilment or otherwise of Gibrat's Law depend, first, on sample size (section 6), and second, on city size (section 7). The work ends with our conclusions.

2. Gibrat's Law for cities. An overview of the literature

Following Xavier Gabaix and Yannis M. Ioannides [2004], Gibrat's law states that the growth rate of an economic entity (firm, mutual fund, city) of size S has a distribution function with mean and variance that are independent of S . Therefore, if S_{it} is the size of city i at the time t and g is its growth rate, then $S_{it} = S_{it-1}(1 + g)$. Taking logarithms and adding that the rate depends on the initial size, we can obtain the following general expression of the growth equation³:

$$\ln S_{it} - \ln S_{it-1} = \mu + \beta \ln S_{it-1} + u_{it}, \quad (1)$$

where $\mu = \ln(1 + g)$ and u_{it} is a random variable representing the random shocks which the growth rate may suffer, which we shall suppose to be identically and independently distributed for all cities, with $E(u_{it}) = 0$ and $Var(u_{it}) = \sigma^2 \forall i, t$. If $\beta = 0$ Gibrat's Law holds and we obtain that growth is independent of the initial size.

In this case, ($\beta = 0$), it is easy to prove that the expected value of the size of city i at the time t depends only on the number of periods which have passed and on size in the first period:

$$E(\ln S_{it}) = \mu \cdot t + \ln S_{i0}, \quad (2)$$

while the variance would be given by:

² Eeckhout [2004] takes data from the United States census of 1990 and 2000, possibly because they are the only ones to be available on line. Moshe Levy (2008) in a comment to Eeckhout (2004) and Jan Eeckhout (2008) in the reply also consider no truncation point, but only for the 2000 US Census data.

³ The size of a city can be defined, according to the literature, in three ways: in levels (S_{it}), in relative values (S_{it}/\bar{S}_t , \bar{S}_t is the mean size) or in shares ($S_{it}/\sum_i S_{it}$). The crucial parameter in (1) is β , which

determines whether Gibrat's Law holds. The specification (1) in logs makes the estimation of β robust to the three different definitions of city size.

$$\text{Var}(\ln S_{it}) = t \cdot \sigma^2. \quad (3)$$

Consequently, the mean grows over time, which seems reasonable for many economic variables, but variance does too, which is more debatable.

Remember that if $\beta = 0$ city growth is proportional, as it does not depend on initial size. Thus, if the estimation of β is significantly different to zero we will reject the fulfilment of Gibrat's Law. In the case of being greater than zero, we will have divergent growth, because city growth would depend directly and positively on initial size. A sustained process of divergent growth of this kind would result in an increasingly asymmetrical distribution, with small cities getting further and further away from large ones. And if β is negative urban growth would be convergent (mean reversion), as the growth-size ratio would be negative; a larger initial population would mean less growth and vice versa, so that in the long term distribution would tend to be concentrated around a median value. It is simple to prove that when $\beta \neq 0$ expressions (2) and (3) change, becoming

$$E(\ln S_{it}) = \mu \cdot \frac{(\beta + 1)^t - 1}{\beta} + (\beta + 1)^t \ln S_{i0}, \quad (4)$$

$$\text{Var}(\ln S_{it}) = \sigma^2 \cdot \frac{(\beta + 1)^{2t} - 1}{\beta^2 + 2\beta}, \quad (5)$$

and it can be demonstrated (see Appendix) that when $t > 1$ and growth is divergent ($\beta > 0$) variance (5) grows even faster than in (3), while if city growth were convergent ($\beta < 0$) variance (5) would be less than in (3).

In the 1990s numerous studies began to appear which empirically tested the validity of Gibrat's Law. Table 1 shows the classification of all the studies on urban economics that we know of. While the countries considered, statistical and econometric techniques used and sample sizes are heterogeneous, the predominating result is the acceptance of Gibrat's Law.

Thus, both Jonathan Eaton and Zvi Eckstein [1997] and Donald R. Davis and David E. Weinstein [2002] accept its fulfilment for Japanese cities, although they use different sample sections (40 and 303 cities, respectively) and time horizons. Davis and Weinstein [2002] affirm that long-run city size is robust even to large temporary shocks and, in studying the effect of Allied bombing in the Second World War, deduce that the effect of these temporary shocks disappears completely in less than 20 years.

Steven Brakman, Harry Garretsen and Marc Schramm [2004] come to the same conclusion when analysing the impact of bombardment on Germany during the Second World War, concluding that, for the sample of 103 cities examined, bombing had a significant but temporary impact on post-war city growth. Nevertheless, nearly the same authors in Maarten Bosker, Steven Brakman, Harry Garretsen and Marc Schramm [2006] obtain a mixed result with a sample of 62 cities in West Germany: correcting for the impact of WWII Gibrat's Law is found to hold only for about 25% of the sample.

Meanwhile, both J. Stephen Clark and Jack C. Stabler [1991] and Marcelo Resende [2004] also accept the hypothesis of proportional urban growth for Canada and Brazil respectively. The sample size used by Clark and Stabler [1991] is tiny (the 7 most populous Canadian cities), although the main contribution of their work is to

propose the use of data panel methodology and unit root tests in the analysis of urban growth. This is also the methodology which Resende [2004] applies to his sample of 497 Brazilian cities.

For the case of the US there are also several works accepting statistically the fulfilment of Gibrat's Law, whether at the level of cities (Eeckhout [2004] is the first to use all the sample without size restrictions), or with MSAs (Yannis M. Ioannides and Harry G. Overman [2003], whose results reproduce Gabaix and Ioannides, 2004). Also for the US, however, Duncan Black and Vernon Henderson [2003] reject Gibrat's Law for any sample section, although their database of MSAs is different⁴ to that used by Ioannides and Overman [2003].

Other works exist rejecting the fulfilment of Gibrat's Law. Thus, France Guérin-Pace [1995] finds that in France for a wide sample of cities with over 2,000 inhabitants during the period 1836-1990 there appears to be a fairly strong correlation between city size and growth rate, a correlation which is accentuated when the logarithm of the population is considered. This result goes against that obtained by Eaton and Eckstein [1997] when considering only the 39 most populated French cities. Kwok T. Soo [2007] and George Petrakos, Prodromos Mardakis and Helen Caraveli [2000] also reject the fulfilment of Gibrat's Law in Malaysia and Greece, respectively.

For the case of China, Gordon Anderson and Ying Ge [2005] obtain a mixed result with a sample of 149 cities of more than 100,000 inhabitants: Gibrat's law appears to describe the situation well prior to the Economic Reform and One Child Policy period, but later Kalecki's reformulation seems to be more appropriate.

What we wish to emphasize is that, with the exception of Eeckhout [2004], none of these studies considers the entire distribution of cities, as all of them impose a truncation point, whether explicitly, by taking cities above a minimum population threshold or implicitly, by working with MSAs⁵. This is usually due to a practical reason of data availability.

3. The databases

We used city population data from three countries: the US, Spain and Italy. The US is an extremely interesting country in which to analyse the evolution of its urban structure, as it is a relatively young country whose inhabitants are characterised by high mobility. On the other hand we have the European countries, with a much older urban structure and inhabitants who present greater resistance to movement; specifically, Paul C. Cheshire and Stefano Magrini [2006] estimate mobility in the US is fifteen times higher than in Europe.

Considering these two types of country gives us information about different urban behaviors, as while Spain and Italy have an already consolidated urban tissue and new cities are rarely created (urban growth is produced by population increase in existing cities), in the US urban growth has a double dimension, as well as increases in

⁴ The standard definitions of metropolitan areas were first published in 1949 by what was then called the Bureau of the Budget, predecessor of the current Office of Management and Budget (OMB), with the designation Standard Metropolitan Area. This means that if the objective is making a long term analysis it will be necessary to reconstruct the areas for earlier periods, in the absence of a single criterion.

⁵ In the US, to qualify as a MSA a city needs to have 50,000 or more inhabitants, or the presence of an urbanised area of at least 50,000 inhabitants, and a total metropolitan population of at least 100,000 (75,000 in New England), according to the OMB definition. In other countries similar criteria are followed, although the minimum population threshold needed to be considered a metropolitan area may change.

city size, the number of cities also increases, with potentially different effects on city size distribution. Thus, the population of cities (incorporated places) goes from representing less than half the total population of the US in 1900 (46.99%) to 61.49% in 2000, at the same time the number of cities increases by 82.11%, from 10596 in 1900 to 19296 in 2000.

The data for the US we are using are the same as those used by Rafael González-Val [2008]. Our base, created from the original documents of the annual census published by the US Census Bureau, www.census.gov, consists of the available data of all incorporated places without any size restriction, for each decade of the twentieth century. The US Census Bureau uses the generic term *incorporated place* to refer to the governmental unit incorporated under state law as a city, town (except in the states of New England, New York and Wisconsin), borough (except in Alaska and New York), or village, and which has legally established limits, powers and functions.

