

Hackmann, Angelina; Klarl, Torben

Article — Published Version

The evolution of Zipf's Law for U.S. cities

Papers in Regional Science

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Hackmann, Angelina; Klarl, Torben (2020) : The evolution of Zipf's Law for U.S. cities, Papers in Regional Science, ISSN 1435-5957, Wiley, Hoboken, NJ, Vol. 99, Iss. 3, pp. 841-852, <https://doi.org/10.1111/pirs.12498>

This Version is available at:

<https://hdl.handle.net/10419/230101>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

FULL ARTICLE

The evolution of Zipf's Law for U.S. cities

Angelina Hackmann^{1,2}  | Torben Klarl^{1,3} 

¹University of Bremen, Faculty of Business Studies and Economics, Germany

²Junior Research Fellow at Institute for Development Strategies, Indiana University, Bloomington, USA

³Research Fellow at O'Neill School of Public and Environmental Affairs, Indiana University, Bloomington, USA

Correspondence

Torben Klarl, University of Bremen, Faculty of Business Studies and Economics,
Max-von-Laue-Straße 1, D-28359 Bremen,
Germany.

Email: tklarl@uni-bremen.de

Abstract

Exploiting the hierarchical structure of cities and based on a dataset for U.S. cities between 1840 and 2016, the aim of this paper is to analyze the evolution of the U.S. city size distribution. For that purpose we estimate a general three-parameter Zipf model, which can be traced back to Mandelbrot (1982), and validate our results by means of the hierarchical scaling law. Especially in the second half of the twentieth century, we find a pronounced departure from the exact Zipf's law. The city size distribution has become more equally distributed over time. Besides, the applied estimation method reveals evidence for leading cities dominating the remaining largest cities. Thus, the growing equality of the city sizes can be explained rather by growing smaller cities than by a loss of importance of the largest ones.

KEYWORDS

city size distributions, Zipf's law, hierarchical scaling law, urban systems

JEL CLASSIFICATION

R11; R12; R15

1 | INTRODUCTION

Gabaix (1999) proposed that “[...] city size processes must have the time to converge to Zipf's law”. Accordingly, city size processes can be described as an evolutionary process where different states of urbanization require different

Preceding versions of this paper greatly benefit from comments by various participants to several workshops and conferences. In particular, we are grateful to the participants of the recent 8th ifo Dresden Workshop on Regional Economics for valuable comments on an earlier draft of this paper circulating under the same title. The authors thank three anonymous referees and the Editor for insightful comments which have significantly improved the overall quality of this article.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2019 The Authors. Papers in Regional Science published by John Wiley & Sons Ltd on behalf of Regional Science Association International.



forms of city size distributions. Stated in other words: Even if the city size process converges to the well-known exact Zipf's law¹, this law must not necessarily fit every stage of urbanization.

Using a novel methodology and based on a dataset for U.S. cities between 1840 and 2016, the aim of this paper is to answer four important questions: First, which form of the presented family of Zipf models can be used to describe the U.S. city size distribution? Second, do we observe that the U.S. city size distribution exhibits a smooth transition towards the exact Zipf's law from the beginning or are there periods showing a pronounced departure from Zipf's law? Third, if we observe periods of departure, will the city sizes be more equally or unequally distributed than predicted by the exact Zipf's law? Fourth, employing information from the hierarchical structure of cities, do we always find evidence for primate cities for a specific period of time?

In order to answer the raised first three questions, we estimate a more general three-parameter Zipf model, which can be traced back to Mandelbrot (1982). To validate these results and to answer the fourth question, we make use of the finding made by Chen (2012b) who shows that Zipf's law can be derived by the hierarchical scaling law based on a hierarchical urban structure. Intuitively, if the top level of a hierarchy is vacant, we can conclude that there is no evidence for primate cities.

The paper makes the following points: First, for the great majority of the examined years between 1840 and 2016, the U.S. city size distribution can be described by a two-parameter Zipf model with a decreasing scaling exponent. From this result we can conjecture that the U.S. city size distribution has become more equally distributed over time thereby diverging from the exact Zipf's law. Moreover, we find evidence for leading cities dominating the remaining largest cities, which indicates that the growing equality of the city sizes is due to the growth of smaller cities instead of a loss of importance of the largest cities. Relating our results with the findings made by Black and Henderson (2003) or Dobkins and Ioannides (2000), we further conclude that especially in the last decades of the twentieth century, the growth of the largest U.S. areas has mainly taken the form of suburbanization.

The next section outlines the one-, two- and three-parameter Zipf model. In section 3 we show the correspondence between Zipf's law and the hierarchical scaling law. Section 4 presents the estimation and validation procedure, followed by section 5 which presents the data. Section 6 discusses the results. Section 7 contrasts our findings with the relevant literature and concludes.

2 | A FAMILY OF ZIPF MODELS

In this section we shortly introduce the three-, two- and one-parameter Zipf models. Suppose that $P(r)$ denotes the size of a city with rank $r = \{1, 2, 3, \dots\}$, where the largest city has rank 1. Further, let k denote a scale-translational parameter and define q as the scaling exponent.

