Zipf's law and city size distribution: a survey of the literature and future research agenda

Sidra Arshad[a], Shougeng Hu[b,c], Badar Nadeem Ashraf[d,e,*]

[a]School of Public Administration, China University of Geosciences (CUG), Wuhan 430074, PR China

[b]Department of Land Resources Management, China University of Geosciences (CUG), Wuhan 430074, PR China

[c]Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, PR China

[d]School of Economics and Management, China University of Geosciences Wuhan, Wuhan 430074, China

[e]International School, East China Jiao Tong University, Nanchang 330013, Jiangxi, PR China

[*]Corresponding author: badarfcma@gmail.com

Abstract

This study provides a systematic review of the existing literature on Zipf's law for city size distribution. Existing empirical evidence suggests that Zipf's law is not always observable even for the upper-tail cities of a territory. However, the controversy with empirical findings arises due to sample selection biases, methodological weaknesses and data limitations. The hypothesis of Zipf's law is more likely to be rejected for the entire city size distribution and in such case the alternative distributions have been suggested. On the contrary, the hypothesis is more likely to be accepted if better empirical methods are employed and cities are properly defined. The debate is still far from to be conclusive. In addition, we identify four emerging areas in Zipf's law and city size distribution research including the size distribution of lower-tail cities, the size distribution of cities in sub-national regions, the alternative forms of Zipf's law, and the relationship between Zipf's law and the coherence property of the urban system.

Keywords: city size distribution; Zipf's law; Gibrat's law; Pareto distribution; lognormal distribution; double Pareto lognormal distribution

1. Introduction

The empirical regularity often named as Zipf's law is famous in city size distribution literature. This regulatory has fascinated urban scientists since the study of Felix Auerbach [1] who first observed that the size distribution of cities fits a Pareto distribution. Later the idea was further refined by the others such as Singer [2] and Zipf [3]. Specifically, Zipf [3] provided an empirical analysis which suggested that city size distribution can be approximated with a Pareto distribution having a shape parameter (i.e., Pareto exponent) exactly equals to one. This rule is often referred as Zipf's law. In more general terms, Zipf's law implies that, in a system of cities, the largest city is roughly twice the size of the second largest city, about three times the size of the third largest city, and so on. If cities are ranked according to their size and drawn on a graph, the plot of the log of the ranks of cities versus the log of the sizes of cities shows a scatter diagram with a regression line having a slope equals to -1. Figure 1 shows a typical log-rank versus log-size plot for 135 largest urban areas of the United States in census year 2010. Slope of the fitted line is -1.019.

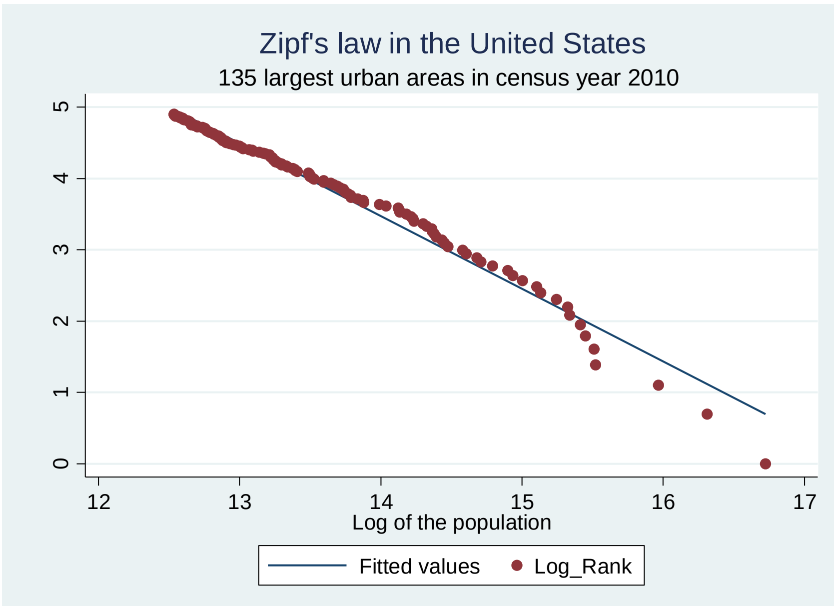

Figure 1: the plot of the log of the ranks versus the log of the sizes of the US cities

For a long time, it has been a common belief that Zipf's law holds for city size distribution (i.e., Zipf's law is universal for cities). Consequently, the urban empiricists have widely used the law as a benchmark to understand urban systems. The estimated value of Pareto exponent shows the hierarchical degree of a system of cities. When the Pareto exponent equals one, Zipf's law holds precisely. The higher the Pareto exponent, the more equally distributed is the city system. On the contrary, the smaller the value of exponent, the more uneven is the system of cities. On the extremes, when exponent=∞, the urban system is very even with all cities of same size. When exponent=0, the urban system is very uneven where one city hosts the entire urban population. The deviations from Zipf's law are considered as the evidence of distortions in urban systems, such as aggregations, segregations and efficiency losses [4]. These distortions then can be traced for possible causes that might be institutional, economic, localization of resources and historical accidents, among others [5-9].

However, a recent parallel literature is more skeptical of the law. For example, Benguigui and Blumenfeld-Lieberthal [10] argue that despite the widespread acceptance of Zipf's law, it is not always observable even as an approximation for city size distribution. In another study, Ausloos and Cerqueti [11] argue that a mere hyperbolic law, like the Zipf's law power function, is often inadequate to describe rank-size relationships. Similar skepticism is evident in empirical literature. For example, Nitsch [12] performed a meta-analysis of 29 publised studies and found that the combined estimate of Pareto exponent is significantly greater than one which suggests a more even city size distribution. Similarly, Soo [9] examined the city size distribution of 73 countries and found that Zipf's law is more often rejected than is expected based on a random chance.

The findings of existing studies have been criticized due to sample selection biases, methodological weaknesses and data limitations. Empirical studies usually use the data of large cities to examine Zipf's law. Opponents however argue that a certain set of cities from the upper-tail of city size distribution can be selected to support the hypothesis of Zipf's law. On the contrary, proponents debate that the fit to Zipf's law improves if appropriate empirical methods and proper definition of cities are used. As a result, some authors have examined Zipf's law for all cities to avoid sample selection bias and show that Zipf's law largely breaks down for the entire city size distribution and in such case the alternative distributions offer a better fit to cities. While at the same time, some others have suggested improvements in empirical methods and city definitions. In addition, several new areas have emerged in Zipf's law and city size distribution research. Due to all these developments, a voluminous literature has mushroomed on city size distribution in last few years. Since the topic is attractive to several fields of studies, the experts such as physicists, statisticians, geographers and economists have contributed to the debate. Though this multidisciplinary approach has helped to shed light on different aspects of the city size distribution, however it also adds to the complexity of the topic. The objective of this paper is to review this diverse literature to consolidate the findings of existing studies, highlight the most important issues in this literature and identify the potential areas for future research. Since the debate of city size distribution has heated up in recent years, our study would be a timely contribution to the debate at its peak.

Our study is different from the review of Nitsch [12] in several aspects. Nitsch [12] used a quantitative approach and did a meta-analysis of 29 empirical studies published from 1969 to 2002 to analyze Zipf's law for overall city size distribution. In contrast, we provide a qualitative review of city size distribution literature. We believe qualitative approach is helpful to shed light on city size distribution literature which in recent years has spread beyond Zipf's law. In this context, our review not only provides a comprehensive and critical analysis of contemporary issues in Zipf's law and city size distribution research, but also includes the studies which provide theoretical underpinnings to Zipf's law or suggest alternative distributions. Moreover, we also identify emerging areas in Zipf's law and city size distribution research. Finally, our review covers the most recent time period from 2000 to 2017 and the number of studies reviewed (114 studies) is comparatively very large.

Previewing the main results, we find that Zipf's law is not always observable even for the upper-tail cities of a territory. This result is not consistent with the common belief that Zipf's law holds precisely at least for upper-tail cities. We further find that the hypothesis of Zipf's law is too often rejected for the entire city size distribution, and in such case the alternative distributions have been suggested. However, the Zipf's law still stands as a strong alternative candidate for the entire city size distribution due to the controversy of how a city is defined and which statistical methods are used. The debate is still far from to be conclusive. Further, we identify several emerging areas in city size distribution research, including the size distribution of lower-tail cities, the size distribution of cities in sub-national regions, the alternative forms of Zipf's law and the coherence property of the urban system.

The paper is organized as follows: Section 2 introduces our literature search methods. In Section 3, we report main findings of literature review. We use a systematic approach in literature review by classifying the studies in specific areas. First, we introduce Zipf's law for the city size distribution in more general terms. Then, we review the theoretical studies which have tried to provide a theoretical explanation to Zipf's law for cities. Next we provide a review of the empirical research. Finally, we examine the specific issues in Zipf's law and city size distribution research. In Section 4, we identify the emerging areas in the city size distribution research. The final section concludes the study.

2. Search of the related literature

We started our review by searching the relevant studies on Zipf's law and city size distribution. We largely restricted our survey to the studies published from 2000 to 2017. This time period was selected for at least two reasons: First, an enormous amount of city size distribution literature has appeared over this period due to the pioneer studies of Gabaix [13] and Eeckhout [14]. Second, the survey conducted by Nitsch [12] can be referred for the literature published before the year 2000.

We used Google Scholar and Scopus as two major databases for our search. The keywords 'Zipf's law', 'city size distribution' and 'city growth' were searched in both databases by limiting the search from 2000 to 2017. To ensure the maximum coverage of published studies, we augmented our search with forward citations and backward references approaches. For the former approach, we searched papers from the citations of major papers published in the early 2000s. For the latter approach, we selected very recent published papers and check their references for previously published studies. In total, we found 327 articles. After a detailed review of these papers, we included 114 most related papers in this study. We tried to include as many studies as we could search and which were related to city size distribution debate. However, we would like to apologize if a significant work of any author has been omitted.

3. Review of the mainstream theoretical and empirical literature

In this section, we briefly summarize the findings of our review. The first subsection introduces the Zipf's law and discusses the empirical methods which have been used to validate the Zipf's law for city size distribution. Second subsection presents the review of theoretical studies. Third subsection sums up the findings of empirical studies. Final subsection discusses how the controversial empirical evidence is reconciled.

3.1 Introduction to Zipf's law

The history of Zipf's law for city size distribution dates back to Felix Auerbach [1] who first observed the Pareto distribution for cities (later Singer [2] also made a similar observation) as.

$$R = \mathrm{A}.P^{-\alpha} \qquad Eq.(1)$$

Or

$$Log\ (R) = \text{Log (A)} - \alpha . Log\ (P) \qquad Eq. (2)$$

Here, P represents the population of a city. R is the rank of a city when cities are ranked from 1 to n by the population size. A and α are constants. α is also referred as Pareto exponent.

This observation later on received greater recognition due to George Kingsley Zipf [3], whom the law also owes its name to. Zipf restated the above relation by positing that the size of objects is inversely proportional to their rank (also referred as 'rank-size rule' or 'power law'). This relation implied that in a system of cities the largest city is roughly twice the size of the second largest city, about three times the size of the third largest city, and so on.

$$P = \frac{\text{K}}{\text{R}^{\text{q}}} \qquad Eq. (3)$$

Or

$$P = \text{K}. R^{-\text{q}} \qquad Eq. (4)$$

Or

$$Log\ (P) = \text{Log (K)} - \text{q}. Log\ (R) \qquad Eq. (5)$$

q is referred as Zipf's exponent. In the special case, when the exponent q equals one (q=1), the city size distribution said to follow Zipf's law. K is a constant.