The number of cities (in brackets) corresponding to each period is: 1900 (10596 cities), 1910 (14135), 1920 (15481), 1930 (16475), 1940 (16729), 1950 (17113), 1960 (18051), 1970 (18488), 1980 (18923), 1990 (19120) and 2000 (19296).

Two details should be noted. First, that all the cities 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 an associated free 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, for the same reason we also exclude all the unincorporated places (concentrations of population which do not form part of any incorporated place, but which are locally identified with a name), which began to be accounted after 1950. However, these settlements did exist earlier, so that their inclusion would again present a problem of inconsistency in the sample. Also, their elimination is not quantitatively important; in fact there were 1430 unincorporated places in 1950, representing 2.36% of the total population of the US, which by 2000 would be 5366 places and 11.27%.

For Spain and Italy the geographical unit of reference is the municipality and the data comes from the official statistical information services. In Italy this is the Servizio Biblioteca e Servizi all'utenza, of the Direzione Centrale per la Diffusione della Cultura e dell'informazione Statistica, part of the Istituto Nazionale di Statistica, www.istat.it, and for Spain we have taken the census of the Instituto Nacional de Estadística⁶, INE, www.ine.es. The de facto resident population has been taken for each city.

We have taken the data corresponding to the census of each decade of the 20th century. For Italy data for the following years have been considered (in brackets, the number of cities for each year): 1901 (7711), 1911 (7711), 1921 (8100), 1931 (8100), 1936 (8100), 1951 (8100), 1961 (8100), 1971 (8100), 1981 (8100), 1991 (8100) and 2001 (8100). No census exists in Italy for 1941, due to its participation in the Second World War, so we have taken the data for 1936. For Spain the following years are considered: 1900 (7800), 1910 (7806), 1920 (7812), 1930 (7875), 1940 (7896), 1950 (7901), 1960 (7910), 1970 (7956), 1981 (8034), 1991 (8077) and 2001 (8077).

⁶ The official INE census have been improved in an alternative database, created by Joaquín Azagra, Prilan Chorén, Francisco J. Goerlich and Matilde Mas [2006], reconstructing the population census for the twentieth century using territorially homogeneous criteria. We have repeated the analysis using this database and the results are not significantly different, so we have presented the results deduced from the official data.

4. A surprising initial result: divergence with all the cities

The first result we wish to present is the estimation of equation (1). We will focus on the analysis of the estimation of parameter β , as whether Gibrat's Law is fulfilled or not depends on its significance and its sign. Table 2 shows the results of the OLS estimation of β for the three countries considering all the cities, without size restrictions. The results of these regressions are usually heteroskedastic, so we have calculated the t-ratios using Halbert White's [1980] Heteroskedasticity-Consistent Standard Errors.

The first conclusion we obtain is that when the entire sample of cities is considered, β is always significantly different to zero, for any period and in the three countries. This result is robust as, while the literature usually admits the possibility of occasional deviations from Gibrat's Law in the short term (with some periods in which urban growth may be convergent or divergent), we are rejecting the fulfilment of Gibrat's Law during all of the 20th century and for three nations. But the really surprising finding is that despite these three countries having such different urban structures and histories, the estimated parameter is always positive (except in the period 1970-1980 in the US), so that the three exhibit divergent behavior throughout the 20th century.

The exception to this process of divergence is the estimation obtained for the US in the decade 1970-1980. The fact that this parameter is negative shows that during this decade the most populous cities grew more slowly. However, this result is atypical, and reflects two demographical circumstances in the United States during this period. First, between 1960 and 1990 there was a decline in the growth of the total population of the US, going from a growth rate of 18.5% in 1950-1960 to 9.8% in 1980-1990⁷. Then, that the total population grew by only 11.4% in 1970 - 1980, the third lowest growth rate in the history of the US since the first census was published in the late 18th century. And in this context of low growth of the total population, the percentage of urban population also fell (understood now as the percentage of the population associated with incorporated places), going from 64.51% of the total population in 1970 to 61.78% in 1980, which is by far the biggest fall in the 20th century. The fact that our estimation of β is negative would reflect the cities in the upper half of the distribution being where growth slowed most.

The overall result we have obtained is immobile and true for the three countries and during the entire twentieth century: when we took all cities without size restrictions, the city growth process was divergent. However, this conclusion can be noticeably qualified and relativised when we examine the importance of sample size (section 6) and city size (section 7) in the fulfilment or not of Gibrat's Law. But first we will analyse in section 5 the consequences on city size distribution of the divergent tendency we have observed.

5. A first note about divergence: lognormality is maintained

5.1 From an empirical viewpoint

In the section above it has been shown that the overall result when the whole distribution is used is divergence. Also, as $\beta > 0$ the variance will grow more than

⁷ Source: <http://www.census.gov/population/censusdata/table-4.pdf>.

linearly (equation (5)), so that in principle the growth process could be explosive and it would be expected that city size distribution would be increasingly asymmetrical.

To corroborate this affirmation with the necessary statistical rigour we carried out Wilcoxon's lognormality test (rank-sum test), which is a non-parametric test for assessing whether two samples of observations come from the same distribution. The null hypothesis is that the two samples are drawn from a single population, and therefore that their probability distributions are equal, in our case, the lognormal distribution. Wilcoxon's test has the advantage of being appropriate for any sample size. The more frequent normality tests –Kolmogorov-Smirnov, Shapiro-Wilks, D'Agostino-Pearson– are designed for small samples, and so tend to reject the null hypothesis of normality for such large sample sizes, although the deviations from lognormality are arbitrarily small.

Table 3 shows the results of the test. The conclusion is that the null hypothesis of lognormality is accepted at 5% for all periods of the 20th century in Spain and Italy. In the US a temporal evolution can be seen; in the first decades lognormality is rejected and the p-value decreases over time, but from 1930 the p-value begins to grow until lognormal distribution is accepted at 5% from 1960 onwards (the same conclusion is reached by González-Val [2008] through a graphic examination of the adaptive kernels corresponding to the estimated distribution of each decade). In fact, if instead of 5% we take a significance level of 1%, the null hypothesis would only be rejected in 1920 and 1930.

However, the form of the distribution in the US for the period 1900-1950 is not far from lognormality, either. Figures 1 and 2 show, respectively, the empirical density functions estimated by adaptive Gaussian kernels for 1900 and for 1950 (the last in which lognormality is rejected). The motive for this systematic rejection appears to be an excessive concentration of density in the central values, higher than would correspond to the theoretical lognormal distribution (in black). Starting in 1900 with a very leptokurtic distribution, with a great deal of density concentrated in the mean value, from 1930 (not shown), when the growth of urban population slows, the distribution loses kurtosis and concentration decreases, accepting lognormality statistically at 5% from 1960.

To sum up, both the test carried out and the visualisation of the estimated empirical density functions seem to corroborate that city size distribution can be approximated correctly as a lognormal (in Spain and Italy during the entire 20th century, and in the US for most decades, depending on the significance level), despite urban growth having been divergent during the entire 20th century for the three countries (with the single exception of the period 1970-1980 in the US).

5.2 From a theoretical viewpoint

The aim of this section is to provide a statistical model capable of explaining the observed behaviors. Specifically, it should justify divergent growth, with the variance of cities growing over time not at a constant rate but faster, and should also be able to maintain the lognormality of the distribution.

For this we will use Kalecki's model [1945], in which we will introduce a variant which will enable us to arrive at the desired final result: a phenomenon which is distributed as a lognormal and whose variance grows over time at a more than constant rate. We refer the reader to Kalecki [1945], whose notation we adopt, for a more detailed explanation.

From equation (3) upwards, we see that, under Gibrat's approach, the variance grows over time, because it moves under the influence of independent cumulative random shocks. On the contrary, Kalecki [1945] considers that the second moment of the variable being studied is governed by economic forces, so that the variance remains constant. We want to emphasize that lognormality is maintained in both cases. However, while with Gibrat growth is unbounded, in Kalecki's framework it is convergent or, at least, not divergent.

In what follows in this section, we try to reconcile our two main empirical outcomes so far: lognormality with divergent growth, divergent in the sense that the variance is growing over time at a more than constant rate (compare (3) and (5) with $\beta > 0$; see the Appendix).

We shall examine this in more detail.

