Cities 149 (2024) 104946

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

A long-term, regional-level analysis of Zipf's and Gibrat's laws in the United States

Rafael González-Val [a,b,*], Domingo P. Ximénez-de-Embún [a], Fernando Sanz-Gracia [a]

[a] Universidad de Zaragoza, Facultad de Economía y Empresa (Campus Paraíso), c/Gran Vía 2, 50005 Zaragoza, Spain

[b] Institut d'Economia de Barcelona (IEB), Facultat d'Economia i Empresa, Universitat de Barcelona, c/John M. Keynes, 1-11, 08034 Barcelona, Spain

ARTICLE INFO

JEL classification:
C12
R11
R12

Keywords:
US states
Zipf's law
Power law
Gibrat's law

ABSTRACT

We test the validity of Zipf's and Gibrat's laws for city size distributions at the regional level from 1900 to 2010 by considering US states. Zipf's law is satisfied for a majority of states, but for the United States as a whole it only held during the first half of the twentieth century. The null hypothesis of a power law is not rejected at the national level or for most states (the maximum number of rejections in one year is 13 states out of 48). There is evidence supporting a weak version of Gibrat's law in the long-term; mean growth is independent of initial population for most city sizes over the entire United States and in 27 states, while the variance of growth is size-dependent.

1. Introduction

The study of power laws or, more concretely, the analysis of the validity of Zipf's law, has a long tradition in many fields such as Urban Economics (Clauset et al., 2009); the same is true in the study of the validity of what is known as Gibrat's law. Zipf's law proposes a linear and stable relationship between the rank and the size (population) of cities; Gibrat's law, or the law of proportionate growth, postulates that the growth rate of cities tends to be independent of their initial population. Excellent surveys on the topic of Zipf's and Gibrat's laws and, in general, city size distribution, can be found in Nitsch (2005), González-Val et al. (2014), Cottineau (2017), and Arshad et al. (2018). The vast majority of the studies consider nations as their geographical unit of reference (the cities that belong to a country) or, at most, a set of nations. Very few studies, conversely, have focused on a subnational area, such as a province or a state, as their geographical unit of reference.

In this context, at a subnational level, the basic reference is the paper by Giesen and Südekum (2011), which analysed the German case. Furthermore, Pérez-Valbuena and Meisel-Roca (2014) analysed Colombia, Subbarayan (2009) and Kumar and Subbarayan (2014) studied a concrete Indian province, Ye and Xie (2012) analysed eight Chinese subregions, Ziqin (2016) considered 26 Chinese provinces, Li and Zhang (2018) analysed subnational Chinese administrative areas, Arshad et al. (2019) studied Pakistan, and Kundak and Dökmeci (2018) focused on Turkish provinces. Surprisingly, the case of city size distribution in the United States (US) has not been analysed in a systematic way at a subnational level. Only Garmestani et al. (2007, 2008) studied the validity of Zipf's and Gibrat's laws for the south eastern and south western US.

Nevertheless, the regional study of city size distributions is an important issue from several perspectives. First, from a theoretical point of view, there is a statistical connection between regional and national city size distributions (Gabaix, 1999; Giesen & Südekum, 2011). Second, from a conceptual point of view, a regional definition of urban systems generally makes more sense than a nationwide urban system (especially for large countries), because most migrations take place between nearby cities within the same region rather than between the largest cities in a country, which are typically located at some distance from one another because of the possible existence of agglomeration shadows (Cuberes et al., 2021; Krugman, 1993). In the context of the United States, Rauch (2014), using US microdata, found that the majority of US citizens (over 68 %) live within 0 and 100 km of their birthplace. Rauch (2014) also found, by estimating a standard gravity equation, that the relationship between the number of people and the distance between their home and place of birth decreases with distance.

This paper largely focuses on studying the validity of Zipf's and Gibrat's laws considering each one of the continental states of the United States from 1900 to 2010, since testing both laws requires long time

* Corresponding author at: Universidad de Zaragoza, Facultad de Economía y Empresa (Campus Paraíso), c/Gran Vía 2, 50005 Zaragoza, Spain.
E-mail addresses: rafaelg@unizar.es (R. González-Val), domingo.perez@unizar.es (D.P. Ximénez-de-Embún), fsanz@unizar.es (F. Sanz-Gracia).

https://doi.org/10.1016/j.cities.2024.104946
Received 25 September 2023; Received in revised form 28 February 2024; Accepted 9 March 2024
Available online 19 March 2024

intervals (Gabaix & Ioannides, 2004). Starting from untruncated data for all cities (incorporated places) by state, a threshold value is endogenously defined for each year and state following the method of Clauset et al. (2009). Then, we focus on the analysis of the upper tail. To the best of our knowledge, this study is the first of its kind.

This research objective is interesting because, as noted above, regional and national city size distributions are connected: Gabaix's second proposition (1999, p. 751) states that if a country is composed of several regions and Gibrat's law holds in each of them, then Zipf's law is verified at both the national and regional levels. However, the converse need not be true: Zipf's law may hold at the national level without Gibrat's law being satisfied in all regions, as long as Gibrat's law is satisfied in the national aggregate (Giesen & Südekum, 2011). Gabaix's work (1999) is one of the most influential theoretical papers in city size distribution published in recent decades but, to date, only one formal test of this hypothesis has been carried out. That study, Giesen and Südekum (2011), empirically corroborated the proposition for the German case.

Our primary results are as follows. Zipf's law holds in the upper tail distribution for most states (the percentage of rejections attains its maximum value, 22.9 %, in 2010). For the United States as a whole, the law holds in the first half of the twentieth century. Moreover, using the test proposed by Clauset et al. (2009), we cannot reject the hypothesis that city size distribution follows a power law in most cases (between 72.9 % and 95.8 % of states, depending on the year). This hypothesis cannot be rejected for the entire US city size distribution in any year except for one (i.e., 1970). Finally, Gibrat's law for mean growth holds for most city sizes in the US, although deviations appear for the largest cities. Gibrat's law also holds in a non-negligible number of states, while the variance of growth rates is never size-independent for either the United States as a whole or for individual states.

The rest of the paper is organised as follows. Section 2 describes the data used, and the methodology is explained in Section 3. The primary results are presented in Section 4, and they are discussed in Section 5, which also concludes the paper.

2. Data

We used the same data set as González-Val (2010) and González-Val et al. (2014). This data set, created from the original documents of the decennial census published by the US Census Bureau, consists of the available data for all incorporated places without any size restriction for each decade of the twentieth century. We also added in population data from the 2010 US census.

We identify cities as places that the US Census Bureau defines as 'incorporated places,' which refers to a governmental unit incorporated under state law as a city, town, borough or village that has legally established limits, powers, and functions. Such incorporated places are administratively defined cities (i.e., legal cities). We excluded Alaska, Hawaii, and Puerto Rico because of data limitations. The District of Columbia is also excluded because it includes only one city. Therefore, we considered the cities in 48 states.

The proportion of the population living in cities (incorporated places) increased from less than half of the total US population in 1900 (47 %) to 63 % in 2010. The population excluded from the sample is what the US Census Bureau calls 'population not in place.' Incorporated places do not cover the entire territory of the United States, and some territories are excluded from any recognised place. Most of these people are part of the rural population. However, the population of incorporated places is almost entirely urban.

Table 1 lists the total number of cities (incorporated places) for each of the 48 states in 1900, 1950 (the date of one of the intermediate census), and 2010. The sample of incorporated places provides comprehensive information about the birth of new cities. In the United States, urban growth has a double dimension: cities increase in size (i.e., population growth), and the number of cities also increases. These two facts have potentially different effects on city size distribution (González-Val, 2010). The number of incorporated places in the sample increased from 9534 in 1900 to 19,124 in 2010. The number of cities also increased over time in all states. But states are quite heterogeneous, and there are states with a small number of cities (e.g., Nevada, New Hampshire, and Rhode Island) and states with a large number of cities (e.g., Illinois, Iowa, and Pennsylvania).

The raw data set is untruncated and includes all incorporated places with no size restrictions. For our purposes, however, the small towns are of little interest because Zipf's law concentrates on the upper tail of the city size distribution (Eeckhout, 2009), although there is empirical evidence indicating that the lower tail of the US city size distribution, the smallest cities, are also Pareto-distributed (Giesen et al., 2010; Giesen & Suedekum, 2014; Luckstead & Devadoss, 2017; Reed, 2001, 2002). This result is confirmed for small cities in other countries, such as India (Devadoss et al., 2016). However, regarding Gibrat's law, Eeckhout (2004) found that, for very small cities and very large cities in the US, from 1990 to 2000, the variation in growth rates was markedly different, although Devadoss and Luckstead (2015) concluded that Gibrat's law held for small cities in the next decade (2000−2010). Nevertheless, from the long-term perspective of 1900 to 2000, González-Val et al. (2014) observed that the smallest cities presented clearly higher variances than the middle-sized and large cities.

Therefore, throughout this paper we focus on an empirical analysis of the upper tail distribution that is defined for each state following the method of Clauset et al. (2009). Table 1 also lists the sample size and population thresholds for the upper tail.

We can distinguish five cases in terms of the number of cities in the upper tail over time: (1) for some states, the number of cities decreases over time (i.e., Maryland, Missouri), (2) for some states, the number of cities increases over time (i.e., Arizona, Connecticut, Delaware, Idaho, Minnesota, Mississippi, Montana, Nebraska, Nevada, New Jersey, Oklahoma, Pennsylvania, Tennessee, Texas, Vermont, and Wyoming), (3) for some states, the number of cities remains fairly constant over time (i.e., Alabama, New Hampshire, and Rhode Island), (4) for some states, the number of cities increases and later decreases over time (i.e., Arkansas, California, Colorado, Florida, Georgia, Illinois, Indiana, Iowa, Kansas, Michigan, New Mexico, New York, North Carolina, North Dakota, Ohio, Oregon, South Dakota, Utah, Washington, West Virginia, and Wisconsin), and (5) for some states, the number of cities decreases and later increases over time (i.e., Kentucky, Louisiana, Maine, Massachusetts, South Carolina, and Virginia).