2.1 | The three-parameter Zipf model

The three-parameter model can be traced back to Mandelbrot (1982):

$$P(r) = \frac{\Theta}{(r + \xi)^q}, \quad (1)$$

where ξ represents an adjustment parameter, and Θ shows a proportionality coefficient. Recently, Chen and Zhou (2008) have shown that Mandelbrot's (1982) law given with equation (1) can be rewritten by capturing the hierarchical urban structure, which we refer to as the three-parameter Zipf model:

$$P(r) = \frac{P_{1-k}}{(r + k)^q}, \quad (2)$$

¹Based on the family of Zipf models, the terms "exact Zipf's law" and "one-parameter Zipf model" are used interchangeably in this paper. If the scaling exponent differs from one, we'll receive a two- or three-parameter Zipf model. See section 2 for more details.



with P_{1-k} showing the size of the $(1 - k)$ -th possible largest city. The scale-translational parameter k indicates a gap between the largest city in the real world (in the data set) and the possible largest city (estimated by the model). If $k = 1$, for example, in the optimal urban ranking predicted by the model, the proportionality coefficient represents the leading city $P_{1-k} = P_0$. It is larger than the largest city in the data P_1 . In contrast to the exact Zipf's law, this model describes a situation of a more evenly ($q < 1$) or a more unevenly ($q > 1$) sized distribution without leading cities in the top level of the urban ranking ($k > 0$).

2.2 | The two-parameter Zipf model

The two-parameter Zipf model directly arrives from equation (2) by letting $k = 0$:

$$P(r) = \frac{P_1}{r^q}, \tag{3}$$

with P_1 showing the size of the largest city in the urban hierarchy.²

2.3 | The one-parameter Zipf model

If we further assume that the scaling exponent is $q = 1$, from equation (3) we directly deduce the exact form of Zipf's law which states that the rank of a city r is inversely related to its size $P(r)$ (Zipf (1949))³:

$$P(r) = \frac{P_1}{r}. \tag{4}$$

3 | THE CORRESPONDENCE BETWEEN ZIPF'S LAW AND THE HIERARCHICAL SCALING LAW

As shown by Chen (2012b), Zipf's law can be transformed into the hierarchical scaling law, which can be applied to reveal the scaling relations of the hierarchical structure of the city sizes. We will use the mathematical relationship between the two models in order to validate the estimated one-, two- or three-parameter Zipf model. In what follows, we briefly show the correspondence between the hierarchical scaling law and Zipf's law.

In a first step, we construct a hierarchy of cities. Suppose there are M levels of cities with $m = \{1, 2, \dots, M\}$. Further, let f_m be the number of cities in the m -th level, whereas f_1 refers to the number of cities in the top level. P_m is the average size of the cities in the m -th level. Following Chen (2012b), the hierarchy of cities can be described by two discrete exponential functions, namely the city number law

$$f_m = f_1 \delta^{m-1}, \tag{5}$$

and the city size law

$$P_m = P_1 \lambda^{1-m}, \tag{6}$$

with parameters $\delta = \frac{f_{m+1}}{f_m} > 1$ and $\lambda = \frac{P_m}{P_{m+1}} > 1$ referring to the number ratio and the size ratio, respectively.

Using these two equations, Chen (2016) shows that the three-parameter Zipf model can be reinterpreted as a fractal model of the rank-size distribution of cities:

$$f_m = \eta P_m^{-D}, \tag{7}$$

²If the leading cities are missing ($k > 0$) and the scaling exponent is $q = 1$, we will receive a special form of the two-parameter Zipf model: $P(r) = P_{1-k}/(r+k)$.

³The Pareto distribution between the cities' ranks r as the dependent variable and the cities' sizes $P(r)$ as the independent variable was initially observed by Auerbach (1913), who found a Pareto exponent close to one. In theory, this Pareto distribution is equivalent to Zipf's law proposed by Zipf (1949), where the cities' ranks form the independent variable and the cities' sizes the dependent variable. We will refer to this distribution as the one-parameter Zipf model or the exact Zipf's law, if the estimated scaling exponent q is equal to one, and as the two-parameter Zipf model, if $q \neq 1$. A large strand of the literature uses the terms "Zipf's law", "Pareto distribution" and "Power law distribution" effectively synonymously (Newmann (2006)).



with $\eta \equiv f_1 P_1^D$ as a proportionality coefficient. The scaling exponent D is directly associated with the fractal dimension of urban hierarchies:

$$D = - \lim_{m \rightarrow \infty} \frac{\ln\left(\frac{N_{m+1}}{N_m}\right)}{\ln\left(\frac{P_{m+1}}{P_m}\right)} = - \frac{\ln\left(\frac{f_{m+1}}{f_m}\right)}{\ln\left(\frac{P_{m+1}}{P_m}\right)} = \frac{\ln(\delta)}{\ln(\lambda)} \equiv \frac{1}{q}, \quad (8)$$

with N_m as the cumulative number of city levels. Taken together, equations (7) and (8) show a direct correspondence between the hierarchical scaling law and the three-parameter Zipf model presented in equation (2).⁴

4 | ESTIMATION AND VALIDATION PROCEDURE

Based on Chen (2016), in this section, we present the estimation as well as the validation procedure that is used to study the evolution of the U.S. city size distribution.

4.1 | First step: Determination of the scaling range

As a first step, the scaling range is determined, which is a straight line on the plot with the logarithmized size of the city on the y-axis and the logarithmized rank of the city on the x-axis. Cities beyond this scaling range represent underdeveloped cities and they are not considered in the analysis. Applying an OLS estimation yields a residual value for each city and standardized residuals can be calculated. As proposed by Chen (2015), if a standardized residual value is smaller than -2 or larger than 2 , then the associated data point will be treated as an outlier based on the significance level $\alpha = 0.05$ and it will be left out of the estimation.

