IMT Institute for Advanced Studies, Lucca

Lucca, Italy

Essays in Regional and Complex Urban Economics

PhD Program in Economics, Markets and Institutions

XXV Cycle

By

Marco Modica

2013

The dissertation of Marco Modica is approved.

Program Coordinator: Prof. Fabio Pammolli, IMT Institute for Advanced Studies Lucca

Supervisor: Prof. Fabio Pammolli, IMT Institute for Advanced Studies Lucca

Tutor: Prof. Jing-Yuan Chiou, IMT Institute for Advanced Studies Lucca

The dissertation of Marco Modica has been reviewed by:

Prof. Massimo Riccaboni, IMT Institute for Advanced Studies Lucca

Prof. Jing-Yuan Chiou, IMT Institute for Advanced Studies Lucca

IMT Institute for Advanced Studies, Lucca

2013

Contents

List of Figures

List of Tables

Acknowledgements

Chapter 2 is a repruduction of the paper already under reviewing “Pareto or log-normal? A recursive truncation approach to the distribution of (all) cities”, co-authored with Giorgio Fazio of University of Palermo and University of Glasgow.

Chapter 3 is the reproduction of the article already under reviewing “Are Gibrat and Zipf Monozygotic or Heterozygotic Twins? A Comparative Analysis of Means and Variances in Complex Urban Systems”, co-authored with Aura Reggiani of University of Bologna and Peter Nijkamp of VU University of Amsterdam.

First of all, I am indebted to two great academicians: Giorgio Fazio (Università degi Studi di Palermo) and Aura Reggiani (Università di Bologna). They have been (and are) two endless sources of ideas, support, and suggestions.

I also want to thank Peter Nijkmap (VU University of Amsterdam) for his general support, comments and contribution to the article reported in chapter 3.

Many thanks to my supervisor Fabio Pammolli (IMT Lucca) for his general support and comments to the development of this dissertation.

I need to thank Simona Losito (IMT Lucca) for the inexhaustible help and effort in the drafting of this dissertation.

I am really grateful to the Maxwell School of Citizenship and Public Affairs at Syracuse University who gave me the possibility to spend a short part of my life in the USA and in the ”real american” (but a bit cold) city of Syracuse in the State of New York. I'm also grateful to the professors who allowed this opportunity: Chihwa Kao and Eleonora Patacchini.

Many thanks also to all other colleagues and friends who helped, supported, criticized and stimulated my work: Alessandro Belmonte (IMT Lucca), Valentino Dardanoni (Università degi Studi di Palermo), Joaquim Oliveira Martins (OECD), Eleonora Patacchini (Syracuse University), Massimo Riccaboni (IMT Lucca), Stuart S. Rosenthal, (Syracuse University).

Finally, many thanks to all the IMT community for the good time I had in Lucca.

Vita

12 April 1984	Born, Palermo (PA), Italy
2006	Bachelor in Economics and Finance (Laurea triennale) Final mark: 98/110 Università degli Studi di Palermo, Palermo (Italy)
2008	MA in Economics Scienze and Finance (Laurea magistrale) Final mark: 110/110 cum laude Università degli Studi di Palermo, Palermo (Italy)
2012-2013	Visiting Research Scholar (September-March) Economics Department with the Maxwell School of Citizenship and Public Affairs University of Syracuse, Syracuse, NY, USA invited by Professor Chihwa Kao and Professor Eleonora Patacchini

Presentations

1. M. Modica (with G. Fazio) "Pareto or log-normal? A recursive truncation approach to the distribution of (all) cities" at:
    - 59th Annual North American Meetings of the Regional Science Association International (Ottawa (Canada), October 2012)
    - Giornate Palermitane di Studi Economici (Ficuzza (Italy), June 2012)
    - 9th World Congress of Regional Science Association International, Changing spatial patterns in a globalising world (Timisoara (Romania), May 2012)
2. M. Modica (with A. Reggiani and P. Nijkamp) " Are Gibrat and Zipf Monozygotic or Heterozygotic Twins? A Comparative Analysis of Means and Variances in Complex Urban Systems" at the 53rd European Regional Science Association Congress (Palermo (Italy), August 2013)
3. M. Modica (with A. Reggiani) "Alternative Interpretation of Regional Resilience: Evidence from Italy" at the XXXIV Conferenza scientifica annuale AISRe (Palermo, Italy, September 2013)

Abstract

Cities are entities that are not “simple” but “complexly organized”. Theories about geographical structure of cities, land use patterns and cities evolution that explain how cities become spatially ordered are expanding to take in consideration this complexity. The conceptual foundation for the existence of central place hierarchies (i.e. the study of agglomeration economies in cities and trasportation and logistic costs) is now completed by the definition of emergent patterns that are not directly linked to the element of their economic processes but included in their “physic mechanisms” (i.e. the study of complex systems). This dissertation explores some of these aspects by performing empirical applications in the fields of regional and complex urban economics.

The dissertation contributes to the long standing debate on the city size distribution. From the empirical standpoint, traditional studies on the distribution of cities typically rely a regularity known as Zipf’s Law. We first investigate some typical shortcomings related to the choiche of the right truncation point to discriminate between upper tail and body of the distribution (chapter 2). Secondly, we invesigate specific conditions leading to a weak form of Gibrat’s law in connection with the different typologies of rank-size distribution (Zipf’s law), by adopting parametric and non-parametric approaches (chapter 3) and, finally, we use both the laws in studying agglomeration forces whithin the European Union (chapter 4).

Chapter 2 (co-authored with Giorgio Fazio)

In the literature, the distribution of the size of cities is a controversial issue with two common candidates: the Pareto and the Log-normal. While the first is most accredited when the distribution is truncated above a certain threshold, the latter is usually considered a better representation when all cities are considered. In this chapter, we reassess the empirical evidence on the city size distribution with respect to the sensitivity of the rank-size rule and a non-parametric alternative test to the choice of truncation point. In particular, we look at US Census data for Census Designated Places and apply a recursive-truncation approach where each possible truncation point of the distribution of all cities is considered. Our results highlight the sensitivity of tests to the truncation point and the difficulty of distinguishing a Pareto tail from the tail of a log-normal.

Chapter 3 (co-authored with Aura Reggiani and Peter Nijkamp)

The regional economics and geography literature has in recent years shown interesting conceptual and methodological contributions on the validity of Gibrat's Law and Zipf's Law. Despite distinct modeling features, they express similar fundamental characteristics in an equilibrium situation. Zipf's law is formalized in a static form, while its associated dynamic process is articulated by Gibrat's Law. Thus, it seems that both Zipf's Law and Gibrat's Law can be conceived of as heterozygote twins. Unfortunately, empirical investigations on the close relationship "Gibrat's Law vs. Zipf's Law" are rather rare. Chapter 3 aims now to answer the following research question: can (a generalization of) Gibrat's Law allow us to infer Zipf's Law, and vice versa? In our conceptual and applied framework, particular attention will be paid to the role of the mean and the variance of city population as key indicators for assessing the (non-)validity of the generalized Gibrat's Law. Our empirical experiments are based on a comparative analysis between the dynamics of the urban population of five countries with entirely mutually contrasting spatial-economic characteristics: Botswana, Germany, Hungary, Luxembourg and Malta. We arrive at the following results: If (i) the mean is independent of city size (first necessary condition of Gibrat's law); and (ii) the coefficient of the rank-size rule/Zipf's Law is different from one, the variance is dependent on city size. This finding suggests an important research implication: in modeling urban growth, Gibrat's law holds only with respect to the condition on the mean, but not on the variance, thus allowing for heterogeneity in the growth of small and big cities. Furthermore, differences in population growth lead to differences in the hierarchy of a city system (with a rank-size coefficient different from one); this phenomenon creates a possibility of asymmetric shocks affecting the distribution of big vs. small cities.

Chapter 4

The creation of the European Union, the distortions caused by the introduction of a single currency in countries structurally so different and the expansion of mobility of people, capital and services due to the constitution of the so-called Schengen Area from the beginning of '90s might have had some impacts on the dynamics of city populations.

This chapter provides a study of the hierarchical structure of the cities within the EU Member States with particular attention on agglomeration forces by means of two very well-known empirical regularities: Zipf's law, as a proxy for agglomeration forces, and Gibrat's law as a test for stationarity. We find that the hierarchical structures of Member States is more even than expected. Moreover, the European Union seems to be split in three distinct areas: an integrated area characterized by the validity of Gibrat's law (temporary shocks might have permanent impacts on the city structure); an area characterized by the presence of mean reversion (any exogenous shock is used up in certain amount of time); a small area where the effects of the shocks is magnified in the big cities.

Finally, we find that only the constitution of the Schengen Area and the share of international trade seem to have a weak impact on the hierarchical structures of Member States.

Chapter 1

Introduction

Cities are entities that are not "simple" but "complexly organized". Theories about geographical structure of cities, land use patterns and cities evolution that explain how cities become spatially ordered are expanding to take in consideration this complexity. The conceptual foundation for the existence of central place hierarchies (i.e. the study of agglomeration economies in cities and transportation and logistic costs) is now completed by the definition of emergent patterns that are not directly linked to the element of their economic processes but included in their "physic mechanisms" (i.e. the study of complex systems). This dissertation explores some of these aspects by performing empirical applications in the fields of regional and complex urban economics.

Content of the dissertation

Chapter 2[1] investigates the city size distribution with respect to the "truncation point". In particular, we propose a recursive-truncation approach to reassess the common Zipf's Law regression and a non-parametric alternative proposed by Clauset, Shalizi and Newman et al. (2009) against

[1]Chapter 2 is the reproduction of the paper already under reviewing "Pareto or log-normal? A recursive truncation approach to the distribution of (all) cities", co-authored with Giorgio Fazio of University of Palermo and University of Glasgow.

all possible truncation points of the entire distribution of cities.

The main challenge of chapter 3[2] is, instead, to empirically explore specific conditions leading to a weak form of Gibrat's law in connection with the different typologies of rank-size distribution (Zipf's law), by adopting parametric and non-parametric approaches. In this sense we recognize features of heterozygotic twins, between Zipf's law and Gibrat's law.

Finally, chapter 4, starting from the consideration that the European Union is still a multifaceted entity, at least in terms of urban structure, investigates whether the deeper and deeper integration of European Union affected the hierarchical structure of the city system of the Member States and, furthermore, if the European Union is an area where urban and regional policies affect Member States in the same way or, on the contrary, if it is still a heterogeneous entity and a heterogeneous area in recovering from exogenous shocks.

Research questions

To summarize, the current dissertation aims at dealing with the following research questions:

1. What is the role played by the choice of the truncation point in discriminating between a Pareto distribution and a log-normal distribution in the context of city size distribution? (Chapter 2)

2. Is it possible to determine the switching point between the two distributions with the use of an "endogenous method"? (Chapter 2)

3. What are the relationships between Zipf's law and Gibrat's law on the light of the so-call "generalized Gibrat's law"? Cordoba (2003) (Chapter 3)

[2]Chapter 3 is the reproduction of the paper already under reviewing "Are Gibrat and Zipf Monozygotic or Heterozygotic Twins? A Comparative Analysis of Means and Variances in Complex Urban Systems", co-authored with Aura Reggiani of University of Bologna and Peter Nijkamp of VU University of Amsterdam.

4. Is it possible to recognize a weak form of Gibrat's law allowing the possibility of asymmetric shocks affecting the distribution of big vs. small cities? (Chapter 3)

5. Does the EU affects the city system hierarchy and agglomeration forces of the Member States? (Chapter 4)

6. Can EU city system be seen as an integrated area, which is an area where urban and regional policies affect Member States in the same way? (Chapter 4)

7. Which are the drivers of agglomeration forces within the EU Member States? (Chapter 4)

Main results

The empirical evidence related to the previous set of research questions of each chapter is summarized as follows.

Chapter 2

First, we estimate the Pareto exponent from the typical rank-size equation for each possible recursive truncation of data. Collecting the recursive estimates, and respective confidence intervals, we find evidence of Zipf's Law with the Pareto exponent equal to one in the very upper tail (above the largest 135 cities). However, we find that the parameter varies (decreases) with respect to the truncation point, a result that is interpreted by Eeckhout (2004) as evidence of a log-normal distribution. In any case, the recursive nature of our regressions, when we start lowering the truncation point, adding smaller cities one at the time, we find that the size of Pareto coefficient first becomes statistically different from one and then starts decreasing, showing a non-monotonic behavior with respect to the truncation point, highlighting a potential non-linearity of Zipf's law with respect to the truncation point. Indeed, the Pareto seems to apply for more than one range of the distribution: when this is truncated to the very upper tail and for an intermediate range of cities.

Second, we apply the method recently suggested by Clauset et al. (2009) to estimate the lower bound of a Pareto distribution and, using a Kolmogorov-Smirnov test, we compare the relative fitness of the data to the Pareto and the log-normal distribution. Using this approach, we find that the Pareto distribution seems to apply for a much longer upper tail than the one traditionally considered in the literature. However, this analysis shows that when the truncation point is extended to include smaller ans smaller cities, it is not possible to disentangle the two on the grounds of statistical significance, but on the grounds of the size of the statistics. In general, the Pareto seems a better fit when we are close to the upper portion of the data and the log-normal seems a better fit when we approach the entire distribution, however the both distribution can apply to the upper portion of the distribution of cities (above around the first 100 cities). Beyond these cities, the upper tail conforms better to the Pareto distribution on the grounds of statistical significance.

Third, we reassess the above methods using simulated data of alternative distributions: a Pareto, a log-normal and a mixture of the two, where the upper tail is Pareto and the main body is log-normal. Our results add to the debate on the distribution of city size, highlighting some novel results in terms of the sensitivity of tests to the truncation point and showing some potential pitfalls in their ability to distinguish between the two distributions. In particular, while the rank size regressions seem to point to the distribution of cities as potentially a false power law with the log-normal simulated data displaying a remarkably similar signature to the real data, the non parametric test seems less conclusive. In this case, the size of the test statistics for the simulated log-normal and the real data are also remarkably similar in the very upper tail. Only when the tail is extended, the test seems to point again in favor of the Pareto. Hence, the non-parametric test seems unable to settle whether the distribution of cities is a weak or a false power law. Moving from the entire distribution towards the upper tail, the log-normal seems to give space to the Pareto. However, moving even further towards the very upper tail, tests seems unable to statistically distringuish among the two.

These results seem to support to the claim by Eeckhout (2009) that an arbitrary choice of the cut-off of the distribution may mislead scholars to conclude in favor of one or the other distribution.

Chapter 3

We show in this chapter that, according to Cordoba (2003), the variance of city growth can be dependent on size if the rank-size coefficient is different from one; in particular, we verified what Cordoba (2003) calls a "generalized Gibrat's law" for different countries with different spatial-economic characteristics: Botswana, Germany, Hungary, Luxembourg and Malta. We found strong evidence of this generalized Gibrat's law for Botswana, Luxembourg and Malta. We found weak evidence of Gibrat's law for Germany and no evidence for Hungary.

Our results confirm the propositions provided by Cordoba (2003). In particular, when the rank-size coefficient is equal to 1, neither the mean nor the variance of growth depend on size; when the rank-size coefficient is greater than 1, the mean is independent of the city size, but not the variance, and small cities face a greater volatility in growth than larger cities; alternatively, when the rank-size coefficient is lower than 1, the mean is independent from the city size, but not the variance, and large cities face a greater volatility in growth than smaller ones.

These results seem to support that Gibrat and Zipf have offered complementary perspectives on city size and systems of cites in a given country. Their contributions are not necessarily identical, but offer new perspectives on the same multi-faceted prism of the space-economy. These laws are part of the same family, but also reveal specific distinct features. In particular, we find that Zipf's law and the rank-size rule behave like"monozygotic" twins, while Gibrat's law seems to show the behaviour of a "heterozygotic" twins.

Chapter 4

We show here that the hierarchical structures of Member States of the European Union is more even than expected. Moreover, EU city system is still far from being an integrated area and in particular, we find that the

European Union seems to be split in three distinct areas: an integrated area characterized by the validity of Gibrat's law where large temporary shocks might have permanent impacts on the city structure; another area that is characterized by the presence of mean reversion and where any exogenous shock is used up in certain amount of time and a small area where the effects of the shocks is magnified in the big cities. Finally, we find that only the constitution of the Schengen Area and the share of international trade seem to have a weak impact on the hierarchical structures of Member States presenting respectively a positive impact on the Zipf coefficient (more even distribution) and a negative impact (more agglomerated distribution) and then indicating two sources, one against and one in favor, of agglomeration forces.

Contribution to the literature

In this section of the introductory chapter aims at stressing the most important innovative contributions to the economic literature of the current dissertation.

Chapter 2

The economic empirical literature on the city size distribution focused on the upper tail without any formal consideration of what is the right truncation point. Moreover, the truncation point is generally arbitrarly choosen. Chapter 2 is, to my knowledge, the first attempt to fill this gap. Moreover, the analysis exposes the sensitivity of existing tests to the choice of a specific truncation and highlights the difficulty of distinguishing between the tail of a log-normal and a power law tail for the population distribution of cities.

Chapter 3

The innovative contributions of chapter 3 are twofold. First of all, our results suggest a new research implication: in modeling urban growth, scholars can allow a certain degree of heterogeneity in the growth of small and big cities. Indeed, we show that, if Zipf's law is different from 1 than Gibrat's law holds only with respect the condition on the mean, but not on the variance. Consequently, this study shows that, the (generalization) of Gibrat's law allows the possibility of asymmetric shocks affecting the distribution of big vs. small cities. Secondly, these results might be useful to "relax" Gibrat's law in its strict interpretation, by reinforcing the hypothesis that small entities face a greater volatility in the growth process.

Chapter 4

The innovative contribution to the economic literature of chapter 4 regards several aspects. To my knowledge, chapter 4 is the first attempt to homogenize the data in performing an international analysis on city size distribution with the use of an endogenous methodology. In this way we can reduce substantially measurement errors arising from arbitrary methodologies typically used. At the same time we use two very well known empirical regularities, Zipf's law and Gibrat's law in an alternative way, where alternative means that aim of the chapter is not to verify if the former laws hold but it means that we use them as a tools to explore other research questions. Finally we analyze agglomeration forces within the European Union.

Chapter 2

Pareto or log-normal? A recursive truncation approach to the distribution of (all) cities

2.1 Introduction[1]

An accurate description of the spatial distribution of population is important for a number of theoretical and policy relevant issues, ranging from a better understanding of firms and people localization decisions to the implementation of national and regional policies in terms of incentives and transport infrastructures. Unfortunately, the literature is still far from reaching consensus on such description. Two specific distributions are most accredited: the Pareto and the log-normal. Disentangling between the two has important theoretical implications. For example, a Pareto distribution implies that cities are the result of agglomeration forces and industry specific productivity shocks. A log-normal

[1]The current chapter is a repruduction of the paper already under reviewing "Pareto or log-normal? A recursive truncation approach to the distribution of (all) cities", co-authored with Giorgio Fazio of University of Palermo and University of Glasgow.

distribution, instead, implies that cities grow proportionally and independently from the initial city size and their distribution results from city-wide rather than industry specific shocks (see Gabaix, 1999, for a discussion).

From the empirical standpoint, traditional studies on the distribution of cities typically rely on the relationship between the log-rank and log-size of cities, a regularity known as Zipf's Law. Several economic interpretations of the estimated Pareto coefficient have been given. For instance, Singer (1930) interprets this coefficient as a measure inequality in the city distribution and, hence, an indicator of the degree of urbanization and population concentration. If the estimated coefficient tends to zero, the entire population of the country tends to be located in one city. On the other end, if the parameter tends to infinity, all cities tend to have the same size (see, also, Soo, 2005). Hence, Singer (1930) considers this coefficient as an index of metropolization, so that the lower the estimated parameter the higher the value of urban land. Parr (1985) proposes a link between the overall level of development and the degree of metropolization, so that more developed countries should show a higher degree of metropolization (and a more unequally distributed city system) because of the improvement of transportation system between cities and regions in those countries. Along this vein, Brackman et al. (2001) interpret the estimated Pareto coefficient as an indicator of industrialization and agglomeration economies. Finally, Reggiani and Nijkamp (2012) interpret the estimated coefficient as an indicator of economic development regarding the urban structure as a socio-economic connected network. Along these lines, they compare this coefficient to the degree of connectivity in a network.

Using the rank-size regression approach, traditional studies favor a Pareto distribution with shape parameter equal to one, i.e. a minus one relationship between the log-rank and the log-size of cities, a regularity known as Zipf's Law. For example, Rosen and Resnick (1980) estimate the value of the Pareto exponent in a sample of 44 countries, finding a mean exponent of 1.136 with most countries falling in the [0.8-1.5] range. They also suggest that larger cities grow faster than smaller cities in most of their sample countries. Soo (2005) updates these results, finding a mean Pareto exponent of 1.105 over a sample of 75 countries, but also concludes for a rejection of Zipf's Law in more than half of cases. Other papers search for historical evidence of Zipf's Law. Guerin-Pace (1995) studies Zipf's Law in France between 1831 and 1990 and shows that the estimated Pareto coefficient may be sensitive to sample selection criteria. Black and Henderson (2003) construct a data set of US metropolitan areas defined over the period 1900-1990 choosing a minimum relative population threshold in each decade. They find a yearly Pareto coefficient around 0.85. Estimated coefficients are again sensitive to the choice of sample size. Glaeser, Ponzetto and Tobio (2011) study almost 200 years of regional changes in the US and show that the Zipf's Law tends to change over time.

However, these studies usually consider only the upper tail of the data, i.e. the largest cities, with a sample truncation point that is usually arbitrarily chosen. Moreover, the evidence in favor of a Pareto has to be reconciled with other empirical evidence that cities grow proportionally, a phenomenon known as Gibrat's Law, which should instead lead to a log-normal city-size distribution.[2] Differently from the above studies, Eeckhout (2004) suggests that the distribution of all cities, rather than just the upper tail, should be considered. He proposes an empirical investigation based on the US Census Designated Places (CDPs) and argues that city growth does not depend on the initial city size, providing evidence in favor of Gibrat's Law. Moreover, he shows that the estimated OLS coefficient of the so-called rank-size rule varies depending on the truncation city size, i.e. the inclusion of smaller (larger) cities in the sample should lead to a smaller (larger) coefficient, a result consistent with an underlying log-normal distribution. Based on these results, Eeckhout concludes that the size distribution of all cities follows a log-normal, rather than a Pareto.

These results have sparked some controversy about the distribution of cities beyond the upper tail. Using the same data of Eeckhout (2004),

[2]Gabaix (1999) shows that Zipf's Law may result as the steady state distribution from Gibrat's Law.

Levy (2009) presents a log-log plot of rank and city size and argues that the distribution of city size can be divided into two parts: a power law fits well the upper part, a log-normal fits better the bottom and middle parts. Eeckhout (2009) highlights the caveats of log-log plots. Instead, he proposes looking at the confidence bands of the log-normal estimates generated by a Lilliefors test and argues that the upper tail is also log-normal.[3]

However, Eeckhout (2009) underlines the difficulty of discriminating between a Pareto upper tail and the tail of a log-normal and hints at what may turn out to be a critical, and yet overlooked, point in the literature: *"With all the data available, and given that one nonetheless does not want to use all data, the question arises what the appropriate truncation point is. The choice of the truncation point becomes endogenous and can be chosen subjectively to favor one hypothesis over another"*. Eeckhout (2009, p. 1682). Hence, it seems that a particular distribution may be favored in empirical studies depending on the chosen truncation point.

The truncation point issue is directly related to the problem of discriminating between different inverse power laws, which can be classified according to Perline (2005) into strong, weak and false, depending on the strength of the Pareto law. A strong Pareto law arises when *"an inverse power law fits the full, untruncated range of the distribution of interest"*; a weak one when *"only some upper portion of the distribution follows an approximate inverse power law"* and a false when *"the largest observations (extremes) of the samples drawn from certain exponential type, and especially log-normal distributions, can closely mimic an inverse power law"* Perline (2005, pp. 75-76). Hence, the point where the sample is truncated may indeed turn out to be critical in discriminating between alternative distributions. Both traditional and recent studies do not thoroughly address this issue.

[3] In an interesting recent paper, Giesen et al. (2010) look at data for all cities in 8 countries and, using non-parametric and parametric goodness of fitness tests, conclude that the distribution of all cities is a Double Pareto Log-Normal (DPLN), i.e. a distribution that is Pareto in the upper and lower tails and log-normal in between. However, the DPLN distribution uses a larger set of parameters compared to the Pareto or the log-normal, which are definitely more parsimonious with only two parameters. Hence, the fitness improvements of alternative distributions, such as the DPLN, should be evaluated in relation to their dependence on a larger set of parameters.

In the light of the above, and in particular of the evidence that the estimated rank-size regression coefficient varies depending on the truncation city size, the typical economic interpretations of the estimated Pareto coefficient have to be considered together with a deeper analysis of the truncation point issue.

In this chapter, we propose a reappraisal of the debate on the city size distribution in relation to the specific issue of the truncation point. To the best of our knowledge, the only other papers attempting to investigate this issue are Bee et al. (2011) and Ioannides and Skouras (2013).

Bee et al. (2011) compare the two distributions by means of the maximum entropy density. This allows them to identify a very long power law spans, between 1205 and 1515 observations. Ioannides and Skouras (2013) try, instead, to determine the switching point between the Pareto distribution and the log-normal using maximum likelihood estimation over a "mixed" distribution, i.e. a distribution with a log-normal "body" and a Pareto "tail". This allows them to identify the threshold for a Pareto distribution around the city with population 60,290, concluding that even if the entire distribution may well be log-normal, most of the population resides in Pareto distribution.

Here, in order to investigate this issue, we propose two alternative approaches based on recursive analysis. Similarly to Eeckhout (2004), we do not constrain the investigation to the upper tail, but look at all cities. However, we look at all possible truncation points of the empirical distribution of all cities in order to discriminate between the two most accredited alternative theoretical distributions: the Pareto and the log-normal.