Let X be the variable being studied, in this case city size. Let Y be the deviation from the mean of $\ln X_0$, the logarithm of the value of the variable X in the period 0. Let M be the second moment of Y . Let y be a change in Y and ΔM a change, at the same time, in M . Let n be the number of observations, which we can identify with the evolution of the variable in time. By definition:

$$\frac{1}{n} \sum (Y + y)^2 = M + \Delta M, \quad (6)$$

and operating we obtain:

$$2 \sum Yy = - \sum y^2 + n \Delta M. \quad (7)$$

In order to get a time-constrained variance, Kalecki hypothesises that the correlation between Y and y is negative. This cannot be applicable to our case. On the contrary, a variance which grows over time requires the relationship between Y and y , between the variable in question and its changes, to be positive, so that $\sum Yy$ is greater than zero. We thus obtain, by construction, a growing variance which, as long as Y is also growing (as is usually the case with cities), does so over time at increasing rates, as is our case⁸. This is the fundamental difference from Kalecki. Now it only remains to demonstrate, in this new scenario, that lognormality is maintained. For this, we will follow the argument of the original article, published in *Econometrica* in 1945.

Mathematically, the content of the previous paragraph takes form in the relationship between Y and y , which we will take as linear for simplicity's sake, being given by:

$$y = \delta Y + z, \quad \delta > 0, \quad z \text{ independent of } Y \quad (8)$$

then:

$$Yy = \delta Y^2 + zY \quad (9)$$

Taking the sums in (9), operating and using (7) we come to:

$$\delta = -\alpha + \frac{\Delta M}{2M}, \quad (10)$$

⁸ It is important to emphasize that variance is also growing in Gibrat's approach, but at a constant rate (see (3)). Our variance grows faster.

where $M = \frac{\sum Y^2}{n}$ and $\alpha = \frac{\sum y^2}{2\sum Y^2}$. Adding the unit to both members of (10),

dividing the total by $\sqrt{1 + \frac{\Delta M}{M}}$ and noting

$$1 - \gamma = \frac{1 + \delta}{\sqrt{1 + \frac{\Delta M}{M}}} \quad (11)$$

the following expression is deduced:

$$1 - \gamma = \frac{1 + \frac{\Delta M}{2M}}{\sqrt{1 + \frac{\Delta M}{M}}} - \frac{\alpha}{\sqrt{1 + \frac{\Delta M}{M}}} \quad (12)$$

coinciding with that found on page 164 in Kalecki [1945]; at the end of the same page and at the top of page 165 the comparisons of orders of magnitudes are established, also valid now, leading to $0 < \gamma < 1$.

Let M_0 be the second moment of Y_0 ; M_1 the second moment of $Y_0 + y_1$; M_2 of $Y_0 + y_1 + y_2$, and so on. From (11) we can immediately see that

$$(1 + \delta_k) = (1 - \gamma_k) \sqrt{\frac{M_k}{M_{k-1}}}. \quad (13)$$

Meanwhile, from (8) we conclude that $Y + y = (1 + \delta)Y + z$, so that successively substituting:

$$Y_0 + y_1 + y_2 + \dots + y_n = Y_0(1 + \delta_1)(1 + \delta_2)\dots(1 + \delta_n) + z_1(1 + \delta_2)\dots(1 + \delta_n) + \dots + z_{n-1}(1 + \delta_n) + z_n \quad (14)$$

If we carry (13) to (14) and operate, we obtain:

$$Y_0 + y_1 + y_2 + \dots + y_n = \frac{Y_0}{\sqrt{M_0}}(1 - \gamma_1)\dots(1 - \gamma_n)\sqrt{M_n} + \frac{z_1}{\sqrt{M_1}}(1 - \gamma_2)\dots(1 - \gamma_n)\sqrt{M_n} + \dots + \frac{z_{n-1}}{\sqrt{M_{n-1}}}(1 - \gamma_n)\sqrt{M_n} + z_n \quad (15)$$

which is exactly the equation (7'') in Kalecki. Given the same circumstances as in the referenced article (first paragraph of page 166), we can conclude that if n is large enough the distribution of $Y_0 + y_1 + y_2 + \dots + y_n$ is approximately normal.

We have thus been able to define a statistical model, following Kalecki's [1945] standard framework, which combines lognormality with a variance growing in time at increasing rates.

6. A second note about divergence: everything depends on sample size

In the overview of the literature which we offered in section 2, we pointed out that empirical studies do not generally use data for all the cities. This is usually due to a

practical reason, the availability of data. For this motive most studies focus on analysing the most populous cities, the upper tail distribution. There are two very reasonable justifications for this approach. First, the largest cities represent most of the population of a country. And second, the growth rate of the biggest cities has less variance than the smallest ones (scale effect).

However, it should be pointed out that any test done on this type of samples will be local in character, and the behavior of large cities cannot be extrapolated to the entire distribution. This type of deduction can lead to the wrong conclusions, as it must not be forgotten that what is being analysed is the behavior of a few cities, which as well as being of a similar size, can present common patterns of growth. Therefore, we could be concluding that Gibrat's Law is fulfilled when what is really happening is that we have focused our analysis on a club of cities which cannot be representative of all urban centres.

Parr and Suzuki [1973] and Eeckhout [2004] demonstrate the importance of choosing sample size in the analysis of city size distribution: the arbitrary choice of a truncation point can lead to skewed results. In the same way, we will analyse if the results for city growth depend on sample size. To do this, we will again estimate (1) for different sample sizes: 50, 100, 200, 500, 1000 and so on, adding groups of 500 cities at a time until they are all considered⁹. We will carry out the estimations starting with the largest cities (from upper tail) and the smallest cities (from lower tail). The smallest cities usually present the highest variance, so that the results may change if the initial sample is made up of only the smallest cities.

Table 4 shows what we have called the *critical sample size* for the US, Spain and Italy, the size from which we reject the null hypothesis $\beta = 0$ with a significance level of 5%. The results show that for small initial sizes the null hypothesis $\beta = 0$ is not rejected, finding empirical evidence favouring Gibrat's Law, whether estimating from the upper tail or the lower tail. However, as the sample size increases this conclusion soon changes and inevitably the parameter becomes significant. While the critical sample size is very variable, for most of the twentieth century it is equal to or less than the 500 biggest cities, while for the smallest cities it tends to be higher, undoubtedly as a consequence of its higher variance. For any lower sample size we accept the local fulfilment of Gibrat's Law.

Three relevant conclusions can also be derived from Table 4. First, we have said that in the three countries we begin by taking 50 cities; when Gibrat's Law is rejected with this sample size we have explored the exact critical sample size for these cases. Obviously, this will be less than 50; this situation occurs six times for Italy and twice for Spain (never for the US). Second, there is a great deal of interannual variability, even between two consecutive periods. Simply as an example, let us take Italy, from the upper tail, from 1911 to 1921 and the next to jumps: we observe that the critical sample size goes from 28 to the substantial figure of 3500 and returns to the low figure of 21. And third, although there are exceptions, critical sample sizes tend to be bigger for the US than for the two Mediterranean countries, which is more evidence in favour of a greater validity of Gibrat's Law for US.

The information in Table 4 should be compared with the sample sizes used in other studies, shown in Table 1. As can be seen, except for Eeckhout [2004] who uses

⁹ For the US, where the sample size is noticeably bigger, we will take cities up to 5000 in batches of 500, and from that number, add them in batches of 1000.

the entire sample, sample sizes tend to be lower (sometimes much lower) than 500 cities, which enables reconciliation of the absolute result we obtained earlier of divergence when considering all the cities with other empirical works. By using low sample sizes they may be positioning themselves below the critical sample size, and accepting the relative fulfilment of Gibrat's Law when the behavior of the entire distribution may be different. Similarly, our conclusion of divergent growth depends on our decision to include all cities without size restrictions. After all, for sample sizes similar to those of other studies, we also accept the validity of Gibrat's Law.

Finally, Table 4 enables us to corroborate the atypical behavior already detected in the period 1970-1980 for the US. While the critical sample size from the lower tail is very high (4,500 cities), only the hundred biggest cities had a parallel growth (Gibrat's Law holds), indicating, therefore, that the cities of the upper half of the distribution were the ones which slowed their growth rate, or in other words, were mainly responsible for the negative sign in the estimation of the beta parameter.

Tables 5, 6 and 7 provide additional information to that offered in Table 4. These three tables present the values of the beta parameter estimated from (1) for the US, Italy and Spain, respectively, always from the upper tail (the tables from the lower tail do not provide qualitatively new results and so are not included). We want to point out two fundamental results.

One, and this is very relevant, is that non-monotonic behaviors are produced regarding the sample size, especially in the US. Effectively, the fifth (1930-1940) to the eighth (1960-1970) column of Table 5 begin by accepting Gibrat's Law, then a convergent behavior is produced (the beta estimation is statistically different from zero and negative), Gibrat's Law is accepted again for two or three intermediate sample sizes, and finally, we have a divergent behavior (the beta estimation is statistically different from zero and positive). We have already mentioned that one of the contributions of this work is that the results regarding Gibrat are largely a function of the number of cities considered, so that when a small or intermediate number of cities is taken, the Law usually tends to be accepted almost systematically. Keeping this in mind, these non-monotonicities only accentuate and strengthen this statement: As the title of this section says, "everything depends on sample size", and all the possible behaviors, including some repeated ones, can be found in the same jump between two contiguous censal periods. This fact is repeated in the last three columns of Table 6 and the last of Table 7.