The relationship between Pareto-form Eq. (2) and Zipf-form Eq. (5) is straightforward, and a simple mathematical manipulation can show that q=1/α. When α→infinity, q→0, and the size of all cities is equal. Zipf's law holds in strict form when α=q=1. This relation implies that Zipf's law holds when city size distribution follows a special form of Pareto distribution with Pareto exponent equals one (α=1).

The empirical literature has used two approaches to confirm Zipf's law for city size distribution. First approach is to estimate any of the Pareto-form Eq. (2) or Zipf-form Eq. (5) to estimate Pareto exponent or Zipf exponent, respectively. If the estimated value of Pareto or Zipf exponent equals one, it confirms that Zipf's law holds precisely. However, the Pareto-form equation has attracted more attention recently. Parameter estimates of Eq. (2) for the 135 largest US urban areas as shown in Figure (1) are $Log\ (R) = 17.745 - 1.019\ Log\ (P)$. Since both equations (i.e., Eq. (2) and Eq. (5)) use the city size data at a specific point of time to examine Zipf's law, this approach is also called a static method.

Second approach is to confirm the Zipf's law by validating the Gibrat's law of proportional growth for city growth process. Gibrat's law implies that the growth rate of a city's population does not depend on the size of the city [15]. That is, we cannot predict a systematic behavior between city growth rates and their size, even though the cities can grow at different rates. More generally, we cannot assert that smaller cities grow faster than larger ones or vice versa. In a seminal study, Gabaix [13] demonstrated that Zipf's law is an outcome of Gibrat's law. He argued that Gibrat's law implies that the growth process of cities have a common mean (equal to the mean city growth rate) and a common variance, that is, both the mean and variance have to be independent from the size of the cities. More in this direction, Córdoba [16] and Córdoba [17] argue that a more generalized Gibrat's law that allows city size to affect the variance of the growth process but not its mean, can result more general rank-size distributions.

Building on these arguments, Modica, Reggiani and Nijkamp [18] focus on bidirectional relationship between Zipf's law and Gibrat's law and show that if the coefficient of Pareto distribution/Zipf's law equals one, the mean and variance of city growth rates are independent of city size. And if the coefficient is different from one, then the mean is still independent while the variance of city growth rates depends on city size.

A number of recent studies have examined the Gibrat's law for city growth process to confirm Zipf's law [18-20]. For example, Modica, Reggiani and Nijkamp [18] and Berry and Okulicz-Kozaryn [19] largely used the following model to explore Gibrat's law.

$$Log\left(P_{i,t}\right) = \beta_0 + \beta_1 Log\left(P_{i,t-1}\right) + \varepsilon_i \qquad Eq. (6)$$

Here, subscripts *i* and *t* represent the city and time, respectively. *P* represents the size (i.e., population) of a city. ε_i is an error term. The coefficient β_1 is the parameter of interest and shows whether size distribution diverges or converges toward its mean. Gibrat's law holds if β_1 is equal to one. When β_1 is greater than one, the size diverges from its mean; that is, the expected growth is larger for large cities. On the contrary, when β_1 is less than one, the size converges toward its mean; that is, the expected growth rate is smaller for large cities. ε_i is an i.i.d random variable with mean μ and variance σ^2. We estimate Eq. (6) for 135 largest US metropolitan areas with the data from latest two census years of 2000 and 2010. We find that β_1= 0.923 (with a robust standard error of 0.0405) which is statistically equals to one; the 95% confidence intervals lie between 0.843 and 1.003. ε_i is also normal. These results confirm Gibrat's law. Since the city size data over a certain time-period is required to calculate city growth rates, this second approach is also referred as dynamic method.

3.2 Review of theoretical studies

Given the widespread acceptance of Zipf's law for cities, several authors have endeavored to provide theoretical underpinnings to the law. This literature can be divivded into three types: First are the studies that use Gibrat's law of proportional growth to explain Zipf's law. Second are the studies which use the human capital accumulation and the central place theories to explain Zipf's law. On the contrary, the third type of studies argues that Zipf's law is a steady state outcome of a self-organizing system and does not need an underlying theory for its justification. Table 1 summarizes these studies.

The underlying principle in the first strand of theoretical models is that in industrialized countries, cities concentrate not only a large part of the population but also the economic activity, and the urban structure is an outcome of dynamic interplay between economic activity and growth process of cities [13, 16, 17, 21-25]. In these models, various economic shocks (i.e., the amenity, productivity or innovations shocks) generate skew in city size distribution. When economic shocks are randomly distributed across cities, there would only be a few large cities due to the lower probability of a city being consistently hit by positive economic shocks which increase city size. When the shocks are independent of city size than city growth will follow Gibrat's law of proportional growth (i.e., the growth rates of cities are independent of their absolute size) and consequently the size distribution would be Pareto with Pareto exponent exactly equals to one. In this context, the seminal study of Gabaix [13] shows that a randomly growing set of cities, which faces random amenity shocks and an impurity that restricts cities to become too small[1], can result in Zipf's law at least in the upper-tail of the city size distribution. Gabaix hypothesizes that growth shocks are independently and identically distributed (i.i.d.) and impact the utilities both positively and negatively. The migration of workers occurs to the cities with higher wages that, in equilibrium, equates the utility-adjusted wages at the margin. By assuming a fixed number of cities and the constant returns to scale for production technology, he shows that all cities exhibit same expected growth rates and the deviations from expected growth rates observe random normal distribution (i.e., city growth follows Gibrat's law). In such a scenario, the coefficient of city size distribution tends to be one and city sizes follow a pure Zipf's law. In this model, only the size distribution of upper-tail cities would follow Zipf's law if new cities emerge in the urban system. Extending this work, Córdoba [17] shows that a generalized form of Gibrat's law where the variance of growth rates can depend on city size but not the mean can produce a distribution with Zipf's exponent different from one. He also shows that under certain conditions, it can produce a pure Zipf's law with Pareto exponent equals to one.

Rossi-Hansberg and Wright [25] develop a model of urban growth that depends on the exogenous productivity shocks specific to each industry. When such a shock occurs, the volume of production and the size of the cities specialized in this industry increases. Urban system, where migration affects the growth, birth and death of cities, would eliminate these local increasing returns to scale at the margins yielding constant returns to scale for the whole urban system in aggregate. In the similar vein, Duranton [22] found that randomly distributed innovation shocks across cities can result in a size distribution of cities which follows Zipf's law. In his model, a city is a collection of industries, and the industries randomly move to the places where research is more successful due to innovations. This model predicts that the size distribution of cities obeys an approximate Pareto distribution. In contrast to above models which assume only one factor at a time (i.e., amenity, productivity or innovation shocks), Lee and Li [24] show that several randomly distributed factors, which might correlate with each other to some extent, can result in a city size distribution which obeys Zipf's law. In a more general framework, Córdoba [16] establishes a standard urban model with localization economies that can produce a size distribution of cities which is Pareto. He shows that to observe Pareto distribution, a system of cities must have a balanced growth path (i.e., all cities have same expected growth rate) and an exogenous driving force whose steady state distribution must be Pareto.

The second strand of studies models the Zipf's law distribution based on alternative approaches, rather than the city growth processes. For example, Behrens, Duranton and Robert-Nicoud [26] model that big cities have higher productivity due to their ability to attract and sort out talented individuals. The individuals move to larger cities for agglomeration economies, however the higher urban costs in big cities let only talented and highly productive individuals to stay. They show if a Pareto distribution is assumed for the talent across cities, then the city size distribution also follows a Pareto distribution similar to the Zipf's law. In another recent study, Hsu [23] models the city size distribution using a hierarchy approach based on central place theory. In this model, the differences in city sizes arise due to the heterogeneity in economies of

[1] Eeckhout [14] shows that if growth rates of cities follow Gibrat's law then it will result a lognormal distribution, which is exponential but not power law even in upper-tail. So to get a power law it is necessary to introduce a truncation point or lower size limit.

scale across goods. If the distribution of scale economies is regularly varying, then the size distribution of cities under a central place hierarchy exhibits a power law.

On the contrary, the third type of studies argues that Zipf's law either breaks down in certain situations or it does not need a theoretical explanation. For instance, Mansury and Gulyás [27] developed a spatial agent-based model to generate a system of cities that exhibits the statistical properties of the Zipf's Law. When agents have the properties of bounded rationality and maximum heterogeneity, their migration to urban areas does not always generate a Zipf's Law city size distribution. In addition, the Zipf's Law breaks down unless the extent of agglomeration economies overwhelms the negative disagglomeration forces. Among the studies which suggest that Zipf's law does not need a theoretical explanation, Axtell and Florida [28] and Semboloni [29] use agent based models where agents interact through probabilistic law for opposing goals and their movements within different places conform to Zipfian process under which the whole urban system converge to power law rule as a steady state. In another study, Gan, Li and Song [30] showed through Monte Carlo simulations that Zipf's law is a statistical occurrence that does not require an economic theory. Similarly, Batty [31] explains that a complex adaptive system can self-organize itself in a distribution which can be described by the Zipf's law. Finally, Corominas-Murtra and Solé [32] argue that the scaling behavior is a common statistical property of stochastic systems which evolve to a steady state between order and disorder.

In this perspective, an important area for future research is to adopt a consolidated approach to settle the issue whether Zipf's law requires a theoretical explanation.

Table 1 Summary of theoretical studies

Findings	Studies
Economic shocks (i.e., the amenity, productivity or innovations shocks) and random city growth process can result a city size distribution which obeys Zipf's law	Gabaix [13] Córdoba [17] Rossi-Hansberg and Wright [25] Duranton [22] Lee and Li [24] Córdoba [16]
Human capital accumulation and the central place theory are used to explain Zipf's law	Behrens, Duranton and Robert-Nicoud [26] Hsu [23]
Zipf's law is a steady state outcome of a self-organizing system and do not need an underlying theory for its justification	Mansury and Gulyás [27] Florida [26] Semboloni [29] Gan, Li and Song [30] Batty [31] Corominas-Murtra and Solé [32]

3.3 Review of empirical studies

This section presents the review of empirical research. As described in Subsection 3.1, Zipf's law can be confirmed either by estimating the log-rank log-size regressions (i.e., Eq. (2) and Eq. (5)) or by validating the Gibrat's law of proportional growth. Recent empirical studies have used any one or both of these approaches to confirm Zipf's law, and we include all these studies in our review.

The studies reviewed are mostly country specific except a few which use multi-country datasets for the empirical analysis. After reviewing the literature, we classified the studies into five groups. First group are the studies which have debated whether Zipf's law applies to a system of cities with mixed answers; some studies support the Zipf's law while others reject it. Second are the studies which accept the hypothesis that Zipf's law precisely holds for cities of territory. Third are the studies which reject the hypothesis of Zipf's law for city size distribution. Fourth type of studies shows that city size distribution approaches to Zipf's law as countries experience urbanization. Finally the last group of studies concludes that city size distribution may evolve away from Zipf's law over time.