4.2 | Second step: Estimation of the Zipf model

Having defined the scaling range, we apply an OLS estimation to the logarithmized version of the three-parameter model

$$\ln(P(r)) = \ln(P_{1-k}) - q \ln(r + k) \quad (9)$$

and increase the scale-translational parameter k until the value of goodness of fit R^2 reaches its maximum. According to the k , we receive a Zipf model describing the city size distribution. If $k = 0$ is optimal and the estimation yields as a scaling exponent the value $q = 1$, then we will receive the one-parameter Zipf model (4). In case that $k = 0$ and $q \neq 1$ or else $k > 0$ and $q = 1$, we obtain the two-parameter Zipf model (3) or a special form of the two-parameter model, respectively. For $k > 0$ and $q \neq 1$, the result is the three-parameter Zipf model (2).

4.3 | Third step: Validation of the Zipf model

The ascertained model can be transformed into the hierarchical scaling law (7), which is based on the hierarchy constructed by the city number law or city size law. In order to validate our Zipf models, we apply the city number law (5). Given a number ratio of e.g. $\delta = 2$, then the number of cities in the different levels will be a geometric sequence such as $1, 2, 4, \dots, 2^{m-1}$. The average city size at each level can be easily calculated, leading to a number based urban hierarchy. We can make a least square calculation to examine whether the hierarchical scaling law can be well fitted to this hierarchical dataset and thereby whether our estimated Zipf model can be validated.

In order to confirm the above, we utilize the empirical relationship between exponents estimated by means of an OLS regression approach and by means of a Reduced Major Axis regression approach. The fractal dimension D is theoretically related to the Zipf scaling exponent q by $q = 1/D$ (see equation (8)). Empirically, q is calculated by

⁴See also Chen (2012a), p. 3295.

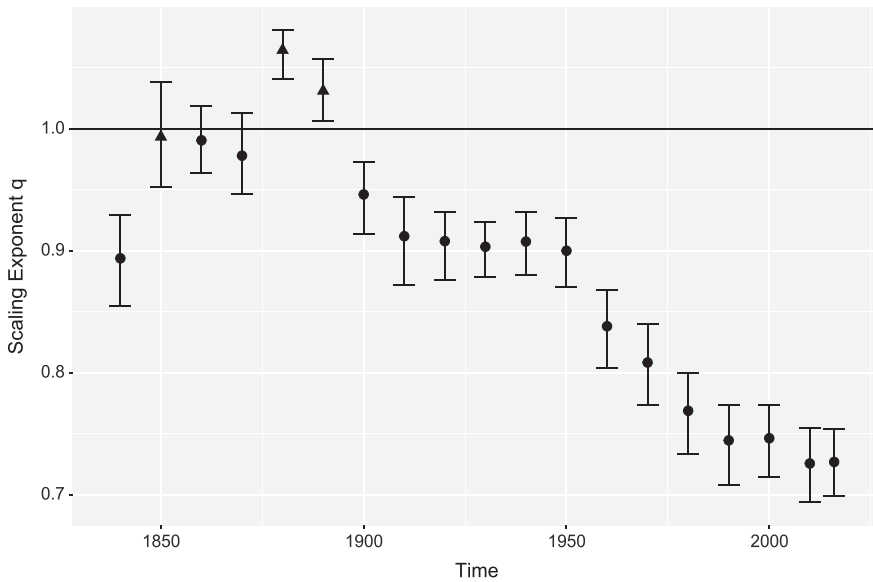


FIGURE 1 Evolution of the scaling exponent q for the 100 largest U.S. cities between 1840 and 2016
Notes: The circles (triangles) show the calculated scaling exponents for $k = 0$ ($k = 1$), when the R^2 reaches its maximum. For each estimated scaling exponent q , the 95%-confidence intervals for the bootstrapped estimate of q with 10,000 replications are depicted.
source: Own illustration based on data by United States Census Bureau.

$q^{OLS} = R^2 / D^{OLS}$. The OLS-estimation can be applied to the Reduced Major Axis regression approach, from which we have

$$D^{RMA} = \sqrt{\frac{D^{OLS}}{q^{OLS}}} = \sqrt{\frac{R^2 / q^{OLS}}{q^{OLS}}} = \frac{R}{q^{OLS}}. \tag{10}$$

If the difference $|D^{OLS} - D^{RMA}| = |(R^2 / q^{OLS}) - (R / q^{OLS})| \rightarrow 0$ ($R \rightarrow 1$), we can be sure that the city structure follows the hierarchical scaling law (7) and the ascertained Zipf model can be validated.

5 | DATASET

In order to study the evolution of the U.S. city size distribution, a dataset from the U.S. Bureau of the Census is applied. It contains the population data of the 100 largest urban places in the U.S., which we refer to as “cities” in this paper.⁵ As a sample selection criterion, we follow Rosen and Resnick’s (1980) number threshold approach and examine a fixed number of cities every ten years from 1840 to 2016.⁶

Besides the size distribution of the 100 largest cities, for the last four dates, we additionally consider larger data samples to check whether the results change when including more cities. The dataset is, therefore, supplemented by the sizes of the 601 largest cities for the years 1990 and 2000 as well as the sizes of the 300 largest cities for 2010 and 2016.