Two, in Italy, and especially in Spain, the predominant behavior is divergence, which does not occur so intensely in the US. Everything seems to indicate, as will be corroborated in the next section, that in the two European countries the biggest cities are the ones that have seen the most growth.

7. And a third note about divergence: non-parametric tests and city size dependence

We are also interested in analysing how growth may depend on city size. In other words, Gibrat's Law can be valid for a club of cities of similar sizes, but not for other clubs of a different representative size. This is not equivalent to the analysis carried out in the previous section. To do so, we will use the non-parametric methodology followed by Eeckhout [2004] and Ioannides and Overman [2003], different to the parametric approach used up to now. It consists of taking the following specification:

$$g_i = m(s_i) + \varepsilon_i, \quad (16)$$

where g_i is the growth rate $(\ln S_{it} - \ln S_{it-1})$ normalised (subtracting the mean and dividing by the standard deviation) and s_i is the logarithm of the i th city size. Instead of making suppositions about the functional relationship m , $\hat{m}(s)$ is estimated as a local mean around the point s and is smoothed using a kernel, which is a symmetrical, weighted and continuous function in s .

To analyse all the 20th century we build a pool with all the growth rates between two consecutive periods. This enables us to carry out long term analysis. And the Nadaraya-Watson method is used, exactly as it appears in Wolfgang Härdle [1990], based on the following expression¹⁰:

$$\hat{m}(s) = \frac{n^{-1} \sum_{i=1}^n K_h(s - s_i) g_i}{n^{-1} \sum_{i=1}^n K_h(s - s_i)}, \quad (17)$$

where K_h denotes the dependence of the kernel K (in this case an Epanechnikov) on the bandwidth h . It is worth pointing out that this procedure is slightly different to directly testing Gibrat's Law, as it estimates a local mean for each point s . The estimation obtained for each point does not then give us information on the fulfilment of Gibrat's Law for all the distribution, but only for that point and the local area centred on it. The size of this area depends largely on the bandwidth h . According to Härdle [1990], pages 25 and 26, if $h \rightarrow 0$ then $\hat{m}(s_i) \rightarrow g_i$. The smaller the bandwidth, the more concentrated are the weights around the observations. When $h = 0$ the estimator would be the corresponding growth rate for each point. While if $h \rightarrow \infty$ then $\hat{m}(s) \rightarrow n^{-1} \sum_{i=1}^n g_i$. The higher the bandwidth the smoother will be the estimated curve, being a straight line in the limit, the sample mean. We use an intermediate bandwidth, 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:

$$\hat{\sigma}^2(s) = \frac{n^{-1} \sum_{i=1}^n K_h(s - s_i) (g_i - \hat{m}(s))^2}{n^{-1} \sum_{i=1}^n K_h(s - s_i)}. \quad (18)$$

The estimator is very sensitive, both in mean and in variance, to atypical values. For this reason we decide to eliminate from the sample the 5% smallest cities, as they usually have much higher growth rates in mean and in variance. This is logical; we are

¹⁰ 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>.

discussing cities of under 200 inhabitants, where the smallest increase in population is very large in percentage terms.

Gibrat's Law implies that growth is independent of size in mean and in variance. As growth rates are normalised, if Gibrat's Law in mean were strictly fulfilled, the estimated kernel would be a straight line on the zero value. Values different to zero involve deviations from the mean. And the estimated variance of the growth rate would also be a straight line in the value one, which would mean that the variance does not depend on the size of the variable analysed. To be able to test these hypotheses, we have constructed bootstrapped 95-percent confidence bands (calculated from 500 random samples with replacement).

Figures 3, 4 and 5 show the estimated kernels of the growth rate of a pool for the entire 20th century for the US, Spain and Italy, respectively. For the US the value zero is always in the confidence bands, so that it cannot be rejected that the growth rates are significantly different for any city size. For Spain and Italy the estimated mean grows with the sample size, although it is significantly different to zero only for the largest cities. One possible explanation is historical: both Spain and Italy suffered wars on their territories during the 20th century, so that for several decades, the largest cities attracted most of the population¹¹.

Therefore, we find evidence in favour of Gibrat's Law for the US throughout the 20th century. Also for Spain and Italy, although the largest cities would present some divergent behavior. However, this evidence does not contradict the initial result of divergence which we obtained in section 4. First, because the analysis we are now carrying out is long term, as it considers all the growth rates of the 20th century jointly. And second, take into account that we are estimating a local mean (in an area whose size depends on bandwidth) for each point s . Conceptually, the idea is, to a certain extent, similar to what we did in the earlier section taking subsamples, although the procedure is very different. Then we were calculating local values by using subsamples, and now we are also calculating local values for each point s and a local area centred on it. Again, and this conclusion is relevant, when considering local results, the evidence in favour of Gibrat's Law increases.

Figures 6, 7 and 8 show the estimated kernels of the variance of growth rate of a pool for the entire 20th century for the US, Spain and Italy, respectively. As expected, while for most of the distribution the value one falls within the confidence bands, indicating that there are no significant differences in variance, the tails of the distribution show differentiated behaviors. In the US the variance clearly decreases with the size of the city, while in Spain and Italy the behavior is more erratic and the biggest cities also have high variance.

To sum up, starting from the estimated kernels of the local means used in this non-parametric section, again the results show that cities of different sizes do not present significant differences in their growth rates, which should lead us to accept Gibrat's Law, with the evidence being somewhat more favourable in the US than in Spain and Italy. In these last two countries, especially in Spain, the growth rate presents some tendency to growth according to city size.

8. Conclusions

¹¹ This result can be related with the "safe harbour effect" of Edward L. Glaeser and Jesse M. Shapiro (2002), which is a centripetal force which tends to agglomerate the population in large cities when there is an armed conflict.

The aim of this work is to test empirically the validity of Gibrat's Law in the growth of cities, using data for all the twentieth century of the complete distribution of cities (without any size restrictions) in three countries: the US, Spain and Italy.

The first result is that when considering the complete samples, we obtain a tendency to divergence. Despite being three countries with very different urban structures and histories, we find a positive relationship between the growth rate of cities and their initial size throughout the 20th century (except in the period 1970-1980 in the US).

However, the empirical evidence shows that this does not impede city size distribution being approximated as a lognormal distribution. The results of Wilcoxon's rank-sum test show that the null hypothesis of lognormality is accepted with a significance level of 5% for all periods of the 20th century in Spain and Italy. In the US an evolution over time is observed; in the first decades the lognormal is rejected and the p-value decreases over time, but from 1930 the p-value begins to grow until lognormal distribution is accepted at 5% from 1960 on. But even in these decades the form of the empirical density functions is similar to a lognormal, while we observe a certain concentration of density in the central values of the distribution. In fact, if instead of 5% we take a significance level of 1%, the null hypothesis would only be rejected in 1920 and 1930 in the US.

To resolve this apparent contradiction, which we deduce from our empirical results, between divergent growth and lognormality, we have defined a statistical model, inspired by Kalecki's [1945] standard framework, which combines lognormality with a variance which grows more than linearly over time.

Moreover, this result of divergent urban growth can and should be qualified in two ways. First, the conclusions obtained in our study depend on the choice of sample size and the inclusion of all cities without size restrictions. When estimating with subsamples and using smaller sample sizes we find there is a critical sample size under which we would always accept the fulfilment of Gibrat's Law. So, for sample sizes similar to those of other studies, we also accept that Gibrat's Law holds. And second, the use of non-parametric methods which relate the growth rate with city size through the estimation of local means enable us to observe that in the long term, the evidence in favour of Gibrat's Law increases, especially in the US.

Appendix

We have two expressions:

$$\text{Var}(\ln S_{it}) = t \cdot \sigma^2 \quad (3)$$

$$\text{Var}(\ln S_{it}) = \sigma^2 \cdot \frac{(\beta + 1)^{2t} - 1}{\beta^2 + 2\beta} \quad (5)$$

If Gibrat's Law is fulfilled ($\beta = 0$), and applying L'Hôpital's rule we obtain that (5)

converges to (3): $\lim_{\beta \rightarrow 0} \left(\sigma^2 \cdot \frac{2t(\beta + 1)^{2t-1}}{2\beta + 2} \right) = \frac{2t\sigma^2}{2} = t\sigma^2$.