We start from the first group of studies which report mixed evidence. Table 2 summarizes the findings of this group of studies. The urban systems of the United States and China are major focus of these studies. For instance, a number of recent studies have supported Zipf's law for the US city size distribution. In this regard, Krugman [33] and Gabaix [13] found that the largest 135 US Metropolitan Statistical Areas (MSAs) follow Zipf's law. Similar evidence is reported by Levy [34] for 150 largest metropolitan areas. Ioannides and Overman [35] analyzed the city growth dynamics over the period 1900-1990 and confirmed the Zipf's law by validating the Gibrat's law of proportional growth. Some studies have shown that the fit to Zipf's law improves if cities are properly defined [19, 36, 37]. For example, Berry and Okulicz-Kozaryn [19] found that Zipf's law holds precisely for the US cities above 500,000 inhabitants if measured as economic areas. Rozenfeld, Rybski, Gabaix and Makse [37] measured cities as area clusters and found that Zipf's law applies to the area clusters with at least 13,000 inhabitants. Jiang and Jia [36] reported that Zipf's law holds remarkably well for all cities (over 2-4 million in total) across the United States when a natural definition of city is used; a city is measured with covered land area.

In contrast, some studies have opposed Zipf's law for the US city size distribution. For example, Black and Henderson [38] analyzed the data of 282 metro cities over the period 1900-1990 and reported a far less than one (0.842) Pareto exponent. They suggested a higher urban concentration in major cities than predicted by the Zipf's law. In another influential contribution, Eeckhout [14] proposed an equilibrium theory to explain the lognormal distribution of cities. Considering the data of all US census places, he showed that the entire city size distribution is lognormal[2], not Pareto. Eeckhout [14] identified the difficulty in distinguishing between the power law in upper-tail and the tail of lognormal distribution. The power law in upper-tail may just be a tail of lognormal distribution and not necessarily a separate Pareto distribution. For a Pareto distribution, the Pareto exponent should not be sensitive to the choice of truncation point. Becaue by choosing a specific truncation point and hence selecting a certain set of cities from the upper-tail, Pareto distribution can be favored over the lognormal. In another recent study, Bee, Riccaboni and Schiavo [40] did a thorough analysis by applying some counterfactual exercises and found that the power-law behavior of the upper-tail is less robust than previously claimed. They argued that the controversy arises due to the limited power of available statistical tests. Rafael González-Val and coauthors find that Gibrat's law only weakly applies in the US and Zipf's law breaks down for the entire city size distribution [41-43]. Some studies have supported this argument when they analyze Zipf's law by validating the Gibrat's law. For instance,

[2] Eeckhout [14]'s observation was consistent with Parr and Suzuki [39] who long ago suggested the lognormal distribution as a good description of the city size distribution.

González-Val, Lanaspa and Sanz-Gracia [43] find that Gibrat's law is valid only for upper-tail cities and breaks down in the long-term for all cities. They conclude the lognormal distribution offers a better fit to the untruncated data over the whole twentieth century.

To settle the controversy, some studies find that upper-tail of the US city size distribution is Pareto while body is lognormal. For example, Levy [34] analyzed the same dataset as used by Eeckhout [14] and found that 150 largest metropolitan areas (0.6% of sample) in Eeckhout's dataset adhere to Zipf's law. Malevergne, Pisarenko and Sornette [20] used uniformly most power unbiased (UMPU) test to distinguish between Pareto and lognormal distribution and found that the largest 1000 places in Eeckhout's data follow a Pareto distribution while lower range cities follow a lognormal distribution. Ioannides and Skouras [44] suggest an 'upper-tail Pareto lognormal distribution' where distribution is robustly Pareto in the upper-tail (top 5%) while it is lognormal in the body. Fazio and Modica [45] use recursive approach and identify similar difficulty in distinguishing between a Pareto upper-tail and the tail of a lognormal distribution. They conclude that 1,000 largest cities follow Zipf's law while lower range is lognormal.

Similarly, whether the city size distribution in China follows Zipf's law is a pending question with different answers. Some studies say "yes" [46-50], while others say "no" [51-57]. For example, Gangopadhyay and Basu [47] used the data of the census years of 1990 and 2000 and find that the largest Chinese cities follow a Pareto distribution. Similarly, Ziqin [50] considered the data of 647 largest cities in the census year 2000 and 655 cities in 2010 and found that city size distribution is exactly Pareto. He reports a Pareto exponent equals to 1.0426. On the contrary, Song and Zhang [57] used the city-level data for the census years 1991 and 1998 and found that Zipf's law does not apply. Anderson and Ge [51] and Li, Wei and Ning [54] favored the lognormal over the Pareto distribution for Chinese cities. In a methodological improvement, Peng [58] employed rolling sample regressions in which sample changes with truncation point. He observed a monotonically decreasing Pareto exponent when lower truncation points are chosen. He reports a 0.84 mean value of the Pareto exponent for his sample. Similarly, Luckstead and Devadoss [56] investigated the data of largest 142 Chinese cities over seven decades from 1950 to 2010. They found that the city size distribution was lognormal from 1950 to 1990, while it is tilting towards Zipf's law (though has not approached exact Zipf's law) in the years 2000 and 2010. More in this direction, Li and Sui [55] observed that the Pareto exponent presents a turning point in 1996 illustrating China's transition from a planned economy to a more market-oriented economy during that period. The specific characteristics of Chinese urban system are devoted to the specific institutional arrangement in China such as the socialist institutions, urban and regional development policies, changes in the urban administrative system, state and local government interests, one child policy and Hukou system [4, 59].

Some studies examine the growth process of Chinese cities and largely fail to validate Gibrat's law. These studies report that Chinese urban system is characterized by parallel growth where small and medium size cities have grown faster than the large cities [48]. Similarly, the particular groups of cities with common location-specific characteristics, such as the similar policy regimes and a natural resource endowment, grow parallel in the long run [60, 61]. However some recent studies find divergence in cities growth process that may result in Zipf's law to hold in future. For example, Fang, Li and Song [62] find that the growth of Chinese cities was size convergent before 2000, but size-independent after 2000. Consequently, the major cities have grown at a rapid pace as compared to small and medium-sized cities after the year 2000.

Some multi-country studies have also reported similar mixed evidence. For example, Rosen and Resnick [63] examined the Zipf's law for 44 countries and found that 32 out of 44 countries had a Pareto exponent equal to or greater than one. Soo [9] used a sample of 73 countries and more recent data from the 1990s. His findings suggest that Zipf's law is more often rejected than is expected based on a random chance; Zipf's law is rejected for 30 out of 73 countries with the Hill estimator and for 53 out of 73 countries with the OLS estimator. Nitsch [12] performed a meta-analysis of 29 studies and showed that the combined estimate of Pareto exponent is significantly greater than one which suggests a more even city size distribution.

Table 2: Summary of empirical Studies (urban systems and studies with mixed evidence)

Findings	Studies	Country	Sample
City size distribution obeys Zipf's law	Krugman [33]	US	130 largest metropolitan areas
	Gabaix [13]		135 largest metropolitan areas from the census year 1990
	Ioannides and Overman [35]		112 cities in the census year 1900 to 334 in the census year 1990
	Berry and Okulicz-Kozaryn [19]		>150,000 inhabitants
	Rozenfeld, Rybski, Gabaix and Makse [37]		1,947 US cities above 12,000 inhabitants
	Jiang and Jia [36]		2-4 million urban agglomerations
City size distribution does not obey Zipf's law	Black and Henderson [38]		282 largest metro cities
	Eeckhout [14]		25,359 places (all places in the year 2000 US census with population from 1 to over 8 million)
	Eeckhout [64]		Same as Eeckhout [14]
	Bee, Riccaboni and Schiavo [40]		28,916 cities; or 17,569 clusters
Upper-tail of city size distribution conforms to Zipf's law, while body and lower-tail are lognormal	Levy [34]		Sample is same as Eeckhout [14]. 150 largest census places follow Zipf's law
	Malevergne, Pisarenko and Sornette [20]		1,000 largest cities with population above 37,000 inhabitants follow Zipf's law
	Ioannides and Skouras [44]		Census Defined Places above 60,290 inhabitants, Metro and Micro areas above 34,853 and Area clusters above 30,635 follow Zipf's law.
	Fazio and Modica [45]		1,000 largest cities follow Zipf's law, lower range is lognormal.
City size distribution obeys Zipf's law	Gangopadhyay and Basu [47]	China	Different sample compositions; minimum threshold is the cities above 50,000 inhabitants in the census year 2000
	Ziqin [50]		655 largest cities in the census year 2010
City size distribution does not obey Zipf's law	Song and Zhang [57]		665 largest cities in the census year 1998
	Anderson and Ge [51]		Cities with more than 100,000 inhabitants
	Luckstead and Devadoss [56]		142 largest cities from 1950 to 2010
	Li, Wei and Ning [54]		657 large in the census year 2010
Average Pareto exponent is greater than	Rosen and Resnick [63]	44 countries	Truncated samples

1			
Zipf's law is more likely to be rejected	Soo [9]	73 countries	Truncated samples

The second group of studies finds that Zipf's law precisely holds (Table 3 summarizes the findings of second to fifth group of studies). In this context, the city size distributions of Germany, Romania, Morocco and Russia are major examples. Giesen and Südekum [65] found that Zipf's law exactly holds for German cities with population above 100,000 inhabitants at both national- and regional-levels. They also find evidence in favour of Gibrat's law. Gligor and Gligor [66] supported the Zipf's law for 265 large and medium urban settlements of Romania for the census year 2002. Ezzahid and ElHamdani [67] explored the Zipf's law for Moroccan cities for the census years 1982, 1994, and 2004. Zipf's law holds for Moroccan cities with more than 50,000 inhabitants[3]. Rastvortseva and Manaeva [69] found that the size distribution of largest Russian cities was Zipf in the census year 2014.

Opposite to the second group, the third group of studies argues that Zipf's law does not hold. City size distributions of Canada and Spain are clear examples. Lalanne [70] examined the size distribution of 152 largest Canadian cities for the census years 1971, 1981, 1991 and 2001. He rejected the Zipf's law for all periods. Urban structure is dominated by few large cities such as Toronto, Montreal and Vancouver. Similarly, the growth process is deterministic, rather than random, where city growth rates depend on city size, previous growth and spatial structure. In another study, Dubé and Polèse [71] analyzed the size distribution of 135 largest urban areas over the period 1971-2011 and found that the initial city size positively affects subsequent growth. Likewise, some studies have estimated much lower Pareto exponent for Spanish urban system [6, 72]. For example, Le Gallo and Chasco [6] analyzed the data of 722 Spanish municipalities over the period 1901-2001 and found that the estimated values of Pareto exponent varies from 0.54 to 0.66 in different years.

The fourth set of studies reports that city size distribution approaches to Zipf's law as countries experience urbanization. In this context, the urban systems of India and Brazil are major examples. For India, Gangopadhyay and Basu [47] supported Zipf's law for upper-tail Indian cities in the census years 1981 and 2001. Similar observation was made by Gangopadhyay and Basu [48] when they considered the data of four census years 1981, 1991, 2001 and 2011. In addition, Gangopadhyay and Basu [48] found that the growth process of major Indian cities follows Gibrat's law. Luckstead and Devadoss [56] examined the size distribution of Indian cities from 1950 to 2010. They found that the distribution was lognormal from 1950 to 1980, while the Pareto in census years 1990 and 2010. Later was the period when India experienced rapid economic growth through industrialization, which caused widespread migration of workers from rural to urban areas. For Brazil, Moura and Ribeiro [73] examined the size distribution of Brazilian cities (with at least 30,000 inhabitants) for the four census years of 1970, 1980, 1991 and 2000. They concluded that upper-tail Brazilian cities follow a power law. Matlaba, Holmes, McCann and Poot [74] looked at the evolution of 185 functionally defined urban areas of Brazil over the period from 1907 to 2008. They found that even though both Zipf's and Gibrat's laws did not formally hold in Brazil's past, however gradually the stochastic urban growth process that is consistent with Gibrat's law has led to a more Zipfian city size

[3] In another study, Schaffar and Nassori [68] focused just on urban growth process of Moroccan urban system from 1994 to 2010 and found that urban system is converging towards more even distribution.

distribution in recent years. In another recent study, Ignazzi [5] reported similar results for the large Brazilian cities for the data from 1871 to 2010.