⁵Before 1950, urban places were defined as incorporated places with at least 2500 inhabitants. Since 1950, the Census Bureau has differentiated between large cities, which are considered in our study, and urbanized areas in order to account for suburban areas in the vicinity of large cities. The data can be accessed online from <https://www.census.gov/population/www/documentation/twps0027/twps0027.html> and <http://demographia.com/db-uscity98.htm> and <https://www.census.gov/data/tables/2016/demo/popest/total-cities-and-towns.html>.

⁶An overview of the number of inhabitants for selected years can be found in Table 1.



TABLE 1 Number of inhabitants of 100 largest U.S. cities

Rank	1860		1910	
1	New York	813,669	New York	4,766,883
2	Philadelphia	565,529	Chicagocity	2,185,283
3	Brooklyn	266,661	Philadelphia	1,549,008
4	Baltimore	212,418	St. Louis	687,029
5	Boston	177,840	Boston	670,585
6	New Orleans	168,675	Cleveland	560,663
7	Cincinnati	161,044	Baltimore	558,485
8	St. Louis	160,773	Pittsburgh	533,905
9	Chicago	112,172	Detroit	465,766
10	Buffalo	81,129	Buffalo	423,715
20	Milwaukee	45,246	Kansas	248,381
30	Syracuse	28,119	Toledo	168,497
40	NewBedford	22,300	Paterson	125,600
50	Petersburg	18,266	Albany	100,253
60	Poughkeepsie	14,726	Springfield	88,926
70	Harrisburg	13,405	St. Joseph	77,403
80	Elizabeth	11,567	Evansville	69,647
90	New London	10,115	Charleston	58,833
100	Wilmington	9,552	South Bend	53,684
Rank	1960		2010	
1	New York	7,781,984	New York	8,175,133
2	Chicago	3,550,404	Los Angeles	3,792,621
3	Los Angeles	2,479,015	Chicago	2,695,598
4	Philadelphia	2,002,512	Houston	2,099,451
5	Detroit	1,670,144	Philadelphia	1,526,006
6	Baltimore	939,024	Phoenix	1,445,632
7	Houston	938,219	San Antonio	1,327,407
8	Cleveland	876,050	San Diego	1,307,402
9	Washington	763,956	Dallas	1,197,816
10	St. Louis	750,026	San Jose	945,942
20	Buffalo	532,759	El Paso	649,121
30	Newark	405,220	Baltimore	620,961
40	St. Paul	313,411	Colorado Springs	416,427
50	Tulsa	261,685	Wichita	382,368
60	Albuquerque	201,189	Lexington – Fayette	295,803
70	Gary	178,320	Newark	277,140
80	Bridgeport	156,748	Laredo	236,091
90	Montgomery	134,393	North Las Vegas	216,961
100	Greensboro	119,574	San Bernardino	209,924

6 | RESULTS

6.1 | The evolution of the U.S. city size distribution

We follow Chen's (2016) procedure to estimate the scaling exponent q for different parameters k and stop when the value of goodness of fit (R^2) reaches its highest value.⁷ Figure 1 shows the evolution of the scaling exponent q over time. For most of the years, this maximum is attained for $k = 0$ (see Table 2), thus rejecting a three-parameter Zipf model.

⁷As an example, the estimation results for the years 1880 and 2016 are depicted in Figure 2.