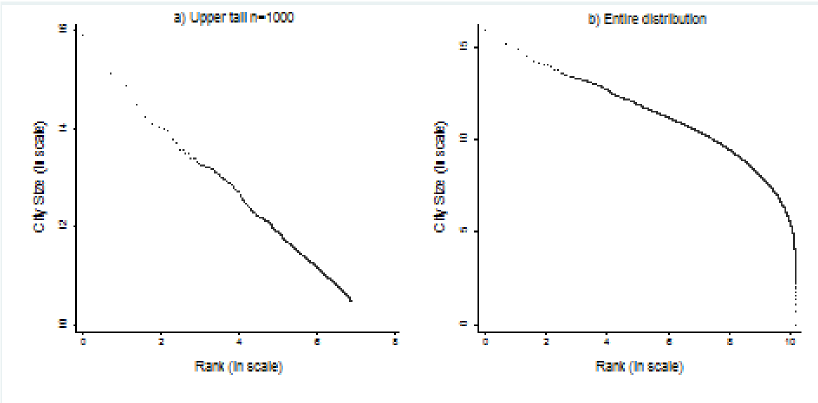
Using this approach, we are also able to provide a more extensive investigation into the city size distribution. First, we estimate the Pareto exponent from the typical rank-size equation for each possible recursive truncation of data. Collecting the recursive estimates, and respective confidence intervals, we can statistically assess the adherence of Zipf's Law for each truncated sample of the distribution of all US cities. We find that the parameter varies (decreases) with respect to the truncation point, a result that is interpreted by Eeckhout (2004) as evidence of a log-normal distribution. Thanks to the recursive nature of our regressions, however, we are also able to highlight a potential non-linearity of Zipf's law with respect to the truncation point. Indeed, the Pareto seems to apply for more than one range of the distribution: when this is truncated to the very upper tail and for an intermediate range of cities.

Second, we apply the method recently suggested by Clauset, Shalizi and Newman (2009) to estimate the lower bound of a Pareto distribution and, using a Kolmogorov-Smirnov test, we compare the relative fitness of the data to the Pareto and the log-normal distribution. Using this approach, we find that the Pareto distribution seems to apply for a much longer upper tail than the one traditionally considered in the literature. In particular, we identify a population threshold of around 57,775, a number very close to that found by Ioannides and Skouras (2013).

Third, we reassess the above methods using simulated data of alternative distributions: a Pareto, a log-normal and a mixture of the two, where the upper tail is Pareto and the main body is log-normal. Our results add to the debate on the distribution of city size, highlighting some novel results in terms of the sensitivity of tests to the truncation point and showing some potential pitfalls in their ability to distinguish between the two distributions. In particular, moving from the entire distribution towards the upper tail, the log-normal seems to give space to the Pareto. However, moving even further towards the very upper tail, tests seems unable to statistically distinguish among the two. These simulations seem to confirm the difficulty to identify the city size distribution as either a weak or a false inverse power law.

The rest of the chapter is organized as follows. The next section presents the empirical strategy and the results of recursive Zipf's Law equations and Kologorov-Smirnov tests. Section 2.3 replicates the methodology using simulated data. Section 2.4 concludes.

Figure 1: Log-log rank size plots (first 1000 largest cities vs entire distribution)



2.2 A recursive approach to the ditribution of all cities

A long tradition of papers underlines the difficulty of discriminating between a Pareto tail and log-normal distribution. For example, in reference to the use of log-log plots, which provide a visual assessment of the rank-size rule, Macauley (1922) states that the linearity of the tail of a frequency distribution charted on a logarithmic scale is not informative of a Pareto distribution, as it is a common feature of various types of frequency distributions. Parr and Suzuki (1973, p. 343) similarly affirm that: "*[...]truncation of the log-normal distribution at an appropriately high level enables the truncated portion to be regarded as not significantly different from the rank size distribution*". This point is illustrated in Figure 1 that compares log-log plots for the upper tail and for the entire distribution. While the left quadrant clearly points to a Pareto, the right seems to point to a log-normal.

More formally, Eeckhout (2004) shows that a variable P obeys a Pareto distribution if its density function, $\phi(P)$, and cumulative density function, $\Phi(P)$, are:

$$\phi(P) = \frac{a\underline{P}^a}{P^{a+1}} \qquad \forall P \geq \underline{P}, \tag{2.1}$$

$$\Phi(P) = 1 - \left(\frac{\underline{P}}{P}\right)^a \; \forall P \geq \underline{P}, \tag{2.2}$$

where q is a positive shape parameter and $\underline{P}$ is the scale parameter or the truncation city size, i.e. the minimum value of population P. The parameter q is also known as the Pareto coefficient and is a tail index. As mentioned above, in a log-log plot the distribution is represented by a straight line and Zipf's law satisfies Pareto with $q = 1$.

According to Clauset, Shalizi and Newman (2009), few phenomena seem to obey the Pareto distribution for all values and, as discussed above, most studies on the city size distribution find that the Pareto distribution is a good representation just for the upper tail, i.e. above a minimum threshold. However, even when a researcher intends to investigate just the upper tail, the choice of $\underline{P}$ may be critical, as a truncation point that is too high (low) may shorten (lengthen) the "right" size of the upper tail biasing tests of the appropriate distribution. The identification of the right truncation point may be interesting also for another issue. If, as sustained in some literature, the upper tail is Pareto and the entire distribution is log-normal, is there a switching point between the two distributions? This issue has received some attention in physics and statistics (see, among the others, Mitzenmacher, 2004; Perline, 2005; Clauset, Shalizi and Newman, 2009), but, with the exception of Eeckhout (2004) and very recently by Ioannides and Skouras (2013), it has been largely ignored in economics, where the choice of threshold is usually arbitrary.

In order to investigate the sensitivity of the distribution to the truncation point, we apply a recursive approach to the distribution of all cities. Following Eeckhout (2004), to consider "all" cities we use US Census data covering almost all the US population in "incorporated" and "unincorporated" places in the year 2000 and, for comparison, the year 2010. An incorporated place is an entity (populated area) with its own municipal government (city, town, village, borough and so on). Unincorporated places are, instead, areas lacking of own municipal government. In the US Census, these take the name of Census Designated Places (CDPs). The CDPs have been included for the first time in the year 2000. For the year 2000, the dataset covers 25,359 places and 208 millions US residents of the total 281 millions and, for the year 2010, 29,494 places and 230 millions US residents of the 308 millions total. The difference in number of places is due to changes introduced by the US Census Bureau: 24,841 are identical in the two years. Even though they may not coincide with the economically more meaningful definition of city, and previous work has considered Metropolitan Areas as the reference unit (see Gabaix, 1999; Ioannides and Overman, 2003), we prefer "places" as reference units in order to make our result comparable with Eeckhout (2004) and account for a larger population size. Here, for robustness we replicate the analysis for the two years. However we have to keep in mind that "*for size distribution studies, the entire metropolitan area is the most desirable choice for an urban unit as it represents an integrated economic unit. Since many workers and consumers in a city often reside in the surrounding suburbs, it seems reasonable to include these areas in the definition of the city*" Rosen and Resnick (1980, p. 170).

2.2.1 Recursive Zipf's Law

As mentioned above, we use a recursive approach to observe the adherence of the data to Zipf's Law for all possible truncation points of the distribution of all cities. As standard in the literature, we estimate the Pareto coefficient using simple rank-size OLS regressions where, following Gabaix and Ibragimov (2011), the rank is shifted by 0.5 to correct for the potential bias in small samples highlighted by Gabaix and Ioannides (2004), so that the estimating equation is:

$$ln(rank - 0.5) = k - q\, lnP, \tag{2.3}$$

where k is a constant and P is the population size. Standard errors are given by $(2/n)^{0.5}\,\hat{q}$. The parameter $\hat{q}$ is estimated for recursively truncated samples of the city size distribution, starting with the ten most pop-

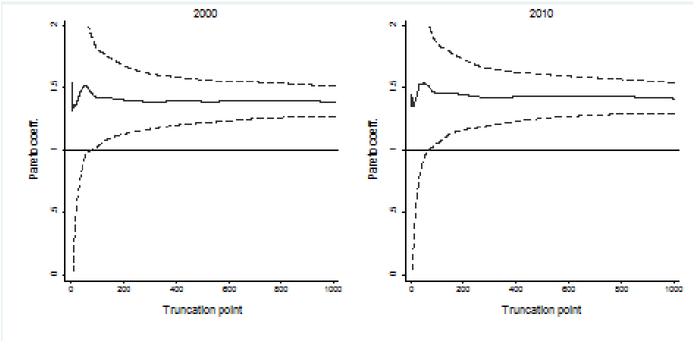
Table 1: Estimated coefficients for chosen truncation thresholds. Dependent variable ln (Rank-0.5)

	2000 Census Data				2010 Census Data			
$\underline{N}$	$\underline{P}$	$\hat{k}$(s.e.)	$\hat{q}$(s.e.) [GI s.e.]	R^2	$\underline{P}$	$\hat{k}$(s.e.)	$\hat{q}$(s.e.) [GI s.e.]	R^2
135	155,554	21.955	1.423	0.992	178,395	22.532	1.460	0.993
		(0.137)	(0.011) [0.173]			(0.139)	(0.011) [0.178]	
2,000	19,383	20.747	1.322	0.997	21,039	20.870	1.322	0.995
		(0.045)	(0.004) [0.042]			(0.057)	(0.005) [0.042]	
5,000	6,592	18.623	1.129	0.984	7,273	18.721	1.137	0.983
		(0.052)	(0.005) [0.023]			(0.057)	(0.006) [0.023]	
12,500	1,378	15.954	0.864	0.960	1,556	16.064	0.866	0.961
		(0.036)	(0.004) [0.011]			(0.037)	(0.004) [0.011]	
25,000	42	13.187	0.553	0.875	193	13.899	0.630	0.922
		(0.021)	(0.003) [0.005]			(0.021)	(0.003) [0.006]	
29,000	-	-	-	-	35	13.136	0.538	0.882
						(0.018)	(0.003) [0.005]	

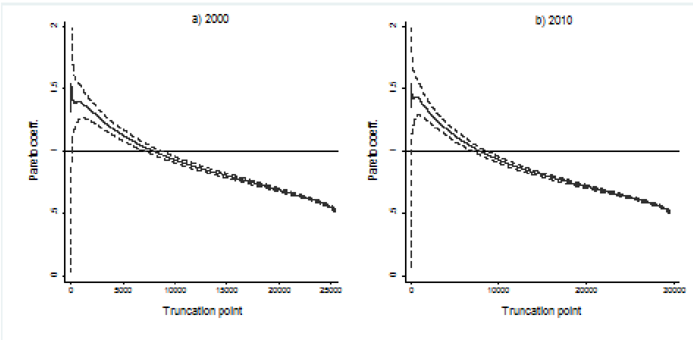
ulated cities and then adding one (less populated) city at the time until, like Eeckhout (2004), we consider all cities. Collecting the estimates of the Pareto exponent together with the respective 95% confidence interval, we can statistically assess the validity of Zipf's Law for each truncated city size distribution. In particular, while the estimated Pareto coefficient should be invariant to the truncation point, it should increase under the log-normal (Eeckhout, 2004).

Table 1 extracts the recursive OLS estimates of equation (2.3) for the six truncation points reported in Eeckhout (2004). These results seem consistent with previous work. The estimated Pareto coefficients seem, indeed, "threshold sensitive": the "longer" the upper tail, the lower the estimated coefficient. As already indicated in Eeckhout (2004), the coefficients decrease together with the truncation point. It is also interesting to compare the estimated parameters for the two different census years. The 2010 Census contains a larger number of observations mostly thanks to the improved accuracy in the definition of unincorporated places, as many CDPs present in the Census 2000 dataset have been split into two or more CDPs especially in the middle and in the lower tail of the distribution. The presence of these new observations does not seem to affect

Figure 2: Recursive Pareto coefficient and 95% Confidence Interval (Gabaix-Ibragimov s.e.)



(a) Upper tail (1000 largest cities)



(b) All cities

our results.

In Figures 2a and 2b we present the full recursive estimates. Figure 2a focuses on the largest 1,000 cities to look more closely at the upper tail.[4] Over this range, the estimated Pareto coefficient looks quasi-constant, indicating a potential Pareto distribution.

The coefficient shows some degree of fluctuation in the very first ob-

[4]We anticipate in this note that we focus on the first 1,000 cities because when lowering the threshold after that point, the estimated Pareto coefficient is decreasing. Moreover we will show in Figure 2b that around that threshold we have a change in the "slope" of the estimated coefficient, indicating, in our opinion, the switch between the log-normal distribution to the Pareto distribution.

servations, probably due to the influence of individual observations in a smaller sample, and then increases. In terms of statistical significance, estimates are indifferent from one for the first observations and are then statistically different from one, settling around the average estimated parameter of 1.4. Hence, Zipf's Law seems to be rejected, if not for the very first observations.

Interestingly, this information was not evident by looking at Table 1, where it was only possible to see the rank-size rule as a diminishing threshold process, but it was not possible to fully gauge the adherence of the data to Zipf's law for the upper tail.

What happens if we extend the analysis from the upper tail (here, the first 1000 cities) to all cities in the sample? Figure 2b shows all the estimated Pareto coefficients (and 95% confidence intervals) against each recursive truncation threshold. A number of results are worth mentioning. First, the coefficient clearly diminishes (increases) as we include smaller cities (larger cities), a result that, contrary to Figure 2a, corroborates the evidence of log-normality. Second, in terms of statistical significance, the recursive coefficients seem to display non-monotonic behavior. The Pareto coefficient is not statistically different from one in the very upper tail, where researchers typically set their cut-off point to estimate Zipf's Law (see Black and Henderson, 2003; Soo, 2005), but also for a second range of cities (between the 7,116th and 8,773rd in the year 2000 and 7,066th and 8,763rd in the year 2010). Hence, the Pareto exponent is not statistically different from 1 for two samples of the same distribution. Clearly, this result could only emerge by looking at all possible truncation points.

Figures 2a and 2b confirm how picking an arbitrary $\underline{P}$ may (mis)lead researchers to conclude in favor of a specific distribution. Finally, comparison of the left and right panels shows that results are robust to the use of different census years and are stable over time, with similar patterns and hierarchy. This also suggests that a more precise specification of the medium-small cities in 2010 does not significantly affect the results.

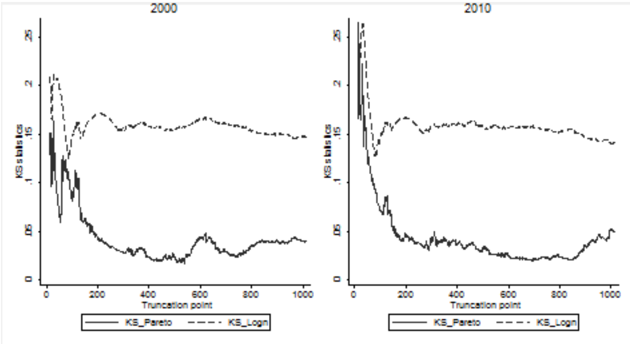
2.2.2 A Non-parametric Test Alternative

In this section, we exploit an alternative non-parametric methodology in order to discriminate between a Pareto distribution and a log-normal distribution. As discussed above, few phenomena seem to obey the Pareto distribution for all values and most studies find that the Pareto distribution is a good representation just for the upper tail, i.e. above a minimum threshold. Here we use the method proposed by Clauset, Shalizi and Newman (2009) to estimate this minimum threshold, $\underline{\hat{P}}$. The authors suggest testing the equality between the theoretical and empirical density functions using Kolmogorov-Smirnov tests over recursively truncated distributions. In this context we test the hypothesis that our distribution is a Pareto distribution. This means that we first run a Kolmogorov-Smirnov goodness of fit test for a Pareto distribution in the same recursive way we have done in the previous section for the estimation of the Pareto coefficient. Namely we get around 25,000 thousands KS tests in 2000 and around 29,000 KS tests in 2010. Then we choose as threshold, indicating the switch between log-normal and Pareto distribution, the population level of that city corresponding to the minimum KS statistics in a given year. For example, suppose the minimum KS appears in a subsample with the top 500 cities. Then, the population of the city with rank 500 will be the level of the population threshold. In sum, our estimate $\underline{\hat{P}}$ is then the value of $\underline{P}$ that minimizes the "recursive" Kolmogorov-Smirnov statistics, D:

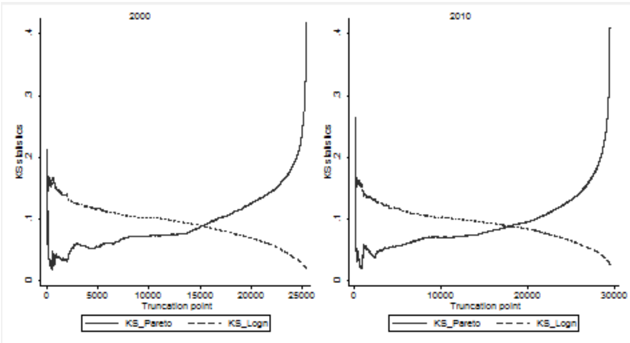
$$D = \sup_{p \geqslant \underline{p}} |\Phi_p(x) - \Phi(x)|, \tag{2.4}$$

where $\Phi_p(x)$ is the empirical cumulative density function for p i.i.d observations, and $\Phi(x)$ is the theoretical cumulative density function of the Pareto distribution. The KS statistic computes the supremum of the absolute value of the set of distances among the two. Under the null, the difference between the two is zero, i.e. the sample is drawn from the reference distribution. Rejection of the null, however, should be considered carefully, as the KS test tends to over-reject the null when the sample is large.

Figure 3: Recursive Kolmogorov-Smirnov test



(a) Upper tail (1000 largest cities)



(b) All cities

Figure 3 reports the recursive Kolmogorov-Smirnov statistics. As before, we begin the recursive analysis with the largest cities and then add smaller ones until we include all cities. In panel a) of Figure 3, we first look at the largest 1000 cities. The overall evidence seems to favor the Pareto distribution. Interestingly, however, for the very upper tail (exactly, 93 cities in 2000 and 90 in 2010), the KS statistics are visually too close and do not allow disentangling between the two distributions. For these observations, the p-values of the KS statistics confirm that both theoretical distributions can equally adapt to the empirical.[5] After this

[5]The authors will provide details on the p-values upon request.

portion of the upper tail, however, the Pareto seems to adapt better to the data with p-values rejecting the null of the KS test for a large portion of the upper tail. In particular, the KS test is rejected for the Pareto up to around the 1500th truncation point in the year 2000 around the 990th truncation point in the year 2010.

Following the approach of Clauset, Shalizi and Newman (2009), we find the minimum of the KS statistic for the 536th city in the year 2000 (D=0.0173) and 695th city in the year 2010 (D= 0.0198), which implies a minimum population threshold of 57,777 and 55,081 inhabitants, respectively. These tests, then, highlight a Pareto upper tail well before the arbitrary threshold of 100,000 inhabitants typically used by scholars for US data (Soo, 2005). In panel b) of Figure 3, we show KS recursive statistics for all truncated samples up to the entire untruncated distribution. Again, the evidence in favor of one or the other distribution changes depending on the truncation point: comparison of the KS statistics shows first the Pareto and then the log-normal as a better fit. Just like for the very upper tail, the two distributions are again indistinguishable half way to the entire distribution. When the distribution of all cities is considered, the log-normal appears as the best fit, as indicated by Eeckhout (2004). These results are in line with those obtained from the Zipf's Law regressions in the previous sub-section. If we do not take into account the problem of the correct cut-off, the KS test, also, could lead to concluding for a Pareto, when the true distribution is log-normal, and viceversa. Further, the evidence presented cannot rule out that a portion of the distribution of cities, the upper tail in particular, may be power law distributed and it confirms the difficulty of disentangling a Pareto and a log-normal in a portion of the upper tail.

2.3 Weak or false inverse Power-law?

The above analysis seems to confirm the sensitivity of test results with respect to the choice of truncation point. Moreover, it seems to highlight the distribution of cities as potentially as either a weak or a false power law, according to the definitions of Perline (2005). Indeed, it is not clear whether *"only some upper portion of the distribution follows an approximate inverse power law"* (weak power law) or *"the largest observations (extremes) of the samples drawn from certain exponential type, and especially log-normal distributions, can closely mimic an inverse power law"* (false power law), Perline (2005, pp. 68-69).

To further investigate this issue, we reassess the rank-size regressions and the non-parametric alternative against simulated data. In particular, we simulate three different random datasets: a log-normal, a Pareto with shape parameter equal to one (so that Zipf's law holds) and a "mixture" of Pareto upper tail (first 1000 observations) and log-normal body. In detail, we draw a log-normal dataset with same mean (7.28) and standard deviation (1.75) of the real data in the year 2000. For the "mixture" data we replace the first 1000 observations of the log-normal distribution with a sample where the first observation is twice the second, thrice the third and so on.

Following the same steps of the previous section, we first report the estimated recursive Pareto coefficients over the range of the 1000 largest cities and then over the entire distribution. Results are presented in panel a) of Figure 4. Looking at the upper tail, we notice a quasi-constant behavior of the coefficient (with different means) for all three simulated datasets. As expected, the estimated coefficients are not significantly different from one for the Pareto and the mixture-distributions. Interestingly, the estimated Pareto coefficients are not significantly different from one in the very upper tail to then become different from one for the simulated log-normal data, exhibiting a similar size and statistical significance to the real data.

When we look at the entire distribution (panel b) of Figure 4, the estimated Pareto coefficients are, unsurprisingly, constant over the entire distribution. They are flatter for the mixture data, displaying a long Pareto tail. Again, the simulated log-normal displays the same signature of the real data.

In Figure 5, we repeat the recursive non-parametric approach on the simulated data. Overall, for the largest 1000 cities in panel a), the KS statistics seem to indicate that the Pareto distribution is better than the log-normal, irrespective of the type of simulated distribution.

However, in the very upper tail the KS statistics for the Pareto and the log-normal are indistinguishable, just like for the real data. When we add smaller and smaller cities beyond the 1000th in panel b) of Figure 5, results show great concordance between the real data, the log-normal

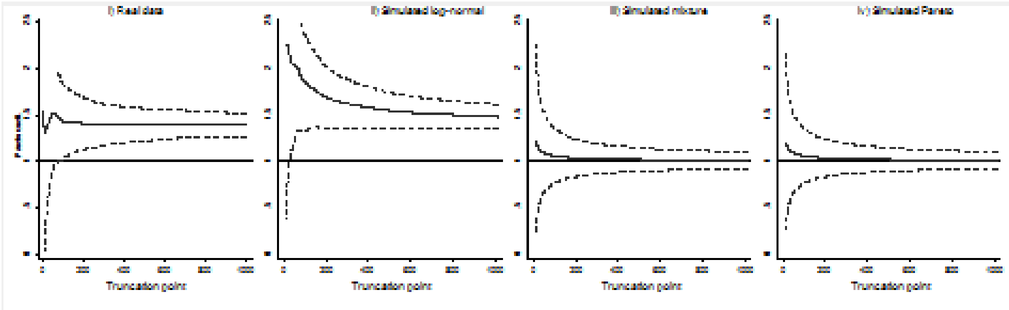
and the "mixture" distribution, with the latter unsurprisingly exhibiting a longer upper tail.

Both Zipf's Law and Kolmogorov-Smirnov tests seem to highlight the simulated log-normal as the most similar to the real data. This result seems to suggest that for a portion of the upper tail, and especially the distribution of the largest cities, the log-normal may be a close representation of the real data, as well as the Pareto. This evidence, however, has to be combined with that from Kolmogorov-Smirnov tests in the previous section, where the log-normal and the Pareto could both apply to the very first observations of the upper tail (around the largest 100 cities), but when the tail is extended, only the Pareto distribution is statistically indifferent from the real data. Hence, we are unable to unambiguously establish whether the distribution of cities falls in the weak or false power law category.

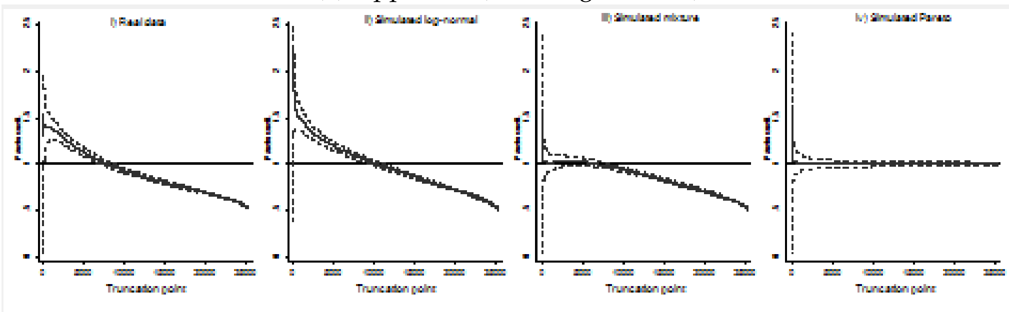
2.4 Conclusions

The identification of the correct city size distribution emerges from the literature as controversy with two most likely candidates: the Pareto and the log-normal distributions. Recently, some commentators (see Eeckhout, 2009, in particular) have suggested the possibility that part of this controversy may be due to the arbitrary choice of truncation of the distribution of all cities. A truncation point that is too high, or too low, may bias tests of the appropriate distribution of the upper tail. Also, a false power law may emerge when the extremes of the samples closely mimic

Figure 4: Rank-Size Regressions over simulated data

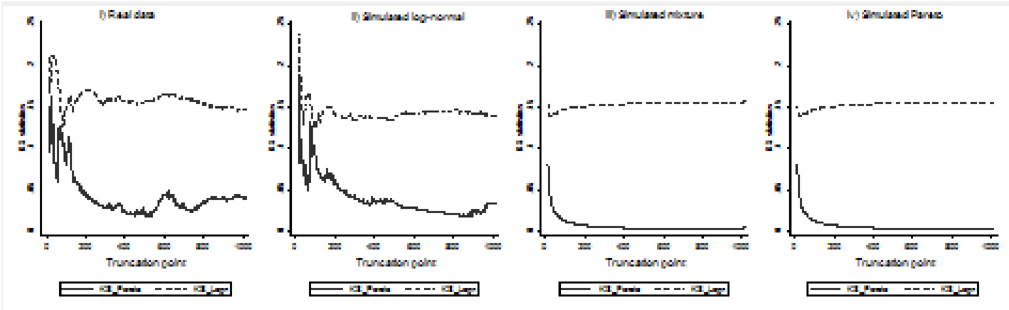


(a) Upper tail (1000 largest cities)

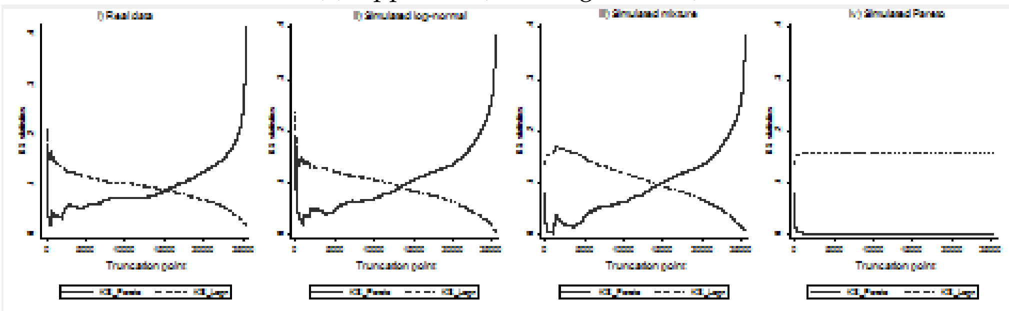


(b) All cities

Figure 5: Simulated KS



(a) Upper tail (1000 largest cities)



(b) All cities

an inverse power law, especially if drawn from a log-normal (Perline, 2005). Yet, this issues has been substantially overlooked in the economics literature and on the light of these facts the typical economic interpretations of the estimated Pareto coefficient lose their importance without an in depth analysis of the truncation point.