Overall, the population at the truncation point increases over time in almost all cases (Table 1). Furthermore, it is the most common for the number of cities at the upper tail to increase when the total number of cities increases (although in recent decades there has been a subsequent decrease in the number of cities in many states). These results confirm that the estimated threshold increases with sample size (Fazio & Modica, 2015).[1]

Note that we use administrative definitions for both cities (incorporated places) and regions (states). We acknowledge that, since the contribution of Rozenfeld et al. (2011), the consideration of the administrative definition of a city is at least debatable. Beginning with the work of Schmidheiny and Südekum (2015), there has been a boom in the development of new methods to delineate and define urban areas using building density, machine learning, personal judgment, and other methods (e.g., Arribas-Bel et al., 2021; Ch et al., 2021; de Bellefon et al., 2021; Galdo et al., 2021; Moreno-Monroy et al., 2021). Regarding regions, Mori et al. (2020) used empirical methods to define regions in a set of countries, including the United States. Giesen and Südekum (2011) used some samples of random regions and spatial groups of cities

[1] This is not a particular feature of the Clauset et al. (2009) method; the thresholds estimated using alternative methodologies (Bee et al., 2011, 2013; Malevergne et al., 2011) also increase with sample size (Fazio & Modica, 2015).

Table 1
Sample size on a state-by-state basis.

State/year	Total number of cities 1900	1950	2010	Upper tail size (cities) 1900	1950	2010	Avg. (1900–2010)	Truncation size (population) 1900	1950	2010
Alabama	155	289	444	58	53	56	87.3	880	4225	11,620
Arizona	12	45	82	12	20	28	31.6	521	3466	25,536
Arkansas	164	370	496	23	192	140	147.8	1748	517	1801
California	110	302	452	33	110	105	90.8	3057	9188	73,732
Colorado	123	233	269	56	109	83	97.5	634	684	3887
Connecticut	24	26	29	11	16	17	14.3	12,681	17,455	27,620
Delaware	33	50	57	13	21	32	25.8	1132	1015	973
Florida	73	283	396	47	90	52	74.4	543	2752	48,452
Georgia	303	481	527	70	97	63	117.3	1150	2424	12,950
Idaho	34	174	201	32	122	142	91.0	283	337	377
Illinois	893	1146	1284	312	636	115	367.9	995	641	21,838
Indiana	377	519	566	258	374	333	301.3	596	479	861
Iowa	668	918	946	383	413	386	410.3	500	523	682
Kansas	346	596	626	196	350	200	238.2	527	349	1010
Kentucky	223	288	419	138	62	69	93.3	519	2926	5723
Louisiana	98	221	302	53	43	56	68.2	688	4666	6112
Maine	21	21	22	16	6	10	10.3	5311	20,913	15,722
Maryland	90	144	156	66	53	52	70.4	410	1420	3844
Massachusetts	39	39	44	25	12	35	26.0	21,766	80,536	35,177
Michigan	370	482	532	190	292	125	206.8	859	877	7088
Minnesota	431	763	849	67	218	555	248.0	1648	1012	363
Mississippi	173	261	294	21	114	158	116.4	2678	1116	1019
Missouri	489	756	938	355	97	75	115.0	331	2836	10,204
Montana	25	120	128	19	41	45	55.3	995	1522	1464
Nebraska	354	524	530	78	132	193	154.2	850	856	539
Nevada	2	14	17	2	9	17	9.6	2100	2400	1130
New Hampshire	11	12	12	10	10	12	9.1	7023	12,811	8477
New Jersey	176	315	319	51	92	114	101.7	3244	6766	9450
New Mexico	9	70	101	6	50	31	24.3	2735	784	6024
New York	425	590	613	191	201	158	176.4	1495	2826	5399
North Carolina	276	406	532	91	135	62	146.5	707	1598	13,656
North Dakota	75	325	357	37	164	90	138.5	576	328	563
Ohio	670	885	938	372	580	107	232.4	613	542	17,288
Oklahoma	124	497	589	77	119	151	136.8	437	1676	1964
Oregon	106	201	238	46	136	40	73.9	506	539	12,883
Pennsylvania	759	969	1013	140	226	236	192.3	3416	4948	4282
Rhode Island	6	7	8	6	6	4	5.4	13,343	37,564	71,148
South Carolina	162	224	267	73	57	128	62.9	536	2688	1697
South Dakota	135	297	309	87	157	110	121.8	311	373	547
Tennessee	84	229	343	49	59	107	75.8	1266	3191	4396
Texas	186	688	1183	45	147	213	146.8	3422	5183	10,400
Utah	68	197	234	42	74	27	64.6	877	974	26,263
Vermont	29	43	46	19	21	31	22.7	954	1267	565
Virginia	132	209	229	115	108	145	116.1	288	1010	923
Washington	82	233	275	40	145	36	82.8	761	710	29,799
West Virginia	113	211	231	78	118	86	85.1	429	1146	1420
Wisconsin	260	526	584	95	260	104	260.9	1584	778	7692
Wyoming	16	85	97	10	23	65	54.3	737	2089	383
US (nationwide)	9534	16,284	19,124	1475	887	451	940.8	2714	13,798	69,772

in addition to the German federal states. However, data for all these new city and region definitions are only available for a small number of recent years; for a long-term perspective (from 1900 onwards), the primary and unique data available correspond to administrative units.

3. Methodology

Let S denote a city's relative size, that is to say, its population in year t divided by the average population of all US cities in the upper tail in that year. From a long-term temporal perspective of steady-state distributions, it is necessary to use a relative measure of size (Gabaix & Ioannides, 2004) because the same population can correspond to different city sizes at different moments in time. For instance, a city of 1000 inhabitants does not have the same city size in 1900 and in 2010. Additionally, a city can grow in absolute population but decrease in relative size. If S is distributed according to a power law, also known as a Pareto distribution, the density function is $p(S) = \frac{a-1}{\underline{S}}\left(\frac{S}{\underline{S}}\right)^{-a} \ \forall S \geq \underline{S}$, where $a > 0$ is the Pareto exponent (or the scaling parameter) and $\underline{S}$ is the population of the city at the truncation point.

We first calculated population thresholds by state and year to define the upper tail city size distribution. Although we had information for all cities without size restrictions, we did not expect Zipf's and Gibrat's laws to hold true for the entire sample of cities[2]; on the contrary, the current mainstream view in the literature is that only the largest cities in the upper tail are Pareto distributed (Ioannides & Skouras, 2013).

We followed the procedure of Clauset et al. (2009) specifically designed to select an optimal truncation point.[3] To select the lower bound, the Pareto exponent is estimated for each possible sample size using the maximum likelihood (ML) estimator, computing the Kolmogorov–Smirnov (KS) statistic for each sample size. The truncation point

[2] We repeated all the analysis using all cities without size restrictions, and Zipf's and Gibrat's laws were rejected for most of the states and time periods. These results are available from the authors upon request.

[3] For a review of the different methods available to define the threshold and their primary properties, see Fazio and Modica (2015).

finally selected (listed in Table 1 for the years 1900, 1950, and 2010) corresponds to the value of the threshold for which the KS statistic is the smallest. These thresholds were then used to define the upper tail samples used throughout our analysis.

3.1. Zipf regression

From the density function $p(S)$, one can obtain the expression $R = A \cdot S^{-a}$, which relates the empirically observed rank R (1 for the largest city, 2 for the second largest, and so on) to the relative city size S. Then, after taking logs, one can add in the correction proposed by Gabaix and Ibragimov (2011) to yield the following equation:

$$ln\left(R - \frac{1}{2}\right) = b - alnS + \varepsilon. \quad (1)$$

Both the standard Zipf regressions and the ML estimator are strongly biased in small samples (Gabaix & Ioannides, 2004; Goldstein et al., 2004). Clauset et al. (2009) argued that small-sample bias can be significant for the ML estimator. Those authors recommended the rule of thumb $N \geq 50$, where N is the number of cities in the upper tail, to obtain reliable parameter estimates. Unfortunately, that sample size is larger than the number of cities in the upper tail in some states (see Table 1), particularly in the first decades of the twentieth century.[4] An additional potential issue is that the ML estimator may be biased when the size distribution of cities does not follow a power law (Soo, 2005).

To correct for the small-sample bias, Gabaix and Ibragimov (2011) proposed subtracting $1/2$ from the rank to obtain an unbiased estimation of the exponent. Their numerical results demonstrate the advantage of this approach over the standard OLS Zipf regressions. In addition, the results of Gabaix and Ibragimov (2011) also suggest that the OLS approaches to tail index estimation are more robust than ML estimator of the tail index under deviations from power laws. That is, Gabaix and Ibragimov (2011)'s Rank-1/2 estimator not only corrects the bias in small samples—it is also more accurate than the ML estimator if the upper tail distribution does not follow a power law.

If $\hat{a} = 1$ then Zipf's law holds, meaning that, ordered from largest to smallest, the size of the second city is half of that of the first, the size of the third is one third city of that of the first, and so on. In any case, a is interpreted as a measure of the degree of inequality in the city size distribution: when a increases (decreases) over time, the distribution becomes more equal (unequal). Standard errors are calculated by applying Gabaix and Ioannides (2004)'s correction: $GI\ s.e. = \hat{a} \cdot (2/N)^{1/2}$, where N is the sample size. We use these corrected standard errors to calculate the confidence intervals of $\hat{a}$ at the 95 % confidence level and to test whether the Pareto exponent is significantly different from 1.

3.2. Power law test

Analysis based on Zipf's regressions can be very useful, but it is also characterised by some limitations. Clauset et al. (2009) pointed out some of those limitations: first, calculations of standard errors are inaccurate; second, the value of the fraction of variance accounted for by the fitted line has very little power as a hypothesis test; and finally, the regression lines are not valid distributions (see Appendix A in Clauset et al. (2009) for details).