Finally, the last group argues that the distribution of a system of cities may evolve away from Zipf's law over time. In this context, the size distributions of Malaysian, Mexican and Turkish cities are key examples. For instance, Soo [75] examined the Zipf's law for Malaysian cities with at least 10,000 inhabitants. Using the data of five census years 1957, 1970, 1980, 1991 and 2000, he showed that Zipf's law held only in 1957 but not in other years. Since the year 1957, the city size distribution in Malaysia has evolved away from Zipf's law. He also rejected Gibrat's law by finding that some cities (i.e., smaller cities, state capitals and the cities in the states of Sabah and Selangor) have grown faster than the rest of the urban system. Pérez-Campuzano, Guzmán-Vargas and Angulo-Brown [7] studied the size distribution of Mexican cities from 1900 to 2000 using the census data collected every ten years. The estimated values of Pareto exponent range between $\alpha \approx 0.7$ and $\alpha \approx 1.1$. The deviations from Zipf's law were more remarkable at the beginning and at the end of the 20th century. For Turkey, Duran and Ozkan [76] found that Zipf's law held in 1965 with a Pareto exponent of 1.004 and the exponent increased to 1.014 till 1980 and started decreasing after that with values of 0.96 in 1990, 0.894 in 2000 and 0.824 in 2010. It was significantly different from one from 2007 onward[4].

Table 3 Summary of empirical studies

Main findings`	Study	Country	Sample
City size distribution exactly conforms to Zipf's law.	Giesen and Südekum [65]	Germany	Cities with more than 100,000 inhabitants
	Gligor and Gligor [66]	Romania	265 large and medium urban settlements
	Ezzahid and ElHamdani [67]	Morocco	Cities with more than 50,000 inhabitants
	Rastvortseva and Manaeva [69]	Russia	
City size distribution does not conform to Zipf's law	Lalanne [70]	Canada	152 largest urban areas
	Dubé and Polèse [71]		135 largest urban areas
	Lanaspa, Pueyo and Sanz [72]	Spain	
	Le Gallo and Chasco [6]		722 Spanish municipalities
City size distribution approaches to Zipf's law as countries experience urbanization	Gangopadhyay and Basu [47]	India	Different samples; minimum threshold is the cities above 10,000 inhabitants in the census year 2001
	Gangopadhyay and Basu [48]		Cities above 212,523 inhabitants in the census year 2011
	Luckstead and Devadoss [56]		58 largest cities from 1950 to 2010
	Moura and Ribeiro [73]	Brazil	Cities with 30,000 inhabitants or more

[4] Deliktas, Önder and Karadag [77] use data from 1980 to 2007 and find strong support in favor of Zipf's law for Turkish cities when the rank-minus-half rule of Gabaix and Ibragimov [78] is used to estimate rank-size regression.

	Matlaba, Holmes, McCann and Poot [74]		185 largest functionally defined urban areas
	Ignazzi [5]		Census years data from 1871 to 2010
City size distribution may evolve away from Zipf's law over time	Soo [75]	Malaysia	Cities with more than 10,000 inhabitants
	Pérez-Campuzano, Guzmán-Vargas and Angulo-Brown [7]	Mexico	Cities with more than 15,000 inhabitants
	Duran and Ozkan [76]	Turkey	Cities with more than 37,522 inhabitants in the year 2012

Above discussion suggests that Zipf's law is not a universal phenomenon even for the upper-tail cities. The validity largely varies from country to county and depends on country characteristics such as the institutional setting, cultures, economic and urban policies and history. One important implication from above findings is that we specifically need to establish the validity of Zipf's law for the city size distribution of an urban system at a specific point of time.

3.4 Reconciling the empirical evidence

Empirical evidence provided by different studies on Zipf's law has been criticized on several grounds. The studies which support Zipf's law are criticized based on sample selection, while the studies which reject Zipf's law are criticized due to weak empirical methods and improper definition of cities.

As shown in Tables 2 and 3, almost all studies which support Zipf's law work with truncated data from the upper-tail of the city size distribution. The choice of truncation point remains with researcher and introduces arbitrariness in sample selection. Due to this shortcoming, the opponents of the Zipf's law argue that the choice of a truncation point biases the results because a sample of cities from the upper-tail can be selected to accept the hypothesis of Zipf's law over other alternative distributions.

One the contrary, the proponents of Zipf's law argue that too often rejection of Zipf's law for city size distribution is a result of improper definition of cities and the use of weak empirical methods. These studies suggest that cities should be measured as economic/functional urban areas rather than the administrative units. Functional definition considers close suburbs while the administrative usually ignore them. The fit to Zipf's law improves if cities are properly defined. Similarly, these authors suggest that the empirical results reported by a number of studies are biased due to inappropriate econometrics techniques. Zipf's law is more likely to hold if an accurate method is used. Below subsections review these opinions.

3.4.1 Sample selection and alternative distributions

Sample selection issues

One main challenge for the authors to confirm Zipf's law for cities is to decide the data truncation point for sample selection. A truncation point can be selected to accept the Zipf's law over another alternative distribution. Though different studies have suggested different methods for data truncation, the arbitrariness remains a concern in all methods.

Rosen and Resnick [63] proposed two alternative methods for choosing the data truncation point: 'the number threshold method' whereby a fixed number of cities are selected and 'the size threshold method' whereby the cities above a fixed size threshold are kept in sample for each time period or for each region/country. Further, Wheaton and Shishido [79] suggested the third 'the urban population threshold method' whereby the cities with a given proportion of a country's total urban population are included into the sample[5]. Recently, Li and Sui [55] proposed the fourth 'the number percentage threshold method' to select sample. This method chooses a fixed percentage of the number of cities.

All these sample selection methods face criticism. For example, the number threshold method will include only a small proportion of the cities in large countries which have higher number of cities, while a large proportion of cities in small countries which have small number of cities. Similarly, the size threshold method will include higher number of cities in a large and populated country, while only few cities from a country with small size and less population. Further, the urban population threshold method is influenced by the degree of metropolitanization of an urban system. Finally, the number percentage threshold method will exclude more cities in absolute terms if an urban system has a large number of cities as compared to the cities excluded from an urban system with small number of cities.

Thus in all methods, the sample selection remains arbitrary. One plausible answer is to use untruncated data. However, when untruncated city data is considered, the Zipf's law largely breaks down. Several studies have shown the sensitivity of Pareto exponent to sample size [14, 45, 58, 75, 80]. For example, Krugman [33] and Gabaix [13] found that Zipf's law applies to 135 largest US Metropolitan Statistical Areas. In contrast, Eeckhout [14] showed that the US city size distribution is lognormal if all census defined places (25,359 places in the census year 2000) are considered. Fazio and Modica [45] specifically examined the data truncation issue for the US cities using a recursive regressions approach. They concluded that Zipf's law may be favored in upper-tail by truncating the data, however the lognormal distribution seems to better fit the entire sample. Similarly, Gangopadhyay and Basu [47] found that Zipf's law can only be favored for Chinese cities if a higher truncation point is chosen. In another study, Peng [58] employed the rolling sample regressions for Chinese cities and discovered that Pareto exponent monotonically decreases if lower truncation points are chosen.

Alternative distributions

Considering that Pareto distribution may break down for untruncated samples, a number of new distributions have been suggested for the entire city size distribution[6]. Table 4 summarizes these distributions. Of these the q-exponential was suggested by Malacarne, Mendes and Lenzi [88]. Eeckhout [14] showed that the lognormal distribution is appropriate. Some studies have combined Pareto and other distributions to formulate composite distributions. Different statistical methods have been used to construct these distributions. In this vein, the

[5] This method is referred as urban population threshold method by the Li and Sui [55].

[6] Other similar distributions are also available in literature. For example, Eliazar and Cohen [81], Eliazar and Cohen [82] and Eliazar and Cohen [83] drive power law distributions to describe incomes of individuals in a country (including other natural phenomena) by using Lorenz asymptotic analyses. In another strand, Eliazar and Cohen [84] and Eliazar and Cohen [85] and Cohen and Eliazar [86] construct composite distributions, such as a distribution with power laws in both tails and log-Gaussian body, using geometric Langevin dynamics. Recently, Eliazar [87] employ an entropy based approach to establish power-law distributions.

prominent distributions are the upper-tail Pareto lognormal distribution [44], the Pareto-tails lognormal distribution [89], the double Pareto lognormal distribution [90-92], the threshold double Pareto Singh–Maddala distribution [93], the Pareto-positive stable distribution [94] and the Pareto ArcTan distribution [95].

Malacarne, Mendes and Lenzi [88] introduced the q-exponential distribution to cities and found that it offers a better fit to the entire city size distribution of the US and Brazil. Subbarayan and Kumar [96] used the untruncated data of all Indian cities and towns over the period 1951-2011 and examined that the which one Pareto, three parameter lognormal or q-exponential distribution is a good fit to empirical data. They find that q-exponential outperforms the other two distributions. On the other hand, Soo [75] found that the q-exponential distribution does not offer a good fit to the data of all administratively defined Malaysian cities.

Ioannides and Skouras [44] found that the US city size distribution is robustly Pareto in the upper-tail (top 5%) while it is lognormal in the body. They introduced it as a new 'upper-tail Pareto lognormal distribution' function which switches between a Pareto upper-tail and a lognormal body and lower-tail[7]. Extending the work of Ioannides and Skouras [44], Luckstead and Devadoss [89] suggested the Pareto-tails lognormal distribution which consists of upper-tail Pareto, middle range lognormal, and lower-tail Pareto. Pareto-tails lognormal distribution defines two switching points, one between a Pareto upper-tail and a lognormal body and the other between a lognormal body and a reverse Pareto lower-tail, and thus clearly delineates between the three behaviours. Luckstead and Devadoss [89] found that the Pareto-tails lognormal distribution outperforms the upper-tail Pareto lognormal distribution for the entire US city size distribution. In another recent study, Luckstead, Devadoss and Danforth [98] found similar evidence for all Indian cities.

Reed [91] and Reed and Jorgensen [99] suggested that the size distribution of lower-tail cities is an inverse power law, in addition to the power law distribution in upper-tail. They modeled that a lognormally distributed initial state, which follows a geometric Brownian motion over an exponentially distributed length of time, results in a new double Pareto lognormal distribution (DPLN). This distribution has power laws in both upper and lower-tails and a lognormal body, but without clearly delineating between the three behaviours. This distribution has shown a good fit to empirical data. For instance, Giesen, Zimmermann and Suedekum [100] found that DPLN distributions outperforms the lognormal for the entire city size distributions of eight countries (Brazil, the Czech Republic, France, Germany, Hungary, Italy, Switzerland and the US) Similarly, González-Val, Ramos, Sanz-Gracia and Vera-Cabello [101] used untruncated city data of Italy, Spain and the US from 1900 until 2010, and the data of last available year for remaining countries of the OECD to examine city size distribution. They compared DPLN, lognormal, log-logistic and *q*-exponential and found that the distribution which best fits the data in most of the cases (86.76%) is the DPLN. In another study, Vitanov and Ausloos [102] observed that DPLN outperforms Zipf's law for Bulgarian cities.