In this chapter, we investigate the city size distribution with respect to the "truncation point". In particular, we propose a recursive-truncation approach to reassess the common Zipf's Law regression and a non-parametric alternative proposed by Clauset, Shalizi and Newman (2009) against all possible truncation points of the entire distribution of cities.

Some interesting results emerge from this analysis. First, in line with previous results, we find evidence of Zipf's Law with the Pareto exponent equal to one in the very upper tail (above the largest 135 cities). However, when we start lowering the truncation point, adding smaller cities one at the time, we find that the size of Pareto coefficient first becomes statistically different from one and then starts decreasing, showing a non-monotonic behavior with respect to the truncation point. Statistically, the coefficient crosses one for different ranges of truncation points: in the upper tail and much later when the sample is truncated around the 8000th city. In this situation, we need to reject the classical Zipf's law (Pareto distribution with shape parameter equal to one) most of the time and when that estimated coefficient is equal to one, we showed that in general it lies in a region where the estimated coefficient is decreasing according to the chosen threshold. Finally, in line with Eeckhout (2004), the log-normal seems the best fit when the entire distribution of cities is considered.

The same recursive approach is also applied using the non-parametric method proposed by Clauset, Shalizi and Newman (2009). This analysis shows that the both distribution can apply to the upper portion of the distribution of cities (above around the first 100 cities). Beyond these cities, the upper tail conforms better to the Pareto distribution on the grounds of statistical significance. When the truncation point is extended to include smaller and smaller cities, it is not possible to disentangle the two on the grounds of statistical significance, but on the grounds of the size of the statistics. In general, the Pareto seems a better fit when we are close to the upper portion of the data and the log-normal seems a better fit when we approach the entire distribution. Also, the Pareto distribution seems to be longer than traditionally postulated by previous studies on the grounds of an arbitrary truncation point.

These results seem to support to the claim by Eeckhout (2009) that an arbitrary choice of the cut-off of the distribution may mislead scholars to conclude in favor of one or the other distribution. While the log-normal seems to best fit the entire sample, truncating the distribution may lead to conclude in favor of a Pareto, especially in the upper tail. Even then, however, the analysis returns a kind of non-monotonic behavior indicating that a Pareto might apply over more than one range of city sizes: in the very upper tail and when the sample is truncated mid-way to the distribution of all cities.

Finally, we assess whether the distribution of cities can potentially fall into the weak or false power law categories defined by Perline (2005). To this end, we replicate the proposed recursive rank-size and the non-parametric alternative test of city size distributions using simulated data drawn from a Pareto, a log-normal and a mixture of the two. While the rank size regressions seem to point to the distribution of cities as potentially a false power law with the log-normal simulated data displaying a remarkably similar signature to the real data, the non parametric test seems less conclusive. In this case, the size of the test statistics for the simulated log-normal and the real data are also remarkably similar in the very upper tail. Only when the tail is extended, the test seems to point again in favor of the Pareto. Hence, the non-parametric test seems unable to settle whether the distribution of cities is a weak or a false power law.

This analysis seems to provide novel methodological insights on the discrimination between alternative city size distributions and the truncation point problem. We are able to provide new and robust empirical evidence in support of the claim of Eeckhout (2009) that the conclusion in favor of the Pareto or log-normal may depend on the truncation point. Similarly, we provide further evidence of the difficulty highlighted in part of the literature of distinguishing between the tail of a log-normal and a power law tail for the population distribution of cities. However, we also provide novel empirical evidence compared to the existing literature. In particular, our recursive rank size regressions show that Zipf's Law applies not just for the part of the distribution involving the largest cities, but also for an intermediate range of cities that is far beyond the upper tail. Importantly, this evidence is confirmed in the non-parametric analysis. Moreover, we find that the so called "upper tail" of the distribution may be considerably longer than what is traditionally found by previous studies, a result that may have interesting implications for the theoretical and policy debate related to the distribution of cities.

Chapter 3

Are Gibrat and Zipf Monozygotic or Heterozygotic Twins? A Comparative Analysis of Means and Variances in Complex Urban Systems

3.1 Gibrat's Law vs Zipf's Law: Preliminary Considerations[1]

3.1.1 Preface

Cities all over the world offer an amazing variety in terms of size and growth rates. Despite these differences, systems of cities do not exhibit a

[1] The current chapter is a repruduction of the paper already under reviewing "Are Gibrat and Zipf Monozygotic or Heterozygotic Twins? A Comparative Analysis of Means and Variances in Complex Urban Systems", co-authored with Aura Reggiani of University of Bologna and Peter Nijkamp of VU University of Amsterdam.

random pattern, but a strict regularity in terms of urban hierarchies and inter-urban connectivity. The genesis of such hierarchical perspectives on city size and urban systems can already be found in the seminal contributions of Christaller (1933) and Lösch (1940). The validity of these frameworks has extensively been tested in subsequent statistical experiments in many countries around the world. The conceptual foundation for the existence of central place hierarchies rests on various pillars: agglomeration advantages in cities (depending on city size), smart specialization of industries (depending on scale advantages in different size classes of cities), and transportation and logistics costs (depending on distance frictions between cities or between cities and their hinterlands). Urban hierarchies and inter-urban connectivity are therefore two sides of the same coin (see Paelinck and Nijkamp 1976).

Clearly, it ought to be added that the spatial range of interurban linkages has extended drastically over recent decades. Whereas a century ago, most cities were at best part of an interlinked regional or national system, nowadays cities are often part of a globally connected network. The underlying globalisation force field is not a random system either, but strictly governed by economic efficiency determinants, in which globally connected service networks and commodity chains play a critical role (see for fundamental contributions Neal (2012), Newman (2010)). There is no doubt that we live in a highly connected world. This is highlighted in a study by Reggiani and Schintler (2005), who assert: "*Our modern world is in a continuous state of flux. Modern transport systems and the emerging new style behaviours have created an unprecedented rise in mobility, at all spatial levels. The ever rising mobility patterns apply to all types of movement work, business, shopping and leisure, as well as to freight transport. Globalisation certainly plays a key role in this dynamic framework*" (p.1). Globalization – which is defined as a broad area of increasing internationalization of markets, changing consumption patterns and shifting of industrial activities all over the world – appears to form a common denominator for consumer/user (economic) activities with an immediate impact, namely, the emergence of a highly interconnected, interdependent and complex system of networks, or the rise of a complex network society. Consequently, mobility and migration – though continuously changing – lead to the increasing phenomenon of urban agglomerations: "*Agglomeration and residential mobility of the population between different geographic locations are tightly connected to economic activities*" Eeckhout, 2004, p. 1429). Surprisingly, despite the complex evolution of current socio-economic spatial networks, two robust empirical regularities seem to hold: Gibrat's law affirming that city growth does not depend on size, and Zipf's law stating the proportionality of a given city size to its rank.[2]

In the field of spatial economics, these two regularities have given rise, especially since the late '90s, to an increasing number of empirical studies, testing cities and economic growth at various spatial levels (national, regional, local), by means of Gibrat's law and Zipf's law. In the majority of urban studies, Zipf's law and Gibrat's law are generally confirmed by empirical data. For example, Eeckhout (2004) and Gonzalez-Val (2010) test the validity of both laws for all US cities: the former in the period 1990-2000, and the latter over the period 1900-2000; Ioannides and Overman (2003) and Gabaix and Ioannides (2004) examine Gibrat's and Zipf's laws in the US metropolitan areas over the same period (1900-1990); and Giesen and Suedekum (2011) investigate both these laws, by considering a sample of German cities with a population greater than 100,000 in the period 1975-1997. In contrast, only a few studies seem to reject these two empirical regularities, in particular, Black and Henderson (2003), who reject Gibrat's law over a sample of US metropolitan areas in the period 1900-1990, and Gonzalez-Val et al. (2012), who in their study of all cities in the US, Italy and Spain over the 20th century, find weak evidence of Gibrat's law. These contrasting results have prompted a continuous debate in the literature, on the (non)validity of Gibrat's law and/or Zipf's law. In this context, recent work has shown that Gibrat's law can be generalized, in the sense that only the mean of the city growth

[2]Another way to refer to Zipf's Law is a Pareto distribution, with a shape parameter equal to 1. It is investigated using the so-called rank-size rule. We note here that the slope coefficient of the rank-size rule represents the inverse form of the parameter of the conventional Pareto distribution. For more details, we refer inter alia to Adamic (2000) and Parr (1985). In this chapter we refer to Zipf's law (Zipf's distribution), when the rank-size coefficient is exactly equal to 1. In all the other cases we refer to the rank-size rule (rank-size distribution).

is independent from city size, while its variance can change according to size (Cordoba, 2003). Gibrat and Zipf have offered strong evidence on population dynamics and hierarchies, respectively, boiling down to simple modeling perspectives on the evolution of urban systems in a given country.

These two laws are often analyzed together, given their possible complementarity. Indeed, Champernowne (1953) and Simon (1955) have shown that rank-size distributions arise naturally, if Gibrat's law is satisfied. Gabaix (1999) has demonstrated that Gibrat's law leads to a Zipf distribution, while Cordoba (2003) argues that a weak version of Gibrat's law leads to more general rank-size distributions. Clearly, the question emerges whether Zipf's law (or its more general rank-size rule) is only a static version of the more dynamic Gibrat's law. Strictly speaking, it might be hypothesized that Zipf's law and the rank-size rule show features of monozygotic twins, while Zipf's law and Gibrat's law show features of heterozygotic twins, as the latter law is able to reveal distinct modeling outcomes, even though in equilibrium it expresses identical fundamental characteristics. A test of this proposition calls for evidence-based research.

Starting from these considerations, the present chapter aims to answer the following research question: can (a generalisation of) Gibrat's law allow us to infer Zipf's law and vice versa, by empirically analyzing the link between these two laws, in the context of urban growth, and, in particular, the dynamics of city size distributions? In this framework, particular attention will be paid to the role of the mean and variance of the city population as a key indicator for assessing the validity (or non-validity) of the generalised Gibrat's law. Consistently with Eeckhout (2004), we focus our empirical investigation on the entire city size distributions of five selected countries (Botswana, Germany, Hungary, Luxembourg and Malta) and not only on the upper tail,[3] as other studies have done (see among others, Giesen and Suedekum, 2011; Guerin-Pace,

[3]Eeckhout (2004) shows that if the city growth does not depend on city size "*then the estimated OLS coefficient of the so-called rank-size rule varies depending on the truncation city size, i.e. the inclusion of smaller (larger) cities in the sample, leads to a smaller (larger) coefficient*" (Fazio and Modica, 2012, p. 3).

1995; Rosen and Resnick, 1980 and Soo, 2005). First, we are able to find evidence of the existence of Gibrat's Law for three out of the five pre-selected countries; we then test the empirical relationship between the (generalised) Gibrat's law and Zipf's law, by considering the dynamics of the hierarchical structure of the various city systems, on the basis of the mean and variance indicators.

The chapter is then organised as follows. In this section (Section 3.1), we first summarise Gibrat's law (Subsection 3.1.2) and Zipf's law (Subsection 3.1.3); we then address the literature on the relationship between these two laws (Subsection 3.1.4). This short review constitutes the basis for constructing our empirical analysis aimed at testing Gibrat's law vs Zipf's law in specific case studies. Section 3.2 describes the rationale underlying the selection of the five countries under analysis (Subsection 3.2.1 and 3.2.2), by focusing on their different spatial economic characteristics and related statistics (Subsection 2.3), while subsequent sections illustrate the results of the empirical analysis devoted to testing Gibrat's law (Section 3.3), as well as the link between Gibrat's law and Zipf's law (Section 3.4). The chapter concludes with some methodological considerations and directions for future research (Section 3.5).

3.1.2 Gibrat's Law

In 1931, Gibrat observed that the growth rate of a citys population does not depend on the size of the city. In other words, although cities can grow at different rates, no systematic behaviour exists between their growth and their size, so that, according to Gibrat (1931), we cannot affirm that larger cities grow faster than smaller ones or vice versa. More formally, we can say that a log-normal distribution (or Gibrat distribution) arises (if certain conditions hold) "*as the limiting distribution of the product of positive random variates and the number of terms in the product tends to infinity*" (Chesher, 1979, p. 403). Analytically, we can write the following logarithmic expression, as in Steindl (1968):

$$\log P(t) = \log P(0) + \varepsilon(1) + \varepsilon(2) + \ldots + \varepsilon(T) \tag{3.1}$$

where $P(t)$ is the size of a certain city at time t, $P(0)$ is the initial population, and $\varepsilon(t)$ is a random variable (indicating random shocks), i.i.d random variable with mean μ and variance σ^2. Equation (3.1) identifies the logarithm of the size of a given city as the sum of the initial size and past growth rates. It should be noted that this stochastic process leads to the log-normal distribution (1) of the variable $P(t)$, only if a sufficiently strong condition holds, namely, the law of proportionate effect. This law can now be interpreted as follows: "*A variate subject to a process of change is said to obey the law of proportionate effect if the change in the variate at any step of the process is a random proportion of the previous value of the variate*" (Chesher, 1979, p. 403). The implication of Gibrat's law is that the growth processes of cities have "*a common mean (equal to the mean city growth rate) and a common variance*" (Gabaix, 1999, p. 741), that is both the mean and variance have to be independent from the size of the cities.

In economic terms, the above formula means that, if we impose noise on a process, after a certain (long) time, the deterministic pattern generated by some fundamental structural variables appears again. This is important, if we want to build models close to empirics; and hence: "*economic models that explain city size distribution by relying on characteristics of hierarchies between cities, demand supply curves, technological considerations, and the like are at best incomplete if they fail to satisfy Gibrat's law in the end*" (Gabaix, 1999, p. 742).

However, Gibrat's law is only a part of the story. Indeed, this law is related to the rank-size rule, and, in particular, to Zipf's law. The link between the two regularities has been widely debated, because the proportionate growth rate process (Gibrat' law) gives rise, as Gibrat (1931) stated, to the log-normal distribution (see (1)); the proportionality of the size to the rank (Zipf's law, or the more general rank-size rule) is associated, instead, with a Pareto[4] distribution. Yet many studies have shown that a random growth process can generate power laws[5] (see Richardson, 1973, for a review).

[4]Note that the slope coefficient of the rank-size rule represents the inverse form of the parameter of the conventional Pareto distribution. See, for more details, Adamic (2007) and Parr (1985).

[5]The Zipf distribution is a particular distribution of the power law family.

In the scientific literature, the link between a log-normal and Pareto distribution is still a matter of controversy. According to several authors, the Gibrat (log-normal) process can generate Zipf's law only at the upper tail distribution (see e.g. Blank and Solomon, 2000; Eeckhout, 2004, 2009; Levy, 2009), as already outlined by Parr and Suzuki in 1973 (p. 343): "*truncation of the log-normal at an appropriately high level enables the truncated portion to be regarded as not significantly different from the rank-size distribution. Furthermore: these distributions all belong to the same family and in the upper tail are similar and consistent with the rank-size rule. Lower parts of these distributions, however, often exhibit significant differences and frequently do not conform to the rank-size rule*" (Carrol, 1982, p. 5). However, Gabaix (1999) has shown that the presence of an infinitesimally small barrier, on an identical growth process across sizes, necessarily leads to a Zipf distribution (see also Subsection 1.4 below). Against this methodological background, we briefly review in the next section the formal definition of Zipf's law and its spatial economic interpretations.

3.1.3 Zipf's Law

The second well-known spatial regularity is given by the so-called Zipf's law (on the basis of a first study by Auerbach[6] in 1913). In 1949, Zipf observed and established that the sizes of the cities in a country are proportional to their rank. This means that for example, in Botswana, the size of Gaborone is roughly twice the size of Francistown, the second largest city, three times the third largest city, Molopolole, and so on. Formally, this can be written as:

$$P = KR_i^{-\rho} \tag{3.2}$$

Equation 3.2 is known as the rank-size rule and is usually expressed in logarithmic form, as follows:

$$\log P_i = \log K - \rho \log R_i \tag{3.3}$$

[6] "*The population of a city is inversely proportional to the number indicating its rank among the cities of a given country*" (Auerbach, 1915, p. 384).

where P_i is the population of city i, R_i is the rank of the ith-city and K is a constant. Zipf's law holds precisely, when the coefficient ρ is equal to one.[7] Several interpretations of the Zipf coefficient, ρ, have been proposed in the literature. In principle, the ρ-coefficient can be seen as an indicator of the hierarchical degree of a system of cities (Singer, 1930). In fact the ρ-coefficient measures how unequal the city distribution is: the higher the ρ-coefficient, the more unequally distributed is the city system. On the contrary, the smaller the value of ρ, the more even is the system of cities (in the extreme, when $\rho = 0$, we have a very even system of cities all of the same size; when $\rho = \infty$, instead, we have only one city hosting the entire population). Related to this interpretation, Singer (1930) considers this ρ-coefficient as an index of metropolisation; in particular, this author affirms that the higher the ρ-coefficient, the more important is the urban land value. In the same vein, Parr (1985) shows the existence of a U-shaped pattern of the Pareto coefficient (the counterpart of the ρ-coefficient) over time; he establishes a link between the overall level of development and the degree of metropolisation, that is, the more developed a country, the greater will be the degree of metropolisation. Interestingly, Parr (1985, p. 208) argues: "*the process of concentration is facilitated by (and ultimately dependent on) improved interurban and interregional transportation*". Against this background, Brakman et al. (2001) interpret ρ as an indicator of industrialisation and agglomeration economies, while more recently, Reggiani and Nijkamp (2012) have considered the urban structure in a country as a socio-economic connected network and interpret ρ as an indicator of economic development by comparing it with the connectivity degree distribution[8] in a network.[9]

[7]Notice that here $\rho = 1/q$ of the previous chapter.

[8] "*The degree distribution, $P(k)$, gives the probability that a selected node has exactly k links*" (Barabasi and Oltvai, 2004, p.102).

[9]In particular, Reggiani and Nijkamp (2012) show the following: high values of the connectivity degree distribution (> 3) match a random network, which corresponds to a homogeneous urban setting, characterized by low values of ρ (< 0.5). Vice versa, small values of the connectivity degree distribution (< 0.5) indicate a hub configuration in the network, which corresponds to a high urban heterogeneity, expressed by a value of $b > 1$.

In summary, Gibrat's law expresses the growth process of a certain variable (firm, city, income, wealth, etc.), independent of its size, while Zipf's law presents the static relationship of the size of this variable with its rank. The main concern in the literature has been whether and how Gibrat's law and Zipf's law are mutually linked, for example, by means of the same formal model of the size distribution of the variable in hand, within a given urban-spatial system. In the next section, we briefly outline the main methodological challenges which constitute the basis for our subsequent empirical analysis.

3.1.4 The Relationship between Gibrats Law and Zipfs Law

"*It has long been noted, in economics since at least Champernowne (1953),..., that random growth process could generate power laws*" (Gabaix, 1999, p. 741). The debate on the two distributions is still very alive, mainly because it is a difficult task to identify the most discriminating between these two, at least in the upper tail of the distribution (Perline, 2005). As mentioned in Subsection 3.1.1, both are similar and consistent with the rank-size rule in the upper tail (Carrol, 1982).

Many arguments have been put forward in the literature for explaining distinct variations between these distributions. Here, we focus on stochastic models, "*perhaps the most influential theories bearing upon the rank-size problem*" (Carrol, 1982, p. 5). The stochastic approach mathematically derives certain distributions of city size by starting from a given stochastic process; more precisely, these models postulate that different distributions arise as the result of steady state outcomes of different stochastic processes describing the underlying economic forces (Steindl, 1968). In particular, these models postulate that rank-size regularities are the steady-state equilibria of the law of proportionate effect, namely Gibrat's law. Then proportional growth can explain the Pareto/Zipf distribution. In this section, we review the main models which have influenced subsequent studies; for an extensive review, the reader can consult Carrol (1982) and Suarez-Villa (1988).

Champernowne (1953) presents a model on income distribution (this can be adapted to city size as well) based on the stochastic process from Markov chains.[10] The model takes into consideration a stochastic matrix, namely, a matrix where all probabilities of transition from a class of income to another are reported. Furthermore, the model makes the assumption of a constant number of incomes through time: "*Under these circumstances, and provided certain other conditions*[11] *are satisfied, the distribution will tend towards a unique equilibrium distribution dependent upon the stochastic matrix but not on the initial size*" (Champernowne, 1953, p. 318). This model is able to explain the Pareto distribution for income size.

Simon (1955) proposes, instead, an interesting study on the class of skew distribution functions. He shows how a preferential attachment process, that is a law of proportionate effect with a constant probability that new small units enter at any time in the process, leads to a class of skewed distributions, incorporating among them both the Pareto and log-normal distribution.

Gabaix (1999, p. 750) presents a model with a random walk with a small barrier "*and, more interestingly, that, as the barrier become lower, the exponent (of the Pareto distribution) converges to 1. So, in essence, all we need is an infinitesimally small barrier, to ensure that the steady state distribution will be Zipf*". Thus, Gabaix's model shows that a proportional growth process can lead to an exact Zipf distribution.

All of the above mentioned studies have provided one important contribution to the literature: under plausible conditions,[12] a proportionate growth process can lead to a Pareto/Zipf distribution. In summary, Champernowne's model (1953) established that a proportional growth process leads to the limit of a Pareto distribution. Simon's model (1955) generalises Champernowne's results, proving that Gibrat's law can lead to different skewed distributions, one of which is Pareto. Finally, Gabaix

[10] Recall that Gibrat's law can be interpreted as a Markov process.

[11] In particular, the stability condition used in this model implies a quite unrealistic negative expected value of a change in income (Steindl, 1968).

[12] In particular, a) negative expected change of income in Champernowne (1955); b) steady inflow of new and small cities in Simon (1955); and c) lower barrier for Gabaix (1999).

(1999) shows that Gibrat's law can lead to an exact Zipf distribution (see also Cordoba, 2003).

However, despite the extensive literature on Gibrat's law, according to Kalecki (1945), the model proposed by Gibrat, although formally correct, has implications that are not very realistic, particularly in relation to the variance.[13]

Kalecki (1945) assumes that the variance of the random variable $\varepsilon(t)$ in Eq. (3.1) changes over time for three reasons: (i) the variance may change due to economic forces only; (ii) the variance may increase purely due to the influence of random shocks, and (iii) the variance may change due to both economic forces and random shocks. He thus proposes a model in which there is a negative correlation between the size P and the random variable $\varepsilon(t)$ (in Eq. (3.1)). In this way, as size increases (as time grows), the random shocks are smaller, thus preventing the tendency of the variance to increase. In this context, Cordoba (2008, p. 1463) proposes a: "*generalization of Gibrat's law that allows size to affect the variance of the growth process but not its mean*". In particular, one of the implications of Cordoba's generalized model is that non-proportionality of the variance is required to take into account a ρ-coefficient different from one (in Eq. (3.3)). More specifically, the larger the ρ-coefficient, the more unequal is the distribution, and this makes a growth process more volatile.[14] On the basis of Cordoba's results, we can outline the following relationships between Zipf's law and Gibrat's law:

- (a) If $\rho = 1$, Zipf's law holds. In order that Gibrat's law applies, neither the mean nor the variance of growth can depend on size.

- (b) If $rho > 1$, the distribution is more unequal. In order that Gibrat's law applies, it is necessary that the mean is independent of the city size, but not the variance; indeed, the associated growth

[13] The law of proportionate effect, indeed, assumes that a given variable $P(t)$ changes by a random and independent effect amount as time goes by; so even though it is realistic that the mean remains constant as t increases, the same is not true for the variance because it should increase instead.

[14] The volatility is a measure of fluctuation of a process. We will use the variance as an indicator of the volatility of an underlying proportionate growth process.

process requires that smaller cities face a greater volatility of growth than larger cities.

- (c) If $rho < 1$, the distribution is more evenly distributed. Again, in order that Gibrat's law applies, it is necessary that the mean is independent of the city size, but not the variance. Here, the associated growth process requires that larger cities face a greater volatility of growth than smaller cities.

Starting from these considerations, the main challenge and aim of this chapter is to empirically explore the above mentioned relationship "Gibrat's law vs Zipf's law", according to the three statements (a)-(c) above. It should be noted that empirical investigations of the relationship "Gibrat's law vs Zipf's law" are still rare. We may refer here, amongst others, to studies on firm growth (see Fujiwara et al. 2003), and studies on city growth (Berry and Okulicz-Kozarin 2012; Dittmar 2011; Giesen and Suedekum 2011). All of these analyses have mainly focused on the first item, (a), the relationship between Zipf and Gibrat. To the best of our knowledge, no empirical application has focused, so far, on how (partial) deviations from either law can affect the other law; in other words, on statements (b) and (c).