Let's see what happens if $\beta > 0$ or $\beta < 0$:

$$(3) - (5) = t \cdot \sigma^2 - \sigma^2 \cdot \frac{(\beta + 1)^{2t} - 1}{\beta^2 + 2\beta} = \frac{\sigma^2}{\beta(\beta + 2)} \left[t(\beta^2 + 2\beta) - (\beta + 1)^{2t} + 1 \right] = \frac{\sigma^2}{\beta(\beta + 2)} [f(\beta)]$$

Considering time t as a continuum beginning with zero, the expression between brackets $f(\beta)$ is only defined if $-1 < \beta$. Also, if $\beta > 0$ then $\frac{\sigma^2}{\beta(\beta + 2)} > 0$, while if $-1 < \beta < 0$

then $\frac{\sigma^2}{\beta(\beta + 2)} < 0$.

Therefore, to find out the total sign of the difference (3)–(5) we must study the behavior of the function $f(\beta) = t(\beta^2 + 2\beta) - (\beta + 1)^{2t} + 1$. The maximum or minimum of this function is obtained by defining the corresponding optimisation problem, whose first order condition is given by:

$$\frac{df(\beta)}{d\beta} = f'(\beta) = 2t(\beta + 1 - (\beta + 1)^{2t-1}) = 0,$$

From which we deduce that at the extreme $1 = (\beta + 1)^{2t-2}$, which means that $f(\beta)$ is maximum or minimum in $\beta = 0$. In order to know if the optimum $\beta = 0$ is maximum or minimum we obtain the second order condition:

$$\frac{d^2 f(\beta)}{d\beta^2} = f''(\beta) = 2t(1 - (2t - 1)(\beta + 1)^{2t-2}),$$

and evaluate the sign in $\beta = 0$: $f''(\beta = 0) = 4t(1 - t) < 0$ as long as $t > 1$.

Thus we already know that the function $f(\beta)$ is concave and reaches its maximum in $\beta = 0$ as long as $t > 1$. Considering that $f(0) = 0$, this function always takes negative values except in the maximum.

The final sign of the difference (3)–(5) will be (maintaining the conditions $-1 < \beta$ and $t > 1$):

1. When $\beta > 0$ we have seen that $\frac{\sigma^2}{\beta(\beta + 2)} > 0$ is fulfilled and city growth is divergent. The variance of the cities will be bigger than if Gibrat's Law were fulfilled: (3) < (5).
2. When $\beta < 0$ city growth is convergent. The variance of the cities will be less than if Gibrat's Law were fulfilled: (3) > (5).
3. When $\beta = 0$ (3) = (5).

Acknowledgements

The authors would like to thank the Spanish Ministerio de Educación y Ciencia (SEJ2006-04893/ECON project and AP2005-0168 grant from FPU programme), the DGA (ADETRE research group) and FEDER for their financial support. The comments of Arturo Ramos and members of the ADETRE research group contributed to improving the paper.

References

- [1] **Anderson, Gordon, and Ying Ge.** 2005. "The Size Distribution of Chinese Cities." *Regional Science and Urban Economics*, 35: 756-776.
- [2] **Azagra, Joaquín, Pilar Chorén, Francisco J. Goerlich, and Matilde Mas.** 2006. *La localización de la población española sobre el territorio. Un siglo de cambios: un estudio basado en series homogéneas (1900-2001)*. Fundación BBVA.
- [3] **Black, Duncan, and Vernon Henderson.** 2003. "Urban Evolution in the USA." *Journal of Economic Geography*, 3: 343-372.
- [4] **Bosker, Maarten, Steven Brakman, Harry Garretsen, and Marc Schramm.** 2006. "A Century of Shocks: the Evolution of the German City Size Distribution 1925 – 1999." CESifo working paper 1728.
- [5] **Brakman, Steven, Harry Garretsen, and Marc Schramm.** 2004. "The Strategic Bombing of German Cities during World War II and its Impact on City Growth." *Journal of Economic Geography*, 4: 201-218.
- [6] **Cheshire, Paul C., and Stefano Magrini.** 2004. "Population Growth in European Cities: Weather Matters- but only Nationally." *Regional Studies*, 40(1): 23-37.
- [7] **Clark, J. Stephen, and Jack C. Stabler.** 1991. "Gibrat's Law and the Growth of Canadian Cities." *Urban Studies*, 28(4): 635-639.
- [8] **Córdoba, Juan C.** Forthcoming. "A generalized Gibrat's Law for Cities." *International Economic Review*.
- [9] **Davis, Donald R., and David E. Weinstein.** 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity." *American Economic Review*, 92(5): 1269-1289.
- [10] **Duranton, Gilles.** 2006. "Some Foundations for Zipf's Law: Product Proliferation and Local Spillovers." *Regional Science and Urban Economics*, 36: 542-563.
- [11] **Duranton, Gilles.** 2007. "Urban Evolutions: The Fast, the Slow, and the Still." *American Economic Review*, 97(1): 197-221.
- [12] **Eaton, Jonathan, and Zvi Eckstein.** 1997. "Cities and Growth: Theory and Evidence from France and Japan." *Regional Science and Urban Economics*, 27(4-5): 443-474.
- [13] **Eeckhout, Jan.** 2004. "Gibrat's Law for (All) Cities." *American Economic Review*, 94(5): 1429-1451.
- [14] **Eeckhout, Jan.** Forthcoming. "Gibrat's Law for (all) Cities: Reply." *American Economic Review*.
- [15] **Gabaix, Xavier.** 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics*, 114(3): 739-767.
- [16] **Gabaix, Xavier, and Yannis M. Ioannides.** 2004. "The Evolution of City Size Distributions." In *Handbook of Urban and Regional Economics*, vol. 4, ed. John V. Henderson and Jean. F. Thisse, 2341-2378. Amsterdam: Elsevier Science, North-Holland.

- [17] **Gibrat, Robert.** 1931. *Les Inégalités Économiques*. Paris: Librairie du Recueil Sirey.
- [18] **Glaeser, Edward L., and Jesse M. Shapiro.** 2002. "Cities and Warfare: The Impact of Terrorism on Urban Form." *Journal of Urban Economics*, 51: 205-224.
- [19] **González-Val, Rafael.** 2008. "The Evolution of the US Urban Structure from a Long-run Perspective (1900-2000)." http://mpra.ub.uni-muenchen.de/9732/1/MPRA_paper_9732.pdf.
- [20] **Guérin-Pace, France.** 1995. "Rank-Size Distribution and the Process of Urban Growth." *Urban Studies*, 32(3): 551-562.
- [21] **Härdle, Wolfgang.** 1990. "Applied Nonparametric Regression." In *Econometric Society Monographs*. Cambridge, New York and Melbourne: Cambridge University Press.
- [22] **Ioannides, Yannis M., and Henry G. Overman.** 2003. "Zipf's Law for Cities: an Empirical Examination." *Regional Science and Urban Economics*, 33: 127-137.
- [23] **Kalecki, Michael.** 1945. "On the Gibrat Distribution." *Econometrica*, 13(2): 161-70.
- [24] **Levy, Moshe.** Forthcoming. "Gibrat's Law for (all) Cities: A Comment." *American Economic Review*.
- [25] **Parr, John B., and Keisuke Suzuki.** 1973. "Settlement Populations and the Lognormal Distribution." *Urban Studies*, 10: 335-352.
- [26] **Petrakos, George, Prodromos Mardakis, and Helen Caraveli.** 2000. "Recent Developments in the Greek System of Urban Centres." *Environment and Planning B: Planning and Design*, 27(2): 169-181.
- [27] **Resende, Marcelo.** 2004. "Gibrat's Law and the Growth of Cities in Brazil: A Panel Data Investigation." *Urban Studies*, 41(8): 1537-1549.
- [28] **Santarelli, Enrico, Luuk Klomp, and Roy Thurik.** 2006. "Gibrat's Law: An Overview of the Empirical Literature." In *Entrepreneurship, Growth, and Innovation: the Dynamics of Firms and Industries*, ed. Enrico Santarelli, 41-73. New York: Springer.
- [29] **Soo, Kwok T.** 2007. "Zipf's Law and Urban Growth in Malaysia." *Urban Studies*, 44(1): 1-14.
- [30] **Sutton, John.** 1997. "Gibrat's Legacy." *Journal of Economic Literature*. 35(1): 40-59.
- [31] **White, Halbert.** 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, 48: 817-38.
- [32] **Zipf, George.** 1949. *Human Behaviour and the Principle of Least Effort*, Cambridge, MA: Addison-Wesley.