As an alternative, and as a complement to Zipf regressions, we implement here the statistical test proposed by Clauset et al. (2009) to test whether or not the data follow a power law. Note that this approach is more general because we test the data for this type of distribution, with Zipf's law being a particular case with a Pareto exponent equal to 1.

The test is based on a measurement of the 'distance' between the empirical distribution of the data and the hypothesised Pareto distribution. This distance is compared with the distance measurements for comparable synthetic data sets drawn from the hypothesised Pareto distribution, and we define the p-value as the fraction of the synthetic distances that are larger than the empirical distance. This semi-parametric bootstrap approach is based on the iterative calculation of the Kolmogorov–Smirnov (KS) statistic for 300 bootstrap data set replications.

The Pareto exponent is estimated for each state and year using the maximum likelihood estimator, and then the KS statistic is computed for the data and the fitted model. The test uses from the observed data and checks how often the resulting synthetic distribution fits the actual data as poorly as the ML-estimated power law. Therefore, the null hypothesis is the power law behaviour of the original sample. Nevertheless, this test has an unusual interpretation because, regardless of the true distribution from which our data were drawn, we can always fit a power law. Clauset et al. (2009) recommend the conservative choice that the power law is ruled out if the p-value is below 0.1: "that is, it is ruled out if there is a probability of 1 in 10 or less that we would merely by chance get data that agree as poorly with the model as the data we have." Therefore, this procedure only allows us to conclude whether a power law is a plausible fit to the data.

As an alternative, we also tested whether the data could be described by a lognormal distribution. A lognormal distribution has, along with the Pareto distribution, been considered in studies of city size for many years. Such a distribution can describe the entire city size distribution (Eeckhout, 2004) or just the upper tail (Eeckhout, 2009; Levy, 2009). The standard test to check the lognormal behaviour of a sample is the KS test, which has been previously applied to city sizes by Giesen et al. (2010) and (González-Val, 2019), among others. The KS test's null hypothesis is that the two samples—the actual data and the fitted lognormal distribution—come from the same distribution.

3.3. A parametric test of Gibrat's law

A first way to test the relationship between growth and initial relative city size is to run the following regression equation (Eeckhout, 2004; Gabaix, 2009; Sutton, 1997):

$$ln(S_{it}) = \alpha + \beta ln(S_{it-1}) + u_{it}. \quad (2)$$

where u_{it} is a random variable representing the random shocks that the growth rate may suffer, which we shall assume are identically and independently distributed for all cities. The parameter of interest here is β, because if the estimated value of β is close to unity ($\hat{\beta} \cong 1$), the growth process is a random walk with drift, which provides statistical evidence that the process obeys Gibrat's law of proportionate growth (Ahundjanov & Akhundjanov, 2019; Ahundjanov et al., 2022; Akhundjanov & Drugova, 2022).

We ran Eq. (2) for each cross-sectional decade in our sample data, and robust standard errors are used to calculate the confidence intervals of $\hat{\beta}$ at the 95 % confidence level and to test whether this coefficient is significantly different from 1. Alternatively, like Ahundjanov and Akhundjanov (2019), we also computed a Wald test of the null hypothesis that $\hat{\beta} = 1$.

Note that Eq. (2) imposes a linear function assumption in the form of conditional expectation of growth rates, which may be viewed as a disadvantage in the analysis of Gibrat's law (Ahundjanov & Akhundjanov, 2019; Akhundjanov & Drugova, 2022). Among others, Ioannides and Overman (2004) have highlighted the advantages of the nonparametric approach over the standard parametric approach. Mainly, nonparametric methods do not impose any structure on underlying relationships that may be nonlinear and may change over time (no need to

[4] There are eight states with fewer than 50 cities in the upper tail distribution at all points in time. Furthermore, the percentage of states with a number of cities below that reference value is equal to or higher than 25 % in all years except for one (i.e., 1930).

restrict the relationship to being stationary); this is especially important when long periods are considered (González-Val, 2023).

3.4. A nonparametric approach to test the validity of Gibrat's law

Like Giesen and Südekum (2011), we conducted a nonparametric analysis of growth rates in the long-term on a state-by-state basis. We computed the growth rates for all of the cities in our upper tail samples; the population growth rate of city i in year t is defined as $(S_{it} - S_{it-1})/S_{it-1}$, with S denoting the city's relative size. Starting with this gross growth rate, we then subtract the mean and divide that quantity by the standard deviation of the growth rates of the upper tail cities for the entire United States that year, to build normalised growth rates. The growth rates need to be normalised because we are considering a pool of all growth rates from 1900 to 2010. Note that this is not a balanced panel: cities enter and exit the upper tail, and the number of cities in the upper tail also changes over time because the population threshold is different each year. Then, to compute the growth rates, we require city i to be part of the upper tail in year t-1 (although it can be out of the tail in year t). We accordingly ensure that the total number of observations in our pool of growth rates from 1900 to 2010 by state is consistent with the sum of the number of cities in the upper tail in the initial year from 1900 to 2000.

To perform the nonparametric analysis, we run the following kernel regression:

$$g_i = m(lnS_i) + \varepsilon_i, \tag{3}$$

where g_i is the normalised growth rate. Instead of making assumptions about the functional relationship m, $\hat{m}(lnS_i)$ is estimated as a local mean around point $ln(S)$ and is smoothed using a kernel, which is a symmetrical, weighted, and continuous function in $ln(S)$. Thus, the nonparametric estimate allows growth to vary with the initial (log) relative population over the entire distribution. Doing so ensured that the estimated nonparametric relationship between growth and size was more accurate than the estimate obtained using standard parametric models (Ioannides & Overman, 2004).

We ran the kernel regression for a pool of 1900–2010 on a state-by-state basis and for the entire United States. To estimate $\hat{m}$, the Nadaraya–Watson method is used, as it appears in Härdle (1990, Chapter 3).[5] Kernel regressions have become popular in the empirical literature in recent decades (Eeckhout, 2004; Giesen & Südekum, 2011; González-Val, 2010, 2023; Ioannides & Overman, 2003), and $\hat{m}$ is typically estimated using the Nadaraya–Watson estimator. Alternative methods, such as the LOcally WEighted Scatter plot Smoothing (LOWESS) algorithm (Cleveland, 1979) based on local polynomial fits, yield similar results (González-Val et al., 2014).

We tested the hypothesis that urban growth is independent of its population's initial size (i.e., the underlying growth model is a multiplicative process), a proposition known as Gibrat's law. Gibrat's law implies that growth is independent of size in both mean and variance. Given that growth rates are normalised, if Gibrat's law was strictly fulfilled and growth was size-independent the estimated kernel of both the mean and variance would be a straight line and any deviation would involve deviations from the mean or variance.

4. Results

4.1. National results

4.1.1. Zipf's law

We ran Eq. (1) for each decennial census from 1900 to 2010 and obtained the Pareto exponent and the corrected standard error to compute the 95 % interval estimation of the exponent. The results are reported in Table 2; one can see that Zipf's law is not rejected for the US upper tail distribution from 1900 to 1950. Nevertheless, in the second half of the twentieth century and 2010, the Pareto exponent is significantly different from 1. The growth of the Pareto exponent after 1920 implies that the size distribution of cities became less unequal over time.

4.1.2. Power law test

The data in Table 2 show that the null hypothesis of a power law cannot be rejected for the United States at the significance level of 10 % for every investigated year except for 1970, using Clauset et al. (2009)'s method. Therefore, the Pareto distribution provides a plausible fit to the data for the upper tail city size distribution in the US. On the other hand, the lognormal distribution is clearly rejected for all years.[6]

4.1.3. Gibrat's law

We ran Eq. (2) for each cross-sectional decade from 1900 to 2010 and obtained the estimated $\hat{\beta}$ coefficient and the robust standard error to compute the 95 % interval estimation of the coefficient. The results are reported in the last columns of Table 2; Gibrat's law is rejected in the short-term in all decades, with the estimated coefficient being close to 0.9 in most periods.

From a long-term perspective, the nonparametric estimates of the mean and the variance of the growth rate depending on the initial relative size of the urban nuclei for the United States on a nationwide basis are shown in Fig. 1. The 95 % confidence intervals are indicated as well. One can see from this figure that Gibrat's law for means holds reasonably well for the small- and medium-sized cities in the long-term: the estimated mean growth is nearly a straight line around the value of zero for log-relative sizes equal to or smaller than zero (i.e., cities with a population equal to or smaller than the contemporary average city size). This finding indicates that growth is independent of initial city size. However, for the largest cities in the upper tail distribution we observed lower-than-average growth (i.e., convergent behaviour), because the 95 % confidence intervals no longer include zero, especially for log-relative sizes >2. Nevertheless, there are many fewer observations in the top sizes than in the other city sizes,[7] as the wider confidence intervals at the top upper tail attest to. Furthermore, the variance of growth is clearly size-dependent: the larger the initial relative size of a city, the lower the variance of growth. That outcome is standard and expected; see González-Val et al. (2014).

4.2. Regional results

4.2.1. Zipf's law

Table 3 summarises the results of the tests carried out on a state-by-state basis. We find strong support for Zipf's law on the state level: Zipf's law cannot be rejected for most states in all decades. The number of rejections is especially low in the first half of the twentieth century (in 1900, Zipf's law cannot be rejected in any state); the evidence against the validity of Zipf's law slightly increases in the latter half of the sample period. The maximum number of rejections, 22.9 % of the states, is attained precisely in 2010.

Fig. 2 shows a map highlighting the 30 states for which Zipf's law cannot be rejected in any year of the sample period 1900–2010. A geographical pattern is also evident in this figure: Zipf's Law cannot be rejected for any of the states in the southern United States (mainly agricultural states); that situation does not hold true for the states in the

[5] We used an Epanechnikov kernel and Silverman's rule of thumb to set the bandwidth.