Given these propositions, we aim to test items (a)-(c) in the context of urban dynamics, by examining urban patterns in countries which exhibit different spatial-economic characteristics. Our methodology includes two main steps: (i) we first analyse Gibrat's law for the set of cities in the selected countries (Section 3.3), in order to (ii) construct a validation framework to assess the relationship Gibrat's vs Zipf's law (Section 3.4). In Section which follows, we will first describe the characteristics of our case studies, in particular the countries to be investigated.

Table 2: Spatial Economic Characteristics of the Five Countries under Analysis

Country	Year	Km^2 (thousands)	Pop. (milions)	Density	% Urban pop.	% Pop. growth	Car (per (1000 people)	Railway (1000 Km	Roadway Km	GDP p.c.	Growth	Total Investment
Botswana	2011	581	1.85	3.19	62.00%	1.47	133	0.90	25.80	9,481	5.1%	21.52%
Germany	2007	357	81.78	229	74.00%	-0.20	564	41.90	644.50	40,403	2.7%	19.26%
Hungary	2011	93	9.99	107.4	69.00%	-0.18	347	8.10	197.50	3,045	1.70%	19.07%
Luxembourg	2011	2,5	0.51	205.6	85.00%	1.13	739	0.27	5.20	106,958	1.00%	21.17%
Malta	2009	0.32	0.42	1,338	94.00%	0.36	679	0	3.10	20,437	-2.70%	15.80%

3.2 Choice of Case Studies: Descriptive Analysis and Statistics

3.2.1 Preface

Given our empirical objective, aiming to analyse Gibrat's vs Zipf's law in different spatial socio-economic landscapes, the rationale underlying the choice of our case studies is the following. We have selected five distinct countries characterised by different typologies, according to five main distinguishing criteria: (a) OECD vs non-OECD country; (b) advanced economy vs non-advanced economy; (c) centrally located vs non-centrally located; (d) size of surface; and (e) increasing vs decreasing population. On this basis, we have identified five countries: Botswana, Germany, Hungary, Luxembourg and Malta for study.[15] In Table 2 we report, for each country, some economic indicators (such a GDP per capita, growth rate and percentage of investment over GDP), as well as some other important indicators for the mobility and transportation system (such as the length of railways and roadways, and the number of cars per thousand people).

Some points are worth noting here. Botswana is the only non-OECD country, while all the others are OECD countries. Botswana shows the features of a non-advanced[16] economy; however, it exhibits a trend towards an increase in population and economic growth. Germany was a founding member of the European Community in 1957 (which became

[15] Alphabetic order.

[16] According to the IMF classification.

the European Union (EU) in 1993); it is central in Europe and is a large country in terms of surface area and population, with an advanced economy. Hungary joined the EU in 2004; it is located in central Europe, but shows a non-advanced economy and a decreasing population. Luxembourg, like Germany, was a founding member of the European Community in 1957; it is a small country, but very central in Europe with a high income per capita. Malta, like Hungary, joined the EU in 2004; it is a small, peripheral country but with an advanced economy; it is also one of the most densely populated countries worldwide. Clearly, other choices could have been made, but the present set of countries should represent a sufficiently interesting collection of cases for in-depth investigation.

We will now concisely outline the major economic and demographic characteristics of these countries (Subsection 3.2.2). We will then show some descriptive statistics and we will offer a descriptive analysis (Subsection 3.2.3).

3.2.2 Profiles of Case Studies

In this section we will describe a few characteristic features of the urban system of five case countries further analysed in our study.

Botswana is the largest country in the sample in terms of size: it is around 581,000 sq. km, but with a very low population density: 3.19 inhabitants per sq. km. Furthermore, it is poor in terms of transport infrastructures and of number of cars per 1,000 people; the urban population covers only 61

Germany, instead, is the most developed country from an economic point of view. It has the largest population in Europe and is also large in terms of surface area, 357,000 sq. km. Moreover, it is the second densest country in our case studies. It also has very good infrastructures.

Hungary is a mid-sized country with a low degree of transport infrastructures and a low average number of cars per inhabitant. The population density is low and also the number of urbanised people is low.

Luxembourg is a small but rich country, highly connected with the rest of Europe. It has one of the largest ratios of number of cars to population size. In terms of GDP per capita it is the richest country in our sample.

Finally, Malta is the smallest country of our data-set in terms of both population and surface area; however, the population density is huge and this is reflected in the number of urbanised people.

3.2.3 Data and Descriptive Statistics

We collected data from the National Institute of Statistics for all five of these countries.[17] In particular, we collected data from the Central Statistics Office of Botswana, the Institute for Employment Research[18] (IAB) in Germany), the Hungarian Central Statistical Office, the STATEC-Institute National de la Statistique et des Etudes Economique of Luxembourg, and the National Statistique Office of Malta.

It should be noted that an extensive debate concerns the type of unit under analysis: several studies have been carried out using metropolitan areas, i.e. by considering the entire population in a given city, as well as all populations of suburban areas. Nevertheless, here we aim to carry out an analysis as comparable as possible between the five countries, by also including, as much as possible, all the cities in a given country. For these reasons, we consider in our analysis the entities legally defined as cities or villages in their countries, although we are aware that the administrative definition given by legal borders might not fulfill our scopes exactly. In order to have a comparable unit, in all countries we have selected those localities which are similar to a municipality.[19]

Another concern is due to the fact that we have different temporal horizons, which, sometimes, are short. This is the case for Botswana, where we have only two census observations (2000 and 2010), as well as for Hungary, where, although the time span is 30 years (1980-2011), we have only four census observations. For Germany, however, although

[17] For all countries we have data over all cities from the biggest to the smallest one.

[18] The authors wish to thank Uwe Blien and Anette Haas (IAB, Germany), for kindly providing the data used in our study on German cities (Sections 3.3 and 3.4).

[19] We encountered some difficultly in making the right choice for Botswana where we also had data for small localities but we chose to collect all localities with ID code 100, namely villages and cities.

Table 3: Descriptive Statistics of the Five Countries under Analysis

Country	Year	N. cities	ln (Mean)	ln (Variance)	ln (Median)	Skewness	Kurtosis
Botswana	2011	461	7.15	1.22	6.97	0.78	2.23
Germany	2007	12,262	7.42	1.50	7.30	0.34	0.17
Hungary	2011	3,154	6.75	1.34	6.70	0.40	0.95
Luxemburg	2011	116	7.81	0.92	7.61	0.86	1.45
Malta	2009	68	8.35	0.93	8.29	-0.59	0.31

the time span is 15 years, we have annual data (1993-2007), so we can conduct a more precise analysis. Finally, Luxembourg and Malta have a long time series considering all census data from 1821 until 2011 for Luxembourg, and from 1901 to 2009 for Malta; unfortunately the time span between two subsequent observations is not constant even in the same country and especially before World War II.

In Table 3 we report some descriptive statistics. First, it is interesting to notice that the median for the five different countries reflects the percentage of urban people in any single country: Malta shows the largest median, 3,983 inhabitants, and we recall that the urbanisation in this country is over 90%; Luxembourg has a median of 2,018 people and an urban population of over 80%. Hungary shows the smallest median, 812 people, with a relatively low urbanization rate, 68%.

Secondly, we report the (log) mean. In all five countries the mean is greater than the median which indicates asymmetry. The asymmetry of the distribution is also confirmed by the skewness. In general, in city size probability distributions, we expect a positive skewness, which denotes the fact that most of the observations lie on the left of the distribution. As a consequence, the left tail is longer: this means that we expect a greater number of small cities than large ones. In our sample this is true for four countries: Botswana, Germany, Hungary and Luxembourg, while Malta shows a negative skewness,[20] thus indicating a greater concentration of

[20]Negative skewness might be seen as a contradiction with a mean greater than the median, but this is not the case. For example, von Hippel (2005) argues: "*the mean is right of the median under right skew, and left of the median under left skew. This rule fails with surprising frequency... It can fail in distributions where one tail is long but the other is heavy. Most commonly, the rule fails in discrete distributions where the areas to the left and right of the median are not*

large cities. The reason for this is likely due to the huge urbanization rate in that country that pushes the concentration of people into large centres and this is magnified by the small size of Malta.

Kurtosis is a measure of peakedness of the distribution. In all our case studies, the value of kurtosis is positive: this indicates a situation in which the distributions show heavy tails and peakedness with reference to a normal distribution (whereas negative kurtosis indicates light tails and flatness). We notice the presence of fat tails and peakedness especially in Botswana (which shows the highest positive kurtosis[21]).

Following on from the above observations, in Section 3.3 we focus our attention on the validity of Gibrat's law, in order to design an analytical framework that is useful for meeting the ultimate goal of our analysis: a comparison of Gibrat's law and Zipf's law.

3.3 Testing Gibrat's Law: Method and Results

3.3.1 Preface

Given our objective first to test the validity of Gibrat's law in the five countries concerned (Botswana, Germany, Hungary, Luxembourg and Malta), in this section we show the adopted methods and the related results. In particular, we consider two different methodologies: (i) parametric analysis (Subsection 3.3.2); and (ii) non-parametric analysis (Subsection 3.3.3). Concerning the parametric technique, we search for deviations of sizes from their mean, first in a cross-sectional setting (Model A in Subsection 3.3.2) and then in a longitudinal setting (Model B in Subsection 3.3.2). Concerning the non-parametric technique, we look at the mean and variance of the size, given the initial size (Subsection 3.3.3).

equal". Our case covers both.

[21] De Carlo (1997, p. 294), in an interesting note on kurtosis, affirms: "*it represents a movement of mass that does not affect the variance. Consider the case of positive kurtosis, where heavier tails are often accompanied by a higher peak. Note that if mass is simply moved from the shoulders of a distribution to its tails, then the variance will also be larger. To leave the variance unchanged, one must also move mass from the shoulders to the centre, which gives a compensating decrease in the variance and a peakl*".

3.3.2 Parametric Analysis

Model A: OLS Regression

In this section we use an OLS regression model and report the results from the parametric analysis. We check dynamic deviations from the proportionality of mean growth and variance to size, by using a method firstly proposed by Kalecki (1945) and subsequently utilized, among others, by Bottazzi et al. (2001). In particular, the adopted model is the following OLS model:

$$g_i^t = \beta_i g_i^{t-1} + \nu_i \tag{3.4}$$

where g_i^t is the deviation of the logarithm of the population of city i from the mean of the logarithms of the city populations at time t, and ν is the error term. β is the parameter to be estimated and "*provides an estimate of the divergence/convergence of the size distribution toward its mean*" (Bottazzi et al., 2001, p. 1184). Gibrat's law holds if β is equal to one.[22] When β is lower than one, this means that size converges towards its mean, namely, the larger a city, the smaller the expected growth. On the contrary, when β is greater than one, the larger a city and the larger the expected growth. We test the model for each time-step.

As an indicator of the volatility of the growth process, we use the variance ratio, θ_t, between the variance of g at time t, $\sigma(g_t)$, and the variance of g at time $t-1$, $\sigma(g_{t-1})$:

$$\theta_t = \frac{\sigma^2(g_i^t)}{\sigma^2(g_i^{t-1})} \tag{3.5}$$

If the variance is stable along two subsequent years, the variance ratio, will be close to unity.

Results of the estimation of Eq. 3.4 and Eq. 3.5 are reported in Table 4 and Table 5 . Two main conclusions arise from here. Firstly, looking

[22] We report here only the condition on the estimated β. However it should be noted that another condition is necessary to affirm that Gibrat's law is in operation, indeed the error terms have to be serially uncorrelated. We, then, add one more lag in Eq. (3.4) to verify this additional condition. In most of the cases the error terms result serially uncorrelated.

Table 4: Model A Estimates (Countries: Botswana, Germany, Hungary and Luxembourg; Different Years)

Country	Year	β	Robust s.e.	θ_t	R2	N. obs.
Botswana	2011	.995**	.0149	1.0782	.92	460
Germany	1994	.999	.0003	1.0009	.99	12,280
	1995	.999*	.0002	.9983	.99	12,291
	1996	.999*	.0003	.9987	.99	12,291
	1997	.998	.0002	.9978	.99	12,291
	1998	.998	.0002	.9984	.99	12,291
	1999	1.00 ***	.0003	1.0011	.99	12,293
	2000	.999***	.0003	.9998	.99	12,294
	2001	1.001	.0001	1.0024	.99	12,294
	2002	1.001	.0001	1.0017	.99	12,294
	2003	1.00*	.0001	1.0009	.99	12,293
	2004	1.001	.0001	1.0027	.99	12,292
	2005	1.001	.0001	1.0031	.99	12,293
	2006	1.001	.0001	1.0027	.99	12,993
	2007	1.001	.0002	1.0039	.99	12,259
Hungary	1990	1.055	.0019	1.1215	.99	3,121
	2001	1.034	.0022	1.0806	.99	3,121
	2011	1.028	.0025	1.0674	.99	3,121
Luxembourg	1851	.941*	.0264	.9557	.93	116
	1871	.993**	.0276	1.0779	.92	116
	1880	1.005 **	.0306	1.0667	.95	116
	1890	1.048	.0183	1.0855	.93	116
	1900	1.109	.0324	1.2833	.96	116
	1910	1.079	.0187	1.1803	.99	116
	1922	1.01**	.0123	1.0473	.97	116
	1930	1.103	.019	1.2406	.98	116
	1935	1.001***	.007	1.0059	.99	116
	1947	1.014*	.0061	1.0355	.99	116
	1960	1.069	.0132	1.1725	.98	116
	1970	1.044	.0156	1.1198	.97	116
	1981	1.016**	.0144	1.0576	.98	116
	1991	.993 **	.0104	.9964	.99	116
	2001	0.955	.0076	.9199	.99	116
	2002	0.994	.0015	.989	.99	116
	2003	.995**	.0034	.9926	.99	116
	2004	.997**	.0018	.9947	.99	116

Table 5: (...continued) Model A Estimates (Countries: Botswana, Germany, Hungary and Luxembourg; Different Years)

Country	Year	β	Robust s.e.	θ_t	R2	N. obs.
Luxembourg	2005	.996**	.0024	.9926	.99	116
	2006	.997**	.0015	.9961	.99	116
	2007	0.992	.0029	.9851	.99	116
	2008	0.992	.0023	.985	.99	116
	2009	.996**	.0019	.994	.99	116
	2010	.996*	.0016	.9924	.99	116
	2011	.999**	.0015	1.00	.99	11
Malta	1921	0.836*	.0770	.7891	.89	51
	1931	1.008**	.0151	1.0306	.99	54
	1948	.912**	.0582	.9684	.86	54
	1957	1.018*	.0249	1.0808	.96	55
	1967	.987**	.0163	.9853	.99	56
	1985	.936**	.0423	.9732	.90	59
	1995	.941*	.0265	.9407	.94	63
	2000	.991**	.0057	.9840	.99	67
	2001	.997**	.0015	.9959	.99	68
	2002	.995*	.0010	.9920	.99	68
	2003	.996	.0008	.9931	.99	68
	2004	.997	.0008	.9950	.99	68
	2005	.992**	.0220	1.0055	.98	68
	2006	.998**	.0013	.9967	.99	68
	2007	1.01	.0035	1.0210	.99	68
	2008	1.001**	.0012	1.0027	.99	68
	2009	1.003*	.0016	1.0069	.99	68

at the parameter β, Gibrat's law does not always hold (over time). Germany, Luxembourg and Malta (especially the latter two) are clear examples of this intermittency: we can see periods where Gibrat's law holds and others where it does not. Secondly, it seems that the effect of the (non)validity of the law is lengthy: namely, Gibrat's law in general holds (or does not hold) continuously for two or three time windows. Consequently, the first result of our analysis is that testing Gibrat's law requires a data-set of considerable length, or as many possible observations as we can. It is also interesting to note that in general, when Gibrat's law does not apply, the variance at time t is higher than the variance at the previous times; this is denoted by the parameter θ in Eq. 3.5 greater than one. This is, of course, consistent with the idea of Gabaix (1999) which, according to Gibrat's law, both the mean and variance of the growth rate have to be independent with respect to the size.

In more detail, by observing β in Tables 4 and 5 , we can see that in Botswana, Luxembourg and Malta, Gibrat's law holds quite often ($\beta = 1$). In Germany it does not apply more than half of the time and in Hungary Gibrat's law never holds.

This analysis, although intuitive, presents some shortcomings due to the fact that it does not allow all the temporal aspects to be included. In particular, even though it is good to have yearly observations, some doubt about the validity of the above method can arise when the time windows are longer. For this reason, as previously anticipated, we utilize a further parametric technique (model B), firstly suggested by Clark and Stabler (1991) and adopted, among others, by Black and Henderson (2003). The emerging results are illustrated in the next section.

Model B: The Unit Root Test Approach

The starting point of the second, complementary, parametric method is that testing for proportional growth implies testing for a unit root process. In other words, we are looking for the presence of mean reversion in the stochastic growth process. Mean reversion is a mathematical concept denoting that a process of both high and low growth is temporarily present and that population growth will tend to move to the average

Table 6: Model B Estimates for the Five Countries

Country	γ	Robust s.e.	Time span	N. obs.
Botswana	.99569**	.0149	10	460
Germany	1.00026	.0002	15	172,049
Hungary	1.03855	.0013	31	9,363
Luxembourg	1.00919**	.0065	190	2,900
Malta	.97780**	.0113	108	1,071

growth over time. Following Black and Henderson (2003) we use:

$$\ln(P_i^t) = \alpha + \delta^{t-1} + \gamma \ln P_i^{t-1} + \epsilon_i^t \tag{3.6}$$

where, α is a constant, δ^{t-1} are fixed time effects, P_i^t is the population of city i at time t and $t-1$. γ is the parameter to be estimated and if Gibrat's law holds, it implies that $\gamma = 1$.[23] Black and Henderson (2003) argue that this null hypothesis does not permit an auto-regressive process to the error so pooling OLS suffices. The results are summarised in Table 6.

In general Gibrat's law holds for those countries where Gibrat applies cross-sectionally more than a half time:[24] Botswana, Luxembourg and Malta. In Germany, although the estimated parameter, γ, is very close to unity, it is significantly different from one because of the very small standard error. Finally, in this case, Hungary does not indicate the presence of Gibrat's law.

In summary, according to both parametric methods A and B, the urban dynamics of Botswana, Luxembourg and Malta seem to capture Gibrat's law, while Gibrat's law does not appear to be appropriate for the urban evolution underlying Germany and Hungary.

Given this preliminary analysis of Gibrat's law, based on parametric methods, we explore the validity of these first results, by means of

[23] A value of γ lower than one implies mean reversion.

[24] See Tables 4 and 5.

a non-parametric analysis. Non-parametric analysis provides an important tool to explore directly the independence of the mean and variance of the growth from the size. In this way we can gain an indication of the behaviour of the mean and variance.

3.3.3 Non-Parametric Analysis

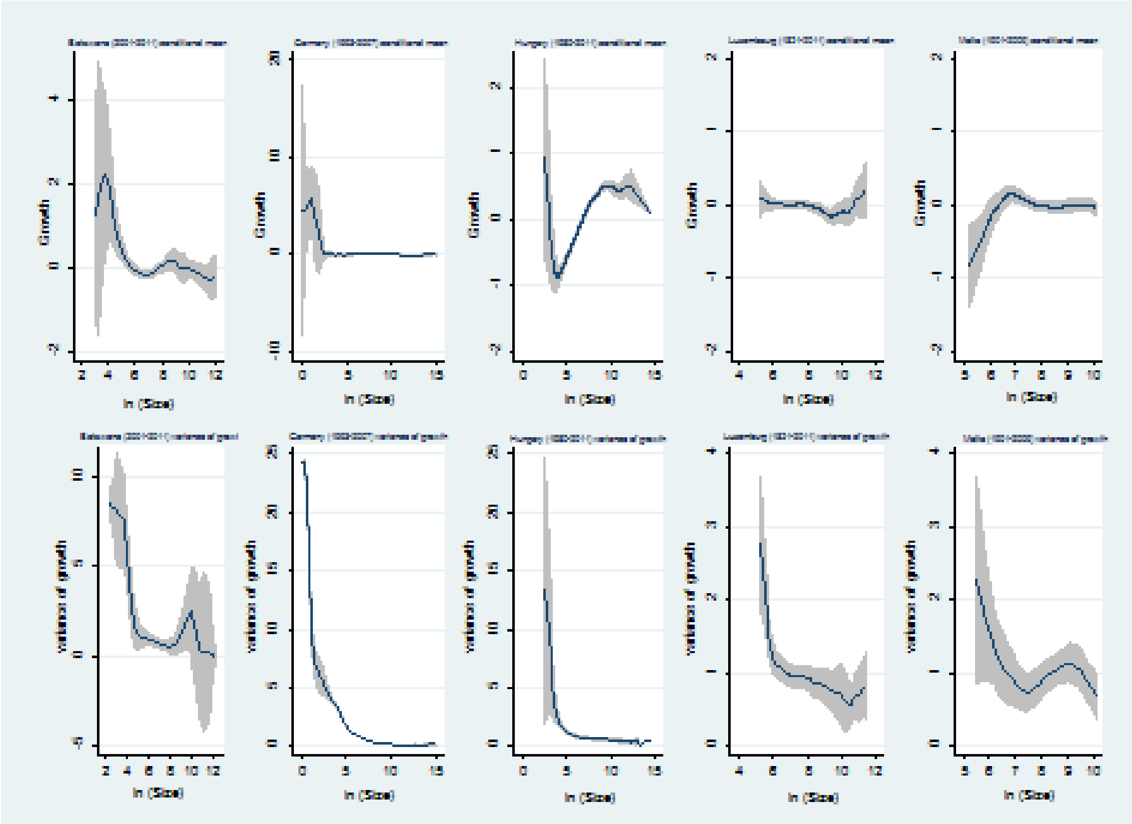
In this section we show the results of the non-parametric analysis strictly based on Ioannides and Overman (2003). We use the normalized growth rate, namely the difference between a city's growth rate and the mean city growth rate, all divided by the standard deviation of growth. The strength of non-parametric estimation is that we do not impose any relationship between the dependent and independent variables. According to Cameron and Trivedi (2005, p. 294), we: "*let the data show the shape of the relationship*"; this is an especially convenient approach when we do not know *a priori* the correct distribution of the data. In our analysis, we will use the Nadaraya-Watson (NW) method (Nadarya 1964; Watson 1964), where the bandwidths are calculated with an optimal rule of thumb.[25]

If Gibrat's law holds, the non-parametric estimation of the conditional mean and variance should be stable across different population sizes. Furthermore, because of normalization, we expect the conditional mean growth to be equal to zero, and the conditional variance of growth equal to one. It should be noted that, while the standard parametric regression methods provide only an aggregate relationship between growth and size which is constrained to hold over the entire distribution of city sizes, the non-parametric estimates allow the growth to vary with size over the distribution.

In Figure 6 we show the NW estimator for conditional mean growth (upper panel) and variance (lower panel) for the entire city size distribution. Following Cordoba (2003), the independence of the expected conditional growth rate always has to be satisfied, while the variance can be affected by the city size. In general, smaller cities face a faster growth than larger ones. However, very quickly (in most cases), the conditional

[25] We refer readers to their papers for a more detailed description.

Figure 6: Non-parametric Estimator for Conditional Mean and Variance (Countries: Botswana, Germany, Hungary, Luxembourg and Malta; Different Years)



mean appears to become stable. This evidence is consistent with the model of Gabaix (1999) where a truncation concerning the small cities is necessary to have stationarity.

By considering the specific countries in Figure 6, Luxembourg and Malta show a striking stationarity in mean and variance, so that we can affirm that in these cases, Gibrat's law holds. Botswana appears to face a huge variability at the bottom of the distribution for both mean and variance, but in the upper tail the evidence of independence of the mean is striking too. Thus, in principle, we can also accept Gibrat's law for this country. We can, on the contrary, reject Gibrat's law for Germany and Hungary, by confirming the results emerging from the parametric analysis (see Subsection 3.3.2).

This non-parametric test suffers from several shortcomings: firstly, outliers have a huge impact on the results, in particular the variance (Eeckhout, 2004; Gonzalez-Val et al., 2012); and secondly, we cannot directly compare the ρ-coefficient with Gibrat's law.

For these reasons, after the previous tests that confirm Gibrat's law for Botswana, Luxembourg and Malta, the final methodological step is the analysis of the relationship between Gibrat's law and rank-size/Zipf's law, by considering, as key indicators, the dynamics of the ρ-coefficient, the β-parameter and the θ-parameter; in other words, the evolution of the dynamics of economic development of the country. This will be reported in the next section.

3.4 Gibrat's Law and Zipf's Law: A Comparative Study

3.4.1 Role of the Adopted Parameters

In the previous sections we have shown that Botswana, Luxembourg and Malta seem to obey Gibrat's law, while this seems not to be the case for Germany and Hungary. The final step in our analysis is then the investigation of the relationship between Gibrat's law and the rank-size/Zipf's law, by means of the rules (a), (b) and (c) (outlined in Subsection 3.1.4).

From the operational viewpoint, we investigate the relationship between the ρ-coefficient in Eq. (3.3) and the estimated parameters β and θ from Eqs. (3.4) and (3.5), on the basis of Cordoba's propositions (a); (b) and (c) (Subsection 3.1.4). In particular, we estimate the ρ-coefficients in the rank-size rule (3.3) by means of a modification proposed by Gabaix and Ibragimov (2011), according the following:

$$\log P_i = \log K - b \log R_i - 0.5 \tag{3.7}$$

where P_i, K, q and R_i are the same as in Eq. (3.3). In Tables 7 and 8 we report the estimated ρ-coefficients and the parameters β and θ, according to Eqs. (3.7), (3.4) and (3.5), respectively, for each of the five countries.

Concerning the coefficient ρ (Eq. 3.7), it should be noted that we interpret the ρ-coefficient as a measure of hierarchy of city size distribution. In this sense a positive change in the estimated ρ-coefficient denotes a situation where larger cities have grown more than smaller ones (in relative terms); thus an increasing ρ-coefficient (see Eq. 3.3) reflects the tendency towards agglomeration economies in the country at hand (see Subsection 3.1.4).