Ramos, Sanz-Gracia and González-Val [93] introduced the threshold double Pareto Singh–Maddala (TDPSM) distribution to city size data. This distribution has Pareto behaviour in the upper and lower-tails, and Singh–Maddala body (see Singh and Maddala [103] for the details of Singh-Maddala function). They found that TDPSM distribution outperforms lognormal and

[7] Calderín-Ojeda [97] find that the size distribution of all French settlements (i.e,. communes) from 1962 to 2012 is best explained by lognormal upper-tail Pareto distribution.

DPLN distributions for the entire US city size data of three types (all places in 2000–2010, incorporated places in 1900–2000 and CCA clusters in 1991–2000). Puente-Ajovín and Ramos [8] compared lognormal, DPLN, TDPSM, and normal-Box–Cox distributions for four European countries (i.e., France, Germany, Italy and Spain). They found that TDPSM outperforms the other three while the DPLN is second best.

Sarabia and Prieto [94] introduced a new Pareto-positive stable (PPS) distribution function to explain the entire city size distribution. In this function, zero and unimodality are possible, and the classical Pareto and Zipf distributions are included as a particular case. They compare this new distribution with the Pareto, lognormal and Tsallis distributions using the city data of Spain over several census years, and observe that new distribution offers a better fit to the data than the other three distributions. This distribution has been examined at sub-national levels also. For instance, Kumar and Subbarayan [104] examined the city size distribution of the Andhra Pradesh state of India and found that Pareto-Positive Stable (PPS) distribution offers a better fit to actual city data than the Pareto and lognormal distributions. In another study, similar evidence has been reported by Vallabados and Arumugam [105] for another Indian state of Kerala.

Gómez-Déniz and Calderín-Ojeda [95] drove a more generalized Pareto distribution, the Pareto ArcTan (PAT) distribution, using the circular inverse of the tangent function. This distribution includes Pareto and Zipf's distributions as limited cases. Using the city data of Australia and New Zealand, they showed that PAT distribution improves the performance of classical Pareto, lognormal and Pareto positive stable distributions.

These all are very interesting developments in the city size distribution literature. Though, several functions have been invented with some empirical validation, however the widespread empirical acceptance of any of these functions is still need to be established and is an interesting area for future research. In addition, future research may endeavor to provide theoretical underpinnings to these alternative distributions.

Table 4 Alternative distributions

Distribution	**Studies**
q-exponential distribution	Malacarne, Mendes and Lenzi [88] Subbarayan and Kumar [96] Soo [75]
Lognormal distribution	Eeckhout [14]
Double Pareto-lognormal (DPLN) distribution	Reed [91] Reed and Jorgensen [99] Giesen, Zimmermann and Suedekum [100] González-Val, Ramos, Sanz-Gracia and Vera-Cabello [101] Vitanov and Ausloos [102]
Threshold double Pareto Singh–Maddala (TDPSM) distribution	Ramos, Sanz-Gracia and González-Val [93] Puente-Ajovín and Ramos [8]
Pareto-positive stable (PPS) distribution	Sarabia and Prieto [94] Kumar and Subbarayan [104] Vallabados and Arumugam [105]
Pareto ArcTan (PAT) distribution	Gómez-Déniz and Calderín-Ojeda [95]
Upper-tail Pareto lognormal distribution	Ioannides and Skouras [44]
Pareto-tails lognormal distribution	Luckstead and Devadoss [89] Luckstead, Devadoss and Danforth [98]

3.4.2 Methodological improvements

As described in Subsection 3.1, the empirical literature has widely used the regression approach in the form of Eq. (2) (restated below) to examine Zipf's law for cities.

$$Log\ (R) = \mathrm{Log}\ (\mathrm{A}) - \ \alpha.Log\ (P) \qquad Eq.(2)$$

Eq. (2) is estimated using ordinary least squares (OLS) estimator. To accept Zipf's law, the estimated value of α from Eq. (2) should be statistically equal to one. A usual t-test is used to test the null hypothesis that the estimated value of α is not different from one. However, a number of recent studies have identified the weaknesses of this approach and suggest some improvements.

In this respect, Gabaix and Ioannides [106] argued that the estimated value of α from Eq. (2) is likely to be downward biased because the size of the largest city appears 'too big'. This effect would particularly be strong when the sample size is relatively small. In the same vein, the standard errors for the Pareto exponent by OLS estimator would also be underestimated and, as a result, Zipf's law can be rejected too often based on the *t*-test. Nishiyama, Osada and Sato [107] introduced a new test and demonstrated by using the same data as used by the Soo [9] that Zipf law is rejected for only 1 of 24 countries under their new test whereas it is rejected for 23 of 24 countries under the usual *t*-test used by the Soo [9].

To control these problems, Gabaix and Ibragimov [78] proposed an alternative version of Eq. (2) by introducing the log of rank-1/2 as dependent variable. Specifically, the following form is suggested.

$$Log\ \left(R - \frac{1}{2}\right) = \mathrm{Log}\ (\mathrm{A}) - \ \alpha.Log\ (P) \qquad Eq.(7)$$

Gabaix and Ibragimov [78] argue that this methodological development eliminates the biases linked with the estimation of Eq. (2). By doing so, the regression specified in Eq. (7) might improve the chances of accepting the hypothesis in favour of Zipf's law. However, from the empirical evidence reviewed in Subsection 3.3, it seems that this methodological improvement cannot account for all the empirical deviation from Zipf's law. Several empirical studies have tried to validate the Zipf's law by confirming the Gibrat's law for city growth process, and still have rejected the Zipf's law[8]. To reject the Zipf's law with the alternative method to some extent suggests that the methodological improvement would not account for all the deviations from Zipf's law in empirical research. However, we do acknowledge that more empirical research is needed by using the new regression model to rule-out this allegation. Another suggestion for future research is to confirm Zipf's law by using the log-rank log-size regression and Gibrat's law approaches together for more robust empirical evidence.

In other methodological improvements, Eq. (7) can be estimated using rolling samples regression [58] and recursive regression [45] approaches. Use of rolling and recursive regression approaches can shed light on how the estimated values of Pareto exponent change if data truncation point is changed. In this regard, recursive regression approach is specifically useful since it has the ability to estimate Pareto exponent by adding lower size cities one-by-one. If the

---

[8] As described in Subsection 3.1, Gibrat's law can be confirmed by using the nonparametric (see, for example, Ioannides and Overman [35]) and parametric approaches (see, for example, Black and Henderson [38]).

estimated Pareto exponent remains invariant after adding lower size cities into the sample then Zipf's law holds. Conversely, if the estimated Pareto exponent changes with the increase in sample size then the city size distribution follows some alternative distribution such as the lognormal.

Several studies have criticized the Lilliefors approach used by Eeckhout [14] to favour lognormal over Pareto distribution. Ioannides and Skouras [44] argue Lilliefors test is not appropriate to test a lognormal specification. This test has very little power to detect deviations from a hypothesized distribution when these deviations occur in the tail. Using a switching approach, Ioannides and Skouras [44] showed that upper tail fits a Pareto distribution. Malevergne, Pisarenko and Sornette [20] used uniformly most power unbiased (UMPU) test to distinguish between Pareto and lognormal distributions and supported the Pareto distribution for 1000 largest places in Eeckhout's data.

3.4.3 Definition of a city and Zipf's law

Another major support in favor of Zipf's law comes from the studies who argue that a proper definition of the city should be used. This literature shows that empirical validation of Zipf's law is sensitive to the definition of cities [12, 14, 19, 36, 37, 63, 108-112]. Though different authors have used different words to define cities, overall the literature has used three types of city definitions: administratively defined cities [14], functionally defined cities [19, 37, 111, 112] and natural cities [36, 109, 110]. This literature has shown that Zipf's law offers a better fit to city size distribution when cities are defined as functionally defined urban areas rather than the administratively defined cities, and a pure form of Zipf's law is observed if the cities are measured as natural cites.

Most of the studies on Zipf's law have used administrative definition to measure cities (incorporated places, municipalities, communes, etc.). There are at least two drawbacks of administrative definition of cities. First, the administrative definition of cities is largely used for political purposes and varies widely across countries. Second, this definition usually excludes close suburbs and do not consider economic integration of the population. To overcome these shortcomings, an alternative approach which defines cities as functional urban areas has been adopted. Recent studies have shown how these alternative definitions lead to opposing results.

For example, Eeckhout [14] used the US data for all census defined places (including cities, towns, and villages) and found that lognormal distribution offers a better fit to the data. While, Rozenfeld, Rybski, Gabaix and Makse [37] applied a bottom-up approach of constructing area clusters from high resolution data on population density in the US. These area clusters are independent from administrative boundaries. They conclude that the size distribution of area clusters with at least 13,000 inhabitants closely follows Zipf's law[9]. In the similar vein, Jiang and Jia [36] introduced the concept of 'natural city' where they define cities based on connectivity of roads and populations rather than administratively defined boundaries. The urban agglomerations are demarcated by clustering street nodes, including intersections and ends. They find that Zipf's law holds quite well for all natural cities of the US.

Some multi-country studies have examined Zipf's law for functionally defined or natural cities. For former definition, the studies have used the functional urban areas data that stems

[9] For upper-tail of the US cities, Berry and Okulicz-Kozaryn [19] find that the fit to Zipf's law improves when the upper-tail cities are defined as economic areas using commutation data.

from a collaboration of the European Commission (EC) and the Organization for Economic Co-operation and Development (OECD) [111]. EC-OECD defined functional urban areas are identified by partitioning the earth surface into 1km^2 grid cells. Using satellite images, high and low density cells are identified where a cell is classified as high density if it has a population density of greater than 1,500 inhabitants. Populated clusters are made by joining high density cells which share at least one border. Low density cells are also added in a populated cluster if these low density cells are encircled by high density cells included in the cluster. A populated cluster is designated as urban centre if it has greater than 50,000 inhabitants. As a final step, commuting flows data is used to add local administrative units (municipalities) into an urban centre. If more than 15% workers from a local administrative unit work in urban centre, than this unit is also added to urban centre. Addition of contiguous local administrative units makes up a larger urban zone, while addition of non-contiguous local administrative units, with 15% workers working in urban centre, makes up a polycentric larger urban zone. Based on this approach, Schmidheiny and Suedekum [111] use data of 692 functional urban areas of 31 European countries and find that the size distribution of these areas does not follow Zipf's law when these areas are consider as one group. They argue that the largest European cities are too small to follow Zipf's law. Contrary, Veneri [112] examine the Zipf's law for 29 OECD countries and find that Zipf's law holds for the cities within each individual country except the three sample countries (i.e., Mexico, Poland and Spain). Zipf's law also holds when cities are aggregated at continent level or at the whole OECD level. He emphasizes that Zipf's law is universal (page-89).

For natural cities, Jiang, Yin and Liu [110] identify 30,000 natural cities around the World using night time imagery data of 1992, 2001 and 2010 and find that Zipf's law exactly holds for all cities. Zipf's law remains valid at the continental level except the Africa. One major concern with this study is that it rejects Zipf's law for several countries when it is estimated at country-level despite the fact that the system of cities is more coherent [113] at country-level due to free within country movements.