[6] Results from the Vuong's model selection test clearly support the power law behaviour of the data compared with the lognormal distribution. These results, not shown in this paper, are available from the authors upon request.

[7] Non-reported bivariate kernels confirm the sparsity of data for the top city sizes.

Table 2
Results of the tests for the entire United States.

		Zipf's law			Power law	Lognormality	Gibrat's law		
Year	Upper tail size (cities)	Pareto exponent	Corrected s. e.	95 % interval estimation	Test p-Value	KS p-Value	$\widehat{\beta}$	Robust s. e.	95 % interval estimation
1900	1475	1.04	0.04	(1.11–0.96)	0.52	0.00			
1910	712	1.07	0.06	(1.18–0.96)	0.98	0.00	0.87	0.02	(0.83–0.91)
1920	2078	1.00	0.03	(1.06–0.94)	0.41	0.00	0.92	0.01	(0.90–0.94)
1930	1118	1.04	0.04	(1.13–0.96)	0.61	0.00	0.84	0.02	(0.81–0.88)
1940	1200	1.06	0.04	(1.14–0.97)	0.71	0.00	0.94	0.01	(0.92–0.96)
1950	887	1.10	0.05	(1.20–1.00)	0.66	0.00	0.89	0.01	(0.86–0.91)
1960	625	1.20	0.07	(1.34–1.07)	0.90	0.00	0.74	0.03	(0.67–0.81)
1970	832	1.24	0.06	(1.36–1.12)	0.07	0.00	0.82	0.03	(0.76–0.87)
1980	790	1.30	0.07	(1.43–1.17)	0.38	0.00	0.83	0.02	(0.79–0.88)
1990	550	1.35	0.08	(1.51–1.19)	0.76	0.00	0.89	0.02	(0.86–0.93)
2000	572	1.37	0.08	(1.53–1.21)	0.94	0.00	0.92	0.01	(0.90–0.95)
2010	451	1.41	0.09	(1.60–1.23)	0.73	0.00	0.91	0.01	(0.88–0.93)

Notes: The lower bound of the upper tail was estimated using the methodology of Clauset et al. (2009). The Pareto exponent was estimated using the Rank-1/2 estimator of Gabaix and Ibragimov (2011). Standard errors were calculated by applying the corrected standard errors of Gabaix and Ioannides (2004). The power law test is a goodness-of-fit test; the null hypothesis is that there is power law behaviour for the data in the upper tail. The KS test's null hypothesis is that the data follow a lognormal distribution. The regression results for the parametric testing of Gibrat's law report the estimated $\widehat{\beta}$ coefficient from the cross-sectional estimation of Eq. (2) by decade at year t (t = year in the first column).

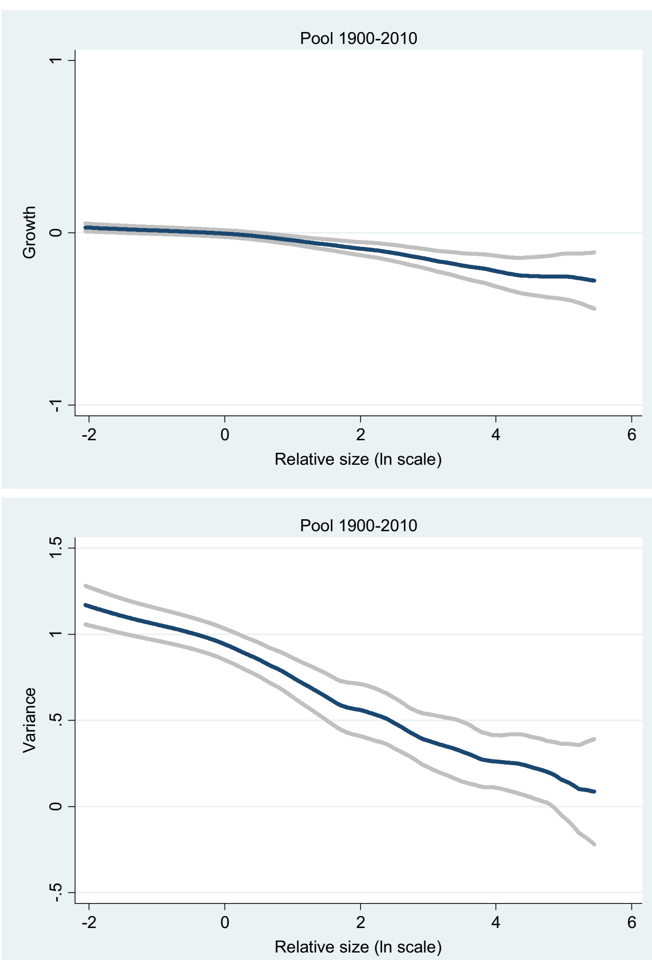

Fig. 1. Nonparametric estimates of the growth rate and its variance for the United States (1900–2010, 10,839 observations).

industrial northern part of the country.

Table 4 presents the estimated Pareto exponent for each year and state. An asterisk indicates cases for which Zipf's law is not fulfilled (i.e., when the null hypothesis $\widehat{a} = 1$ is rejected at the 5 % confidence level, using Gabaix and Ioannides's (2004) corrected standard errors). There is clearly no systematic behaviour: one can find states in which the Pareto exponent clearly decreases in all the periods from 1900 to 2010 (e.g., North Dakota), states in which it increases (e.g., California), states in which it is stable around a value of 1 (e.g., West Virginia) and states in which it exhibits great variation over time (e.g., New Mexico).

4.2.2. Power law

The results in Table 3 reveal that the null hypothesis of a power law is rejected only for a few states at a significance level of 10 %. That is, the data exhibit power law behaviour in the majority of states in all time periods. The percentage of rejections ranges from a minimum of 4.2 % of the states in 1910 to a maximum of 27.1 % in 1970.[8]

Compared with the power law results, the percentages of rejections of the lognormal distribution (also listed in Table 3) are significantly larger. The lognormal distribution is rejected for most states in all years except 1900: the percentage of rejections then is 43.8 %. Therefore, the Pareto distribution fits the data better than a lognormal distribution at both the national and regional levels.

4.2.3. Gibrat's law

The last columns in Table 3 summarise the results of the parametric testing of Gibrat's law in the short-term on a state-by-state basis. We find support for Gibrat's law on the state level because it cannot be rejected for most states in all decades except 1970–1980. The number of rejections is close to one-third of the number of states in most periods, although the evidence against the validity of Gibrat's law slightly increases in the first decades of the second half of the sample period

[8] One might expect that rejection of a power law should imply a rejection of Zipf's Law because Zipf's Law is a specific type of power law. However, a power law is rejected in more instances than Zipf's Law in many years (Table 3). Note that the significance level considered in these two tests is different. If the 5 % level is considered in both tests, the number of rejections of Zipf's Law is equal to or larger than the number of rejections of a power law in all years except for two (i.e., 1900 and 1930). Data supporting these results are available from the authors upon request.

Table 3
Summary of the results of the tests on a state-by-state basis.

Year	Zipf's law rejections (5 %)		Power law rejections (10 %)		Lognormality rejections (5 %)		Gibrat's law rejections (5 %)	
	Number	Percentage	Number	Percentage	Number	Percentage	Number	Percentage
1900	0	0 %	5	10.4 %	21	43.8 %		
1910	1	2.1 %	2	4.2 %	33	68.8 %	15	31.3 %
1920	3	6.3 %	5	10.4 %	30	62.5 %	5	10.4 %
1930	2	4.2 %	7	14.6 %	33	68.8 %	21	43.8 %
1940	5	10.4 %	6	12.5 %	33	68.8 %	18	37.5 %
1950	4	8.3 %	7	14.6 %	35	72.9 %	15	31.3 %
1960	6	12.5 %	6	12.5 %	32	66.7 %	23	47.9 %
1970	9	18.8 %	13	27.1 %	31	64.6 %	24	50.0 %
1980	9	18.8 %	9	18.8 %	30	62.5 %	26	54.2 %
1990	9	18.8 %	12	25.0 %	32	66.7 %	15	31.3 %
2000	9	18.8 %	7	14.6 %	28	58.3 %	15	31.3 %
2010	11	22.9 %	7	14.6 %	31	64.6 %	21	43.8 %

Notes: Number of states (and percentages) with rejections for the different statistical city size distributions and the parametric testing of Gibrat's law.

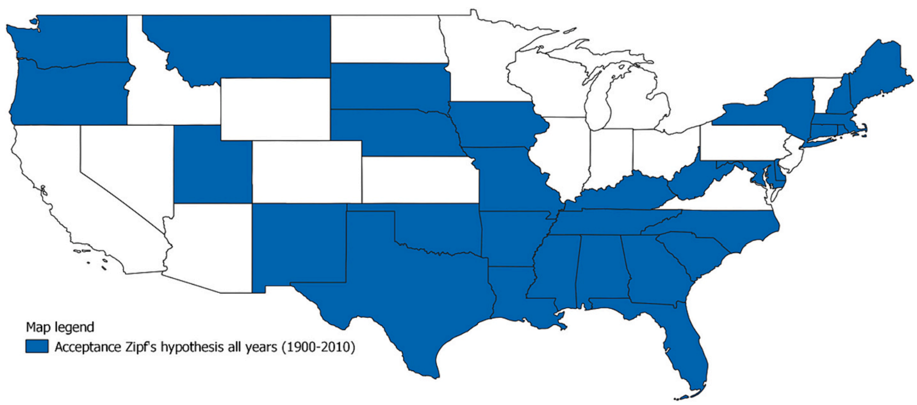

Fig. 2. Zipf's law by state, 1900–2010.
Notes: The 30 states for which Zipf's law cannot be rejected in any year of the sample period 1900–2010 are shown in blue. The 18 states for which Zipf's law is rejected at least once from 1900 to 2010 are shown in white. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(1950–1960, 1960–1970, and 1970–1980).