It is interesting to also pay attention to the β-parameter (see Eq. (3.4)), which is the degree of divergence of the size distribution from its mean: a β-value lower than one indicates that larger cities have an expected growth lower than smaller ones. We note here that the condition β=1 indicates the validity of Gibrat's law (Bottazzi et al. 2001).

Finally, the parameter θ reflects the ratio between variance at time t and $t-1$ (see Eq. (3.5)). This provides a measure of the volatility of the growth process as a θ-value equal to one indicates the stability of the variance between years.

Overall, by means of these three parameters, we can experiment with the propositions (a), (b) and (c) in Subsection 3.1.4. For example, an increasing/decreasing ρ-coefficient – indicating changes in the growth rate between large and small cities – should lead to a generalized Gibrat's law. It appears then that the ρ-coefficients, together with the β- and θ-parameters, offer insights into different aspects of the same growth process: the ρ-coefficient captures the output of the growth process, while the β- and θ-parameters take into account the mean and variance of the growth process, respectively. In the latter context (regarding the role of the mean and variance), it seems worthwhile to test the different dynamics of the large cities vs the small cities, in order to explore in more detail where a greater volatility shows up. We can then split, for each country, our sample into two halves by defining two sub-samples; one for the large cities and the other one for the smaller cities. We then estimate the parameters β and θ, according to Eqs. (3.4) and (3.5) respectively, for these two sub-samples. In this way we can analyse, firstly, whether Gibrat's law holds separately for large and small cities; and, secondly, whether the growth process is more volatile. We recall that, according to Cordoba (2003), in order to preserve a Pareto/Zipf coefficient different from one, the underlying growth process has to be different for smaller and larger cities (Subsection 3.1.4).

In the next sections, the role of the various parameters ρ, β and θ in capturing the relationship "rank-size rule vs Gibrat's law" will be illustrated with reference to the empirical analyses in each of the five countries.

3.4.2 Botswana

Starting with Botswana, we can see that the estimated ρ-coefficient is greater than one for both the years 2001 and 2011, indicating a predominance of larger cities. In particular, in 2011 the estimated ρ-coefficient is slightly greater than that one in 2001, thus showing a tendency – in the last decade – towards a higher economic development. By considering the relationship with Gibrat's law, we then investigate condition b) of Subsection 3.1.4. Considering the entire sample, we have already shown that Gibrat's law holds in 2011 with an estimated β-parameter not significantly different from one ($\beta = 0.995^{**}$).[26] However, considering the two sub-samples, we find evidence of Gibrats law for large cities ($\beta_{BIG} = 0.979^{**}$) but not for small ones ($\beta_{small} = 0.761$). This indicates that larger cities of the sub-sample of small cities (i.e. medium size cities)

[26]Where ** indicates a significance level at 5%.

Table 7: The Zipf and Gibrat Parameters (Countries: Botswana, Germany, Hungary and Luxembourg; Different Years)

Country	Year	ρ-coeff.	Robust s.e.	β	θ	β_{BIG}	θ_{BIG}	β_{small}	θ_{small}	N. obs.
Botswana	2001	1.137	.0746	-	-	-	-	-	-	464
	2011	1.173	.0746	.995**	1.0078	.979**	1.000	0.761	1.011	461
Germany	1993	1.399	.0178	-	-	-	-	-	-	12,280
	1994	1.397	.0178	.999	1.0009	.996	.994	1.001**	1.009	12,291
	1995	1.396	.0178	.999*	.9983	.996	.994	1.00**	1.004	12,291
	1996	1.394	.0177	.999*	.9987	.997	.995	1.001**	1.006	12,291
	1997	1.392	.0177	.998	.9978	.997	.994	.999**	1.002	12,291
	1998	1.390	.0177	.998	.9984	.997	.994	.997**	1.000	12,293
	1999	1.390	.0177	1.00 ***	1.0011	.998	.996	1.004	1.012	12,294
	2000	1.390	.0177	.999***	.9998	.998	.998	.998**	.998	12,294
	2001	1.391	.0177	1.001	1.0024	.999	.999	1.002	1.006	12,294
	2002	1.392	.0177	1.001	1.0017	.999**	.999	1.00**	1.002	12,294
	2003	1.392	.0177	1.00*	1.0009	.999**	1.001	.998	.998	12,293
	2004	1.393	.0177	1.001	1.0027	1.000**	1.001	1.001	1.004	12,293
	2005	1.396	.0178	1.001	1.0031	1.001	1.002	1.001**	1.003	12,293
	2006	1.398	.0178	1.001	1.0027	1.001	1.001	.999**	1.000	12,293
	2007	1.401	.0178	1.001	1.0039	1.002	1	1.001	1.005	12262
Hungary	1980	1.129	.0285	-	-	-	-	-	-	3,121
	1990	1.186	.0300	1.055	1.12	1.018	1.046	1.076	1.194	3,121
	2001	1.223	.0309	1.034	1.08	0.993	.999	1.054	1.151	3,121
	2011	1.258	.0316	1.028	1.06	1.005	1.020	1.010**	1.070	3,154
Luxembourg	1821	.5031	.0660	-	-	-	-	-	-	116
	1851	.4965	.0652	.941*	.9557	.949**	.978	.730	.764	116
	1871	.5154	.0676	.993**	1.0779	.962**	1.070	.887**	.981	116
	1880	.5350	.0702	1.005 **	1.0667	1.003**	1.060	.811**	.871	116
	1890	.5881	.0772	1.048	1.0855	.995**	1.160	.944**	.965	116
	1900	.6744	.0885	1.109	1.2833	1.115	1.340	.947**	.969	116
	1910	.7350	.0965	1.079	1.1803	1.091	1.210	.986**	1.035	116
	1922	.7500	.0984	1.01**	1.0473	0.998**	1.025	1.042**	1.140	116
	1930	.8377**	.1100	1.103	1.2406	1.090	1.240	1.01**	1.048	116
	1935	.8391**	.1101	1.001***	1.0059	.986**	.976	1.033*	1.081	116
	1947	.8543**	.1121	1.014*	1.0355	1.013**	1.034	.982**	1.009	116
	1960	.9252**	.1214	1.069	1.1725	1.011**	1.059	1.006**	1.109	116
	1970	.9735**	.1278	1.044	1.1198	0.965	.963	1.016**	1.162	116
	1980	.9923**	.1302	1.016**	1.0576	0.942	.922	1.022**	1.150	116
	1991	.9803**	.1287	.993 **	.9964	0.949	.920	1.053**	1.145	116
	2001	.9409**	.1235	0.955	.9199	0.949	.907	.910	.871	116
	2002	.9365*	.1229	0.994	.989	.998**	.997	.990**	.983	116

Table 8: The Zipf and Gibrat Parameters (Country: Malta; Different Years)

Country	Year	ρ-coeff.	Robust s.e.	β	θ	β_{BIG}	θ_{BIG}	β_{small}	θ_{small}	N. obs.
Luxembourg	2003	.9330**	.1225	.995**	.9926	1.00**	1.00	.997**	1.00	116
	2004	.9302**	.1221	.997**	.9947	1.00**	1.00	1.00**	1.020	116
	2005	.9270**	.1217	.996**	.9926	.995**	.992	.992**	.989	116
	2006	.9253**	.1215	.997**	.9961	1.001	1.00	.994**	.991	116
	2007	.9195**	.1207	0.992	.9851	0.994	.988	.974**	.955	116
	2008	.9140**	.1200	0.992	.985	.999**	.998	.976	.955	116
	2009	.9120**	.1197	.996**	.994	1.00**	1.00	.989**	.981	116
	2010	.9094**	.1194	.996*	.9924	1.00**	1.00	.986*	.974	116
	2011	.9094**	.1194	.999**	1.00	1.00**	1.00	.997**	.996	116
Malta	1901	.9883**	.1957	-	-	-	-	-	-	51
	1921	.8799**	.1693	0.836*	.7891	.893**	0.874	.487	.624	54
	1931	.8838**	.1700	1.008**	1.0306	.959**	0.965	1.067**	1.184	54
	1948	.8591**	.1638	.912**	.9684	.714	0.852	.689**	.678	55
	1957	.8760**	.1655	1.018*	1.0808	.769	0.673	1.073**	1.238	56
	1967	.8939**	.1645	.987**	.9853	.961**	0.955	.954**	.940	59
	1985	.8305**	.1479	.936**	.9732	.573	0.510	.915**	1.02	63
	1995	.8038**	.1388	.941*	.9407	.654	0.656	.923**	.912	67
	2000	.8005**	.1372	.991**	.9840	.996**	1.00	1.003**	1.009	68
	2001	.7985**	.1369	.997**	.9959	.996**	0.993	.997**	.996	68
	2002	.7956**	.1364	.995*	.9920	.994**	0.988	.993*	.986	68
	2003	.7932**	.1360	.996	.9931	.998**	.996	.993*	.987	68
	2004	.7914**	.1357	.997	.9950	.997**	.996	.995*	.990	68
	2005	.7883**	.1352	.992**	1.0055	.954**	.999	.969**	.985	68
	2006	.7881**	.1351	.998**	.9967	1.00**	1.010	.994*	.988	68
	2007	.7940**	.1361	1.01	1.0210	1.01	1.016	1.02	1.042	68
	2008	.7944**	.1362	1.001**	1.0027	.999**	1.00	1.00**	1.00	68
	2009	.7963**	.1365	1.003*	1.0069	.997**	.996	1.01	1.016	68

have an expected growth lower than smaller ones.[27]

Thus, large and small cities face two different underlying growth processes; however, this is still consistent with proposition b) of Subsection 3.1.4 predicting that the associated growth process requires that smaller cities face a greater volatility of growth than larger cities. For this reason we now analyse the behaviour of the variance.

The variance ratio for the entire sample in Botswana is greater than one ($\theta = 1.078$), indicating a greater volatility of the process in 2011. This

[27] This fact confirms the results in Subsection 3.3.3 in Figure 6: our estimated conditional mean growth is below zero in the mid-range of cities (city size in logarithmic terms between 6 and 8).

latter condition is not enough to investigate our statements b) of Subsection 3.1.4, because it only refers to the temporal non-stability of the variance, without considering the spatial aspect, namely the (non)independence of the variance with respect to the size of the cities.[28]

For this reason, we analyse the two sub-samples separately, as previously anticipated. The variance, θ, for large cities shows a striking stability ($\theta_{BIG} = 1$), while, for the small cities, it is slightly greater than one ($\theta_{small} = 1.011$), implying an (increasing) change in the underlying volatility of the growth process for small cities. Given this fact, we can affirm that at time t (2011), the variance is unchanged for large cities but increases for the small ones, indicating a dependence of variance with respect to size; in particular, smaller cities face a greater volatility than large cities. In summary, statement b) (Subsection 3.1.4), which affirms: if $\rho > 1$, in order that Gibrat's law occurs, it is necessary that the mean is independent from the city size but not the variance, indeed the associate growth process requires that smaller cities face a greater volatility of growth than larger cities, is satisfied for the whole sample.

3.4.3 Germany

Germany shows an U-shaped ρ-coefficient: it decreases until 1999 and then it increases. In fact Germany shows a lower degree of agglomeration between 1993 and 1999, namely larger cities become less heavy in the city system. After 1999, Germany shows again a process of concentration indicated by the increasing ρ-coefficient. By considering the relationship with Gibrat's law, we then investigate condition b) of Subsection 3.1.4. This case is somewhat different from the previous case for one main reason: the ρ-coefficient also shows a decreasing trend. Considering the entire sample, we have five years in which Gibrat's law holds. In particular, in 1995, 1996, 1999, 2000 and 2003, the estimated β-parameters are not significantly different from one.

Now if we focus on the period 1993-1999, where a decreasing ρ-coeffic-

[28]It suggests a change in the variance over the time and then this might also imply changes in the dependence of the variance over the size.

ient applies, we can note that β_{BIG} is significantly lower than one, whereas β_{small} is not significantly different from one in most cases, indicating a situation in which the larger the city, the lower the expected growth. On the contrary, in the period 2000-2007,[29] where an increasing ρ-coefficient applies, we can notice that β_{BIG} are often not significantly different from one, while β_{small} are (most of the time) significantly greater than one, indicating a situation in which the larger the city, the larger the expected growth. In this situation we can figure out the following growth processes: when ρ-coefficient is decreasing, we have modifications on the growth process of large cities; in particular, the larger the city, the lower the expected growth.

On the other hand, when ρ is increasing, small cities present a different growth process, namely the larger the city, the larger the expected growth. However, note that both cases should lead to the same effect on the underlying growth process (i.e. a greater volatility of the variance for small cities), in order to satisfy condition b). For this reason, we analyze the variance ratio, θ. By considering the entire sample, the variance ratio, θ, is often close to unity, but slightly lower than one until 1999 when it becomes stable (and equal to one). Again, this latter condition says few things about the independence of the variance from the size. We then analyze the two sub-samples separately and in particular we analyze those years where Gibrat's law holds (according to proposition b) of Subsection 3.1.4). Unfortunately, we do not have enough observations to make any inference about proposition b) in the period 2000-2007 because Gibrats law holds only in 2003. Focusing on the period 1993-1999, it should be noted, firstly, that when Gibrat's law holds, the variance ratios for large cities, $\theta_{_{BIG}}$, are less than one, that is the variance at time t is lower than that at time $t-1$. Instead, the variance ratios for small cities, $\theta_{_{small}}$, are greater than one. At (any) time t, large cities face a lower volatility while small cities face a greater (or almost stable) volatility. We can then affirm that, at (any) time t, small cities face a greater volatility of growth. This

[29]Notice that in the years in which the ρ-coefficient is equal to that of previous year (i.e. 1998, 1999 and 2000), Gibrat's law holds. The stability of the process should not lead to any changes in the hierarchical structure of cities.

is again consistent with proposition b) of Subsection 3.1.4. Again we can affirm that for those years in which Gibrat's law holds ($\beta = 1$), statement b) (Subsection 3.1.4), which affirms: "if $\rho > 1$, in order that Gibrat's law occurs, it is necessary that the mean is independent from the city size but not the variance. In particular the latter should be greater for small cities", is satisfied for the whole sample.

3.4.4 Hungary

In Section 3.3, we have shown that Gibrat's law does not hold in Hungary. We have already mentioned the role of the capital (Budapest) in this country, which attracts most of the population and most of the investment on infrastructure. In percentage terms, more than half of the total population of Hungary lives in the Budapest urban area. Moreover, it faces a migration flow from its rural area to the centre of the city. The evolution of the ρ-coefficient in this country reflects the tendency to agglomeration in the large cities: indeed the estimated ρ-coefficient is increasing and greater than one. In this situation we can check for statement b) of Subsection 3.1.4. Unfortunately Gibrat's law never holds, since the β-coefficients are always significantly greater than one. This means that size diverges towards its means, namely, the larger a city, the larger the expected growth. Then it is straightforward that the variance ratio, θ, increases for both large and small cities. However, it should be noted that the variance ratio of small cities, θ, is always greater than the variance ratio of large cities; thus, although we cannot formally show evidence of generalised Gibrat's law (in particular regarding statement b)), we again show a greater volatility for small cities when $\rho > 1$.

The analysis carried out over these three countries provides an important first conclusion. On the basis of statement b), we have shown that when $\rho > 1$ and Gibrat's law holds ($\beta = 1$), the variance ratio (θ-parameter) is actually greater for small cities. Moreover, when the ρ-coefficient is greater than one but decreasing, we have modifications on the growth process of large cities, but not on those of small cities: in particular, the larger the city, the lower the expected growth. On the other hand, when the ρ-coefficient is greater than one, but increasing, small cities present the opposite growth process: namely the larger the city, the larger the expected growth. Consequently, we find evidence of the generalised Gibrat's law – as in statement b) of Subsection 3.1.4 – in the countries displaying $\rho > 1$, where the independence of the mean with respect to the size is in operation, while the same is not true for the variance.

A reasonable criticism, at this point in the analysis, could be that generally, small entities (cities, firms and so on) present a greater volatility than larger ones. We can anticipate that we will find the opposite evidence in the case of $\rho < 1$.

We now move to the situation where $\rho < 1$. By considering the relationship with Gibrat's law, we investigate condition c) of Subsection 3.1.4 which predicts that the associate growth process requires that smaller cities face a lower volatility of growth than larger cities.

3.4.5 Luxembourg

Luxembourg shows an estimated ρ-coefficient lower than one. It increases until 1930, and after that it is not significantly different from one. Between 1821-1922, we are in condition c) of Subsection 3.1.4. Considering the entire sample, we have already shown that Gibrat's law holds in the first three years of the sample (1851-1880) and in 1922 with estimated β-parameters not significantly different from one. Moreover, considering the two sub-samples, most of the time we find evidence of Gibrat's law for both large and small cities. At a first glance it seems that large and small cities face the same underlying growth processes. However, to test statements c) of Subsection 3.1.4, we need to take into account the behaviour of the variance.

The variance ratio for the entire sample in Luxembourg in the period 1821-1922 is always greater than one, indicating a greater volatility of the process as time goes by. Indeed, when we split the sample in two halves, the variance ratios for the large cities show values always greater than one, while, for the small cities, they are always below one. This implies an (increasing) change in the underlying volatility of the growth process for large cities, in contrast to a (decreasing) change in the underlying volatility of the growth process for small cities. Given this fact, we can affirm that at time t, the variance is increased for the large cities but decreased for the small cities, indicating a dependence of variance with respect to size; in particular, smaller cities face a lower volatility than large cities. In summary, statement c) (Subsection 3.1.4), which affirms: "if $\rho < 1$, In order that Gibrat's law occurs, it is necessary that the mean is independent from the city size but not the variance, indeed the associate growth process requires that smaller cities face a lower volatility of growth than larger cities", is satisfied for the whole sample in 1821-1930.

In the period 1935-2011, however, the ρ-coefficient is not statistically different from one. By considering the relationship with Gibrat's law, we then investigate condition a) of Subsection 3.1.4 which predicts that the associated growth process requires that smaller cities face the same growth as larger cities. Considering the entire sample, we have already shown that Gibrat's law holds most of the time (estimated β-parameters not significantly different from one). Considering the two sub-samples, we find similar evidence for both large and small cities. However it is interesting to note that in those years where Gibrat's law does not hold, the estimated parameters β and θ show very different behaviour (i.e θ_{BIG}=.965 and θ_{small}=1.162 in 1970), but, in general, in those years where Gibrat's law holds, the differences between the estimators are not so large (i.e. θ_{BIG}=1.00 and θ_{small}=1.02 in 2004). In summary, statement a) (Subsection 3.1.4), which affirms: "if $\rho = 1$, then in order that Gibrat's law occurs neither the mean nor the variance of growth can depend on size" is satisfied for the whole sample.

3.4.6 Malta

Finally, Malta presents a ρ-coefficient not significantly different from one. Again, by considering the relationship with Gibrat's law, we investigate condition a) of Subsection 3.1.4. Taking into account the entire sample, Gibrat's law holds most of the time (estimated β-parameters not significantly different from one). Considering the two sub-samples, Gibrat's law fails more in the large cities. In these cases, differences between the estimated parameters β and θ – in the two sub-samples – arise. Clearly, when Gibrat's law holds for both sub-samples, no major differences between the two parameters occur. In summary, statement a) (Subsection 3.1.4) is satisfied.

3.4.7 Synthesis

A synthesis of the above results – confirming the hypotheses by Cordoba (2003) – is presented here. The analysis carried out in this section prompts several interesting conclusions. We have been able to empirically verify the presence of a "generalized Gibrat's law", as theoretically predicted by Cordoba (2003). In particular, we have verified statements a), b) and c) of Subsection 3.1.4. In more detail, we have shown that when $\rho > 1$ (statement b)) and Gibrat's law holds (β=1), the variance-ratio (θ-parameter) is actually higher for the small cities, in comparison to that for large cities, indicating a larger volatility for small cities. On the contrary, when $\rho < 1$ (statement c)) and Gibrat's law holds (β = 1), the variance-ratio (θ-parameter) is actually lower for the small cities, in comparison to that for large cities, indicating a larger volatility for the smaller ones. When $\rho = 1$ (statement a)) and Gibrat's law holds, our findings agree with previous research, as both the mean and variance appear to be independent from the size.

Moreover, when the ρ-coefficient is greater than one but decreasing, we have modifications on the growth process of large cities, but not on those of small cities; in particular, the larger the city, the lower the expected growth. On the other hand, when $\rho > 1$ but increasing, small cities present the opposite growth process, namely the larger the city, the larger the expected growth. We have, of course, an opposite behaviour when the ρ-coefficient is less than one.

3.5 Conclusion

The aim of our research work was to explore specific conditions leading to a generalization of Gibrat's law in connection with the different typologies of rank-size distribution. For this purpose we empirically explored the link between the rank-size exponent, ρ, with the necessary conditions for Gibrat's law (that is mean and variance of the growth have to be independent from the size). We started our analysis based on the conclusion of Cordoba (2003, p. 3): *"Pareto distributions with larger exponents (more unequal distributions) require more volatile growth processes"*. As far as we know, the conventional methodologies (Section 3.3) used to test Gibrat's law do not address this issue. In particular, a greater (lower) volatility of the variance is usually not empirically envisaged. We showed, instead, that, according to Cordoba (2003), the variance can be dependent on size if the rank-size coefficient is different from one; in particular, we verified what Cordoba (2003) calls a "generalized Gibrat's law" for different countries with different spatial-economic characteristics: Botswana, Germany, Hungary, Luxembourg and Malta. We found strong evidence of this generalized Gibrat's law for Botswana, Luxembourg and Malta. We found weak evidence of Gibrat's law for Germany and no evidence for Hungary.

Our results confirm the propositions provided by Cordoba (2003). In particular, when ρ=1, neither the mean nor the variance of growth depend on size; when $\rho > 1$, the mean is independent of the city size, but not the variance, and small cities face a greater volatility in growth than larger cities; alternatively, when $\rho < 1$, the mean is independent from the city size, but not the variance, and large cities face a greater volatility in growth than smaller ones. Gibrat and Zipf have offered complementary perspectives on city size and systems of cites in a given country. Their contributions are not necessarily identical, but offer new perspectives on the same multi-faceted prism of the space-economy. These laws are part of the same family, but also reveal specific distinct features. In particular, we find that Zipf's law and the rank-size rule behave like "monozygotic" twins, while Gibrat's law seems to show the behavious of a "heterozygotic" twin. These results might be useful to "relax" Gibrat's law in its strict interpretation, by reinforcing the hypothesis that small entities face a greater volatility in the growth process.

Our analysis prompts various intriguing research questions in the future. While Gibrat's law and Zipf's law mirror important organised structures in the topology of systems of cities, other relevant structural patterns may be investigated as well, such as the existence of fractal structures in urban systems (based, for example, on Mandelbrot's principles) or the persistent existence of spatial population or socio-economic disparities (based, for example, on Herfindahl's index). Clearly, the dynamics of such processes deserve due attention. In addition, the above applied investigation also calls for more fundamental research into the functional or behavioural backgrounds of such regularities. Three research directions are important here; (a) the interdependence between population indicators and broader socio-economic indicators for a system of cities; (b) the degree of various cities in the same national system; (c) the relationship between recent strong evolutionary trends in the digital world and the development of cities (and systems of cities).

Chapter 4

Does the EU have a homogeneous urban structure area? The role of agglomeration and the impact of (hypothetical) shocks on the EU urban structure

4.1 Introduction

" Cities and metropolitan areas are the engines of economic development. They are also at the frontline when it comes to tackling obstacles to growth and employment, such as social exclusion and environmental degradation " European Commission (2010, p. 3).

The effort of the European Union in (sustainably) developing urban areas has been, at least by the late 1980s and early 1990s,[1] high by means of a range of policies over many areas of activity (Backer et al., 1997). Currently, indeed, the EU aims to reduce regional disparities through the so called Regional Policy. In particular, the EU has two funds for this purpose: the Structural Funds and the Cohesion Funds, that play a key role in the reduction of regional disparities in terms of income, wealth and opportunities and in the development of Europe's towns and cities.

However, in spite of the effort of the EU in the reduction of regional disparities, "*no single blueprint of sustainability will be found, as economic and social systems and ecological conditions differ widely among countries*", (World Commission on Environment and Development, 1987, p. 52). Furthermore, as noted by Aldskogius (2000), the urban structures between the Member States of the European Union is very different for several reasons, from historical ones to the administrative subdivision through the economic structure of the Member States. Table 10 shows selected characteristics of the Member States in 2011 and gives an interesting picture of the differences between countries in term of urban structure.

Total population and land size vary consistently among countries within the EU, however Member States show big differences also in term of population density, percentage of rural population, people who live in the largest city and also in other variables as agricultural land, road density and rail density.

It is interesting to notice that different countries show different patterns in terms of urban structure. For instance Belgium has a population density of 364 people per sq. km and it also shows a low percentage of rural population (2.5 over the entire population) but it presents an agricultural land equal to the 45% of the entire surface. On the contrary, the Netherlands, a country with almost same land size as Belgium, shows a higher population density than Belgium but also a higher rural population and a higher percentage of agricultural land, indicating a different

[1] At least from the publication of the Brundtland Report from the United Nations World Commission on Environment and Development (WCED) in 1987 which provided an analysis for a sustainable course of development within societies and suggested broad remedies and recommendations.