From this discussion, it seems that one major cause of conflicting empirical evidence on Zipf's law is the extreme heterogeneity in the samples of cities that have been used to perform the tests. The doubt lies in the quality of the city definitions and the delineations used to measure city size. More empirical work is needed to settle the issue of city definition and is most promising area for future research. In this context, future research may consider different city definitions within a country/region for comparison. Similarly, multi-country studies can use city data with consistently defined cities across countries for comparative purposes. Another concern with above studies is that a very high truncation point (50,000 inhabitants) has been used to define a functional area. Again the question arises whether Zipf's law applies if lower size functional areas are included into the sample.

4. Emerging areas in city size distribution research

Several new areas have emerged in Zipf's law and city size distribution research. We collected and classified the published studies and found four major areas in this regard. First are the studies that examine Zipf's law for the systems of cities at sub-national or supranational levels. Second group examines the size distribution of lower-tail cities. Third are the studies which suggest alternative forms of Zipf's law. Finally, the fourth group links the Zipf's law with the coherence property of the urban system. Following sub-sections review the studies in each of these four areas.

4.1 Zipf's law at sub-national or supranational levels

Virtually, the debate on Zipf's law has revolved around the size distribution of cities at country-level. However, a number of recent contributions have examined the Zipf's law at sub-national (i.e., province, state etc.) or supranational (i.e., OECD, EU etc.) levels. In addition, some studies have examined Zipf's law for the total urban populations of provinces or states within a single country or for the total populations of the World countries.

At sub-national levels, a few recent studies have examined Zipf's law for city size distribution within a province, state or any other sub-regional classification. For instance, in an important study Giesen and Südekum [65] examined the size distribution of all German cities as well as the size distribution of cities within German regions. Building on the theory developed by Gabaix [13] that Zipf's law follows from a stochastic urban growth process, they show that if urban growth follows random growth process then the city size distribution both at national- and regional-levels tend to follow the Zipf's law. They also show that any random sample drawn from the sample of all cities also obeys Zipf's law if urban growth is random. Similarly, Subbarayan [114] used the city data of Tamilnadu state of India from 1901 to 2001 and found that overall Zipf's law holds with a value of Zipf's exponent of 1.1424 for cities inhabited by a population of 10,000 and above, with a value of 1.0623 for cities of population of 5000 and above, and with a value of 0.9725 for all cities. For another Indian state Andhra Pradesh, Kumar and Subbarayan [115] found that city size distribution follows Zipf's law with a coefficient of 1.002 in the year 2001. Contrary, Ye and Xie [49] examined the city size distributions for Chinese urban system as a whole and for eight sub-regions (east, central-south, north, northeast, northwest, and southwest). In contrast to the studies on German and Indian sub-regions, they found that although Zipf's law clearly captures the main trends of the urban system in China as a whole, however, there are dramatic variations when Zipf's law is applied to sub-regions. They attribute this sub-regional variance in city-size distribution to differences in socioeconomic environment and urban developmental strategies in sub-regions. Similarly, Ziqin [50] examined the Zipf's law for each of the 26 provinces of China using data from 1990 to 2010. Similar to Ye and Xie [49], he documents complicated city size distribution at province level; the city size distribution in some provinces is more even ($\alpha > 1$), in some provinces exactly obeys Zipf' law ($\alpha = 1$), in some provinces is more uneven ($\alpha < 1$), and in some provinces has primate characteristics.

At supranational levels, Luckstead and Devadoss [116] rejected the Zipf's law for the size distribution of world's largest 600 cities. They attribute this finding to a lower cross-country mobility of population. On the other hand, Jiang, Yin and Liu [110] examined the Zipf's law in a global setting considering about 30,000 natural cities [10] around the world. In contrast to Luckstead and Devadoss [116], they found that Zipf's law holds remarkably well for all cities at global level. Similarly, two other studies have used the data of functional urban areas (FUA) from multi-countries and report different results. For instance, Veneri [112] examined Zipf's law for 29 OECD countries considering different sample compositions. Specifically, Zipf's law holds for the cities within each individual country except the three sample countries including Mexico, Poland and Spain. Zipf's law also holds when cities are aggregated at continent level or at the whole OECD level. He emphasizes that Zipf's law is universal (page-89). On the contrary, Schmidheiny and Suedekum [111] analyzed the 692 functional urban areas of 31 European

[10] Jiang, Yin and Liu [110] identify natural cities using satellite night imagery data.

countries and observe that the size distribution of these urban areas as a group do not follow Zipf's law. They conclude that largest European cities are too small to follow Zipf's law.

Some studies have measured population size at a level other than the city to examine Zipf's law. For instance, Rose [117] examined the Zipf and Gibrat's laws for countries and found that similar to cities, the population distribution across countries also follows these two laws. In another study, González-Val and Sanso-Navarro [118] reassessed the findings of Rose [117] through Gibrat's law. They also found evidence (although weaker than Rose [117]) that the growth process of countries is similar to that of cities.

Another strand of research has examined Zipf's law for the sizes of states or provinces within a country. In this context, Soo [119] explored the Zipf and Gibrat's laws for the size distribution of US states. He observed that lognormal distribution offers a better fit to state sizes rather than the Pareto distribution, though he could not reject Pareto distribution statistically. He also rejects Gibrat's law because small states grew faster than the large states. In another study, Soo [120] examined the both laws for provinces of Brazil, China and India. He found complex evidence; the size distribution of Chinese provinces conforms to Zipf's law, the size distribution of Indian states do not conforms to Zipf's law, while he was unable to reject statistically any of the Pareto or lognormal distribution for Brazilian states. He also finds that Gibrat's law holds for Brazilian states. He argues that different characteristics of three countries have led to different state size distributions in three countries.

From this discussion, it is hard to establish that Zipf's law is universal for cities at sub-national or supranational levels. However, these are emerging areas in Zipf's law and city size distribution research and it is too early to make a conclusion and more empirical research is needed by considering sub-national or supranational samples.

4.2 The size distribution of lower-tail cities

In contrast to the long tradition of size distribution of upper-tail cities, recently some scholars have specifically considered the size distribution of lower-tail cities. Again, the Pareto (reverse) and lognormal distributions remain two controversial candidates for the size distribution of lower-tail cities. For example, Reed (2001, 2002) first observed an inverse power law in the lower-tail of city size distribution using the data of smallest 5,000 settlements of the US in 1998. Building on this work, a number of recent empirical studies have found that the lower-tail cities also follow a Pareto distribution but in a reverse direction. In this respect, Devadoss and Luckstead [121] analyzed a sample of US small cities and supported the reverse Pareto. In another study, Devadoss, Luckstead, Danforth and Akhundjanov [122] found similar evidence for small Indian cities.

Similarly, other studies have examined Gibrat's law of proportionate growth for small cities. Based on the theoretical model of Gabaix [13], these studies predict that if lower-tail cities obey a Pareto distribution, then these cities must also obey the Gibrat's law. In the similar vein, Devadoss and Luckstead [123] used the data of lower-tail small US cities for the census years of 2000 and 2010 and found that the growth process of these small cities follows Gibrat's law.

For lognormal distribution, Calderín-Ojeda [97] examined the size distribution of French settlements (communes) from 1962 to 2012 and found that upper-tail cities follow a Pareto distribution while lognormal distribution offers a better fit to untruncated settlements data of medium and small cities.

Like the size distribution of upper-tail cities, the Pareto and lognormal distributions are most likely candidates for the size distribution of lower-tail cities. It is also an emerging area for future empirical research.

4.3 Alternate Zipf's laws

Some studies suggest that the estimation of rank-size rule specified in Eq. (2) provides an approximation of the Zipf's law and acceptance or rejection of Zipf's law should not depend on the condition of Zipf's exponent to be statistically equal to one. Zipf's law still can hold if rank-size rule is only partially verified. These studies suggest that the focus should not be on to accept or reject the Zipf's law, but be on how well Zipf's law fits to the empirical city size distribution [106]. Building on this concept, recent contributions have introduced different forms of Zipf's law depending on how well it fits to the actual city size data.

For example, Perline [124] introduce three forms of Zipf's law: strong inverse power law (SIPL), weak inverse power law (WIPL) and false inverse power law (FIPL). SIPL refers to a situation where the power law exactly fits to the sample data for all ranges of the variable of interest. WIPL refers to a situation where the sample data fit a distribution that has an approximate inverse power form only in some upper range of values. Finally FIPL refers to a situation where a highly truncated sample from an exponential-type, especially from the lognormal, distribution can closely mimic a power law.

Benguigui and Blumenfeld-Lieberthal [10] suggest to study the city size distribution in three classes: class 1, class 2 and class 3. Class 1 refers to a situation where Zipf's exponent is exactly equal to one and size distribution follows a linear Zipf's law. Class 2 refers to a situation when Zipf's exponent is greater than one with a parabolic shape. Similarly, Class 3 refers to a situation where Zipf's exponent is less than one, also with a parabolic shape. The upper-tail cities have a parabolic shape in class 2, while they might have a power law shape in class 3 though it is a tail of lognormal distribution.

Chen [125] argue that city size distribution observes an evolutional urban process and different forms of Zipf's model are needed at different stages of urbanisation. He identifies that there are one parameter, two parameters and three parameters models. If the size distribution of a set of cities does not obey one parameter model, then the two parameter model can fit it well, and if the size distribution does not obey two parameter model then the three parameter model can fit it well. He defines these models as follows.

$$\mathrm{P(r)} = \frac{\mathrm{P_1}}{\mathrm{r}} \qquad Eq. (8)$$

Here, P(r) refers to the size of a city. r is the rank of a city (r = 1, 2, 3, . . .) when the size is used to rank cities. This is one parameter model with the only parameter P_1 called proportional coefficient and represents the size of the largest city. This model is pure form of Zipf's law with Zipf's exponent equal to one, and it is same as the one introduced above in Section 3.

The pure form of Zipf's law can be generalized as two parameter model as below.

$$\mathrm{P(r)} = \mathrm{P_1 r^{-q}} \qquad Eq. (9)$$

Here, q is a scaling exponent. This model contains two parameters: the proportionality coefficient (P_1) and the scaling exponent (q). This model allows a Zipf's coefficient different from one.

Three parameter model is as follows:

$$\mathrm{P(r)} = \mathrm{P_{1-k}(r+k)^{-q}} \qquad Eq. (10)$$

Here k is a scale-translational parameter of city rank, and P_{1-k} is a parameter that shows the size of the (1−k)th city that is defined in the possible world rather than the real world. This model can fit well to more even size distribution without leading cities in the top level of urban hierarchy.

4.4 The coherence property of the urban system

The controversy in the empirical evidence can also be explained in terms of the coherence property of the urban system. To observe strict power laws (i.e., Zipf's law), a system of objects should be coherent [113]. A system of cities does not obey true power law behavior because it is either incomplete or inconsistent with the conditions under which one might expect power laws to emerge. Power laws can only be applied to a group of cities which are integrated institutionally (i.e., common rules, common culture, common language, etc.) and economically and have co-evolved over time. The group of cities, which historically has observed an integrated evolution, converges to an organic economic unit. Consequently, the size distribution of cities becomes internally consistent for the group as a whole and obeys statistical properties of power laws. For example, the urban system of the United States is more coherent at the national-level as compared to the urban systems in each state. On the other hand, the urban system in the European Union is less coherent at union-level as compared to the urban systems in individual European countries. Underlying reason is that historically the US urban system is integrated and has coevolved over long time period at national level, while the urban system in the European Union is more integrated and has coevolved at individual country-level.