Table 5 presents the estimated $\widehat{\beta}$ coefficient for each decade and state from the cross-sectional estimation of Eq. (2). An asterisk indicates cases for which Gibrat's law does not hold (i.e., when the null hypothesis $\widehat{\beta} = 1$ is rejected at the 5 % confidence level, using a Wald test). One can find states in which Gibrat's law holds in all the decades from 1900 to 2010 (e.g., Vermont), states in which it is rejected in only one or two decades (e.g., Delaware, New Hampshire, Connecticut, North Carolina, and West Virginia), and states in which it is rejected in most periods (e.g., Minnesota, New Jersey, and Ohio).

If we focus on long-term growth, Figs. 3 and 4 show the nonparametric estimates of the mean and variance of the growth rate for several selected states for a pool of all of the decade-by-decade growth-size pairs from 1900 to 2010; the results for each of the 48 states are reported in Figs. A1 and A2 in Appendix A. Note that the normalised growth rates were obtained by subtracting the mean and dividing by the standard deviation of all US cities in the upper tail of each year, not only the cities corresponding to that state and, therefore, the interpretation is that, if the line is above (below) zero then, on average, the cities in that state have grown with more (less) intensity than the rest of the US cities. The same procedure was adopted by Giesen and Südekum (2011). The mean and variance of growth satisfy Gibrat's law if the estimates are represented by an approximately horizontal line, independent of initial size.

A visual inspection of Figs. 3 and A1 in Appendix A confirms that Gibrat's law for mean growth is reasonably satisfied for 27 states (i.e., Alabama, Arkansas, Colorado, Connecticut, Delaware, Georgia, Indiana, Iowa, Kentucky, Louisiana, Maine, Massachusetts, Michigan, Mississippi, Missouri, Montana, New Hampshire, New Mexico, New York, North Carolina, Pennsylvania, South Carolina, Vermont, Virginia, West Virginia, Wisconsin, and Wyoming).[9] The map shown in Fig. 5 highlights these states for which Gibrat's law holds for most city sizes. The figure reveals a clear geographical pattern: most of the states are located in the eastern part of the United States, in the Northeast, Southeast and Midwest regions. On the other hand, Gibrat's law is rejected for all states along the West Coast and surrounding areas. Furthermore, there is an isolated column of states in which Gibrat's law for mean growth is fulfilled: Montana, Wyoming, Colorado, and New Mexico.

In many cases the nonparametric estimate of mean growth is a flat line for most city sizes, but deviations can be observed for the largest cities of the upper tail. The results for Kansas, Minnesota, New Jersey, and South Dakota are ambiguous, and in the remaining states Gibrat's law for means clearly does not hold. In those cases, the estimated mean growth exhibits a convergent pattern (i.e., growth decreases with initial size). That situation holds true in some states such as Florida. However, in other states (e.g., Idaho or Nevada) a divergent pattern emerges, with the largest cities growing more than the smaller cities; see Fig. A1 in Appendix A.

[9] In many of these states the parametric testing of Gibrat's law in the short-term revealed that Gibrat's law holds in most decades.

Table 4
Pareto exponents estimated by state and decade.

State	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
Alabama	1.04	1.04	1.02	1.00	1.00	1.04	0.98	1.06	1.11	1.16	0.85	1.25
Arizona	0.99	1.05	1.33	1.14	0.89	1.06	0.80	0.74	0.76	**0.71***	**0.66***	0.91
Arkansas	1.22	1.40	1.37	1.08	1.07	0.94	0.88	0.88	0.88	0.89	0.95	0.93
California	0.88	0.90	0.90	0.90	0.92	1.01	1.19	1.33	**1.38***	**1.44***	1.44	**1.49***
Colorado	0.91	0.96	0.96	0.90	0.92	0.88	0.95	**0.68***	**0.69***	**0.68***	0.80	0.79
Connecticut	1.16	1.05	1.05	1.15	1.16	1.10	1.40	1.52	1.51	1.55	11.06	1.57
Delaware	0.78	0.81	0.85	0.85	0.80	0.88	0.91	0.86	0.85	0.84	0.83	0.85
Florida	0.97	1.00	0.98	0.98	0.95	0.96	1.05	1.13	1.02	1.50	1.62	1.59
Georgia	1.13	1.16	1.12	0.96	1.06	1.08	1.06	1.06	1.11	1.09	1.17	1.27
Idaho	1.17	1.30	1.00	0.98	0.94	0.90	1.05	0.83	1.02	0.98	**0.74***	**0.69***
Illinois	1.08	1.01	0.99	0.90	**0.88***	**0.87***	1.20	**0.74***	1.23	1.36	1.39	**1.51***
Indiana	0.97	0.92	**0.85***	**0.82***	**0.83***	**0.80***	**0.81***	**0.80***	**0.79***	1.26	**0.81***	**0.77***
Iowa	1.15	1.10	1.07	1.02	1.01	0.97	0.94	0.93	0.95	0.93	0.92	0.90
Kansas	1.11	1.05	1.01	0.97	0.94	0.90	1.00	0.94	0.92	0.89	0.87	**0.83***
Kentucky	1.05	1.10	1.31	1.01	1.06	1.29	1.27	1.25	1.07	1.41	1.53	1.35
Louisiana	0.92	0.98	0.85	0.97	1.00	0.96	0.96	0.98	0.96	0.97	0.97	1.00
Maine	1.48	1.36	1.61	1.63	1.70	1.74	1.38	2.07	2.13	2.14	2.07	2.10
Maryland	0.82	0.82	0.83	0.84	0.80	0.84	0.82	0.80	0.84	0.87	0.88	0.96
Massachusetts	1.29	1.35	1.37	1.41	1.40	1.39	1.37	1.47	1.59	1.60	1.65	1.65
Michigan	1.00	0.92	**0.85***	0.92	**0.84***	**0.81***	**0.80***	**0.79***	**0.80***	0.93	1.48	1.07
Minnesota	1.01	1.05	1.06	1.01	1.03	1.01	0.96	0.91	1.43	**0.74***	0.89	**0.69***
Mississippi	1.71	1.06	1.06	1.03	1.08	1.01	0.98	1.00	1.02	0.97	1.68	0.85
Missouri	0.98	0.93	0.91	0.96	0.93	0.99	1.06	1.06	1.13	1.06	1.17	1.21
Montana	0.94	0.81	0.97	0.91	1.03	1.02	0.96	0.92	0.89	0.86	0.89	0.85
Nebraska	1.15	1.21	1.17	1.14	1.09	1.05	0.99	0.95	0.94	0.91	0.88	0.86
Nevada	1.44	1.04	1.25	1.16	1.18	0.93	0.76	**0.58***	1.18	**0.51***	**0.49***	**0.43***
New Hampshire	1.40	1.36	1.37	1.46	1.50	1.57	1.60	1.51	1.45	1.41	1.28	1.11
New Jersey	0.86	0.86	0.91	1.00	1.03	1.10	1.23	**1.32***	**1.35***	**1.38***	**1.37***	1.35
New Mexico	2.47	2.19	1.98	0.89	1.67	0.85	0.78	0.74	1.06	1.08	1.07	0.93
New York	0.87	0.86	0.79	0.87	0.88	0.89	0.93	0.94	0.99	1.01	1.02	1.02
North Carolina	1.16	1.08	1.08	0.99	0.94	0.96	0.94	0.93	0.94	0.92	0.91	1.05
North Dakota	1.39	1.26	**1.26***	1.21	1.16	1.08	0.95	0.91	0.94	0.90	0.87	0.83
Ohio	0.91	**0.86***	0.94	0.92	**0.79***	**0.78***	1.08	1.14	1.27	1.29	1.35	**1.43***
Oklahoma	1.27	1.31	1.27	1.11	0.90	1.05	1.08	0.94	0.98	0.94	0.93	0.91
Oregon	0.98	0.93	0.94	0.93	1.02	0.90	0.85	0.84	1.19	1.13	1.11	1.13
Pennsylvania	1.06	1.12	1.13	1.15	1.16	1.18	**1.22***	1.24	**1.27***	**1.28***	**1.28***	**1.28***
Rhode Island	0.96	0.99	1.03	1.06	1.11	1.25	1.61	1.95	2.24	2.24	2.07	1.95
South Carolina	1.02	1.18	1.20	1.14	1.14	1.14	1.22	1.31	1.36	1.32	1.00	0.92
South Dakota	1.26	1.31	1.32	1.20	1.14	1.05	0.99	0.98	0.98	0.92	0.93	0.89
Tennessee	0.94	0.95	0.90	0.88	0.90	0.95	0.99	0.96	0.98	0.99	1.00	0.95
Texas	1.25	1.11	1.04	1.04	1.05	1.03	0.99	0.99	1.00	1.00	1.03	0.95
Utah	1.18	1.10	1.14	1.07	1.08	1.01	0.92	0.89	1.21	1.00	0.97	1.79
Vermont	1.00	1.57	0.99	0.99	0.97	0.98	0.95	0.89	0.93	1.51	**0.69***	0.81
Virginia	0.80	0.89	0.85	0.86	0.87	0.82	**0.75***	**0.69***	**0.70***	**0.66***	**0.66***	**0.67***
Washington	0.86	0.87	0.87	0.82	0.84	0.82	0.84	1.11	1.21	0.90	1.38	1.43
West Virginia	1.02	1.02	1.05	0.99	1.00	1.00	0.98	1.02	1.08	1.08	1.07	1.07
Wisconsin	1.04	0.93	0.92	**0.86***	0.87	0.88	**0.84***	**0.81***	**0.85***	**0.83***	1.21	1.23
Wyoming	0.89	0.94	1.22	0.78	**0.75***	1.18	**0.69***	**0.67***	**0.70***	0.73	**0.66***	**0.71***

Notes: The lower bound of the upper tail for each state was estimated using the methodology of Clauset et al. (2009). The Pareto exponent was estimated using the Rank-1/2 estimator of Gabaix and Ibragimov (2011). * The null hypothesis $\hat{a} = 1$ can be rejected at the 5 % confidence level, using the corrected standard errors of Gabaix and Ioannides (2004).