Table 9: Spatial Economics Characteristics of the 27 EU Member States in 2009

Country	Population (thousands)	Land (km^2)	Population Density (people $*km^2$)	Rural Population	Population in in the largest city (% tot. urban)	Agricultural land (% of tot)	Road Density (length $*1000$ km^2)	Rail Density (length $*1000$ km^2)
Austria	8,419	82,430	102	32.3	30.0	38.4	1,278.9	76.3
Belgium	11,008	30,280	364	2.5	18.0	45.0	4,987.4	105.9
Bulgaria	7,476	108,560	69	26.9	22.0	46.3	362.5	37.3
Cyprus	1,117	9,240	121	29.5	31.4	13.5	1,586.1	0
Czech Republic	10,546	77,250	137	26.6	15.1	54.9	1,619.4	122.1
Denmark	5,574	42,430	131	13.1	24.4	62.1	1,698.7	61.9
Estonia	1,340	42,390	32	30.5	42.8	22.0	1,283.1	26.4
Finland	5,387	303,900	18	16.3	32.6	7.6	230.9	17.5
France	65,437	547,660	119	14.3	20.7	53.4	1,870.4	54.5
Germany	81,726	348,610	234	26.1	5.7	48.4	1,805.2	117.6
Greece	11,304	128,900	88	38.5	47.1	63.6	890.7	19.5
Hungary	9,971	90,530	110	30.6	25.0	63.9	2,123.2	99.0
Ireland	4,487	68,890	65	37.8	39.5	60.8	1,366.6	46.1
Italy	60,770	294,140	207	31.6	8.2	47.3	1,618.4	67.2
Latvia	2,220	62,180	36	32.3	46.2	29.5	1,131.4	34.7
Lithuania	3,203	62,670	51	32.9	24.4	42.9	1,240.9	27.1
Luxembourg	517	2,590	200	14.6	21.9	50.6	2,018.2	106.2
Malta	419	320	1309	5.2	50.9	28.1	696	0
Netherlands	16,696	33,730	495	16.9	7.7	56.8	3,294.7	69.7
Poland	38,216	304,200	126	39.1	7.3	53.0	1,356.0	62.1
Portugal	10,637	91,470	116	39.0	44.0	40.3	899.9	36.0
Romania	21,390	230,060	93	47.2	16.5	58.8	342.8	45.2
Slovak Republic	5,440	48,090	113	45.2	14.1	40.1	892.5	73.9
Slovenia	2,052	20,140	102	50.1	26.4	23.2	1,920.3	60.6
Spain	46,235	498,800	93	22.6	16.2	55.5	1,348.1	30.3
Sweden	9,453	410,340	23	14.8	16.3	7.5	1,272.2	25.8
United Kingdom	62,641	241,930	259	20.4	15.5	71.6	1,619.1	67.6
Euro Area	332,990	2,551,580	131	24.5	15.4	45.5	1,445.1	54.1
European Union	503,680	4,181,730	120	26.0	15.6	45.1	1,355.7	53.4

allocation of people between urban and rural areas. This fact is also underlined by the amount of population living in the largest city: 18% for Belgium (Anvers) and 7.7 % for the Netherlands (Amsterdam). Germany and Italy, instead, although presenting differences in size in terms both of population and land size, show similar urban patterns characterized by (among others characteristics) a low percentage of population in the largest city. This fact could be due to historical reasons, indeed since a long time both Germany and Italy have been politically fragmented. The competition between numerous states was high and led to the creation of several market centers and capitals, large and small. As a result of this process, the two countries present a large numbers of towns and cities where - especially in Italy - the memory of city-state remains alive (Le Gales, 2002). Instead, Spain, that is a country of size similar to Germany and Italy, shows a slightly different pattern: the population density, 93, and the rural population, 22.6, are lower than in those countries. The population living in the largest city, instead, is relatively higher. Finally, it should be noticed that the transition countries show a higher rural population than those ones denoting a higher level of economic development.

This first rough picture of the differences of urban structures among countries might denote different underlying economic forces of EU's Member States. For instance, the degree of urbanization is strictly related to differences in regional per capita income (Mera, 1975); odds in the population density can lead to different conditions in the labor market (Armington and Acs, 2002) and/or investments in infrastructure (Fay and Yepes, 2003; Randolph, Bogetic and Hefley, 1999) and so on. On this regards, the countries of Benelux (Belgium, the Netherlands and Luxembourg) present population densities that are the highest among EU countries and, related to that, they all show a high density level both in terms of road and rail networks. On the contrary, Bulgaria, Finland and Romania, countries with a population density lower than 100 people per km^2 present a low road and rail density.

On the other way round, the impact of a given policy can lead to a change in the urban structure. For instance, Yang (1999) shows that,

Table 10: Spatial Economics Characteristics of the 27 EU Member States in 2009

Country	Population (thousands)	Land (km^2)	Population Density (people $*km^2$)	Rural Population	Population in in the largest city (% tot. urban)	Agricultural land (% of tot)	Road Density (length $*1000$ km^2)	Rail Density (length $*1000$ km^2)
Austria	8,419	82,430	102	32.3	30.0	38.4	1,278.9	76.3
Belgium	11,008	30,280	364	2.5	18.0	45.0	4,987.4	105.9
Bulgaria	7,476	108,560	69	26.9	22.0	46.3	362.5	37.3
Cyprus	1,117	9,240	121	29.5	31.4	13.5	1,586.1	0
Czech Republic	10,546	77,250	137	26.6	15.1	54.9	1,619.4	122.1
Denmark	5,574	42,430	131	13.1	24.4	62.1	1,698.7	61.9
Estonia	1,340	42,390	32	30.5	42.8	22.0	1,283.1	26.4
Finland	5,387	303,900	18	16.3	32.6	7.6	230.9	17.5
France	65,437	547,660	119	14.3	20.7	53.4	1,870.4	54.5
Germany	81,726	348,610	234	26.1	5.7	48.4	1,805.2	117.6
Greece	11,304	128,900	88	38.5	47.1	63.6	890.7	19.5
Hungary	9,971	90,530	110	30.6	25.0	63.9	2,123.2	99.0
Ireland	4,487	68,890	65	37.8	39.5	60.8	1,366.6	46.1
Italy	60,770	294,140	207	31.6	8.2	47.3	1,618.4	67.2
Latvia	2,220	62,180	36	32.3	46.2	29.5	1,131.4	34.7
Lithuania	3,203	62,670	51	32.9	24.4	42.9	1,240.9	27.1
Luxembourg	517	2,590	200	14.6	21.9	50.6	2,018.2	106.2
Malta	419	320	1309	5.2	50.9	28.1	696	0
Netherlands	16,696	33,730	495	16.9	7.7	56.8	3,294.7	69.7
Poland	38,216	304,200	126	39.1	7.3	53.0	1,356.0	62.1
Portugal	10,637	91,470	116	39.0	44.0	40.3	899.9	36.0
Romania	21,390	230,060	93	47.2	16.5	58.8	342.8	45.2
Slovak Republic	5,440	48,090	113	45.2	14.1	40.1	892.5	73.9
Slovenia	2,052	20,140	102	50.1	26.4	23.2	1,920.3	60.6
Spain	46,235	498,800	93	22.6	16.2	55.5	1,348.1	30.3
Sweden	9,453	410,340	23	14.8	16.3	7.5	1,272.2	25.8
United Kingdom	62,641	241,930	259	20.4	15.5	71.6	1,619.1	67.6
Euro Area	332,990	2,551,580	131	24.5	15.4	45.5	1,445.1	54.1
European Union	503,680	4,181,730	120	26.0	15.6	45.1	1,355.7	53.4

among the others, urban policies, investments in the urban sector, institutions and financial policies are responsible for the long-term income differential between urban and rural areas. Moreover, urban places expand their influence on a much larger hinterland than rural areas and the difference between urban and rural places goes behind the simple difference in population size, but also in a different concentration of economic activities, people and cultures. (Pacione, 2001).

These facts also have an impact on the definition of agglomeration: how many inhabitants are necessary before a given agglomeration can be defined as "urban"? Across countries this definition varies substantially and also within the same countries this definition can change from time to time (United Nations, 1974). Moreover as reported in Manual VIII (United Nations, 1974) it is very difficult if not impossible to prescribe a uniform limit for all countries because qualitative characteristics or even residential densities in settlements of a given size can vary significantly among the countries.

However, European countries are trying to adopt a 'common' definition by which units with 10,000 or more inhabitants are considered as urban, those with 2,000-9,999 inhabitants as semi-urban, and those with less than 2,000 as rural. Despite this definition, still a considerable number of Member States use different thresholds and/or rules aiming to discriminate what is rural and what is urban. As an example, Table 6 of the Demographic Yearbook 2009-2010 of the United Nations reports the definition of "urban" across countries over the world. It can be noticed that between EU Member States the definition of urban is not homogenous, i.e. Austria defines urban areas all the localities with a population above 2,000; France adopts a population threshold of 10,000 and Bulgaria gives just a legal definition of what might be considered urban (United Nations, 2009, Table 6).

Then, the European Union is still a multifaceted entity, at least in terms of urban structure, where the harmonization between countries is still going on and, given those stylized facts, we can ask whether the deeper and deeper integration of European Union affected the hierarchical structure of the city system of the Member States and, furthermore, if the European Union is an area where urban and regional policies affect Member States in the same way or, on the contrary, if it is still a heterogeneous entity and a heterogeneous area in recovering from exogenous shocks.

The study of the hierarchical structure is important for several reasons, first of all because the population is spread across geographic areas in a way that, although continuously changing, is not possible to define as random. Indeed, countries have faced a strong tendency toward agglomeration, namely population gathers within proper areas like cities, and currently the agglomeration within cities "*is an extremely complex amalgam of incentives and actions taken by millions of individuals, businesses, and organizations*" Eeckhout (2004, p. 1429). Moreover we can affirm that the principal determinants of the dynamics of city populations are economic and institutional factors.

Then the creation of the European Union, the distortions caused by the introduction of a single currency in countries structurally very different and the expansion of mobility of people, capital and services due to the constitution of the so-called Schengen Area from the beginning of '90s, with the intent to build a common/integrated area, might have had some impacts on the dynamics of city populations.

In this chapter we use an alternative approach to examine whether, i) the EU affects the city system of the Member States and ii) EU city system can be seen as an integrated area. Specifically, we use, in an alternative way, two very well known empirical regularities, Zipf's law (Auerback, 1915; Zipf, 1949) and Gibrat's law (Gibrat, 1931). Alternative means that aim of this chapter is not to verify if the former laws hold but it means that we use them as a tools to explore the research questions i) and ii).

Zipf's law predicts the degree of hierarchization of a system of cities and it can be seen as an indicator of the strength of the agglomeration forces in a system of cities (Singer, 1930; Brakman et al., 2001). Gibrat's Law suggests that the growth of a city is independent of size, a condition that would hold if temporary shock has a permanent impact on the pattern of growth of the cities (Brakman et al., 2004). The present chapter then, takes the moves from the huge literature on Zipf's and Gibrat's law.

In particular, this work has its roots on the research of Rosen and Resnick (1989) and Soo (2005) for the part related to Zipf's law and on the works of Gabaix (1999), Brakman et al. (2004) and Giesen and Suedekum (2012) for the part related to Gibrat's law.

Rosen and Resnick (1989) provide an international analysis of Zipf's law over 44 countries. They found the validity of Zipf's law in most of the cases and moreover they provide empirical evidence that the Zipf exponent is positively influenced by per capita GNP, total population and railroad density, and negatively related to land area.[2] Soo (2005) updates these results but concludes for a rejection of Zipf's Law in more than half of cases. Moreover he found that political variables matter more than economic variables in determining the size distribution of cities. In details the total government expenditure and war dummy enter with a very strong positive coefficient. Instead, among economic variables, the degree of scale economies seems to significantly and positively influence the Zipf coefficient, whereas trade as a percentage of GDP weakly and negatively influence it.

Gabaix (1999) and Giesen and Suedekum (2011) address the same issue, i.e. if urban growth in all regions follows Gibrat's law, then we should observe the Zipfian rank-size rule among large cities both at the regional and national level. Gabaix (1999) provides theoretical rationale while Giesen and Suedekum (2011) provide empirical evidence. Finally, Brakman et al. (2004) analyze whether war shock has an impact on post-war German city growth finding that war shock has only a temporary impact on it since German city growth faces a tendency toward mean reversion.

The present chapter, then, sets out to do three things: the first is to provide an accurate description of agglomeration. This involves a static analysis, supported by Zipf's law, accounting for the way the population is gathered over different geographic locations. The second is to provide an accurate description of population mobility. This involves a dynamic

[2]Positively means that greater per capita GNP, total population and railroad density are associated with a more even distribution of cities. Instead, negatively means that greater land area is associated with a more agglomerated distribution of cities.

analysis, supported by Gibrat's law, accounting for the evolution over time of the distribution of the population over different locations. Here we are able to determine if EU city system can be seen as an integrated area and, in particular, if temporary shock has a permanent impact on the pattern of city growth. Finally, once population mobility is understood, we explore the impact of the underlying economic mechanism and of the constitution of the EU, the Euro and the Schengen Area on agglomeration forces.

We find that, generally, the hierarchical structures of Member States of EU is more even than expected. Moreover, EU city system is still far from being an integrated area and in particular, we find that the European Union seems to be split in three distinct areas: an integrated area characterized by the validity of Gibrat's law where large temporary shocks might have permanent impacts on the city structure; another area that is characterized by the presence of mean reversion and where any exogenous shock is used up in certain amount of time and, finally, a small area where the effects of the shocks is magnified in the big cities. Finally, we find that only the constitution of the Schengen Area and the share of international trade seem to have a weak impact on the hierarchical structures of Member States presenting respectively a positive impact on the Zipf coefficient (more even distribution) and a negative impact (more agglomerated distribution) and then indicating two sources, one against and one in favor, of agglomeration forces. Those results are in line with previous findings of Soo (2005) and Alperovich (1993).

The next section outlines Zipf's law and Gibrat's law. Section 4.3 describes the data and Section 4.4 presents the results. Then Section 4.5 concludes.

4.2 Zipf's law, Gibrat's law: definitions, methodologies and economic interpretations

The Member States of the EU have a long history of conflict. Only in the last century Europe has been the main theater of the two World Wars. However, after WWII, European Countries started an integration project that led to the creation of European Union that, in some way, covered up old tensions (Nahoi, 2011). The creation of the European Union with free movement of goods, services people and capital might have changed the population distribution across geographic areas. Nevertheless, as mentioned before, people tend to concentrate within common restricted areas like cities in a way that is not random (Eeckhout, 2004).

On this regard, Brakman, Garretsen and van Marrewijk (2001) affirm that a simple and plausible way to explain birth and growth of cities should be definitively linked to the existence of agglomeration and congestion forces. More specifically, following Kooij (1988), urbanization follows three distinct periods: Pre-industrialization, characterized by high transport cost that discourages agglomeration; Industrialization, where transportation cost starts to decrease and the importance of industrial plants becomes higher since an increases of return to scale; Post-industrialization, where congestion and declining importance of industrial production become established. Zipf's law, although it does not directly address the characterization presented by Kooij, is often described in term of tension between congestion and agglomeration forces, (see Sutton, 1997; Gabaix and Ioannides, 2004 for a survey) and it provides a proxy for urbanization.

In more detail, Zipf's law is an empirical regularity stating the proportionality of the cities to their rank. This means that for example, in Italy, the size of Rome (the largest city in Italy) is roughly twice the size of Milan, the second largest city, three times the third largest city, Naples, and so on. Formally, this can be written as:

$$R_i = KP_i^{-q}, \tag{4.1}$$

Equation 4.1 is known as the rank-size rule and is usually expressed in logarithmic form, as follows:

$$ln(R_i) = k - q\, lnP_i, \tag{4.2}$$

where R_i is the rank of the cities assigned as 1 to the largest city, 2 the second largest city and N the smallest one, k is a (ln) constant, P_i is the population size and q is the parameter of interest, the so-called rank-size coefficient that here will be used as a proxy for agglomeration forces.[3]

Several interpretations of the rank-size coefficient, q, have been given in the literature. The reader can consult Modica, Reggiani and Nijkamp (2013) for a brief summary. In this chapter we interpret the q-coefficient in the simplest way, that is as an indicator of the strength of the agglomeration forces in a system of cities (Singer, 1930; Brakman et al., 2001). On this vein, the rank-size coefficient might be considered an index of metropolization, in the sense that the lower the q-coefficient, the more important is the urban land value. In fact the q-coefficient measures how (un)even the city system is, namely this means that the higher the q-coefficient, the more equally distributed the city system (in the extreme, when q=∞, the system of cities is very even: all cities of the same size; when q=0, instead, the system will be composed just by one city hosting the entire population).

In the above interpretation (i.e. the tension between agglomeration and congestion) of Zipf's law the size of cities matters, however, there is another way to explain Zipf's law in which the (relative) size of cities does not matter, this is the case of the use of Gibrat's law in explaining Zipf's law (Gabaix, 1999).

Gibrat's law is a rule stating that the (relative) growth of a given entity (city, firm, income and so on) does not depend on its size.[4] This means that, albeit an entity can grow at different rates, it can't be found any systematic behavior between their growth and their size. Then, according to Gibrat (1931), we cannot affirm that larger entities grow faster than smaller ones or vice versa. Analytically, following Steindl (1968):

$$\log P(t) = \log P(0) + \varepsilon(1) + \varepsilon(2) + \ldots + \varepsilon(T), \tag{4.3}$$

[3]Notice that we estimate the rank-size coefficient using simple rank-size OLS regressions where, following Gabaix and Ibragimov (2011), the rank is shifted by 0.5 to correct for the potential bias in small samples highlighted by Gabaix and Ioannides (2004), so that the estimating equation is $ln(R_i - 0.5) = k - q\, lnP$.

[4]The reader can consult Modica et al. (2013) for a deep discussion on Gibrat's law and its relationship with Zipf's law.

where, $P(t)$ is the size of a certain entity at time t,$P(0)$ is the initial size, and $\varepsilon(t)$ is a random variable (indicating random shocks). Equation 4.3 defines the logarithm of the size of a given entity as the sum of the initial size and past growth rates. Moreover as Gabaix (1999) notices, Gibrat's law implies that the growth process of a given entity presents a common mean and a common variance, or, in other terms, mean and variance of the growth have to be independent from the size of the entity.

Several authors have proposed economic interpretations of Gibrat's law that differ only in some shades. For instance, Black and Henderson (2003), state that a shock affects in the same way big and small entities whereas, Brakman, Garretsen and Schramm (2004), state that a large temporary shock can have a permanent impact. This means that a shock can change the growth path toward another size equilibrium. In this work, we adhere to the latter interpretation and then we will use Gibrat's law to test whether a (hypothetical) shock (will) influence the Member States of EU in the same way or, on the contrary, if it (will) has heterogeneous effects.

Several methodologies have been proposed to test Gibrat's law. However, it is commonly tested looking at the non-stationarity of the growth series. The rationale underlying this test is the following: if the series is stationary it will converge toward a constant mean and, therefore, large values (big entities) will be followed at any further periods by smaller and smaller values, while small values (small entities) will be followed by larger and larger ones. In this way, the level of the series at any periods can significantly predict the one of the next period and then the growth of the entities are not independent on the initial state. On the other hand, if the series is integrated, then large values and small values will occur with probabilities that do not depend on the current size of the entities and then the growth of the entities are independent on the initial state.

Finally, it has to be noticed that the studies on Zipf's and Gibrat's law usually consider only the upper tail of the data, i.e. the largest cities, with a sample truncation point that is usually arbitrarily chosen. Here to avoid any sources of arbitrariness we use a method proposed by Clauset et al. (2009) and it will be deep presented in the next section. [5]

To sum, Zipf's law shows the static relationship of the size of a certain entity with its rank. Gibrat's law, instead, presents the dynamic growth process of this entity. We will use both the laws to show how the EU tends toward agglomeration and if it can be seen as a homogeneous area, but first we will introduce the data in the next section.

4.3 The data

In this chapter, we collect city population data from the National Statistics Office of the Member States. The time windows vary according to the availability of the data from any Statistical Office, however when we present the cross-countries analysis we adopt a unique time span namely from 1991 to 2011. This is due because of two main reasons. First of all, the time period of interest for this work is the one covering the period of greater strength of integration among Member States, i.e. between 1991 and 2011. Second, that's the only time span that has observations for all the 27 Member States.

One main concern is that the method used to define the borders of the statistical units can change dramatically and, often, the definition of cities changes in relation to the official definition of city boundaries given by the statistical authorities (Soo, 2005). As an example the UK administrative geography is organized as follows: Country -> Region $->$ County/Unitary Authority $->$ Local Authority $->$ Ward and - some case Parish (mainly rural areas). Italy, instead, is based on a legal administrative division still strictly related to historical subdivision in State $->$ Region $->$ Province $->$ Municipality. It should be noticed that UK Local Authority is something slightly different from an Italian Municipality, since the latter, typically, do not include surrounding suburbs, i.e. the places where the workers of a city reside are considered as other municipalities.

Moreover, the debate on the unit of observations is still alive. Since

[5] For a thorough analysis of this issue the reader can consult Gabaix (1999), Fazio and Modica (2012) and Ioannides and Skouras (2013).

a long time, it has been noticed that defining the unit of study as urban places, legal cities or urban agglomeration may affect the estimates of the rank-size coefficient (Rosen and Resnick ,1980; Cheshire, 1999). On this regard, the literature proposed several solutions: metropolitan areas (among the others Ioannides and Overman, 2003; Dobkins and Ioannides, 2000); economic areas (Berry and Okulicz-Kozaryn, 2012); clustered populated areas (Rozenfeld, Rybski, Gabaix, and Makse, 2011) and, finally, natural cities (Jing and Jia, 2010).

Due to the availability of data, to overcome this issue, we choose the most homogenous unit of observations among Member States we have found, namely "locality" that is the smallest administrative level of aggregation available for any country. [6]

Another main concern related to the data, already mentioned in the previous section, is the choice of the appropriate truncation point for the upper tail. The methods generally adopted on this regards are three: the choice of a fixed number of cities between countries (i.e. the 100 larger cities); the choice of a fixed population threshold (i.e. all the cities with a population above 100,000); all the cities for which the sum of the population accounts for some given proportion of a country's population (i.e. all the largest cities for which the sum of the population accounts the 10% of the country's population), Cheshire (1999).

However, all these proposed methodologies discount a certain amount of arbitrariness (Fazio and Modica 2012; Ioannides and Skouras, 2013). In order to avoid this arbitrariness, we exploit an alternative non-parametric methodology to estimate the appropriate minimum threshold, firstly proposed by Clauset et al. (2009).

The rationale behind this methodology lies on the idea that Zipf's Law should led to a Pareto distribution with a shape parameter (that is our q) equal to 1 (Eeckhout, 2004). Then, Clauset et al. (2009) propose to test the equality between the theoretical and empirical density functions

[6]It should be noted that EUROSTAT provides "Nomenclature of territorial units for statistics" (hereinafter referred to as NUTS). Although it might be an alternative and perhaps better measure of agglomeration, it is still based on administrative units within the Member States, where the only discriminant is based on the population size, i.e. NUTS3 has a minimum of 150,000 and a maximum of 800,000 inhabitants.

using a Kolmogorov-Smirnov tests over recursively truncated distributions and choosing the sample showing the best fit to a Pareto distribution (the minimum of the statistics D below), indicating the sample of cities where agglomeration forces are stronger (Fazio and Modica, 2012).

In sum, our estimate population threshold, $\underline{\hat{P}}$, is the value of population, $\underline{P}$, that minimizes the Kolmogorov-Smirnov statistics, D:

$$D = \sup_{p \geqslant \underline{p}} |\Phi_p(x) - \Phi(x)|, \tag{4.4}$$

where $\Phi_p(x)$ is the empirical cumulative density function for any city of population p above $\underline{p}$, and $\Phi(x)$ is the theoretical cumulative density function of the Pareto distribution. The KS statistic just computes the supremum of the absolute value of the distances between the two. Under the null, the distance between the two is zero and then the sample is drawn from the reference distribution, that in our case is a Pareto.

In this way we are able to collect a unique data-set avoiding some of the common issues, namely we identify an upper-tail strictly related to the agglomeration characteristics of the Member States allowing homogeneity in the data. Table 11 summarizes the main characteristics of the data-set.

Data for the second stage regression, which seeks to find the impact of the constitution of the European Union on the city hierarchical structures, are obtained from the World Bank World Development Indicators, the UNIDO Industrial Statistics Database and EUROSTAT. In more details, the variable road density is constructed using EUROSTAT data on the length of the roads. GDP per capita in constant US dollars, total land, total population, trade as a percentage of GDP, and number of internet users are from the World Bank World Development Indicators. The data for public expenditure are those from the World Bank World Development Indicators with the exception of year 2011 (still not available) and then they are retrieved from the original source, namely the single national statistical offices. Scale economies are constructed starting from information presented in the UNIDO Industrial Statistics Database.

More details on the data used in the second stage regressions will be presented in the next section as well as the results.

Table 11: Description of the data and estimated threshold of the 27 EU Member States

Country	Population (thousands)	N. of Localities	Average population per locality	Threshold	Population Size truncation
Austria	8,431	2,357	3,577.00	654	2,497
Belgium	10,928	589	18,553.48	184	17,691
Bulgaria	7,327	5,059	1,448.31	1,405	571
Cyprus	839	388	2,162.37	67	1,966
Czech Republic	10,335	6,216	1,662.64	1,891	728
Denmark	4,832	1,469	3,289.31	392	1,528
Estonia	1,324	226	5,858.41	187	962
Finland	5,375	336	15,997.02	81	14,067
France	64,304	36,674	1,753.39	3,377	3,052
Germany	81,752	11,421	7,158.04	838	17,164
Greece	10,934	1,034	10,574.49	492	4,638
Hungary	9,986	3,153	3,167.14	1,012	1,526
Ireland	2,318	858	2,701.63	293	945
Italy	59,571	8,092	7,361.72	1,170	10,300
Latvia	2,261	523	4,323.14	372	853
Lithuania	2,171	102	21,284.31	30	11,623
Luxembourg	512	115	4,452.17	90	1242
Malta	412	67	6,149.25	13	10,770
Netherlands	14,432	2,025	7,126.91	175	14,885
Poland	38,200	2,478	15,415.66	1,099	8,293
Portugal	10,132	3,867	2,620.12	1,196	1,485
Romania	19,600	3,182	6,159.65	1,659	2,979
Slovak Republic	5,435	2,888	1,881.93	1,051	938
Slovenia	1,915	3,074	622.97	1,496	222
Spain	47,190	8,115	5,815.16	513	15,851
Sweden	8,003	1,912	4,185.67	249	4,518
United Kingdom	52,518	3,121	16,827.30	206	46,357

4.4 Results

In this section we present the results. First we provide an accurate description of agglomeration. This involves a static analysis, supported by Zipf's law, accounting for the way the population is gathered over different geographic locations. Second, we move on the analysis of the homogenous/heterogeneous areas within the EU and we provide the dynamic part of the story. This second part provides the dynamic analysis, supported by Gibrat's law, accounting for the evolution over time of the distribution of the population over different locations. Here we are able to determine if EU city system can be seen as an integrated area and in particular if a temporary shock has a permanent impact on the pattern of city growths. Finally, once population mobility is understood, we explore the impact of the underlying economic mechanism and of the constitution of the EU, the Euro and the Schengen Area on the agglomeration forces.