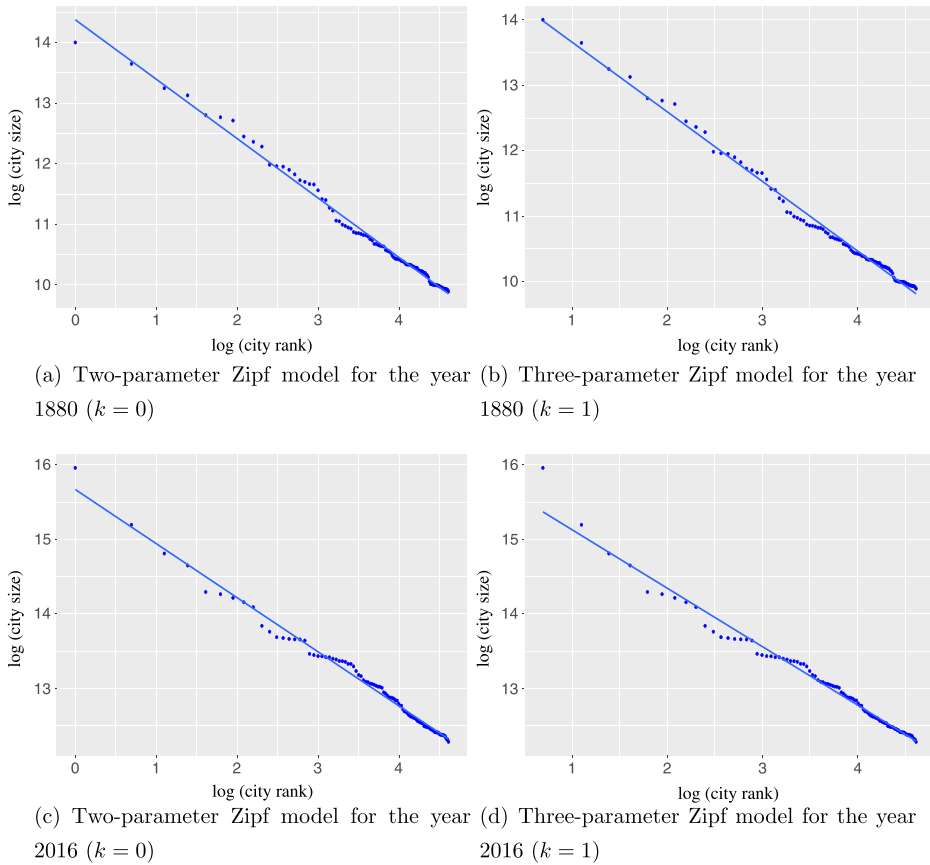


FIGURE 2 Estimation Results for the 100 largest U.S. cities in 1880 and 2016

Notes: The city size distribution in 1880 can be described most accurately by means of the three-parameter model ($K = 1$). The city size distribution in 2016 can be described most accurately by means of the two-parameter model ($k = 0$).

Source: Own illustration based on data by United States Census Bureau.

Further, the scaling exponent q decreases over the time horizon (see Figure 1). For the first 60 years (from 1840 to 1900) q fluctuates around the value one. From 1910 to 1950, we observe that q remains constant taking a value slightly above $q = 0.9$. Starting with the year 1950, the calculated scaling exponent distinctly decreases to $q = 0.75$ in 1990, followed by a further reduction with a value of $q = 0.72$ until 2016. Estimating a scaling exponent q , which is significantly lower than one, indicates a city size distribution which is more equally distributed than expected by the exact Zipf's law.

However, we also find exemptions from this behavior. In particular, we cannot reject the exact Zipf's law ($q = 1$) for the years 1860 and 1870. In the years 1880 and 1890 the value of goodness of fit R^2 did not reach its maximum for $k = 0$, but for $k = 1$. Hence, the U.S. city size distribution followed the three-parameter Zipf model in these years. For the dataset from 1850, $k = 1$ is optimal and $q = 1$ cannot be rejected. We receive a special form of the two-parameter model. As an example, Figure 2 compares the estimation results for $k = 0$ and $k = 1$ in the years 1880 and 2016. For 1880, the city size distribution is most accurately described by a three-parameter model ($k = 1$), as the largest city is too small to dominate the remaining cities. Hence, there is a gap between the real largest city in the data and the possible largest city predicted by the model. For the year 2016, on the contrary, the two-parameter model ($k = 0$) fits the data set most accurately. We find evidence for a leading city dominating the remaining largest cities.

**TABLE 2** Results of OLS estimation

year	scaling range	Zipf model	k	P_{1-k}	q	rse	R^2	D
1840	100	two	0	12.4715	0.8937***	0.0192	0.9911	1.1090
1850	100	two	0	13.1002	0.9174***	0.0192	0.9888	1.0779
1850	100	two'	1	13.4222	0.9933***	0.0235	0.9895	0.9962
1860	100	one	0	13.6844	0.9904***	0.0138	0.9944	1.0040
1870	100	one	0	14.0502	0.9779***	0.0171	0.9921	1.0146
1880	100	one	0	14.3784	0.9829***	0.0186	0.9920	1.0093
1880	100	three	1	14.7240	1.0644***	0.0093	0.9929	0.9328
1890	100	two	0	14.7002	0.9510***	0.0233	0.9921	1.0432
1890	100	three	1	15.0385	1.0310***	0.0131	0.9951	0.9652
1900	100	two	0	14.9145	0.9461***	0.0150	0.9936	1.0502
1910	100	two	0	15.1230	0.9119***	0.0189	0.9912	1.0870
1920	100	two	0	15.3544	0.9078***	0.0146	0.9949	1.0959
1930	100	two	0	15.5427	0.9032***	0.0121	0.9948	1.1014
1940	100	two	0	15.6018	0.9075***	0.0135	0.9940	1.0954
1950	100	two	0	15.7199	0.9000***	0.0142	0.9896	1.0996
1960	100	two	0	15.6553	0.8381***	0.0161	0.9850	1.1753
1970	100	two	0	15.6316	0.8083***	0.0173	0.9800	1.2124
1980	100	two	0	15.5046	0.7688***	0.0172	0.9857	1.2822
1990	100	two	0	15.4963	0.7445***	0.0173	0.9873	1.3261
2000	100	two	0	15.5892	0.7463***	0.0157	0.9889	1.3252
2010	100	two	0	15.5968	0.7256***	0.0161	0.9820	1.3563
2016	100	two	0	15.6694	0.7268***	0.0147	0.9902	1.3624
Larger datasets								
1990	587(601)	two	0	15.5562	0.7650***	0.0045	0.9996	1.3026
2000	601(601)	two	0	15.5728	0.7423***	0.0042	0.9981	1.3446
2010	299(300)	two	0	15.6133	0.7294***	0.0071	0.9961	1.3656
2016	300(300)	two	0	15.6779	0.7298***	0.0065	0.9968	1.3659

Notes: k: scaling parameter, P_{1-k} : size of the (1-k)th-city or size of the largest city P_1 for $k = 0$, q: scaling exponent, rse: robust standard error, R^2 : value of goodness of fit, D: fractal dimension of urban hierarchies, **: special form of the two-parameter Zipf model, ***: statistically significant at a 1% level.