Results pertaining to the variance of growth rates are easier to interpret: Gibrat's law simply does not hold.[10] Figs. 4 and A2 in Appendix A exhibit a decreasing relationship between the variance of growth and the initial relative size in most cases; larger variance corresponds, as might be expected, to the smaller cities.

5. Discussion and conclusions

We have studied the US city size distribution at the national and regional levels; we considered 48 states from 1900 to 2010. Empirical studies of this topic at a regional level are quite scarce, and for the United States they are almost non-existent. We focused on upper tail cities, defined endogenously according to the methodology of Clauset et al., 2009. We used Zipf regressions to conclude that the support for Zipf's law decreases over time. However, Zipf's law is satisfied for a majority of states in all years (the maximum number of rejections is 22.9 % in 2010).

We have employed two complementary techniques to test whether or not individual American states follow a power law in the upper tail (the methodology of Clauset et al. (2009)) or comply with Zipf's law in the same upper tail (Zipf's regressions with corrected standard errors). Cristelli et al. (2012, p. 7) argued that, for Zipf's law to hold, the urban system must be integrated and the sample must be coherent in the sense of being the "result of some kind of optimization in growth processes or of an optimal self-organization mechanism". The same authors also claimed, although without providing any empirical support, that "the size of US cities compose a near Zipfian set, in contrast to the sets composed of the cities from a single state such as California, New York

[10] It might be argued that a weak fulfilment of the law can be observed for Connecticut, Delaware, Idaho, Kentucky, Louisiana, Maine, Massachusetts, Mississippi, Nevada, New Hampshire, Tennessee, and Virginia; see Figure A2 in Appendix A.

Table 5
Regression results for the parametric testing of Gibrat's law.

State	1900–1910	1910–1920	1920–1930	1930–1940	1940–1950	1950–1960	1960–1970	1970–1980	1980–1990	1990–2000	2000–2010
Alabama	1.07	0.96	1.01	0.99	0.96	1.01	0.89*	0.92*	0.95*	1.00	0.87*
Arizona	0.91	0.89	1.10	0.98	0.85	1.25*	1.06	0.99	1.03	1.02	0.71*
Arkansas	1.04	0.95	1.10*	1.00	1.09*	1.06*	0.99	0.99	1.01	0.96*	0.99
California	1.06	1.04	0.85*	0.95	0.98	0.80*	0.85*	0.97	0.95*	0.97*	0.99
Colorado	0.77*	1.00	1.05*	0.97*	0.99	0.83	1.08*	0.96*	1.00	0.92*	0.90*
Connecticut	1.03	1.06	0.90*	0.98	1.01	0.84*	0.82	0.94	0.97	0.42	0.98
Delaware	1.00	0.98	0.97	0.90	0.95*	0.99	1.05	0.99	0.99	0.98	0.97
Florida	0.96	0.99	0.83*	1.01	0.96	0.86*	0.87*	0.98	0.75*	0.87*	0.99
Georgia	1.00	1.01	1.02	0.98	0.96	0.97	0.99	0.95*	0.97	0.83*	0.79*
Idaho	0.73	1.08	1.00	1.01	1.05*	0.91	1.05*	0.91*	1.01	1.04	1.06*
Illinois	1.05*	1.01	1.05*	0.99	0.99	0.66*	1.03*	0.78	0.91*	1.01	1.02
Indiana	1.04*	1.07*	1.03*	0.99	1.01	0.99	1.00	0.99*	0.94	0.99	1.02
Iowa	1.04*	1.01	1.05*	1.01*	1.03*	1.02*	1.00	0.98*	1.01*	1.01	1.01
Kansas	1.04	1.02	1.03*	1.02*	1.03*	0.88*	1.01	1.00	1.04*	1.02	1.03*
Kentucky	0.95*	0.94	1.04	0.96*	0.90*	0.87*	0.97	1.01	0.92	0.94	0.97
Louisiana	0.97	1.00	0.96	0.94*	0.98	0.99	0.94*	1.00	0.98	1.00	0.95*
Maine	1.06*	0.93	0.86	0.96*	0.88	0.96	0.91*	0.97*	1.00	1.03	0.98
Maryland	0.97	1.02	1.01	0.99	0.92*	0.99	0.96	0.94*	0.95*	0.94*	0.95*
Massachusetts	0.93	0.97	1.03	1.00	1.00	0.94*	0.92*	0.95*	0.99	0.98	1.00
Michigan	1.06*	1.03	0.89	0.98*	1.00	0.98	0.98*	0.95*	0.97*	0.97	0.97*
Minnesota	1.02	1.01	1.03*	0.97*	1.01	1.02	1.03*	0.77*	1.05*	0.96*	1.01*
Mississippi	1.05	0.97	1.01	0.96	1.01	1.02*	0.96*	0.96*	0.99	0.69*	1.01
Missouri	0.93	0.99	1.03	0.95*	0.97	0.92*	0.87*	0.95*	0.98	0.91*	0.92*
Montana	0.85	0.84*	1.02	0.91*	1.00	1.02	1.04*	1.01	1.03*	1.00	1.04*
Nebraska	1.00	1.02	1.05*	1.00	1.02	1.06*	1.04*	1.00	1.03*	1.03*	1.02*
Nevada	1.95*	1.08	0.96	0.90	1.26	1.13	1.17*	0.82	1.03	1.06	1.06*
New Hampshire	1.01	0.98	0.93*	0.98	0.97	0.97	1.03	1.00	1.01	1.07	1.00
New Jersey	1.01	0.97*	0.91*	0.96*	0.93*	0.88*	0.93*	0.98*	0.97*	1.00	1.01
New Mexico	1.19	0.88	1.17*	0.32*	1.14*	1.03	1.03	1.06	1.04	1.01	1.05
New York	1.03*	1.00	0.98	0.99*	0.99*	0.96*	0.97*	0.97*	0.99	0.98	1.00
North Carolina	0.99	0.98	1.06*	1.01	0.99	1.02	0.98	0.98	1.02	1.00	0.88*
North Dakota	1.09	1.02	1.02	1.01	1.05*	1.09*	1.05*	0.97	1.03	1.02*	1.03*
Ohio	1.06*	1.00	0.96	0.99*	1.00	0.81*	0.84*	0.93*	0.97*	0.95*	0.95*
Oklahoma	0.88	0.98	1.05	1.04	0.99	0.93	1.02	0.95*	1.01	1.00	1.01
Oregon	1.04	0.96	0.94	0.93*	0.97	1.03	1.00	0.90*	1.02	1.02	0.97
Pennsylvania	0.95*	0.98	0.97*	1.01	0.99	0.96*	0.95*	0.98*	0.99	0.99	1.01
Rhode Island	0.96	0.95	0.98	0.95	0.86	0.75*	0.76	0.88	0.99	1.08	1.05
South Carolina	0.81*	1.05	1.04	0.99	0.99	0.88*	0.91*	0.92*	1.02	1.04	1.06*
South Dakota	0.96	0.97	1.09*	1.04*	1.07*	1.05*	1.01	0.98	1.04*	1.01	1.02
Tennessee	1.01	1.01	1.06*	0.97*	0.94*	0.97	1.03	0.93	0.95*	1.01	1.01
Texas	1.16*	1.09*	0.85*	0.95*	1.04*	1.01	0.97	0.99	0.99	0.94*	1.02
Utah	1.07	0.95	1.04*	0.99	1.03	1.03	1.01	0.72*	0.96	0.96	0.77*
Vermont	0.99	0.99	1.00	1.02	0.97	1.01	1.03	0.94	0.97	1.01	1.00
Virginia	0.94	1.04	0.96	0.98	1.02	1.03	1.02	0.97*	1.01	0.99	0.98
Washington	1.13*	1.00	1.03	0.99	1.00	0.95*	0.78*	0.88*	0.96	0.92*	0.86*
West Virginia	0.97	1.00	1.03	1.00	1.01	1.00	0.96*	0.92*	0.99	1.01	1.00
Wisconsin	1.03	1.01	1.05*	0.99	0.98*	1.02*	1.01	0.95*	1.00	0.95*	0.96*
Wyoming	0.67*	0.68*	1.10	1.01	0.99	1.01	1.03	0.93*	1.00	1.00	0.98

Notes: Each row represents the estimated $\widehat{\beta}$ coefficient for each state from the cross-sectional estimation of Eq. (2) using data for the indicated decade. * The null hypothesis $\widehat{\beta} = 1$ can be rejected at the 5 % confidence level using a Wald test.