Tables

Table 1. - Empirical studies on city growth

Study	Country	Period	Truncation point	Sample size	GL	EcIss
Eaton and Eckstein [1997]	France and Japan	1876-1990 (F) 1925-1985 (J)	Cities > 50,000 inhabitants (F) Cities > 250,000 inhabitants (J)	39 (F), 40 (J)	A	non par (tr mat, lz)
Davis and Weinstein [2002]	Japan	1925-1965	Cities > 30,000 inhabitants	303	A	par (purt)
Brakman et al. [2004]	Germany	1946-1963	Cities > 50,000 inhabitants	103	A	par (purt)
Clark and Stabler [1991]	Canada	1975-1984	7 most populous cities	7	A	par (purt)
Resende [2004]	Brazil	1980-2000	Cities > 1,000 inhabitants	497	A	par (purt)
Eeckhout [2004]	US	1990-2000	All cities	19361	A	par (gr reg); non par (ker)
Ioannides and Overman [2003]	US	1900-1990	All MSAs	112 (1900) to 334 (1990)	A	non par (ker)
Gabaix and Ioannides [2004]	US	1900-1990	All MSAs	112 (1900) to 334 (1990)	A	non par (ker)
Black and Henderson [2003]	US	1900-1990	All MSAs	194 (1900) to 282 (1990)	R	par (purt)
Guérin-Pace [1995]	France	1836-1990	Cities > 2,000 inhabitants	675 (1836) to 1782 (1990)	R	par (corr)
Soo [2007]	Malaysia	1957-2000	Urban areas > 10,000 inhabitants	44 (1957) to 171 (2000)	R	par (purt)
Petrakos et al. [2000]	Greece	1981-1991	Urban centres > 5,000 inhabitants	150	R	par (gr reg)
Bosker et al. [2006]	West Germany	1925-1999	Cities > 50,000 inhabitants	62	M	par (purt); non par (ker)
Anderson and Ge [2005]	China	1961-1999	Cities > 100,000 inhabitants	149	M	par (rank reg); non par (tr mat)
Gibrat's Law: GL	Econometric Issues: EcIss		gr reg: growth regressions	corr: coefficient of correlation (Pearson)		
A: Accepted	parametric methods: par		ker: kernels	lz: Lorenz curves		
R: Rejected	non parametric methods: non par		rank reg: rank regressions			
M: Mixed Results	purt: panel unit root tests		tr mat: transition matrices			

Table 2. - Estimated beta coefficients for the entire sample size

US										
Period	1900-1910	1910-1920	1920-1930	1930-1940	1940-1950	1950-1960	1960-1970	1970-1980	1980-1990	1990-2000
β estimated	0.008	0.022	0.042	0.009	0.048	0.051	0.027	-0.005	0.042	0.015
t-ratio	2.875	10.543	21.568	7.958	30.998	28.674	18.029	-3.934	35.152	12.352
SPAIN										
Period	1900-1910	1910-1920	1920-1930	1930-1940	1940-1950	1950-1960	1960-1970	1970-1981	1981-1991	1991-2001
β estimated	0.010	0.020	0.023	0.028	0.012	0.048	0.115	0.115	0.047	0.013
t-ratio	7.424	14.282	15.512	18.641	6.647	31.734	44.743	49.699	24.524	7.259
ITALY										
Period	1901-1911	1911-1921	1921-1931	1931-1936	1936-1951	1951-1961	1961-1971	1971-1981	1981-1991	1991-2001
β estimated	0.010	0.022	0.019	0.014	0.033	0.042	0.066	0.046	0.025	0.017
t-ratio	5.907	18.507	7.757	10.787	24.031	25.329	33.739	32.319	16.858	16.541

t-ratios calculated using White Heteroskedasticity-Consistent Standard Errors

Table 3. - Wilcoxon rank-sum test of lognormality

US											
Year	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
p-value	0.0252	0.017	0.0078	0.0088	0.0208	0.0464	0.1281	0.1836	0.2538	0.323	0.4168
SPAIN											
Year	1900	1910	1920	1930	1940	1950	1960	1970	1981	1991	2001
p-value	0.5953	0.6144	0.6233	0.6525	0.4909	0.5792	0.6049	0.522	0.5176	0.622	0.7212
ITALY											
Year	1901	1911	1921	1931	1936	1951	1961	1971	1981	1991	2001
p-value	0.2081	0.2205	0.2352	0.291	0.2864	0.3118	0.2589	0.272	0.382	0.4671	0.5287

Ho: The distribution of cities follows a lognormal

Table 4. - Critical sample size from which Gibrat's Law is rejected

US										
Period	1900-1910	1910-1920	1920-1930	1930-1940	1940-1950	1950-1960	1960-1970	1970-1980	1980-1990	1990-2000
From upper tail	500	1000	3000	500	500	200	500	100	3000	7000
From lower tail	1000	1000	9000	2000	3000	2500	2000	4500	1500	500
SPAIN										
Period	1900-1910	1910-1920	1920-1930	1930-1940	1940-1950	1950-1960	1960-1970	1970-1981	1981-1991	1991-2001
From upper tail	1500	200	100	200	200	100	100	500	1000	200
From lower tail	1000	4500	4000	5000	1	2000	500	1000	200	12
ITALY										
Period	1901-1911	1911-1921	1921-1931	1931-1936	1936-1951	1951-1961	1961-1971	1971-1981	1981-1991	1991-2001
From upper tail	18	28	3500	21	500	100	1000	24	29	25
From lower tail	6500	3000	5000	2500	2000	1500	1000	500	500	500

Table 5. - Local estimations of (1) from upper tail, US

N	1900-1910	1910-1920	1920-1930	1930-1940	1940-1950	1950-1960	1960-1970	1970-1980	1980-1990	1990-2000
50	-0.044	0.013	-0.014	0.002	-0.044	-0.054	-0.026	-0.014	-0.008	0.006
100	-0.038	-0.01	-0.005	0.006	-0.003	-0.05	-0.039	-0.046*	-0.031	-0.006
200	-0.005	-0.004	-0.017	-0.002	-0.004	-0.028*	-0.02	-0.034*	-0.022	-0.015
500	0.023*	0.014	0.001	-0.012*	-0.015*	-0.04*	-0.032*	-0.025*	-0.014	-0.007
1000	0.025*	0.02*	0.009	-0.012*	-0.018*	-0.046*	-0.031*	-0.026*	-0.01	-0.017
1500	0.018*	0.025*	0.004	-0.013*	-0.017*	-0.04*	-0.028*	-0.033*	-0.006	-0.007
2000	0.022*	0.021*	0.002	-0.016*	-0.014*	-0.027*	-0.029*	-0.03*	-0.001	-0.006
2500	0.024*	0.025*	0.007	-0.016*	-0.015*	-0.023*	-0.023*	-0.028*	0.002	-0.004
3000	0.024*	0.028*	0.013*	-0.018*	-0.015*	-0.022*	-0.02*	-0.03*	0.008*	-0.004
3500	0.028*	0.025*	0.02*	-0.017*	-0.012*	-0.018*	-0.014*	-0.027*	0.008*	-0.004
4000	0.027*	0.028*	0.021*	-0.016*	-0.007*	-0.015*	-0.017*	-0.028*	0.011*	-0.003
4500	0.026*	0.028*	0.025*	-0.013*	-0.003	-0.01*	-0.011*	-0.027*	0.011*	-0.002
5000	0.024*	0.03*	0.03*	-0.011*	-0.0007	-0.007	-0.007*	-0.026*	0.012*	-0.001
6000	0.028*	0.032*	0.038*	-0.007*	0.002	0.001	-0.001	-0.024*	0.014*	0.002
7000	0.025*	0.032*	0.041*	-0.004*	0.006*	0.007*	0.002	-0.023*	0.017*	0.006*
8000	0.027*	0.033*	0.044*	-0.0005	0.011*	0.012*	0.002	-0.022*	0.019*	0.006*
9000	0.023*	0.034*	0.046*	0.003	0.016*	0.015*	0.005*	-0.021*	0.022*	0.007*
10000	0.018*	0.035*	0.047*	0.005*	0.021*	0.021*	0.008*	-0.017*	0.026*	0.01*
11000		0.031*	0.048*	0.006*	0.026*	0.027*	0.01*	-0.016*	0.029*	0.012*
12000		0.031*	0.048*	0.007*	0.03*	0.032*	0.012*	-0.014*	0.03*	0.011*
13000		0.027*	0.048*	0.009*	0.035*	0.037*	0.015*	-0.015*	0.031*	0.012*
14000			0.046*	0.011*	0.04*	0.042*	0.016*	-0.013*	0.032*	0.014*
15000			0.042*	0.011*	0.044*	0.046*	0.02*	-0.011*	0.034*	0.015*
16000				0.009*	0.048*	0.049*	0.022*	-0.01*	0.036*	0.015*
17000							0.025*	-0.008*	0.038*	0.016*
18000								-0.005*	0.04*	0.017*
19000										0.015*