To check the robustness of our findings, we considered larger datasets, yielding rather similar estimation results.⁸ We find that the scaling parameter $k = 0$ is optimal and the scaling exponent is slightly decreasing from $q = 0.7650$ in 1990 to $q = 0.7298$ in 2016.

To sum up, and with the exemption of the years 1850 to 1890, the U.S. city size distribution significantly (at a 5% level of significance) follows a two-parameter Zipf model in the years 1840-2016 even when considering larger samples. Hence, we can clearly reject the exact form of Zipf's law for U.S. city data. We find that for most of the years the city size distributions are more equally distributed than expected by the exact Zipf's law and that they have become more equally distributed over time.

6.2 | The evolution of the hierarchy of the U.S. cities

We exploit the above mentioned dual relationship between the hierarchical scaling and Zipf's law to obtain a more precise understanding of the structure of urban hierarchies. In particular, we want to explore whether or not the

⁸For the year 1990, the 587 largest cities and for 2000 all of the 601 cities are within the scaling range. In 2010, 299 cities and in 2016 all 300 cities are included in the estimation. Section 4.1 precisely describes how to determine the scaling range.



Level	1880		2016	
	City Number	Average City Size	City Number	Average City Size
1			1	8537673.00
2			2	3340640.00
3	4	780829.25	4	1744720.25
4	8	258424.13	8	1020569.88
5	16	99894.31	16	678376.50
6	32	41654.16	32	410409.28
7	40	24470.05	37	256965.24

TABLE 3 Classification of 100 largest cities in levels (1880 and 2016)

year	range	q^{OLS}	R^2	D^{OLS}	D^{RMA}	$D^{RMA} - D^{OLS}$
1840	1to6	1.0032	0.9860	0.9829***	0.9899	0.0069
1850	3to6	1.2783	0.9914	0.7756***	0.7789	0.0033
1860	1to6	1.0935	0.9992	0.9137***	0.9141	0.0004
1870	1to6	1.0238	0.9960	0.9728***	0.9748	0.0019
1880	3to6	1.4057	0.9971	0.7094***	0.7104	0.0010
1890	3to6	1.3299	0.9899	0.7443***	0.7481	0.0038
1900	1to6	1.0670	0.9874	0.9253***	0.9313	0.0059
1910	1to6	1.0509	0.9782	0.9308***	0.9411	0.0103
1920	1to6	1.0285	0.9844	0.9571***	0.9647	0.0075
1930	1to6	1.0299	0.9893	0.9606***	0.9658	0.0052
1940	1to6	1.0283	0.9879	0.9607***	0.9666	0.0059
1950	1to6	0.9942	0.9877	0.9935***	0.9996	0.0062
1960	1to6	0.9422	0.9806	1.0408***	1.0510	0.0103
1970	1to6	0.9147	0.9730	1.0637***	1.0784	0.0147
1980	1to6	0.8919	0.9773	1.0957***	1.1083	0.0127
1990	1to6	0.8800	0.9747	1.1076***	1.1218	0.0143
2000	1to6	0.8777	0.9766	1.1128***	1.1260	0.0132
2010	1to6	0.8540	0.9752	1.1419***	1.1564	0.0145
2016	1to6	0.8448	0.9775	1.1572***	1.1704	0.0132
Larger datasets						
1990	1to9	0.8225	0.9893	1.2028***	1.2093	0.0065
2000	1to9	0.8120	0.9891	1.2181***	1.2248	0.0067
2010	1to8	0.8052	0.9856	1.2240***	1.2329	0.0089
2016	1to8	0.8022	0.9875	1.2309***	1.2387	0.0078

TABLE 4 Validation Results

q^{OLS} : scaling exponent (OLS-estimation), R^2 : value of goodness of fit, D^{OLS} : fractal dimension of urban hierarchies (OLS-estimation), D^{RMA} : fractal dimension of urban hierarchies (RMA-estimation), ***: statistically significant at a 1% level.

existence of primate cities is a time invariant pattern that describes the U.S. city size distribution. In order to answer this question, we have to make sure that the city structure follows a hierarchical scaling law.⁹

According to the city number law (5), the cities are ranked into 7 levels. If the one- or the two-parameter Zipf model fits the data, the first level in the hierarchical structure consists of the largest city. The next level comprises the second and third largest cities, the third level consists of the fourth to the seventh largest cities and so on. The last level is supposed to comprise 64 cities, but because our dataset only contains 100 cities, the last level comprises 37 cities. Hence, it is not included in the estimation (see Table 3).

⁹Detailed information on the construction of urban hierarchies and the validation procedure are given in section 4.3.



The estimation of a three-parameter or the special form of the two-parameter model with $k > 0$ suggests an absence of leading cities. That is why the first two levels are absent when constructing the city hierarchy. So, the four largest cities are classed with the third level. Again, the last level comprising the forty smallest cities is not included in the estimation, as it is a lame-duck class.¹⁰

Looking at Table 4, we see that the city structure follows a hierarchical scaling law from 1840 to 1950 as well as for 1990-2016 when larger datasets are used. For these years, our estimated Zipf models can be validated. We can confirm the absence of leading cities for the years 1850, 1880 and 1890, in which a three- or special form of the two-parameter model was estimated ($k > 0$), by finding that the hierarchical scaling law fits an urban structure without top levels. For the remaining years, in which we estimated a Zipf model with $k = 0$, we find that the hierarchical scaling law fits an urban structure with the largest cities at the top levels. Hence, we can confirm the existence of leading cities for most of the years. For the 100 largest cities, we observe a pronounced divergence from the hierarchical scaling law starting with 1960 until the year 2016. This can also be seen by comparing the log-rank/log-size plot with the hierarchical scaling relation between the average sizes in the hierarchies of the U.S. cities and the city numbers.