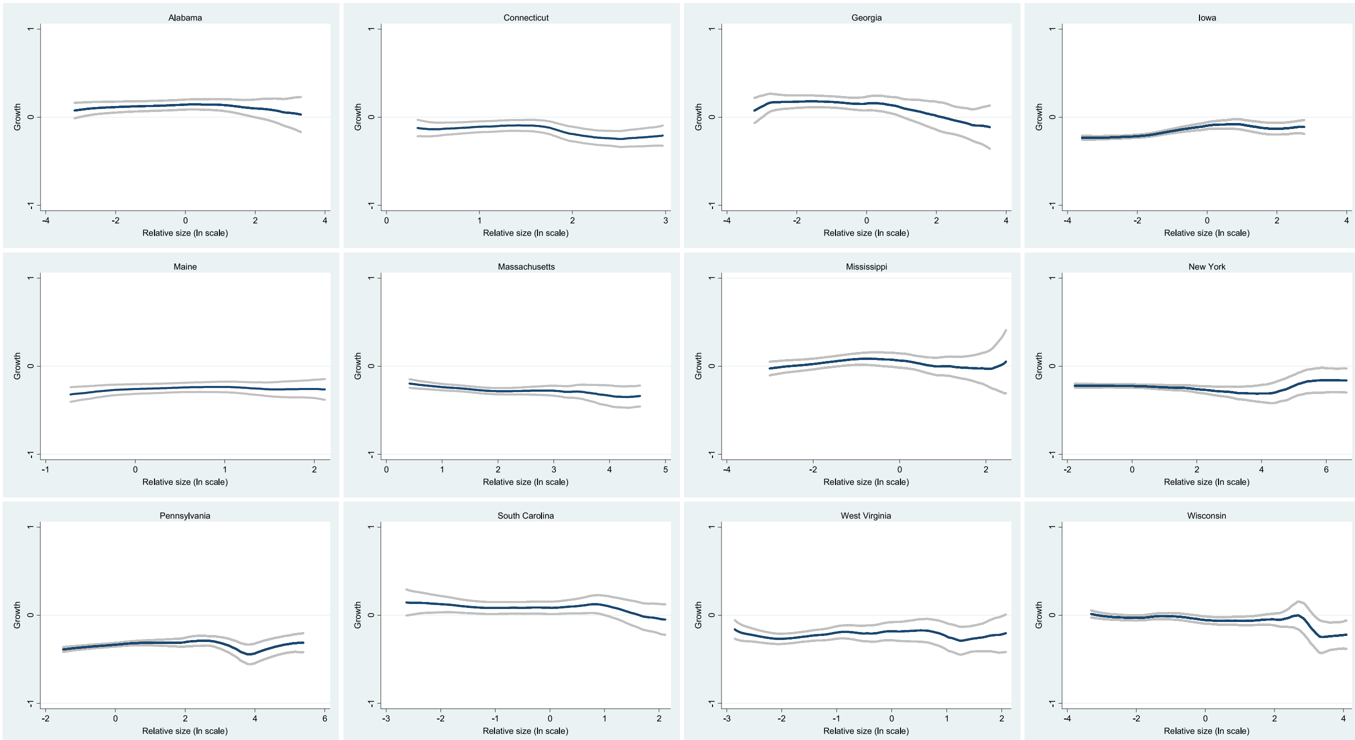

Fig. 3. Nonparametric estimate of the mean growth rate for several selected states, 1900–2010.

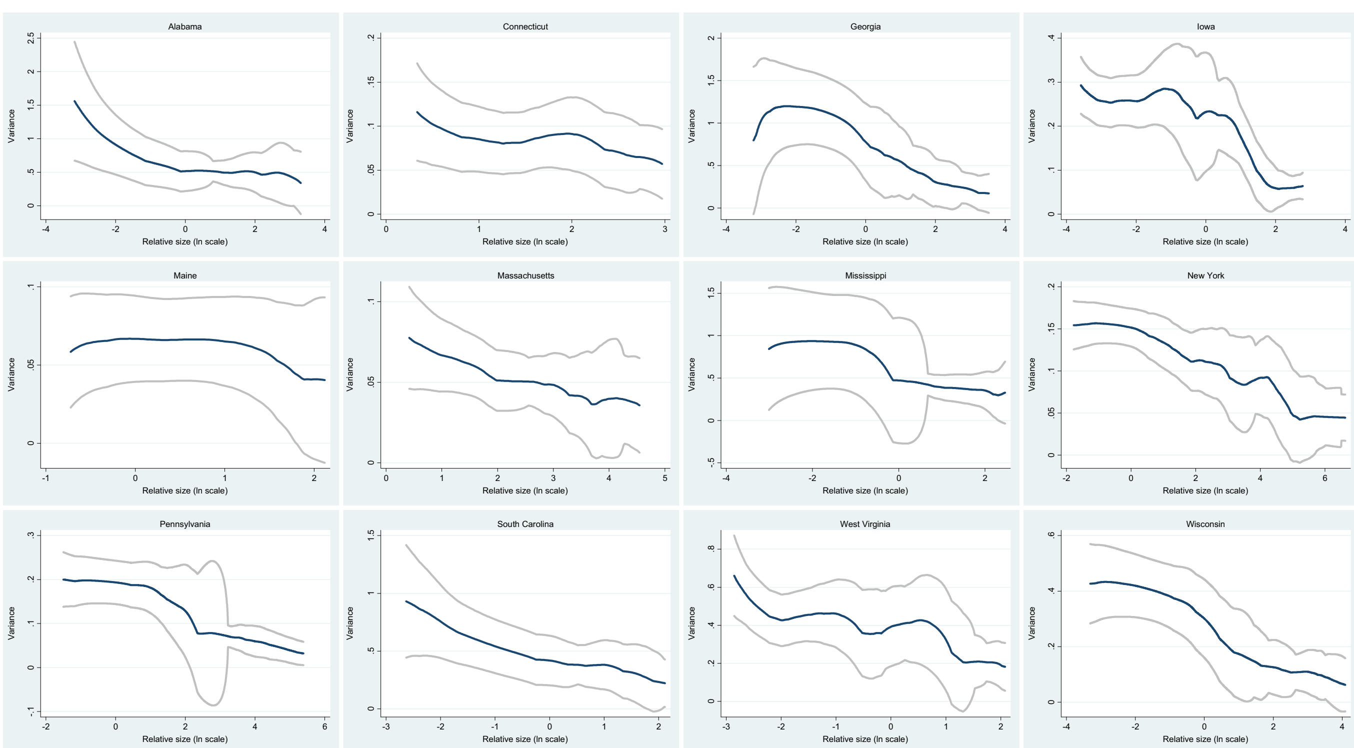

Fig. 4. Nonparametric estimate of the variance of the growth for several selected states, 1900–2010.

State, Illinois, Massachusetts. These cannot be represented by a Zipf's Law". However, our results, based on Zipf's regressions, contradict the hypothesis of Cristelli et al. (2012): the sample that consists of the United States as a whole is not more integrated and organic than the sample of its individual states given that Zipf's law is rejected at the national level from 1960 to 2010.

We also considered the results obtained using the power law test by Clauset et al. (2009), which is considered to be an improvement over Zipf's regressions. Using that methodology, we found weak support for the hypothesis of Cristelli et al. (2012). The power law null is only

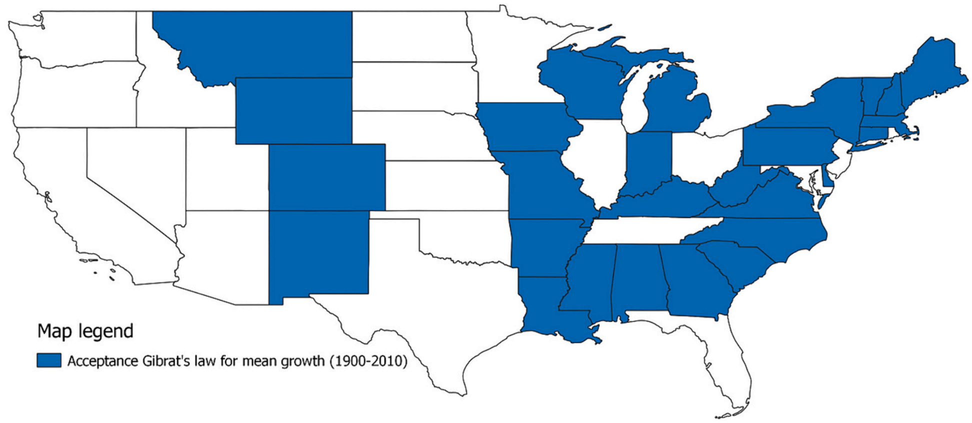

Fig. 5. Gibrat's law for mean growth by state, 1900–2010.
Notes: The 27 states for which Gibrat's law holds for most city sizes over the sample period 1900–2010 are shown in blue. The estimated relationship between growth and size for each of the 48 states considered is reported in Appendix A. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

rejected at 10 % at the national level in one year (1970) and, at the regional level, it is not rejected for most states in all years (although we detect several rejections—a maximum of 27.1 % in 1970 and a minimum of 4.2 % in 1910). Therefore, although we might reject the proposition that Zipf's law holds from 1960 onwards for the entire United States using Zipf's regressions, the methodology of Clauset et al. (2009) at least allows us to conclude that the city size distribution follows a Pareto distribution. Overall, our results support a power law, but they reject the special case of Zipf's law at the national level since 1960.

We have demonstrated that Gibrat's law holds for most city sizes in the long-term at the national level, and for a non-negligible number of states in mean growth both in the short- and long-term. In the short-term, a parametric approach revealed that Gibrat's law cannot be rejected for most states in any decade except one (i.e., 1970–1980). In the long-term, nonparametric estimates showed that there are 20 states in which both Zipf's and Gibrat's laws in means are fulfilled over the entirety of the sample period: Alabama, Arkansas, Connecticut, Delaware, Georgia, Iowa, Kentucky, Louisiana, Maine, Massachusetts, Mississippi, Missouri, Montana, New Hampshire, New Mexico, New York, North Carolina, South Carolina, Vermont (Zipf's law is only rejected in 2000), and West Virginia. That finding partly corroborates the second proposition of Gabaix (1999). On the other hand, Gibrat's law for the variance of growth rates is not satisfied at either the level of the entire United States or the state level; the variance of the growth rates is systematically larger for smaller cities. This result, however, is not unprecedented: González-Val (2010) found that Gibrat's Law held weakly for all US cities from 1900 to 2000. Under this weak version of Gibrat's law, growth is proportionate on average but not in variance. We found that this same pattern of urban growth was reproduced at the state level for these 20 states.

The implications of our state-level empirical findings for regional science and urban economics have already been stated in the previous paragraphs, but we can delve deeper into the analysis. There is a certain consensus in the literature regarding the validity of a power law, and the particular case that Zipf's law represents, for adequately describing the upper tail of city size distributions. Furthermore, the fulfilment of Zipf's law is often associated with a regular and stable distribution in which all city sizes are reasonably represented. In this context, our results corroborate these stylised facts since, at the US state level, both a power law and Zipf's law are the dominant descriptions for data. If Gibrat's law holds in means, the implication is that the urban hierarchy of cities is stable over time. Here, the evidence is not so clear, because for a non-negligible number of states, Gibrat's law in means is not valid in the long-term. Moreover, as expected, Gibrat's law in variances simply does not hold.