* Significant coefficients at 5%

Table 6.- Local estimations of (1) from upper tail, Italy

N	1901-1911	1911-1921	1921-1931	1931-1936	1936-1951	1951-1961	1961-1971	1971-1981	1981-1991	1991-2001
50	0.046*	0.034*	0.044	0.025*	0.018	0.023	-0.021	-0.024*	-0.028*	-0.02*
100	0.038*	0.026*	0.022	0.024*	0.007	0.038*	-0.005	-0.026*	-0.034*	-0.033*
200	0.03*	0.031*	0.023	0.02*	0.011	0.047*	0.009	-0.032*	-0.045*	-0.033*
500	0.024*	0.028*	0.008	0.018*	0.026*	0.052*	0.008	-0.03*	-0.036*	-0.035*
1000	0.017*	0.03*	0.008	0.011*	0.033*	0.06*	0.036*	-0.017*	-0.033*	-0.029*
1500	0.015*	0.027*	0.007	0.011*	0.037*	0.062*	0.044*	-0.013*	-0.025	-0.023*
2000	0.012*	0.023*	0.004	0.011*	0.038*	0.063*	0.054*	-0.004	-0.02*	-0.02*
2500	0.013*	0.022*	0.004	0.011*	0.033*	0.066*	0.058*	0.001	-0.014*	-0.014*
3000	0.014*	0.024*	0.006	0.01*	0.031*	0.064*	0.064*	0.007*	-0.008*	-0.008*
3500	0.013*	0.023*	0.012*	0.01*	0.03*	0.061*	0.067*	0.014*	-0.002	-0.003
4000	0.013*	0.025*	0.015*	0.01*	0.031*	0.06*	0.068*	0.02*	0.0007	-0.0007
4500	0.013*	0.026*	0.015*	0.01*	0.03*	0.057*	0.071*	0.025*	0.004*	0.003
5000	0.014*	0.027*	0.015*	0.01*	0.03*	0.055*	0.073*	0.03*	0.007*	0.004*
5500	0.014*	0.025*	0.015*	0.01*	0.03*	0.052*	0.076*	0.032*	0.01*	0.006*
6000	0.015*	0.026*	0.016*	0.01*	0.03*	0.05*	0.074*	0.036*	0.013*	0.008*
6500	0.014*	0.025*	0.02*	0.011*	0.032*	0.048*	0.072*	0.038*	0.016*	0.009*
7000	0.014*	0.024*	0.02*	0.012*	0.032*	0.046*	0.069*	0.041*	0.018*	0.012*
7500	0.013*	0.023*	0.02*	0.014*	0.035*	0.045*	0.068*	0.044*	0.022*	0.014*
7700	0.011*	0.022*	0.021*	0.013*	0.033*	0.043*	0.066*	0.046*	0.025*	0.017*

* Significant coefficients at 5%

Table 7.- Local estimations of (1) from upper tail, Spain

N	1900-1910	1910-1920	1920-1930	1930-1940	1940-1950	1950-1960	1960-1970	1970-1981	1981-1991	1991-2001
50	0.012	0.022	0.028	-0.012	0.01	0.001	0.002	-0.04	-0.036	-0.033
100	0.004	0.02	0.045*	0.033	0.022	0.047*	0.052*	-0.02	-0.036	-0.017
200	0.009	0.038*	0.037*	0.05*	0.062*	0.044*	0.058*	0.007	-0.008	-0.046*
500	0.007	0.037*	0.028*	0.065*	0.048*	0.046*	0.075*	0.041*	-0.006	-0.042*
1000	0.007	0.028*	0.025*	0.06*	0.033*	0.044*	0.092*	0.069*	0.016*	-0.022*
1500	0.009*	0.031*	0.018*	0.054*	0.026*	0.037*	0.097*	0.082*	0.025*	-0.006
2000	0.008*	0.028*	0.016*	0.051*	0.018*	0.043*	0.088*	0.086*	0.035*	0.003
2500	0.008*	0.021*	0.014*	0.048*	0.017*	0.046*	0.093*	0.094*	0.037*	0.008*
3000	0.008*	0.023*	0.018*	0.045*	0.015*	0.046*	0.092*	0.098*	0.038*	0.008*
3500	0.01*	0.022*	0.02*	0.043*	0.013*	0.047*	0.095*	0.1*	0.042*	0.012*
4000	0.01*	0.02*	0.018*	0.042*	0.013*	0.049*	0.095*	0.102*	0.045*	0.018*
4500	0.012*	0.022*	0.02*	0.04*	0.014*	0.05*	0.1*	0.104*	0.048*	0.02*
5000	0.012*	0.023*	0.02*	0.04*	0.014*	0.05*	0.1*	0.106*	0.051*	0.024*
5500	0.014*	0.023*	0.023*	0.037*	0.015*	0.05*	0.1*	0.108*	0.053*	0.025*
6000	0.013*	0.023*	0.025*	0.035*	0.015*	0.052*	0.102*	0.107*	0.057*	0.026*
6500	0.013*	0.023*	0.025*	0.033*	0.016*	0.051*	0.105*	0.109*	0.057*	0.026*
7000	0.012*	0.024*	0.024*	0.031*	0.016*	0.05*	0.108*	0.111*	0.057*	0.025*
7500	0.012*	0.022*	0.024*	0.029*	0.015*	0.05*	0.111*	0.114*	0.055*	0.023*
7800	0.009*	0.02*	0.023*	0.028*	0.012*	0.048*	0.115*	0.115*	0.047*	0.013*

* Significant coefficients at 5%

t-ratios calculated using White Heteroskedasticity-Consistent Standard Errors for all tables

Figures

Figure 1.- Comparison of the estimated density function (ln scale) and the theoretical lognormal in black (US, 1900)

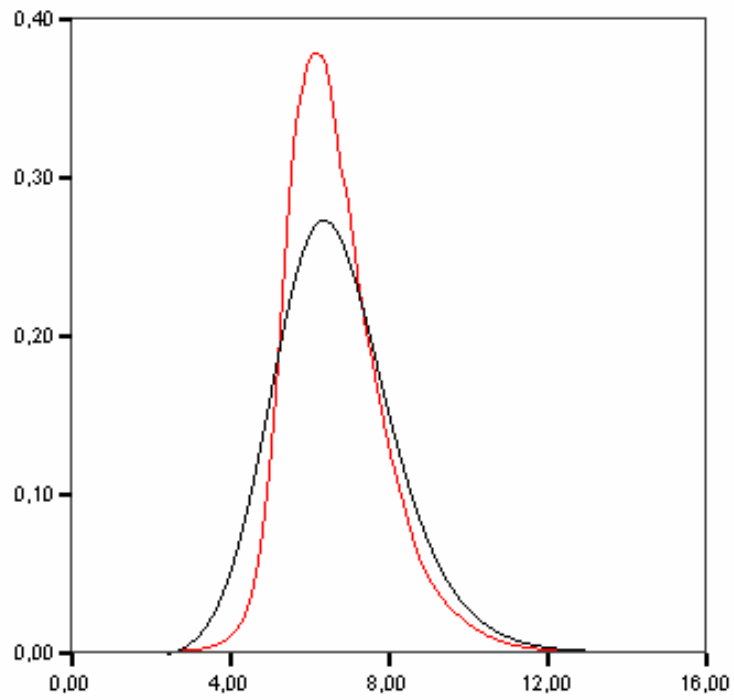


Figure 2.- Comparison of the estimated density function (ln scale) and the theoretical lognormality in black (US, 1950)

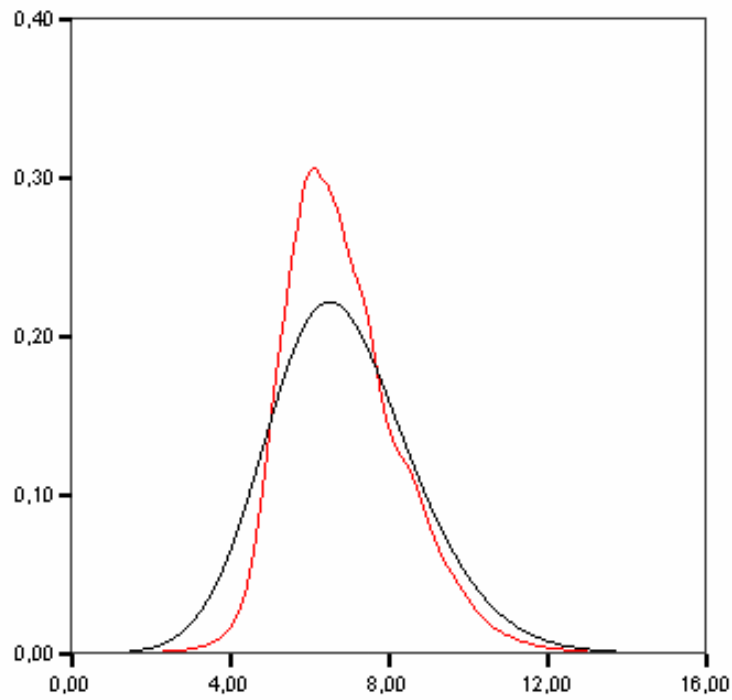


Figure 3.- Estimated kernel (bandwidth 0.5) of growth rate (US, 1900-2000), 152,475 observations