From this perspective, the question arises whether the rejection of Zipf's law hypothesis for some urban systems is due to the coherence property. In this respect, Arshad, Hu and Ashraf [126] examined the Zipf's law and the coherence property of the urban system for Pakistani cities. The authors find that Zipf's law is rejected at national-level, while it is more likely to be held for the city size distribution at province-level (i.e., for each of the four provinces of the country). They argue that the results are driven by the level of coherence of the urban system; the urban systems within Pakistani provinces are more coherent in terms of common language, common culture and common rules as compared to the urban system at national-level. To date, little attention has been paid to the coherence property in Zipf's law and city size distribution research and is an interesting area for future research.

5. Conclusion

In this paper, we provide a systematic review of the existing literature on Zipf's law and city size distribution. The review includes 114 studies published from 2000 to 2017.

Our qualitative review discovers several stylized facts. Existing empirical evidence suggests that Zipf's law is not always observable even for the upper-tail cities of a country/region. This finding is opposite to the conventional wisdom which suggests that Zipf's law for at least upper-tail cities is universal. However, we further find that existing empirical evidence is not robust. The hypothesis of Zipf's law is too often rejected when the data of all cities is considered, and in this case the alternative distributions have been suggested. Contrary, the hypothesis is more likely to be accepted due to improvement in empirical methods and city definitions. These findings suggest that the debate is still far from to be conclusive.

We identify several potential areas for future research. First, future studies may use bias corrected empirical methods and better definition of cities to examine Zipf's law. Second, recently several alternative distributions have been suggested for cities. Future research may endeavour to provide theoretical foundations to these distributions. In addition, more empirical research is also needed for their widespread acceptance. Third, we identify several emerging areas in Zipf's law and city size distribution research which need further consideration. These areas include the size distribution of lower-tail cities, the size distribution of cities in sub-national regions, the alternative forms of Zipf's law and the coherence property of the urban system.

References

[1] F. Auerbach, Das gesetz der bevölkerungskonzentration, Petermanns Geographische Mitteilungen, 59 (1913) 74-76.
[2] H.W. Singer, The "Courbe des Populations." A Parallel to Pareto's Law, The Economic Journal, 46 (1936) 254-263.
[3] G.K. Zipf, Human behavior and the principle of least effort, Reading, MA: Addison-Wesley (1949).
[4] M. Lu, G. Wan, Urbanization and Urban Systems in the People's Republic of China: Research Findings and Policy Recommendations, Journal of Economic Surveys, 28 (2014) 671-685.
[5] C.A. Ignazzi, The Brazilian Urban System: the trajectories of Brazilian cities between general dynamics and specific peculiarities, Cybergeo: European Journal of Geography, (2015) document 754.
[6] J. Le Gallo, C. Chasco, Spatial analysis of urban growth in Spain, 1900–2001, Empirical Economics, 34 (2008) 59-80.
[7] E. Pérez-Campuzano, L. Guzmán-Vargas, F. Angulo-Brown, Distributions of city sizes in Mexico during the 20th century, Chaos, Solitons & Fractals, 73 (2015) 64-70.
[8] M. Puente-Ajovín, A. Ramos, On the parametric description of the French, German, Italian and Spanish city size distributions, Ann Reg Sci, 54 (2015) 489-509.
[9] K.T. Soo, Zipf's Law for cities: a cross-country investigation, Regional Science and Urban Economics, 35 (2005) 239-263.
[10] L. Benguigui, E. Blumenfeld-Lieberthal, The end of a paradigm: is Zipf's law universal?, Journal of Geographical Systems, 13 (2011) 87-100.
[11] M. Ausloos, R. Cerqueti, A universal rank-size law, PloS one, 11 (2016) e0166011.
[12] V. Nitsch, Zipf zipped, Journal of Urban Economics, 57 (2005) 86-100.
[13] X. Gabaix, Zipf's law for cities: an explanation, Quarterly journal of Economics, (1999) 739-767.
[14] J. Eeckhout, Gibrat's Law for (All) Cities, The American Economic Review, 94 (2004) 1429-1451.
[15] R. Gibrat, Les inégalités économiques, Libraire du Recueil Siray, Paris, 1931.
[16] J.-C. Córdoba, On the distribution of city sizes, Journal of Urban Economics, 63 (2008) 177-197.
[17] J.C. Córdoba, A generalized Gibrat's law, International Economic Review, 49 (2008) 1463-1468.
[18] M. Modica, A. Reggiani, P. Nijkamp, Methodological Advances in Gibrat's and Zipf's Laws: A Comparative Empirical Study on the Evolution of Urban Systems, in: H. Shibusawa, K. Sakurai, T. Mizunoya, S. Uchida (Eds.) Socioeconomic Environmental Policies and Evaluations in Regional Science: Essays in Honor of Yoshiro Higano, Springer Singapore, Singapore, 2017, pp. 37-59.
[19] B.J. Berry, A. Okulicz-Kozaryn, The city size distribution debate: Resolution for US urban regions and megalopolitan areas, Cities, 29 (2012) S17-S23.
[20] Y. Malevergne, V. Pisarenko, D. Sornette, Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities, Physical Review E, 83 (2011) 036111.
[21] G. Duranton, Some foundations for Zipf's law: Product proliferation and local spillovers, Regional science and urban Economics, 36 (2006) 542-563.
[22] G. Duranton, Urban evolutions: The fast, the slow, and the still, The American Economic Review, 97 (2007) 197-221.
[23] W.T. Hsu, Central place theory and city size distribution, The Economic Journal, 122 (2012) 903-932.
[24] S. Lee, Q. Li, Uneven landscapes and city size distributions, Journal of Urban Economics, 78 (2013) 19-29.
[25] E. Rossi-Hansberg, M.L. Wright, Urban structure and growth, The Review of Economic Studies, 74 (2007) 597-624.
[26] K. Behrens, G. Duranton, F. Robert-Nicoud, Productive cities: Sorting, selection, and agglomeration, Journal of Political Economy, 122 (2014) 507-553.

[27] Y. Mansury, L. Gulyás, The emergence of Zipf's Law in a system of cities: An agent-based simulation approach, Journal of Economic Dynamics and Control, 31 (2007) 2438-2460.
[28] R. Axtell, R. Florida, Emergent cities: a microeconomic explanation for Zipf's law, in, Society for Computational Economics, 2001.
[29] F. Semboloni, Agents with dycotomic goals which generate a rank-size distribution, CASA working paper no. 33. Centre for Advanced Spatial Analysis: London, (2001).
[30] L. Gan, D. Li, S. Song, Is the Zipf law spurious in explaining city-size distributions?, Economics Letters, 92 (2006) 256-262.
[31] M. Batty, Rank clocks, Nature, 444 (2006) 592-596.
[32] B. Corominas-Murtra, R.V. Solé, Universality of Zipf's law, Physical Review E, 82 (2010) 011102.
[33] P. Krugman, Confronting the Mystery of Urban Hierarchy, Journal of the Japanese and International Economies, 10 (1996) 399-418.
[34] M. Levy, Gibrat's Law for (All) Cities: Comment, The American Economic Review, 99 (2009) 1672-1675.
[35] Y.M. Ioannides, H.G. Overman, Zipf's law for cities: an empirical examination, Regional Science and Urban Economics, 33 (2003) 127-137.
[36] B. Jiang, T. Jia, Zipf's law for all the natural cities in the United States: a geospatial perspective, International Journal of Geographical Information Science, 25 (2011) 1269-1281.
[37] H.D. Rozenfeld, D. Rybski, X. Gabaix, H.A. Makse, The Area and Population of Cities: New Insights from a Different Perspective on Cities, The American Economic Review, 101 (2011) 2205-2225.
[38] D. Black, V. Henderson, Urban evolution in the USA, Journal of Economic Geography, 3 (2003) 343-372.
[39] J.B. Parr, K. Suzuki, Settlement populations and the lognormal distribution, Urban Studies, 10 (1973) 335-352.
[40] M. Bee, M. Riccaboni, S. Schiavo, The size distribution of US cities: Not Pareto, even in the tail, Economics Letters, 120 (2013) 232-237.
[41] R. González ‐ Val, The Evolution Of US City Size Distribution From A Long ‐ Term Perspective (1900 – 2000), Journal of Regional Science, 50 (2010) 952-972.
[42] R. González ‐ Val, A nonparametric estimation of the local Zipf exponent for all US cities, Environment and Planning B: Planning and Design, 39 (2012) 1119-1130.
[43] R. González-Val, L. Lanaspa, F. Sanz-Gracia, New Evidence on Gibrat's Law for Cities, Urban Studies, 51 (2014) 93-115.
[44] Y.M. Ioannides, S. Skouras, US city size distribution: Robustly Pareto, but only in the tail, Journal of Urban Economics, 73 (2013) 18-29.
[45] G. Fazio, M. Modica, Pareto or Log-Normal? Best Fit and Truncation in the Distribution of All Cities, Journal of Regional Science, 55 (2015) 736-756.
[46] Y. Chen, Y. Zhou, Scaling laws and indications of self-organized criticality in urban systems, Chaos, Solitons & Fractals, 35 (2008) 85-98.
[47] K. Gangopadhyay, B. Basu, City size distributions for India and China, Physica A: Statistical Mechanics and its Applications, 388 (2009) 2682-2688.
[48] K. Gangopadhyay, B. Basu, Evolution of Zipf's Law for Indian Urban Agglomerations Vis-à-Vis Chinese Urban Agglomerations, in: Econophysics of Systemic Risk and Network Dynamics, Springer, 2013, pp. 119-129.
[49] X. Ye, Y. Xie, Re-examination of Zipf's law and urban dynamic in China: a regional approach, Ann Reg Sci, 49 (2012) 135-156.
[50] W. Ziqin, Zipf Law Analysis of Urban Scale in China, Asian Journal of Social Science Studies, 1 (2016) 53.