To sum up, for the 100 largest cities and for most of the time span 1840-2016 we find evidence for leading cities dominating the remaining largest U.S. cities and we find a divergence from the hierarchical scaling law.

7 | DISCUSSION AND CONCLUSION

This paper reveals the following aspects of the evolution of the U.S. city size distribution: (1) The 100 largest U.S. cities can mostly be described by a two-parameter Zipf model between 1840 and 2016. (2) For most of the years, the examined scaling exponent q is lower than one and it has decreased, especially during the second half of the twentieth century. (3) The U.S. city size distribution has become more even over time and diverged from the exact Zipf's law. (4) For most of the years, we find evidence for leading cities dominating the remaining largest U.S. cities.

When relating our findings to the existing relevant literature, it is striking that the great majority of studies uses cross-sectional data to check whether or not Zipf's law holds exactly.¹¹ For instance, Krugman (1996) and Gabaix (1999) use data for U.S. Metropolitan Statistical Areas (MSAs) and find that the one-parameter Zipf model holds exactly for a minimum threshold of 280,000 inhabitants. These findings are recently confirmed by Schmidheiny and Suedekum (2015) using novel data from an EC-OECD project. Zipf's law also occurs when applying other city definitions, like economic areas (Berry and Okulicz-Kozaryn (2012)), natural cities (Jiang and Jia (2011)) or geographic clusters (Rozenfeld et al. (2011)). Some studies found opposing results for the U.S. city size distribution (Eckhout (2004)) or found that Zipf's law only holds for the upper tail of the distribution while the body and lower tail are lognormal (Levy (2009), Malevergne et al. (2011) and Ioannides and Skouras (2013)). Using U.S. census data, Soo (2005) found that the largest cities are more evenly and the largest urban agglomerations are more unevenly distributed than predicted by the exact Zipf's law (also see Gan et al. (2006) and Ioannides and Overman (2003)).

Focusing on the long-term perspective of Zipf's law, again, the results depend on the employed city definition. For Metropolitan Statistical Areas, Black and Henderson (2003) or Dobkins and Ioannides (2000, 2001) find an increasing urban concentration, which is higher than predicted by the exact Zipf's law. Other authors focused on states (Soo (2012)), counties (Beeson et al. (2001) and Desmet and Rappaport (2017)) or minor civil divisions (Michaels et al. (2012)) in the U.S.. Closest to our study is González-Val (2010). Comparing U.S. incorporated places, the author finds that the city sizes are lognormally distributed and more unequally distributed than predicted by the exact Zipf's law.

Regarding the upper tail of the city size distribution, the author finds that the cities become more equally distributed over time. This is in line with our results, which clearly show that since 1960, the scaling exponent significantly drops year by year until 2016, indicating more evenly distributed city sizes and a departure from Zipf's

¹⁰The classification for the years 1880 and 2016, when a three-parameter model and a two-parameter model hold, is presented in Table 3.

¹¹A detailed literature review on the theoretical and empirical findings on Zipf's law is given by Arshad et al. (2018).



law for the 100 largest cities. In contrast to González-Val (2010), who explains the convergence of the city sizes with a loss of importance of the largest cities, for most of the time span, we find evidence for leading cities dominating the remaining largest cities. Our results indicate that the growth of the smaller cities plays the main role in the convergence process.

At the same time, Black and Henderson (2003) and Dobkins and Ioannides (2000) found that U.S. MSAs have become more unequally distributed during the twentieth century. Connecting these results to the convergence of city sizes, we found, confirms an increasing suburbanization in the growth process of the largest U.S. urban areas starting in the 1960s (Soo (2005)).¹²

The main point this paper makes is that the U.S. city size distribution has moved away from the exact Zipf's law, especially in the second half of the twentieth century. While for the years 1850, 1880 and 1890, leading cities are missing, they exist for each Census year from 1900 onwards. The scaling exponent decreased, indicating more equally distributed city sizes. In turn, different regimes of Zipf models imply different conditions of city development. Thus, the main deficiency of this paper is that we cannot identify which are the driving forces leading to this evolutionary development away from the exact Zipf's law over time. A more elaborated investigation is definitely needed, but beyond the scope of this paper. Besides that, the rather subjective definition of a city might influence the results this paper makes. Insofar, the results cannot be generalized to other countries or to the same country but with a different city definition. However, this problem is common to every study dealing with city-level data.