4.4.1 Hierarchical structure and its determinants

In this section, we discuss the results from the following equation (rank-size rule with a Gabaix and Ibragimov, 2011 modification):

$$ln(R_i - 0.5) = k - q\, lnP, \tag{4.5}$$

where k is a constant and P is the population size. Standard errors are given by $(2/n)^{0.5}\,\hat{q}$. The parameter $\hat{q}$ is the rank-size coefficient that will be used in the second step. Table 12 and Table 13 present the detailed results of the OLS regressions of Eq. 4.5. We show just the first and the latter results for any Member States, however full details are available from the author upon request.

The rank-size coefficient, *q*, is an index of metropolization, in the sense that the lower the *q*-coefficient, the more important the urban land value. In fact the *q*-coefficient measures how (un)even the city system is, namely this means that the higher the *q*-coefficient, the more equally distributed the city system (in the extreme, when $q=\infty$, the system of cities is very even, all cities have the same size; when $q=0$, instead, the system will be composed just by one city hosting the entire population. Finally, when $q=1$ the city system is said to obey Zipf's law).

The largest value of the rank-size coefficient in 2011 (4.049) is obtained for Malta, followed by Belgium (1.724) results in line with those found by Soo (2005), whereas the lowest value is obtained for Denmark (0.949).

Table 12: Rank-size coefficients and city thresholds of the 27 EU Member States (first and last observation available)

Country	Period	Rank-size coefficient (first year available	Upper Tail	Rank-size coefficient (last year available	Upper Tail
Austria	1981-2011	1.469 (0.0645)	1036	1.440 (0.0796)	654
Belgium	1990-2011	1.691 (0.1713)	195	1.724 (0.1798)	184
Bulgaria	1985-2011	1.114 (0.0383)	1686	1.021** (0.0385)	1405
Cyprus	2001-2011	0.773 (0.0756)	209	0.994** (0.1717)	69
Czech Republic	1996-2011	1.014** (0.0318)	2025	1.066** (0.0347)	1891
Denmark	1976-2011	0.876 (0.0377)	1078	0.949** (0.0676)	392
Estonia	2001-2011	1.096** (0.1142)	184	1.141** (0.1929)	70
Finland	1990-2010	1.226** (0.1380)	158	1.250** (0.1977)	81
France	1975-2009	1.039** (0.0174)	7167	1.200 (0.0292)	3337
Germany	1991-2011	1.284 (0.0619)	859	1.322 (0.0646)	838
Greece	1991-2001	1.182 (0.0061)	544	1.164 (0.0077)	492
Hungary	1980-2011	1.137 (0.0490)	1075	1.110* (0.0493)	1012
Ireland	1991-2011	0.787 (0.0453)	602	0.987** (0.0815)	293
Italy	1991-2011	1.300 (0.0465)	1563	1.400 (0.0579)	1170
Latvia	2001-2009	1.159* (0.0805)	415	1.079** (0.0791)	372
Lithuania	1989-2011	0.904** (0.2506)	26	0.948** (0.2449)	30
Luxembourg	1821-2011	2.139 (0.4322)	49	1.197** (0.1784)	90
Malta	1901-2011	1.084** (0.2395)	41	4.049 (0.2597)	13
Netherlands	2001	-	-	1.199** (0.1282)	175
Poland	1988-2010	1.351 (0.0485)	1553	1.365 (0.0582)	1099
Portugal	2001-2011	1.159 (0.0466)	1239	1.115* (0.0456)	1196
Romania	2011	-	-	1.402 (0.0487)	1659

* Significant at 1% ** significant at 5%; for rank-size coefficient significantly not different from 1

Table 13: (...continued) Rank-size coefficients and city thresholds of the 27 EU Member States (first and last observation available)

Country	Period	Rank-size coefficient (first year available	Upper Tail	Rank-size coefficient (last year available	Upper Tail
Slovakia	1991-2010	1.135 (0.0478)	1130	1.169 (0.0510)	938
Slovenia	2001-2011	1.203 (0.0420)	1639	1.204 (0.0440)	1496
Spain	1991-2011	1.110** (0.0615)	651	1.198 (0.0748)	513
Sweden	1990-2010	1.136** (0.1026)	245	1.085** (0.0972)	249
United Kingdom	1991-2001	1.459 (0.1636)	159	1.467 (0.1445)	206
Euro Area	2001	-	-	1.377 (0.00528)	2660
European Union (excluded Romania)	2001	-	-	1.401 (0.0526)	1419

* Significant at 1% ** significant at 5%; for rank-size coefficient significantly not different from 1

In more details, the rank-size coefficient, q, is significantly greater than 1 for 13 of our 27 countries, while a further 14 observations are significantly not different from 1. These results indicate a situation where the half (in general small countries, both in terms of economic and demographic characteristics) of the EU Member States presents a hierarchical structure strictly following a Zipf's distribution (i.e. the largest city is double of the second largest city, three times of the third and so on) and the other halve presents a more even distribution. Indeed, this means that for 13 countries agglomeration forces are stronger than the other 14 Member States (or on the same way, congestion forces are stronger in the other 14 countries), indicating a straight hierarchy. Notice that there is no cases where the rank-size coefficient is lower than 1.

These results are stronger than those found by Soo (2005) who presents 17 Member States with a coefficient significantly greater than 1 and two Member States with a coefficient not different from 1.[7] The differences in the results, in our opinion, come from two sources: first of all, the data-set

[7] In Soo (2005) Cyprus, Estonia, Ireland, Latvia, Lithuania, Luxembourg, Malta, Slovenia are not in the international analysis.

in the present chapter is shaped to take directly in consideration agglomeration forces using the Clauset et al. (2009) method. This method is able to recognize the cut-off where agglomeration forces are stronger and this can lead to an increase of the cases where the Zipf's law strictly holds. Secondly, we use a modified rank-size rule that solves the problem of the downward biasness in small samples (Gabaix and Ibragimov, 2011).

To sum up, this first picture presents a situation where the European Union is exactly split into two halves. Small countries (in terms of lands) as Cyprus, Denmark, Luxembourg, Netherlands and Portugal, or transition countries as Bulgaria, Czech Republic, Estonia and Hungary are experimenting strong forces toward agglomeration and present a highly hierarchical city system. Instead, big (in terms of land size) countries as Poland, Romania and developed (in economic terms) countries as Belgium, France, Germany, Italy, Spain and the United Kingdom present a more even distribution of cities where the congestion forces play an important role.

Finally, considering the EU as a whole (country), the estimated rank-size coefficient is 1.40 indicating a quite even city distribution. It is interesting to compare this result with the estimated q coefficient of the aggregation of countries using the Euro. The estimated coefficient is almost the same (1.38), although it should be noticed that the estimation of the upper tail with the method proposed by Clauset et al. (2009) returns a longer upper tail for EMU, 2660, than EU, 1419.

Until now we have presented just a static draw of the European situation. We will explore the impact of the underlying economic mechanism and of the constitution of the EU, the Euro and the Schengen Area on the agglomeration forces in the sub-section 4.4.3. In the next section, instead, we address the homogeneity issue.

4.4.2 Is the EU an Integrated Area? The role of hypothetical shocks

In this section we analyze the homogeneity/heterogeneity of city systems in term of growth within the EU using Gibrat's law. Gibrat's law in assessing heterogeneity is not conventional, however Black and Henderson (2003), state that if Gibrat's law holds, then a shock affects in the same way big vs small entities and, Brakman, Garretsen and Schramm (2004), state that if Gibrat's law holds, then a large temporary shock can have a permanent impact. In other words, this means that hypothetical shocks can change (permanently) the growth path toward another size equilibrium.

In general we can interpret Gibrat's law as follows: "*A variate subject to a process of change is said to obey the law of proportionate effect if the change in the variate at any step of the process is a random proportion of the previous value of the variate*" (Chesher, 1979, p. 403). Given this definition, Gabaix (1999) affirms that a growth process that has a common mean and a common variance follows Gibrat's law or, in other words, that both the mean and variance have to be independent from the initial city size.

Moreover, proposition 2 of Gabaix (1999) affirms that if a country is composed of several regions and Gibrat's law holds in each of those regions, then in the whole country Zipf's law is satisfied both at regional and national level. This means that if the growth processes are (in the upper tail) identical within each region, but not necessarily across regions, then in the whole country the strength of the agglomeration forces in the system of cities is necessarily the same. This issue has been firstly addressed empirically by Giesen and Suedekum (2011) for Germany. In this section, we take the moves from those papers and we provide a countrywide test of Gibrat's law from a EU Member States' perspective.

Operationally, a typical way to assess the validity of Gibrat's law is using non-parametric analysis (Ioannides and Overman, 2003; Eeckhout, 2004; Giesen and Suedekum, 2011; Gonzalez-val et al., 2012; Modica et al., 2013). In doing that, we use the normalized growth rate, that is the difference between a city's growth rate and the mean city growth rate, all divided by the standard deviation of growth.

The strength of non-parametric analysis is that we do not impose any relationship between the dependent and independent variables and, moreover, we can see the behavior of mean and variance at any possible truncation, avoiding in this way the concern about the right selection of the truncation point arose by Fazio and Modica (2012). This is also stressed by the following sentence in Cameron and Trivedi (2005, p. 294), "[non-parametric analysis] *let the data show the shape of the relationship*". In fact, the standard parametric regression methods provide an aggregate relationship between growth and size that holds over the entire support of city sizes. Instead, the non-parametric method allow the growth to vary with size over the support.

In this analysis, we use the Nadaraya-Watson (NW) method (Nadarya 1964; Watson 1964), where the bandwidths are calculated with an optimal rule of thumb.[8] If Gibrat's law holds, the NW method provides stable estimated conditional mean and variance across different population sizes, that on the light of the normalization are respectively 0 and 1. We also calculate 5% bootstrapped confidence interval based on 500 samples (see Hardle, 1990). If Gibrat's law does not hold, instead, the series of growth rates is stationary and it will converge toward a constant mean. Therefore, large values (big entities) will be followed at any further periods by smaller and smaller values, while small values (small entities) will be followed by larger and larger ones. In this way, the level of the series at any periods can significantly predict the one of the next period.

The figures below provide the non-parametric estimates of conditional mean and variance for all the Member States.

Figures 7-9 show that, although the EU is still far from being an integrated area, we are able to recognize a homogenous area (in terms of city growth) where exogenous shocks might affect the city structure in the same way (Gibrat's law holds). In details, this area is composed by the following Member States: Belgium, Cyprus, Denmark, Estonia, Finland, France, Germany, Ireland, Lithuania, Malta. In all these cases, indeed, the conditional mean and the conditional variance look independent from the city size.

In the rest of the countries Gibrat's law does not hold and in particular the growth series presents mean reversion and stationarity namely big cities are expected to grow less than small ones (Black and Hendreson, 2003) and this fact is characterized by a conditional mean growth

[8] We refer readers to their papers for a more detailed description.

Figure 7: NW estimator of mean normalized growth rate and normalized variance of the growth rate (different countries and years)

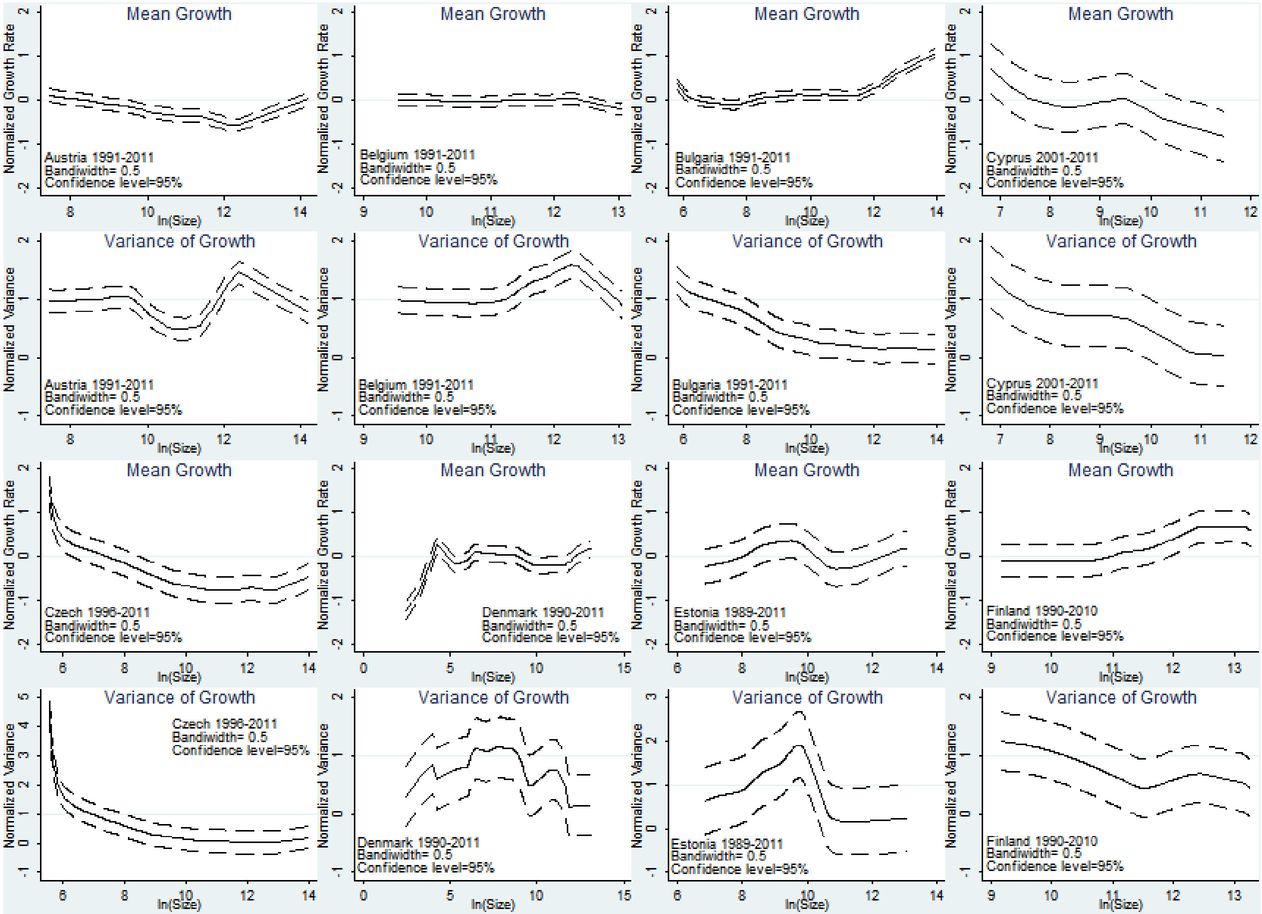


Figure 8: NW estimator of mean normalized growth rate and normalized variance of the growth rate (different countries and years)

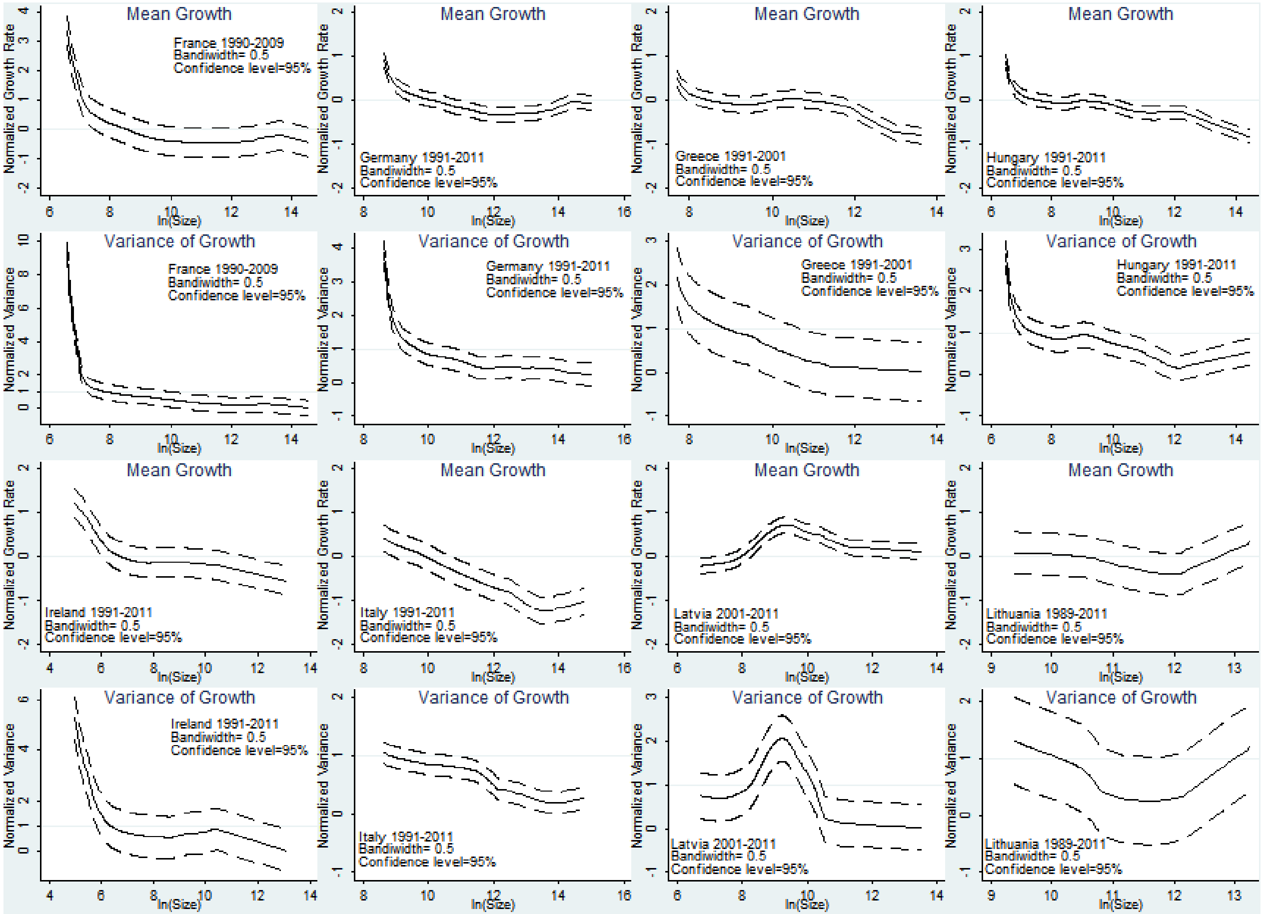
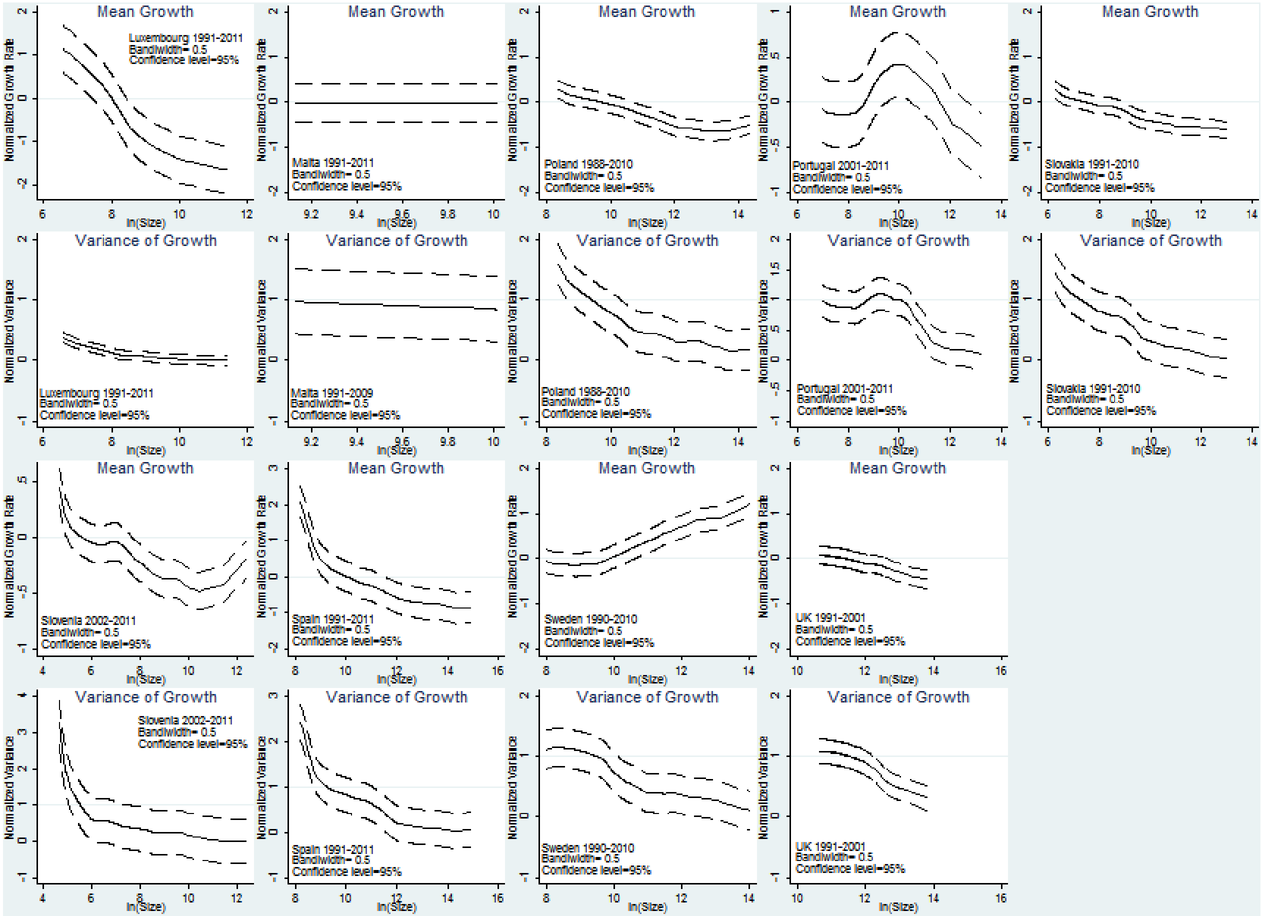


Figure 9: NW estimator of mean normalized growth rate and normalized variance of the growth rate (different countries and years)



and variance that are respectively lower than 0 and 1 for larger size of cities. Those countries are Austria, Czech, Greece, Hungary, Italy, Luxembourg, Poland, Portugal, Slovakia, Slovenia, Spain and UK, and all present mean reversion in the growth process. In this situation, then, any exogenous shock has to be though as temporary and the countries will converge toward the pre-shock conditions in a certain period of time.

Finally, very few countries show a situation where there exists an "explosive" growth process, that is a situation in which big cities are expected to grow more than small ones. In more details those Member States are Bulgaria, Latvia and Sweden. They present a situation where people tend to moves from the small city to the big ones and it is shown by a conditional mean growth and variance that are respectively greater than 0 and 1 for larger size of cities.[9]

To sum, the European Union seems to be split in three distinct areas: an integrated area characterized by the validity of Gibrat's law where large temporary shocks might have permanent impacts on the city structure; another area that is characterized by the presence of mean reversion and where any exogenous shock is used up in certain amount of time and a small area where the effects of the shocks is magnified in the big cities.

Given the static and dynamic picture of the European Union, it remains to explore the impact of the underlying economic mechanism and of the constitution of the EU, the Euro and the Schengen Area on the agglomeration forces. This is studied in deep in the next sub-section.

4.4.3 The impact of economic/institutional variables on agglomeration forces

In this sub-section we try to answer to the following question: does the creation of the European Union with free movement of goods, services people an capital, change the population distribution across geographic areas? The rank-size coefficient, indeed, can be seen as a measure of urbanization: the larger the value of the coefficient, the more even the population of cities in the urban system.

[9]Unfortunately we don't have enough data to explore Gibrat's law in the Netherlands and Romania

There are several potential explanations for variations in its value, one of these can be found in a model of economic geography a la Krugman (1991) and Fujita et al. (1999). These models can be viewed as models of unevenness in the distribution of economic activity and moreover, they state that for certain parameter values, economic activity is agglomerated, while for other parameter values, economic activity is dispersed (i.e. a city system will be more agglomerated the greater are scale economies, the lower are transport costs and the lower the share of international trade in the economy). Henderson and Wang (2007), instead, emphasize on the role of the institutions. The constitution of the EU, then, might had affected the urban structure of the Member States.

In this section we look for both the impact of institutional and economic variables on the level of the hierarchical structure. We also control for other variables that could influence the size distribution of cities, including the size of the country as measured by population, land area or GDP. Following Soo (2005) our estimated equation is:

$$\mathrm{q}_{it} = \beta_0 + \beta_1 CONTROL + \beta_2 ECON + \beta_3 POLITIC + \beta_4 DUMMY + u_{it} \tag{4.6}$$

where q is is the rank-size coefficient estimated in the previous step, CONTROL is the vector of controls for the size of the country, including the log of per capita GDP in constant US dollars, the log of the total land of the country, and the log of population. ECON is a group of economic-geography and connectivity variables: scale economies, transport costs (the inverse of road density), trade as a percentage of GDP, number of internet users per 100 people.

Scale economies is the degree of scale economies, and it is constructed following Soo (2005). They are the share of industrial output in high-scale industries. Transport cost is measured as the inverse of road density.

POLITIC is a set of political variables as government expenditure as a share of GDP in education and foreign direct investment as net inflow (% of GDP). Finally, DUMMY indicates the variables strictly related to the creation of European Union (belonging to EU, adhesion to Schengen and interaction between the two).