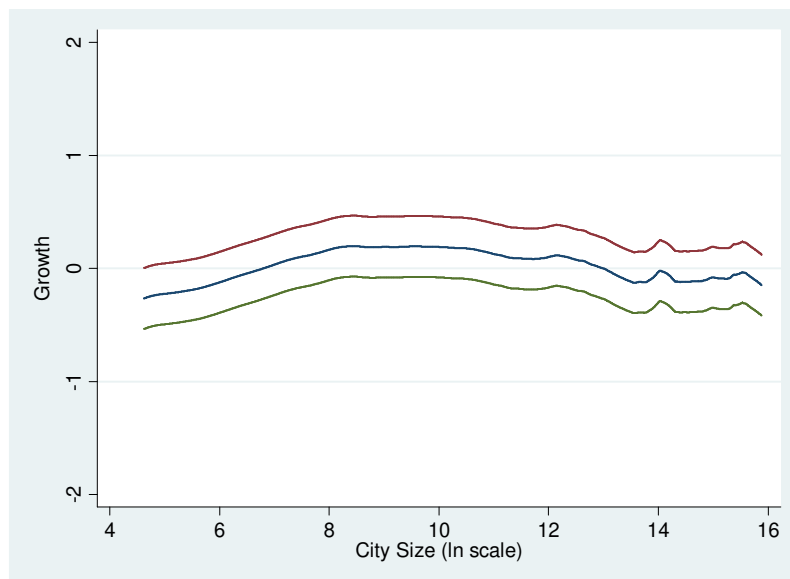


Figure 4.- Estimated kernel (bandwidth 0.5) of growth rate (Spain, 1900-2001), 74,100 observations

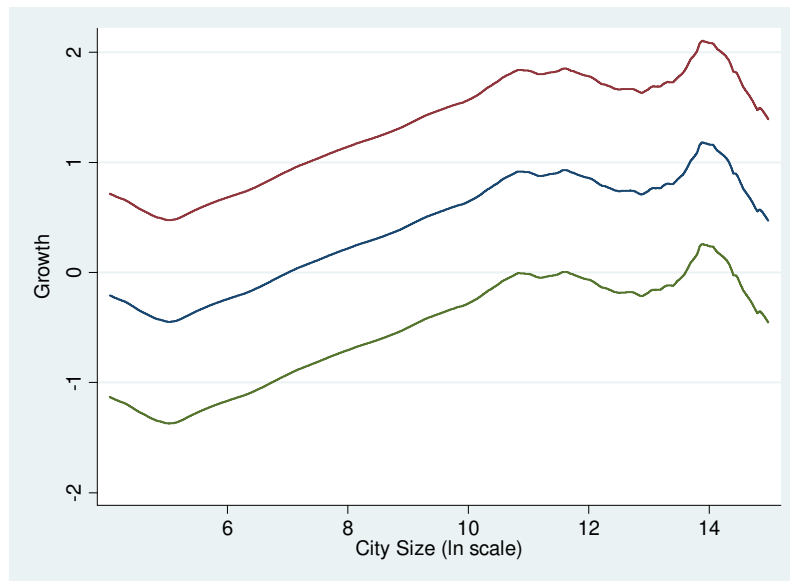


Figure 5.- Estimated kernel (bandwidth 0.5) of growth rate (Italy, 1901-2001), 73,260 observations

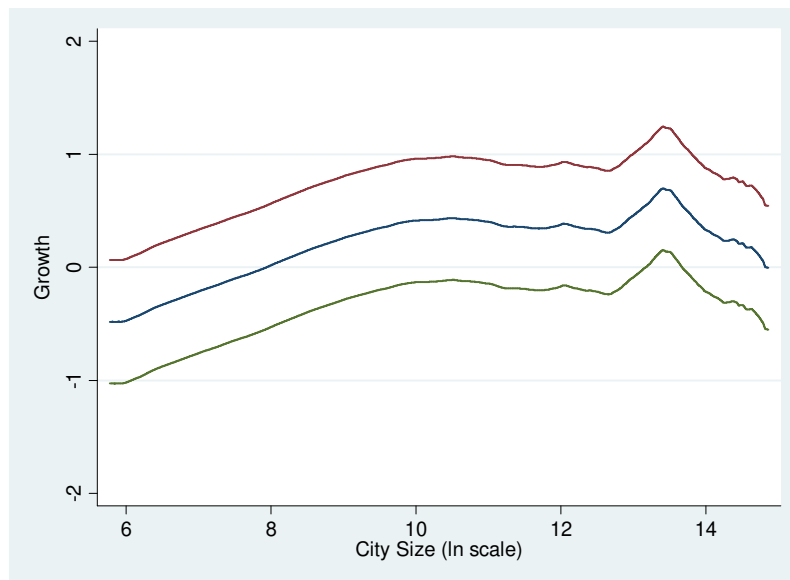


Figure 6.- Estimated kernel (bandwidth 0.5) of the variance of growth rate (US, 1900-2000), 152,475 observations

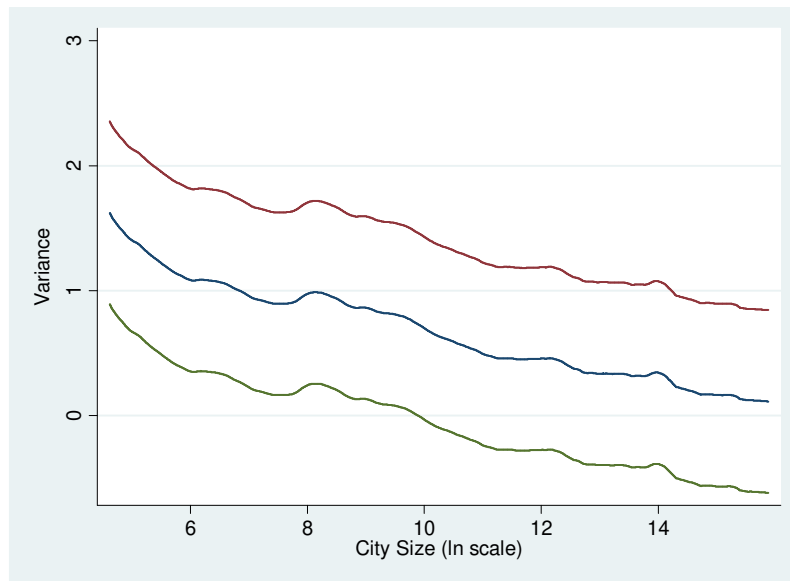


Figure 7.- Estimated kernel (bandwidth 0.5) of the variance of growth rate (Spain, 1900-2001), 74,100 observations

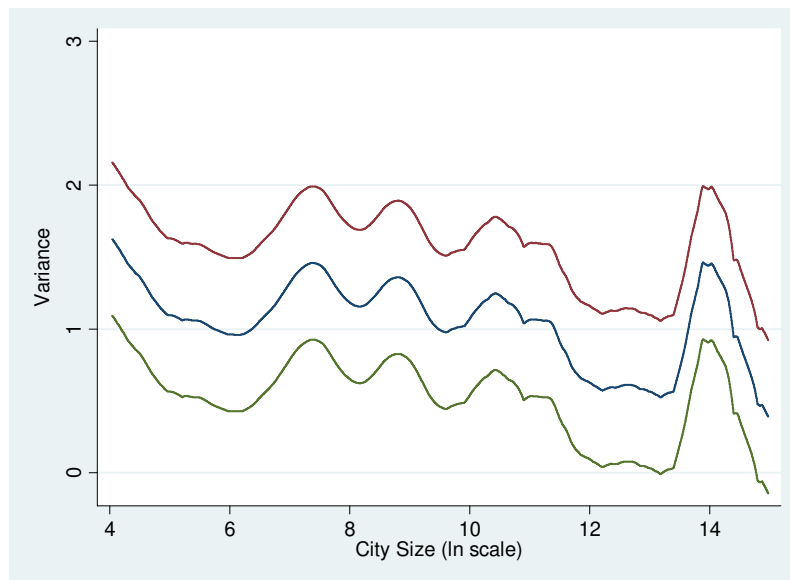


Figure 8.- Estimated kernel (bandwidth 0.5) of the variance of growth rate (Italy, 1901-2001), 73,260 observations