ORCID

Angelina Hackmann  <https://orcid.org/0000-0002-9683-6763>

Torben Klarl  <https://orcid.org/0000-0002-3059-1150>

REFERENCES

- 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: Statistical Mechanics and its Applications*, 492, 75–92.
- Auerbach, F. (1913). Das Gesetz der Bevölkerungskonzentration. *Petermann's Geographische Mitteilungen*, 59(1), 74–76.
- Bayoh, I., Irwin, E. G., & Haab, T. (2006). Determinants of residential location choice: How important are local public goods in attracting homeowners to central city locations. *Journal of Regional Science*, 46(1), 97–120.
- Beeson, P. E., DeJong, D. N., & Troesken, W. (2001). Population growth in U.S. counties, 1840–1990. *Regional Science and Urban Economics*, 31, 669–699.
- Berry, B., & Okulicz-Kozaryn, A. (2012). The city size distribution debate: Resolution for US urban regions and megapolitan areas. *Cities*, 29, 17–23.
- Black, D., & Henderson, V. (2003). Urban evolution in the USA. *Journal of Economic Geography*, 3(4), 343–372.
- Boustan, L., & Shertzer, A. (2013). Population Trends as a Counterweight to Central City Decline, 1950–2000. *Demography*, 50(1), 125–147.
- Chen, Y. (2012a). Zipf's law, 1/f noise, and fractal hierarchy. *Chaos, Solitons and Fractals*, 45(1), 63–73.
- Chen, Y. (2012b). The mathematical relationship between Zipf's law and the hierarchical scaling law. *Physica A: Statistical Mechanics and its Applications*, 391(11), 3285–3299.
- Chen, Y. (2015). The distance-decay function of geographical gravity model: Power law or exponential law. *Chaos, Solitons and Fractals*, 77, 174–189.
- Chen, Y. (2016). The evolution of Zipf's law indicative of city development. *Physica A: Statistical Mechanics and its Applications*, 443, 555–567.
- Chen, Y., & Zhou, Y. (2008). Scaling laws and indications of self-organized criticality in urban systems. *Chaos, Solitons and Fractals*, 35, 85–98.

¹²According to Boustan and Shertzer (2013), a large portion of suburbanization in the U.S. over the twentieth century can be explained by factors associated with the natural evolution process of urbanization, like rising incomes, which led to a larger demand for housing and land, as well as transportation improvements, especially the growing network of interstate highways. Furthermore, the authors state that factors associated with the flight-from-blight theory of suburbanization, like school quality, taxes, crime-rates and socioeconomic factors of the population, reinforced the spatial dispersion. Also see Mieszkowski and Mills (1993), Bayoh et al. (2006) and Kim (2000).



- Desmet, K., & Rappaport, J. (2017). The settlement of the United States, 1800-2000: The long transition towards Gibrat's law. *Journal of Urban Economics*, 98, 50–68.
- Dobkins, L. H., & Ioannides, Y. M. (2000). Dynamic Evolution of the U.S. City Size Distribution, *The Economics of Cities*. Cambridge: Cambridge University Press, 217–260.
- Dobkins, L. H., & Ioannides Y. M. (2001). Spatial interactions among US cities: 1900-1990. *Regional Science and Urban Economics*, 31, 701–731.
- Eckhout, J. (2004). Gibrat's Law for (All) Cities. *The American Economic Review*, 94(5), 1429–1451.
- Gabaix, X. (1999). Zipf's Law for cities: An explanation. *The Quarterly Journal of Economics*, 114, 739–767.
- Gan, L., Li, D., & Song, S. (2006). Is the Zipf law spurious in explaining city-size distributions. *Economics Letters*, 92, 256–262.
- González-Val, R. (2010). The evolution of U.S. city size distribution from a long-term perspective (1900-2000). *Journal of Regional Science*, 50(5), 952–972.
- 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., & Skouras, S. (2013). US city size distribution: Robustly Pareto, but only in the tail. *Journal of Urban Economics*, 73, 18–29.
- Jiang, B., & Jia, T. (2011). Zipf's Law for all the natural cities in the United States: A geospatial perspective. *International Journal of Geographical Information Science*, 25(8), 1269–1281.
- Kim, S. (2000). Urban Development in the United States, 1690-1990. *Southern Economic Journal*, 66(4), 855–880.
- Krugman, P. (1996). Confronting the Urban Mystery. *Journal of the Japanese and International Economies*, 10, 399–418.
- Levy, M. (2009). Gibrat's Law for (All) Cities: Comment. *American Economic Review*, 99(4), 1672–1675.
- 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*, 83(3), 1–11.
- Mandelbrot, B. B. (1982). *The fractal nature of geometry*. New York: W. H. Freeman and Company.
- Michaels, G., Rauch, F., & Redding, S. J. (2012). Urbanization and Structural Transformation. *The Quarterly Journal of Economics*, 127(2), 535–586.
- Mieszkowski, P., & Mills, E. S. (1993). The Causes of Metropolitan Suburbanization. *Journal of Economic Perspectives*, 7(3), 135–147.
- Newmann, M. E. J. (2006). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46, 323–351.
- Rosen, K. T., & Resnick, M. (1980). The Size Distribution of Cities: An Examination of the Pareto Law and Primacy. *Journal of Urban Economics*, 8, 165–186.
- Rozenfeld, H. D., Rybski, D., Gabaix, X., & Makse, H. A. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review*, 101(5), 2205–2225.
- Schmidheiny, K., & Suedekum, 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.
- Soo, K. T. (2012). The size and growth of state populations in the United States. *Economics Bulletin*, 32(2), 1238–1249.
- Zipf, G. (1949). *Human Behaviour and the Principle of Least Effort*. Cambridge MA: Addison-Wesley.

How to cite this article: Hackmann A, Klarl T. The evolution of Zipf's Law for U.S. cities. *Pap Reg Sci*. 2020;99:841–852. <https://doi.org/10.1111/pirs.12498>