Finally, our empirical outcomes have policy implications. First, space matters: we found two clear geographical patterns regarding the validity of Zipf's law (mainly in the southern states) and Gibrat's law in means (mainly in the states located in the east). This geographical pattern characterising the distribution of city sizes in some states must be taken into account in any policy measure attempting to exert some influence on city size distribution, such as policies that tend towards convergence and strive for territorial cohesion. Second, in the states where Gibrat's law in means is not satisfied, the evolution of the distribution is convergent in most cases, which implies that the differences in city sizes are diminishing and, therefore, the distribution is becoming less unequal. An urban structure of cities of similar populations invites an egalitarian treatment by the public bodies in charge of investment in transport infrastructure, education, or healthcare.

CRediT authorship contribution statement

Rafael González-Val: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Domingo P. Ximénez-de-Embún:** Visualization, Investigation, Formal analysis, Data curation, Conceptualization. **Fernando Sanz-Gracia:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Financial support was provided by the Spanish Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación, MCIN/AEI/10.13039/501100011033 (projects PID2020-114354RA-I00 and PID2020-112773GB-I00), the DGA (project S39_20R, ADETRE research group) and European Regional Development Fund (ERDF).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cities.2024.104946.

References

Ahundjanov, B. B., & Akhundjanov, S. B. (2019). Gibrat's law for CO2 emissions. *Physica A: Statistical Mechanics and its Applications, 526*, Article 120944.

Ahundjanov, B. B., Akhundjanov, S. B., & Okhunjanov, B. B. (2022). Power law in COVID-19 cases in China. *Journal of the Royal Statistical Society: Series A, 185*(2), 699–719.

Akhundjanov, S. B., & Drugova, T. (2022). On the growth process of US agricultural land. *Empirical Economics, 63*, 1727–1740.

Arribas-Bel, D., Garcia-López, M.-À., & Viladecans-Marsal, E. (2021). Building(s and) cities: Delineating urban areas with a machine learning algorithm. *Journal of Urban Economics, 125*, Article 103217.

Arshad, S., Hu, S., & Ashraf, B. N. (2018). Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A, 492*, 75–92.

Arshad, S., Hu, S., & Ashraf, B. N. (2019). Zipf's law, the coherence of the urban system and city size distribution: Evidence from Pakistan. *Physica A, 513*, 87–103.

Bee, M., Riccaboni, M., & Schiavo, S. (2011). Pareto Versus Lognormal: a Maximum Entropy Test. *Physical Review E, 84*(2), Article 026104.

Bee, M., Riccaboni, M., & Schiavo, S. (2013). The size distribution of U.S. cities: Not Pareto, even in the tail. *Economics Letters, 120*, 233–237.

Ch, R., Martin, D. A., & Vargas, J. F. (2021). Measuring the size and growth of cities using nighttime light. *Journal of Urban Economics, 125*, Article 103254.

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review, 51*, 661–703.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association, 74*, 829–836.

Cottineau, C. (2017). MetaZipf. A dynamic meta-analysis of city size distributions. *PLoS One, 8*, 1–22.

Cristelli, M., Batty, M., & Pietronero, L. (2012). There is more than a power law in Zipf. *Scientific Reports, 2*, 1–7.

Cuberes, D., Desmet, K., & Rappaport, J. (2021). Urban growth shadows. *Journal of Urban Economics, 123*, Article 103334.

de Bellefon, M.-P., Combes, P.-P., Duranton, G., Gobillon, L., & Gorin, C. (2021). Delineating urban areas using building density. *Journal of Urban Economics, 125*, Article 103226.

Devadoss, S., & Luckstead, J. (2015). Growth process of U.S. small cities. *Economics Letters, 135*, 12–14.

Devadoss, S., Luckstead, J., Danforth, D., & Akhundjanov, S. (2016). The power law distribution for lower tail cities in India. *Physica A: Statistical Mechanics and its Applications, 442*, 193–196.

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

Eeckhout, J. (2009). Gibrat's law for (all) cities: Reply. *American Economic Review, 99*(4), 1676–1683.

Fazio, G., & Modica, M. (2015). Pareto or log-normal? Best fit and truncation in the distribution of all cities. *Journal of Regional Science, 55*(5), 736–756.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics, 114*, 739–767.

Gabaix, X. (2009). Power laws in economics and finance. *Annual Review of Economics, 1*, 255–294.

Gabaix, X., & Ibragimov, R. (2011). Rank-1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business Economics and Statistics, 29*, 24–39.

Gabaix, X., & Ioannides, Y. M. (2004). The evolution of city size distributions. In V. John, J. V. Henderson, & J. F. Thisse (Eds.)*, 4. Handbook of urban and regional economics* (pp. 2341–2378). North-Holland: Amsterdam: Elsevier Science.

Galdo, V., Li, Y., & Rama, M. (2021). Identifying urban areas by combining human judgment and machine learning: An application to India. *Journal of Urban Economics, 125*, Article 103229.

Garmestani, A. S., Allen, C. R., & Gallagher, C. M. (2008). Power laws, discontinuities and regional city size distributions. *Journal of Economic Behavior and Organization, 68*, 209–216.

Garmestani, A. S., Allen, C. R., Gallagher, C. M., & Mittelstaedt, J. (2007). Departures from Gibrat's law, discontinuities and city size distribution. *Urban Studies, 44*, 1997–2007.

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

Giesen, K., & Suedekum, J. (2014). City age and City size. *European Economic Review, 71*, 193–208.

Giesen, K., Zimmermann, A., & Suedekum, J. (2010). The size distribution across all cities-double Pareto lognormal strikes. *Journal of Urban Economics, 68*, 129–137.

Goldstein, M. L., Morris, S. A., & Yen, G. G. (2004). Problems with fitting to the power-law distribution. *The European Physical Journal B - Condensed Matter, 41*(2), 255–258.

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

González-Val, R. (2019). Lognormal city size distribution and distance. *Economics Letters, 181*, 7–10.

González-Val, R. (2023). Parametric, semiparametric and nonparametric models of urban growth. *Cities, 132*, Article 104079.

González-Val, R., Lanaspa, L., & Sanz-Gracia, F. (2014). New evidence on Gibrat's law for cities. *Urban Studies, 51*(1), 93–115.

Härdle, W. (1990). *"Applied nonparametric regression". Econometric society monographs*. Cambridge, New York and Melbourne: Cambridge University Press.

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

Ioannides, Y. M., & Overman, H. G. (2004). Spatial evolution of the US urban system. *Journal of Economic Geography, 4*(2), 131–156.

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

Krugman, P. (1993), "On the number and location of cities", European Economic Review 37(2–3), 293–298.

Kumar, G., & Subbarayan, A. (2014). The temporal dynamics of regional city size distribution: Andhra Pradesh (1951-2001). *Journal of Mathematics and Statistics, 10*, 221–230.

Kundak, S., & Dökmeci, V. (2018). A rank-size rule analysis of the city system at the country and province level in Turkey. *International Journal of Architecture and Planning, 6*, 77–98.

Levy, M. (2009). Gibrat's law for (all) cities: A comment. *American Economic Review, 99* (4), 1672–1675.

Li, W., & Zhang, Y. (2018). Triangle law or power law? *International Review for Spatial Planning and Sustainable Development A: Planning Strategies and Design Concepts, 6*, 203–215.

Luckstead, J., & Devadoss, S. (2017). Pareto tails and lognormal body of US cities size distribution. *Physica A: Statistical Mechanics and its Applications, 465*, 573–578.

Malevergne, Y., Pisarenko, V., & Sornette, D. (2011). Testing the Pareto against the lognormal distributions with the uniformly Most powerful unbiased test applied to the distribution of cities. *Physical Review E, 83*(3), Article 036111.

Moreno-Monroy, A. I., Schiavina, M., & Veneri, P. (2021). Metropolitan areas in the world. Delineation and population trends. *Journal of Urban Economics, 125*, Article 103242.

Mori, T., Smith, T. E., & Hsu, W.-T. (2020). Common power laws for cities and spatial fractal structures. *PNAS, 117*(12), 6469–6475.

Nitsch, V. (2005). Zipf zipped. *Journal of Urban Economics, 57*, 86–100.

Pérez-Valbuena, G. J., & Meisel-Roca, A. (2014). Ley de Zipf y de Gibrat para Colombia y sus regiones: 1835-2005. *Revista de Historia Económica, 32*, 247–286.

Rauch, F. (2014). Cities as spatial clusters. *Journal of Economic Geography, 14*(4), 759–773.

Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics Letters, 74*, 15–19.

Reed, W. J. (2002). On the rank-size distribution for human settlements. *Journal of Regional Science, 42*(1), 1–17.

Rozenfeld, H. D., Rybski, D., Gabaix, X., & Makse, H. A. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review, 101*(5), 2205–2225.

Schmidheiny, K., & Südekum, J. (2015). The pan-European population distribution across consistently defined functional urban areas. *Economics Letters, 133*, 10–13.

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

Subbarayan, A. (2009). The size distribution of cities in Tamilnadu (1901-2001). *International Journal of Agricultural and Statistical Sciences, 5*, 373–382.

Sutton, J. (1997). Gibrat's legacy. *Journal of Economic Literature, 35*(1), 40–59.

Ye, X., & Xie, Y. (2012). Re-examination of Zipf's law and urban dynamic in China: A regional approach. *The Annals of Regional Science, 49*, 135–156.

Ziqin, W. (2016). Zipf law analysis of urban scale in China. *Asian Journal of Social Science Studies, 1*, 53–58.