[51] G. Anderson, Y. Ge, The size distribution of Chinese cities, Regional Science and Urban Economics, 35 (2005) 756-776.
[52] L. Benguigui, E. Blumenfeld-Lieberthal, Beyond the power law – a new approach to analyze city size distributions, Computers, Environment and Urban Systems, 31 (2007) 648-666.
[53] L. Benguigui, E. Blumenfeld-Lieberthal, A dynamic model for city size distribution beyond Zipf's law, Physica A: Statistical Mechanics and its Applications, 384 (2007) 613-627.
[54] H. Li, Y.D. Wei, Y. Ning, Spatial and Temporal Evolution of Urban Systems in China during Rapid Urbanization, Sustainability, 8 (2016) 651.
[55] S. Li, D. Sui, Pareto's law and sample size: a case study of China's urban system 1984–2008, GeoJournal, 78 (2013) 615-626.
[56] J. Luckstead, S. Devadoss, A comparison of city size distributions for China and India from 1950 to 2010, Economics Letters, 124 (2014) 290-295.
[57] S. Song, K.H. Zhang, Urbanisation and city size distribution in China, Urban Studies, 39 (2002) 2317-2327.
[58] G. Peng, Zipf's law for Chinese cities: Rolling sample regressions, Physica A: Statistical Mechanics and its Applications, 389 (2010) 3804-3813.
[59] C.C. Fan, The vertical and horizontal expansions of China's city system, Urban Geography, 20 (1999) 493-515.
[60] Z. Chen, S. Fu, D. Zhang, Searching for the parallel growth of cities in China, Urban Studies, 50 (2013) 2118-2135.
[61] J.-X. Wu, L.-Y. He, How do Chinese cities grow? A distribution dynamics approach, Physica A: Statistical Mechanics and its Applications, 470 (2017) 105-118.
[62] L. Fang, P. Li, S. Song, China's urban development policies and city growth dynamics: An analysis based on the Gibrat's Law, International Review of Economics & Finance, (2017).
[63] K.T. Rosen, M. Resnick, The size distribution of cities: An examination of the Pareto law and primacy, Journal of Urban Economics, 8 (1980) 165-186.
[64] J. Eeckhout, Gibrat's Law for (All) Cities: Reply, The American Economic Review, 99 (2009) 1676-1683.
[65] K. Giesen, J. Südekum, Zipf's law for cities in the regions and the country, Journal of Economic Geography, 11 (2011) 667-686.
[66] L. Gligor, M. Gligor, The fractal city theory revisited: new empirical evidence from the distribution of Romanian cities and towns, Nonlinear dynamics, psychology, and life sciences, 12 (2008) 15-28.
[67] E. Ezzahid, O. ElHamdani, Zipf'S Law In The Case Of Moroccan Cities, Review of Urban & Regional Development Studies, 27 (2015) 118-133.
[68] A. Schaffar, D. Nassori, Urban Growth in Morocco: Convergence versus Concentration, Revue économique, 67 (2016) 207-226.
[69] S. Rastvortseva, I. Manaeva, Zipf'S Law Appearance In The Russian Cities, Regional Science Inquiry, 8 (2016) 51-59.
[70] A. Lalanne, Zipf's Law and Canadian Urban Growth, Urban Studies, 51 (2014) 1725-1740.
[71] J. Dubé, M. Polèse, À propos du rôle de la taille dans la croissance urbaine: Une analyse pour 135 agglomérations canadiennes entre 1971 et 2011, The Canadian Geographer/Le Géographe canadien, 60 (2016) 541-555.
[72] L. Lanaspa, F. Pueyo, F. Sanz, The evolution of Spanish urban structure during the twentieth century, Urban Studies, 40 (2003) 567-580.
[73] N.J. Moura, M.B. Ribeiro, Zipf law for Brazilian cities, Physica A: Statistical Mechanics and its Applications, 367 (2006) 441-448.
[74] V.J. Matlaba, M.J. Holmes, P. McCann, J. Poot, A Century of the Evolution of the Urban System in Brazil, Review of Urban & Regional Development Studies, 25 (2013) 129-151.

[75] K.T. Soo, Zipf's Law and Urban Growth in Malaysia, Urban Studies, 44 (2007) 1-14.
[76] H.E. Duran, S.P. Ozkan, Trade Openness, Urban Concentration And City-Size Growth In Turkey, Regional Science Inquiry, 7 (2015) 35-46.
[77] E. Deliktas, A.Ö. Önder, M. Karadag, The size distribution of cities and determinants of city growth in Turkey, European Planning Studies, 21 (2013) 251-263.
[78] X. Gabaix, R. Ibragimov, Rank − 1 / 2: A Simple Way to Improve the OLS Estimation of Tail Exponents, Journal of Business & Economic Statistics, 29 (2011) 24-39.
[79] W.C. Wheaton, H. Shishido, Urban concentration, agglomeration economies, and the level of economic development, Economic development and cultural change, (1981) 17-30.
[80] F. Guerin-Pace, Rank-Size Distribution and the Process of Urban Growth, Urban Studies, 32 (1995) 551-562.
[81] I.I. Eliazar, M.H. Cohen, Power-law connections: From Zipf to Heaps and beyond, Annals of Physics, 332 (2012) 56-74.
[82] I. Eliazar, M.H. Cohen, Hierarchical socioeconomic fractality: The rich, the poor, and the middle-class, Physica A: Statistical Mechanics and its Applications, 402 (2014) 30-40.
[83] I.I. Eliazar, M.H. Cohen, Rank distributions: A panoramic macroscopic outlook, Physical Review E, 89 (2014) 012111.
[84] I.I. Eliazar, M.H. Cohen, A Langevin approach to the Log–Gauss–Pareto composite statistical structure, Physica A: Statistical Mechanics and its Applications, 391 (2012) 5598-5610.
[85] I.I. Eliazar, M.H. Cohen, Topography of chance, Physical Review E, 88 (2013) 052104.
[86] M.H. Cohen, I.I. Eliazar, Econophysical visualization of Adam Smith's invisible hand, Physica A: Statistical Mechanics and its Applications, 392 (2013) 813-823.
[87] I. Eliazar, Power-law and exponential rank distributions: A panoramic Gibbsian perspective, Annals of Physics, 355 (2015) 322-337.
[88] L.C. Malacarne, R. Mendes, E.K. Lenzi, q-Exponential distribution in urban agglomeration, Physical Review E, 65 (2001) 017106.
[89] J. Luckstead, S. Devadoss, Pareto tails and lognormal body of US cities size distribution, Physica A: Statistical Mechanics and its Applications, 465 (2017) 573-578.
[90] W.J. Reed, The Pareto, Zipf and other power laws, Economics Letters, 74 (2001) 15-19.
[91] W.J. Reed, On the Rank-Size Distribution for Human Settlements, Journal of Regional Science, 42 (2002) 1-17.
[92] W.J. Reed, The Pareto law of incomes—an explanation and an extension, Physica A: Statistical Mechanics and its Applications, 319 (2003) 469-486.
[93] A. Ramos, F. Sanz-Gracia, R. González-Val, On the parametric description of US city size distribution: new empirical evidence, (2014).
[94] J.M. Sarabia, F. Prieto, The Pareto-positive stable distribution: A new descriptive model for city size data, Physica A: Statistical Mechanics and its Applications, 388 (2009) 4179-4191.
[95] E. Gómez-Déniz, E. Calderín-Ojeda, On the use of the Pareto ArcTan distribution for describing city size in Australia and New Zealand, Physica A: Statistical Mechanics and its Applications, 436 (2015) 821-832.
[96] A. Subbarayan, G. Kumar, The size distribution of cities and towns in the states of India (1951-2011) : An evaluation of pareto, lognormal and q-Exponential distributions, International Journal of Agricultural and Statistical Sciences, 11 (2015) 481-486.
[97] E. Calderín-Ojeda, The distribution of all French communes: A composite parametric approach, Physica A: Statistical Mechanics and its Applications, 450 (2016) 385-394.
[98] J. Luckstead, S. Devadoss, D. Danforth, The size distributions of all Indian cities, Physica A: Statistical Mechanics and its Applications, 474 (2017) 237-249.

[99] W.J. Reed, M. Jorgensen, The Double Pareto-Lognormal Distribution—A New Parametric Model for Size Distributions, Communications in Statistics - Theory and Methods, 33 (2004) 1733-1753.
[100] K. Giesen, A. Zimmermann, J. Suedekum, The size distribution across all cities–double Pareto lognormal strikes, Journal of Urban Economics, 68 (2010) 129-137.
[101] R. González-Val, A. Ramos, F. Sanz-Gracia, M. Vera-Cabello, Size distributions for all cities: Which one is best?, Papers in Regional Science, 94 (2015) 177-196.
[102] N.K. Vitanov, M. Ausloos, Test of two hypotheses explaining the size of populations in a system of cities, Journal of Applied Statistics, 42 (2015) 2686-2693.
[103] S. Singh, G.S. Maddala, A function for size distribution of incomes, Econometrica, 44 (1976) 963–970.
[104] G. Kumar, A. Subbarayan, The Size Distribution of Cities in a Region: An Evaluation of Pareto, Lognormal and PPS Distributions, International Journal of Agricultural and Statistical Sciences, 9 (2013) 493-500.
[105] C.A. Vallabados, S.A. Arumugam, An Evaluation of Pareto, Lognormal and PPS Distributions: The Size Distribution of Cities in Kerala, India, Journal of Modern Applied Statistical Methods, 15 (2016) 41.
[106] X. Gabaix, Y.M. Ioannides, The evolution of city size distributions, Handbook of regional and urban economics, 4 (2004) 2341-2378.
[107] Y. Nishiyama, S. Osada, Y. Sato, OLS estimation and the *t* test revisited in rank-size rule regression, Journal of Regional Science, 48 (2008) 691-716.
[108] P. Cheshire, Trends in sizes and structures of urban areas, Handbook of regional and urban economics, 3 (1999) 1339-1373.
[109] B. Jiang, X. Liu, Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information, International Journal of Geographical Information Science, 26 (2012) 215-229.
[110] B. Jiang, J. Yin, Q. Liu, Zipf's law for all the natural cities around the world, International Journal of Geographical Information Science, 29 (2015) 498-522.
[111] K. Schmidheiny, J. Suedekum, The pan-European population distribution across consistently defined functional urban areas, Economics Letters, (2015).
[112] P. Veneri, City size distribution across the OECD: Does the definition of cities matter?, Computers, Environment and Urban Systems, 59 (2016) 86-94.
[113] M. Cristelli, M. Batty, L. Pietronero, There is more than a power law in Zipf, Scientific reports, 2 (2012).
[114] A. Subbarayan, The size distribution of cities in Tamilnadu (1901-2001), International Journal of Agricultural and Statistical Sciences, 5 (2009) 373-382.
[115] G. Kumar, A. Subbarayan, The temporal dynamics of regional city size distribution: Andhra pradesh (1951-2001), Journal of Mathematics and Statistics, 10 (2014) 221-230.
[116] J. Luckstead, S. Devadoss, Do the world's largest cities follow Zipf's and Gibrat's laws?, Economics Letters, 125 (2014) 182-186.
[117] A. Rose, Cities and Countries, Journal of Money, Credit, and Banking, 38 (2006) 2225-2245.
[118] R. González-Val, M. Sanso-Navarro, Gibrat's law for countries, Journal of population economics, 23 (2010) 1371-1389.
[119] K.T. Soo, The size and growth of state populations in the United States, Economics Bulletin, 32 (2012) 1238-1249.
[120] K.T. Soo, Zipf, Gibrat and geography: Evidence from China, India and Brazil, Papers in Regional Science, 93 (2014) 159-181.
[121] S. Devadoss, J. Luckstead, Size distribution of U.S. lower tail cities, Physica A: Statistical Mechanics and its Applications, 444 (2016) 158-162.

[122] S. Devadoss, J. Luckstead, D. Danforth, S. Akhundjanov, The power law distribution for lower tail cities in India, Physica A: Statistical Mechanics and its Applications, 442 (2016) 193-196.
[123] S. Devadoss, J. Luckstead, Growth process of U.S. small cities, Economics Letters, 135 (2015) 12-14.
[124] R. Perline, Strong, weak and false inverse power laws, Statistical Science, (2005) 68-88.
[125] Y. Chen, The evolution of Zipf’s law indicative of city development, Physica A: Statistical Mechanics and its Applications, 443 (2016) 555-567.
[126] S. Arshad, S. Hu, B.N. Ashraf, Zipf's law, the coherence of the urban system and city size distribution: evidence from Pakistan, Available at SSRN: https://ssrn.com/abstract=2963736, (2017).