One potential concern, stressed by Soo (2005), is the fact that using an estimated coefficient from a former stage regression as a dependent variable in a subsequent stage regression it might lead to inefficient estimates and heteroskedasticity in the second stage because of likely measurement error in the first stage, Lewis and Linzer (2005). We could use feasible generalized least squares (FGLS) to overcome this problem. However, Baltagi (1995) and Beck and Katz (1995) show that: i) FGLS yields consistent estimates of the variances only if T goes to infinity and, ii) FGLS tends to underestimates standard errors.

In the present chapter, following Soo (2005), we use, first, panel corrected standard errors with OLS, as proposed by Beck and Katz (1995) because it does not underestimate standard errors and second we run time fixed effect and random effect estimations. The regressions are those that are reported below.

Table 14 presents the results using the OLS estimate of the Pareto exponent as the dependent variable. The number of observations is somewhat less than the full sample because data is not available for all countries in all years. Column (1) is the panel corrected standard errors with OLS model without dummy variables. Column (2) includes all the variables and Column (3) and (4) are, respectively, time fixed effect and random effect models. Two factors seem to have a significant impact on the city hierarchical structures of the countries: the extent of international trade and the adhesion to the Schengen Area. Moreover, population and land size respectively positively and negatively influence the city structure of the Member States.

In more detail international trade enters with a (small) negative impact on the rank-size coefficient, that is something not predicted from the theory since, following Fujita et al. (1999), a greater extent of international trade weakens the force for agglomeration and leads to a more even distribution of economic activity. However this result is in line with previous findings of Soo (2005) and Alperovich (1993).

Column (4) instead shows that the adhesion to the Schengen Area

Table 14: Panel estimation of Eq. (6) (dependent variable = OLS coefficient of q)

Dependent Variable	(1) OLS	(2) OLS	(3) FE	(4) RE
Trade (% of GDP)	-.00398**	-.00382**	-.00228***	-.00173
	(.00167)	(.00157)	(.00134)	(.00193)
Scale Economies	-.00153	.01735	-.00215	.00237
	(.01150)	(.01336)	(.02064)	(.02807)
Expenditure in education (% of GDP)	-.00944	-.01972	-.04112	.02409
	(.04172)	(.04647)	(.03071)	(.05024)
Transport cost	.63581	.71197	-.82159	.789131
	(1.6369)	(2.0571)	(1.9224)	(4.06137)
Foreign direct investment	-.00071	-.00113	-.00022	-.00015
	(.00083)	(.00094)	(.00068)	(.00093)
Internet users	.00340	.00268	.00379	-.00005
	(.00222)	(.00205)	(.00385)	(.00228)
ln (population)	.21367*	.19841*	.06791	.24237*
	(.04678)	(.05319)	(.04653)	(.08341)
ln (land area)	-.41339*	-.40515*	-.13959*	-.36686
	(.10989)	(.10916)	(.04491)	(.07509)
ln(GDP)	.020184	-.00061	.03152	-.00688
	(.0546245)	(.05700)	(.05631)	(.09203)
Dummy Euro		-.23996	-.13857	.00029
		(.16537)	(.17491)	(.24136)
Dummy Schengen		.12124	-.13399	.22971***
		(.16815)	(.12181)	(.13477)
Interaction		.18028	.31286	-.13661
		(.25287)	(.19222)	(.26264)
Constant	2.68183**	3.0277*	1.81866***	1.5775
	(1.0473)	(1.0788)	(.92899)	(1.3881)
R-squares	0.5626	0.5724	0.3616	0.5370
N. obs.	66	66	66	66

* Significant at 1% ** significant at 5% *** significant at 10%

leads to a more even distribution (higher rank-size coefficient). This result can be interpreted as the results of the increased freedom of mobility between Member States that leads to the movements from the big cities (especially in the transition countries) to the small cities.

Comparing our results to previous findings, we find that our results are broadly in line with those of Rosen and Resnick (1980) as they find that the rank-size coefficient is positively related to the total population and negatively related to land area. However, we found lack of significance for most of the economic and political variables presented by Soo (2005).

In conclusion, it seems that the constitution of the European Union *per se* does not have any influence on the change of the population distribution across geographic areas. The only variables that have influenced the city structure of the Member States are those strictly related the increases of the mobility of people, capitals and services (the constitution of the Schengen area) and a variable closed to the idea of globalization, that is the share of international trade.

4.5 Conclusion

The aim of our research work was to draw the current city system of the Member States of the European Union firstly for any single country and subsequently as a whole state. Given this picture we aim to explore if and how the creation of European Union affects the structure of the system of cities of the Member States and primarily if EU city system can be seen as an integrated area. The main objective was the study of the agglomeration forces within all the Member States and the EU as a whole and for this reason we used two very well-known empirical regularities that address (indirectly) this issue, namely Zipf's law and Gibrat's law.

For this purpose we used the rank-size exponent, q, as a proxy for agglomeration forces in a system of cities (Singer, 1930; Brakman et al., 2001) since the lower the q-coefficient, the more agglomerated is the city distribution (Soo, 2005). Gibrat's law, instead, is used as a tool to test whether a (hypothetical) shock (will) influence the EU Member States in the same way or, on the contrary, if it (will) has heterogeneous effects. Indeed, if Gibrat's law holds a large temporary shock might have a permanent and similar impact on the growth path of the cities, this means that a shock can change the growth path toward another size equilibrium.

We started our analysis providing an accurate description of agglomeration. This has involved a static analysis, supported by Zipf's law, accounting for the way the population is gathered over different geographic locations and, a dynamic analysis, supported by Gibrat's law, accounting for the evolution over time of the distribution of the population over different locations.

We showed that the hierarchical structures of Member States of EU is more agglomerated than expected.

Given this picture, the dynamic analysis, supported by Gibrat's law, showed that EU city system is still far from being an integrated area and in particular, we found that the European Union seems to be split in three distinct areas: an integrated area characterized by the validity of Gibrat's law where large temporary shocks might have permanent impacts on the city structure; another area that is characterized by the presence of mean reversion and where any exogenous shock is used up in certain amount of time and a small area where the effects of the shocks is magnified in the big cities.

Finally, by means of a regression analysis we explored how the creation of European Union affects the structure of the system of cities and then the agglomeration forces of the Member States. Our results showed that only the constitution of the Schengen Area and the share of international trade seem to have a weak impact on the hierarchical structures of Member States presenting respectively a positive impact on the Zipf coefficient (more even distribution) and a negative impact (more agglomerated distribution).

In conclusion Gibrat's law and Zipf's law are here used to reflect important organized structures in the topology of systems of cities. However this study can be extended in several directions taking into account for instance the persistent existence of socio-economic disparities between

Member States (based, for example, on Gini index) or considering the role of the new communication technologies on the system of cities or finally considering the transport infrastructure in a integrated system. Clearly, the dynamics of such processes deserve due attention.

References

Chapter 2

Bee, M., Riccaboni, M., Schiavo, S., 2011. Pareto Versus Lognormal: a Maximum Entropy Test. Phys. Rev. E. 84, 026-104.

Black, D., Henderson, V., 2003. Urban Evolution in the USA. Journal of Economic Geography. 3(4), 343–372.

Brakman, S., Garretsen, H., van Marrewijk, C., 2001. An Introduction to Geographical Economics. Trade, Location and Growth. Cambridge: Cambridge University Press.

Clauset, A., Shalizi, C.R., Newman, M. E. J. , 2007. Power-law Distributions in Empirical Data. ArXiv: 0706.1062.

Eeckhout, J., 2004. Gibrat's Law for (All) Cities. American Economic Review. 94(5), 1429–1451.

Eeckhout, J., 2009. Gibrat's Law for (All) Cities: Reply. American Economic Review. 99(4), 1676–83.

Gabaix, X., 1999. Zipf'S Law For Cities: An Explanation. The Quarterly Journal of Economics. 114(3), 739–767.

Gabaix, X., Ibragimov, R., 2011. Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents. Journal of Business & Economic Statistics. 29(1): 24–39.

Gabaix, X. Ioannides, Y.M., 2004. The Evolution of City Size Distributions in a J. V. Henderson & J. F. Thisse (eds.), Handbook of Regional and Urban Economics. Amsterdam: Elsevier, pp. 2341–2378.

Gibrat, R., 1931. *Les inégalités économiques*. Paris: Libraire du Recueil Siray.

Giesen, K., Zimmermann, A. Suedekum, J., 2010. The Size Distribution Across All Cities - Double Pareto Lognormal Strikes. Journal of Urban Economics. 68(2), 129–137.

Glaeser, E. L., Ponzetto, G.A M., Tobio, K., 2011. Cities, Skills, and Regional Change. Working Paper No. 16934, NBER.

Guerin-Pace, F., 1995. Rank-size Distribution and the Process of Urban Growth. Urban Studies. 32(3), 551–562.

Ioannides, Y.M., Overman, H.G., 2003. Zipf's Law for Cities: An Empirical Examination. Regional Science and Urban Economics. 33(2), 127–137.

Ioannides, Y.M., Skouras, S., 2013. US City Size Distribution: Robustly Pareto but Only in the Tail. Journal of Urban Economics. 73(1), 18–29.

Krugman, P., 1996. Confronting the Mystery of Urban Hierarchy. Journal of the Japanese and International Economies. 10(4), 399–418.

Levy, M., 2009. Gibrat's Law for (All) Cities: Comment. American Economic Review. 99(4), 1672–75.

Macauley, F., 1922. Pareto's law and the General Problem of Mathematically Describing the Frequency Distribution of Income in a Wesley Clair Mitchell (eds.), *Income of the United States. Its Amount and Distribution 1909-1919*. New York: National Bureau of Economic Research, pp. 344–394.

Mitzenmacher, M., 2001. A Brief History of Generative Models for Power Law and Lognormal Distributions. Internet Mathematics. 1, 226–251.

Parr, J.B., 1985. A Note on the Size Distribution of Cities over Time. Journal of Urban Economics. 18, 199-212.

Parr, J.B., Suzuki, K., 1973. Settlement Populations and the Lognormal Distribution. Urban Studies. 10(3), 335–352.

Peng, G., 2010. Zipf's Law for Chinese Cities: Rolling Sample Regressions. Physica A: Statistical Mechanics and its Applications. 389(18), 3804 – 3813.

Perline, R., 2005. Strong, Weak and False Inverse Power Laws. Statistical Science. 20(1), 68–88.

Reggiani, A. Nijkamp, P., 2012. Did Zipf Anticipate Socio-economic Spatial Networks? Working Paper No. 816, Quaderni DSE, University of Bologna.

Rosen, K.T., Resnick, M., 1980. The Size Distribution of Cities: An examination of the Pareto Law and Primacy. Journal of Urban Economics. 8(2), 165 – 186.

Singer, H.W. ,1930. The Courbe Des Populations. A Parallel to Pareto's law. The Economic Journal. 48(182), 254-263.

Soo, K.T., 2005. Zipf's Law for Cities: A Cross-Country Investigation. Regional Science and Urban Economics. 35(3), 239–263.

Zipf, G.K., 1949. *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.

Chapter 3

Adamic, L. A., 2000. Zipf, power-Laws, and Pareto a ranking tutorial, retrieved from www.hpl.hp.com

Auerbach, F., 1913. Das Gesetz Der Bevolkerungs Konzentration, Petermanns Geographische Mitteilungen. 59, 73-76.

Auerbach, F., 1915. The law of population concentration. Scientific American Supplement. 76, 384.

Barabsi, L.A., Oltvai, Z. N., 2004. Networks biology: understanding the cells functional organization. Nature Reviews-Genetics. 5, 101-113.

Berry, B. J. L., Okulicz-Kozarin, A., 2012. The city size distribution debate: resolution for US urban regions and megalopolitan areas. Cities. 29, 17-23.

Black, D., Henderson V., 2003. Urban evolution in the USA. Journal of Economic Geography. 3(4), 343-372.

Blank, A., Solomon S., 2000. Power laws in city population, financial markets and internet sites (scaling in systems with a variable number of components). Physica A. 287, 279-288.

Bottazzi, G., Dosi, G., Lippi, M., Pammolli, F., Riccaboni, M., 2001. Innovation and corporate growth in the evolution of the drug industry. International Journal of Industrial Organization. 19(7), 1161-1187.

Brakman, S., Garretsen, H., van Marrewijk, C., 2001. An Introduction to Geographical Economics. Trade, Location and Growth, Cambridge University Press, Cambridge.

Cameron, A. C., Trivedi, P. K., 2005. Microeconometrics, Cambridge University Press, Cambridge.

Carrol, G. R., 1982. National city-size distribution: what do we know after 67 years of research?. Progress in Human Geography. 6, 1-43.

Champernowne, D. G., 1953. A model of income distribution. The Economic Journal. 63(250), 318-351.

Chesher, A., 1979. Testing the law of proportionate effect. Journal of Industrial Economics. 27, 403-411.

Christaller, W., 1933. Die Zentrale Orte in Sddeutschland. *Eine konomisch-Geografische Untersuchung ber Die Gesatzmssigkeit Der Verbreitung Und Entwicklung Der Siedlungen Mit Stdtischen Funktionen*, Gustav Fischer, Jena.

Clark, J. S., Stabler, J. C., 1991. Gibrats law and the growth of Canadian cities. Urban Studies. 28(4), 635-639.

Cordoba, J., 2003. On the distribution of city sizes. Urban/Regional. 0302002, EconWPA.

Cordoba, J., 2008. A generalized Gibrats law. International Economic Review. 49(4), 1463-1468.

DeCarlo, L. T., 1997. On the meaning and use of kurtosis. Psychological Methods. 2(3), 292-307.

Dittmar, J., 2011. Cities, markets and growth: the emergence of Zipfs law. Working Paper, retrieved from http://www.jeremiahdittmar.com.

Eeckhout, J., 2004. Gibrats law for (all) cities. American Economic Review. 94(5), 1429-1451.

Eeckhout, J., 2009. Gibrats law for (all) cities: reply. American Economic Review. 99(4), 167-683.

Fazio, G., Modica, M., 2012. Pareto or log-normal? A recursive truncation approach to the distribution of (all) cities. Working Papers 2012_10, Business School, Economics, University of Glasgow, Glasgow.

Fujiwara, Y., Di Guilmi, C., Aoyama, H., Gallegati, M., Souma, W., 2003. Do Pareto-Zipf and Gibrats law hold true? An analysis with European firms, ArXiv e-prints.

Gabaix, X., 1999. Zipfs law for cities: an explanation. The Quarterly Journal of Economics. 114(3), 739-767.

Gabaix, X., Ibragimov, R., 2011. Rank-1/2: a simple way to improve the OLS estimation of tail exponents. Journal of Business Economic Statistics. 29(1), 24-39.

Gabaix, X., Ioannides, Y. M., 2004. The evolution of city size distributions, in J.V. Henderson and J.-F. Thisse (eds), Handbook of Regional and Urban Economics, Vol. 4. Elsevier, Amsterdam, 23412378.

Gibrat, R., 1931. Les Ingalits Economiques, Libraire du Recueil Siray, Paris.

Giesen, K., Suedekum, J., 2011. Zipfs law for cities in the regions and the country. Journal of Economic Geography. 11(4), 667-686.

Gonzlez-Val, R., 2010. The evolution of U.S. city size distribution from a long-term perspective (1900-2000). Journal of Regional Science. 50(5), 952-972.

Gonzlez-Val, R., Lanaspa, L., Sanz, F., 2012. New evidence on Gibrats law for cities. Working Papers 2012/18. Institut de Economia de Barcelona (IBE), Barcelona.

Guerin-Pace F., 1995. Rank-size distribution and the process of urban growth. Urban Studies. 32(3), 551-562.

Ioannides, Y. M., Overman, H. G., 2003. Zipfs law for cities: an empirical examination. Regional Science and Urban Economics. 33(2), 127-137.

Kalecki, M., 1945. On the Gibrat distribution. Econometrica. 13(2): 161-170.

Levy, M., 2009. Gibrats law for (all) cities: comment. American Economic Review. 99(4), 1672-75.

Lösch, A., 1940. The Economics of location. Yale University Press, New Haven (translated by W. Woglam and W. Stolper in 1954).

Nadaraya, E. A., 1964. Some new estimates for distribution functions. Theory of Probability Its Applications. 9(3), 497-500.

Neal, Z.P., 2012. The connected city. Routledge, London.

Newman, M.E.M., (2010). Networks: An Introduction, Oxford University Press, Oxford.

Paelinck, J.H.P., Nijkamp, P., 1976. Operational Theory and Method in Regional Economics, Saxon House, Farnborough.

Parr, J. B., 1985. A note on the size distribution of cities over time. Journal of Urban Economics. 18, 199-212.

Parr, J. B., Suzuki, K., 1973. Settlement populations and the log-normal distribution. Urban Studies. 10(3), 335-352.

Perline, R., 2005. Strong, weak and false inverse power laws. Statistical Science. 20(1), 68-88.

Reggiani, A., Nijkamp, P., 2012. Did Zipf anticipate socio-economic spatial networks?, Working Paper No. 816, Quaderni DSE, University of Bologna, Bologna.

Reggiani, A., Schintler, L. A., 2005. Methods and models in transport and telecommunications, Cross Atlantic Perspectives, Springer-Verlag, Berlin.

Richardson, H., 1973. Theory of the distribution of city sizes: review and prospects. Regional Studies. 8, 239-251.

Rosen, K. T., Resnick, M., 1980. The size distribution of cities: An examination of the Pareto law and primacy. Journal of Urban Economics. 8(2), 165 -186.

Simon, H., 1955. On a class of skew distribution functions. Biometrika. 42(3/4), 425-440.

Singer, H. W., 1930. The Courbe Des Populations. A parallel to Paretos law. The Economic Journal. 48(182), 254-263.

Soo, K. T., 2005. Zipfs law for cities: a cross-country investigation. Regional Science and Urban Economics. 35(3), 239-263.

Steindl, J., 1968. Size distributions in Economics, in International Encyclopedia of the Social Sciences, Vol. 14, Silks, ed. (New York, Macmillan and the Free Press).

Suarez-Villa, L., 1988. Metropolitan evolution, sectoral economic change, and the city size distribution. Urban Studies. 25, 1-20.

von Hippel, P. T., 2005. Mean, median and skew: correction a textbook rule. Journal of Statistics Education. 13(2).

Watson, G. S., 1964. Smooth regression analysis. The Indian Journal of Statistics, Series A. 26(4), 359-372.

Zipf, G., 1949. *Human behavior and the principle of least effort: an introduction to human ecology*, Addison-Wesley, Cambridge.

Chapter 4

Aldskogius, G., 2000. Urban policy in the structural policy of the European Union. CERUM Working Paper, Umea University, Faculty of Social Sciences, Centre for Regional Science, Umea.

Alperovich, G.A., 1993. An explanatory model of city–size distribution: evidence from cross–country data. Urban Studies. 30, 1591–1601.

Armington, C., Acs, Z., 2002. The determinants of regional variation in new firm formation. Regional Studies. 36(1), 33–45.

Auerbach, F., 1915. The law of population concentration. Scientific American Supplement. 76, 384.

Baker, S., Kousis, M., Richardson D., Young, S. (eds). 1997. The politics of sustainable development: theory, policy and practice within the EU. London: Routledge.

Baltagi, B.H., 1995. Econometric Analysis of Panel Data. Chichester: Wiley.

Beck, N., Katz, J.N., 1995. What to do (and not to do) with time-series cross-section data. American Political Science Review. 89, 634–647.

Berry, B.J.L., Okulicz-Kozarin, A., 2012. The city size distribution debate: resolution for US urban regions and megalopolitan areas. Cities. 29, 17–23.

Black, D., Henderson, V., 2003. Urban evolution in the USA. Journal of Economic Geography. 3(4), 343–372.

Brakman, S., Garretsen, H., van Marrewijk, C., 2001. An introduction to geographical economics. Trade, location and growth. Cambridge: Cambridge University Press.

Brakman, S., Garretsen, H., Schramm, M., 2004. The strategic bombing of German cities during World War II and its impact on city growth. Journal of Economic Geography. 4, 201–218.

Cameron, A.C., Trivedi, P.K., 2005. *Microeconometrics*. Cambridge: Cambridge University Press.

Chesher, A., 1979. Testing the law of proportionate effect. Journal of Industrial Economics. 27, 403-411.

Cheshire, P., 1999. *Trends in sizes and structures of urban areas*, in a Cheshire, P. and Mills, E.S. (eds.). Handbook of Regional and Urban Economics, vol. 3. Amsterdam: Elsevier, pp. 1339–1372.

Clauset, A., Shalizi, C.R., Newman, M.E.J., 2007. Power-law distributions in empirical data. arXiv: 0706.1062.

Dobkins, L., Ioannides, Y.M., 2001. Spatial interactions among U.S. cities: 1900-1990. Regional Science and Urban Economics. 31(6),701–731.

Eeckhout, J., 2004. Gibrat's law for (all) cities. American Economic Review. 94(5), 1429–1451.

European Commission, 2010. The urban dimension in European Union policies 2010. Introduction and part I. Brussels: Inter-Service Group on Urban Development.

Fay, M., Yepes, T., 2003. Investing in infrastructure: what is needed from 2000 to 2010?. Working Paper No. 3102, Policy Research Working Paper Series, The World Bank.

Fazio, G., Modica, M., 2012. Pareto or log-normal? A recursive truncation approach to the distribution of (all) cities. Working Papers 2012_10, Business School, Economics, University of Glasgow, Glasgow.

Fujita, M., Krugman, P., Venables, A.J., 1999. *The Spatial Economy: Cities, Regions and International Trade*. Cambridge: MIT Press.

Gabaix, X., 1999. Zipf'S law for cities: an explanation. The Quarterly Journal of Economics. 114(3), 739–767.

Gabaix, X., Ibragimov, R., 2011. Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents. Journal of Business & Economic Statistics. 29(1), 24–39.

Gabaix, X., Ioannides,Y.M., 2004. The Evolution of City Size Distributions in a J. V. Henderson & J. F. Thisse (eds.). *Handbook of Regional and Urban Economics*. Amsterdam: Elsevier, pp. 2341–2378.

Gibrat, R., 1931. *Les inégalités économiques*. Paris: Libraire du Recueil Siray.

Giesen, K., Suedekum, J., 2011. Zipf's law for cities in the regions and the country. Journal of Economic Geography. 11, 667–686.

Gonzlez–Val, R., Lanaspa, L., Sanz, F., 2012. New evidence on Gibrat's law for cities. Working Papers 2012/18. Institut de Economia de Barcelona (IBE), Barcelona.

Hardle, W., 1992. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.

Henderson, J.V., Wang, H.G., 2007. Urbanization and city growth: The role of institutions. Regional Science and Urban Economics. 37, 283–313.

Ioannides, Y.M., Overman, H.G., 2003. Zipf's law for cities: an empirical examination. Regional Science and Urban Economics. 33(2), 127–137.

Ioannides, Y.M., Skouras, S., 2013. US city size distribution: robustly Pareto but only in the tail. Journal of Urban Economics. 73(1), 18–29.

Jiang, B., Jia, T., 2010. Zipf's Law for All the Natural Cities in the United States: A Geospatial Perspective. arXiv: 1006.0814v2.

Kooij, P., 1988. Peripheral cities and their regions in the Dutch urban system until 1900. The Journal of Economic History. 48(2), 357–371.

Krugman, P., 1991. *Geography and Trade*. Cambridge: MIT Press.

Le Gals., 2002. Cities: Social Conflicts and Governance. New York: Oxford University Press.

Lewis, J.B., Linzer, D.A., 2005. Estimating Regression Models in Which the Dependent Variable Is Based on Estimates. Political Analysis. 13, 345–364.

Mera, K., 1975. On the Urban Agglomeration and Economic Efficiency. Economic Development and Cultural Change. 21(2), 309–324.

Modica, M., Reggiani, A., Nijkamp, P., 2013. Are Gibrat and Zipf Monozygotic or Heterozygotic Twins? A Comparative Analysis of Means and Variances in Complex Urban Systems. Working Paper.

Nadaraya, E.A., 1964. Some new estimates for distribution functions. Theory of Probability & Its Applications. 9(3), 497-500.

Nahoi, O., 2011. War in Europe? Not so impossible. Eurozine.

Pacione,, M., 2001. *Urban Geography: A Global Perspective*. London: Routledge.

Randolph, S., Bogetic, Z. Hefley, D., 1999. Determinants of Public Expenditure on Infrastructure: Transportation and Communication. Working Paper No. 1661, Policy Research Working Paper Series, The World Bank.

Rosen, K.T., Resnick, M., 1980. The size distribution of cities: an examination of the Pareto law and primacy. Journal of Urban Economics. 8(2), 165 – 186.

Rozenfeld, H.D., Rybski, D., Gabaix, X., Makse, H.A., 2011. The area and population of cities: new insights from a different perspective on cities. American Economic Review. 101(5), 2205–25.

Singer, H. W., 1930. "The Courbe Des Populations". A parallel to Pareto's law. The Economic Journal. 48(182), 254–263.

Soo, K.T., 2005. Zipf's Law for cities: a cross-country investigation. Regional Science and Urban Economics. 35(3), 239–263.

Steindl, J., 1968. Size distributions in Economics, in *International Encyclopedia of the Social Sciences*, Vol. 14, Silks, ed. New York: Macmillan and the Free Press.

Sutton, K.T., 1997. Gibrat's legacy. Journal of Economic Literature. 35(1), 40–59.

United Nations, 1974. Manual VIII: Methods for projections of urban and rural population. United Nations publication, Sales No. E.74.XIII.3.

United Nations Statistical Division (UNSTAT), 2011. United Nations Demographic Yearbook 2009-2010. United Nations Statistical Division (UNSTAT).

Watson, G.S., 1964. Smooth regression analysis. The Indian Journal of Statistics, Series A. 26(4), 359–372.

World Commission on Environment and Development, 1987. *Our common future*. Oxford: Oxford University Press.

Yang, D.T., 1999. Zipf's Law for cities: a cross-country investigation. American Economic Review. 89(2), 306–310.

Zipf, G.K., 1949. *Human Behavior and the Principle of Least Effort*. Cambridge, MA: Addison-Wesley.