CESifo WORKING PAPERS

7232
2018

September 2018

The Evolution of Zipf's Law for U.S. Cities

Angelina Hackmann, Torben Klarl

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7232*
Category 6: Fiscal Policy, Macroeconomics and Growth

The Evolution of Zipf's Law for U.S. Cities

Abstract

Exploiting the cascade structure of cities and based on a dataset for U.S. cities between 1840 and 2016, the aim of this short paper is to answer three important questions: First, do we observe that the U.S. city size distribution exhibits a smooth transition to Zipf's law from the beginning or are there periods showing a pronounced departure from Zipf's law? Second, if we observe periods of departure, which alternative laws instead should be used to accurately describe the city size distribution? Third, employing information from the cascade structure of cities, do we always find evidence for primate cities for a specific period of time? Inter alia, we find that the exact Zipf's law has evolved over time from the more general, so-called three-parameter Zipf's law which can be traced back to Mandelbrot (1982).

JEL-Codes: R110, R120, R150.

Keywords: city size distributions, Zipf's law, hierachical scaling law, urban systems.

Angelina Hackmann
University of Bremen
Faculty of Business Studies & Economics
Hochschulring 4
Germany – 28359 Bremen
angelina.hackmann@uni-bremen.de

Torben Klarl
University of Bremen
Faculty of Business Studies & Economics
Hochschulring 4
Germany – 28359 Bremen
tklarl@uni-bremen.de

This version: August 2018

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 forms of city size distributions. Stated in other words: Even if the city size process converges to the well-known Zipf's law, this law must not be necessarily fit to 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 short paper is to answer three important questions: First, do we observe that the U.S. city size distribution exhibits a smooth transition to Zipf's law from the beginning or are there periods showing a pronounced departure from Zipf's law? Second, if we observe periods of departure, which laws instead should be used to accurately describe the city size distribution? Third, employing information from the cascade structure of cities, do we always find evidence for primate cities for a specific period of time?

In order to answer the raised first two questions, according to Chen (2016), we argue that the exact Zipf's law has evolved over time from the more general, so-called three-parameter Zipf's law which can be traced back to Mandelbrot (1982). Thus, we employ the three-parameter Zipf's law to study the rank-size distribution of U.S. cities between 1840 and 2016. To validate these results and to answer the third question, we employ the finding made by Chen (2012b) who shows that Zipf's law can be derived by the hierarchical scaling law based on a cascade structure of cities. 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's law 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. Relating our results with the findings made by Black and Henderson (2003) or Dobkins and Ioannides (2001), 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 and explain the methodological approach to validate the calculated Zipf models. Section 4 presents the empirical strategy, 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's law

In this section we shortly introduce the three-, two- and one-parameter Zipf's law. Suppose that $P(r)$ denotes the size of a city with rank $r = \{1, 2, 3, ...\}$, 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's law

The three-parameter model, which can be traced back to Mandelbrot (1982), reads as:

$$P(r) = \frac{\Theta}{(r + \xi)^{-q}}, \tag{2.1}$$

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 (2.1) can be rewritten by capturing the cascade structure of urban hierarchies:

$$P(r) = \frac{P_{1-k}}{(r + k)^{-q}}, \tag{2.2}$$

with P_{1-k} showing the size of the $(1 - k)$-th possible largest city. In contrast to the exact Zipf's law, this law 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's law

The two-parameter Zipf's law directly arrives from equation (2.2) by letting $k = 0$:

$$P(r) = \frac{P_1}{r^{-q}}, \tag{2.3}$$

with P_1 showing the size of the largest city in the urban hierarchy.[1]

[1] If 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's law.

2.3 The one-parameter Zipf's law

If we further assume that the scaling exponent is $q = 1$, from equation (2.2) 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 $P(r)$ (also see Auerbach (1913) and Zipf (1949)):

$$P(r) = \frac{P_1}{r}. \tag{2.4}$$

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. 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 $m = \{1, 2, ..., M\}$. 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 with two discrete exponential functions, $f_m = f_1\delta^{m-1}$ and $P_m = P_1\lambda^{1-m}$, 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's law can be reinterpreted as a fractal model of the rank-size distribution of cities:

$$f_m = \eta P_m^{-D}, \tag{3.1}$$

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:

$$D = -\lim_{m\to\infty} \frac{\ln(\frac{N_{m+1}}{N_m})}{\ln(\frac{P_{m+1}}{P_m})} = -\frac{\ln(\frac{f_{m+1}}{f_m})}{\ln(\frac{P_{m+1}}{P_m})} = \frac{\ln(\delta)}{\ln(\lambda)} \equiv \frac{1}{q}, \tag{3.2}$$

with N_m as the cumulative number of city levels. Taken together, equation (3.1) and (3.2) shows a direct correspondence between the hierarchical scaling law, Pareto's law and the three-parameter Zipf's law presented with equation (2.2).[2]

[2]See also Chen (2012a), p. 3295.

4 Empirical strategy

To check whether or not the structure of cities follows the hierarchical scaling law, we employ the so called Reduced Major Axis (RMA) regression approach, from which we have (Zhang and Yu (2010)):

$$D^{RMA} = \sqrt{\frac{D^{OLS}}{q^{OLS}}}, \tag{4.1}$$

where D^{OLS} and q^{OLS} are the OLS-estimates from running a standard OLS regression of the logarithmic version of equation (2.2). Thus, if the difference between $|D^{RMA} - D^{OLS}| \rightarrow 0$, we can be sure that the city structure follows a hierarchical scaling law. Appendix 1 explains the further steps of the estimation procedure in more detail.

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. every ten years from 1840 to 2016. This dataset is 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.[3]

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.[4] Figure (6.1) shows the evolution of the scaling exponent q over time. For most of the years this maximum is attained for $k = 0$ (see Table A.2 in the Appendix), rejecting a

[3]Before 1950 urban places were defined as any 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#urban`, `http://demographia.com/db-uscity98.htm` and `https://www.census.gov/data/tables/2016/demo/popest/total-cities-and-towns.html`.

[4]As an example, the estimation results for the years 1880 and 2016 are depicted in Figure A.1 in the Appendix.

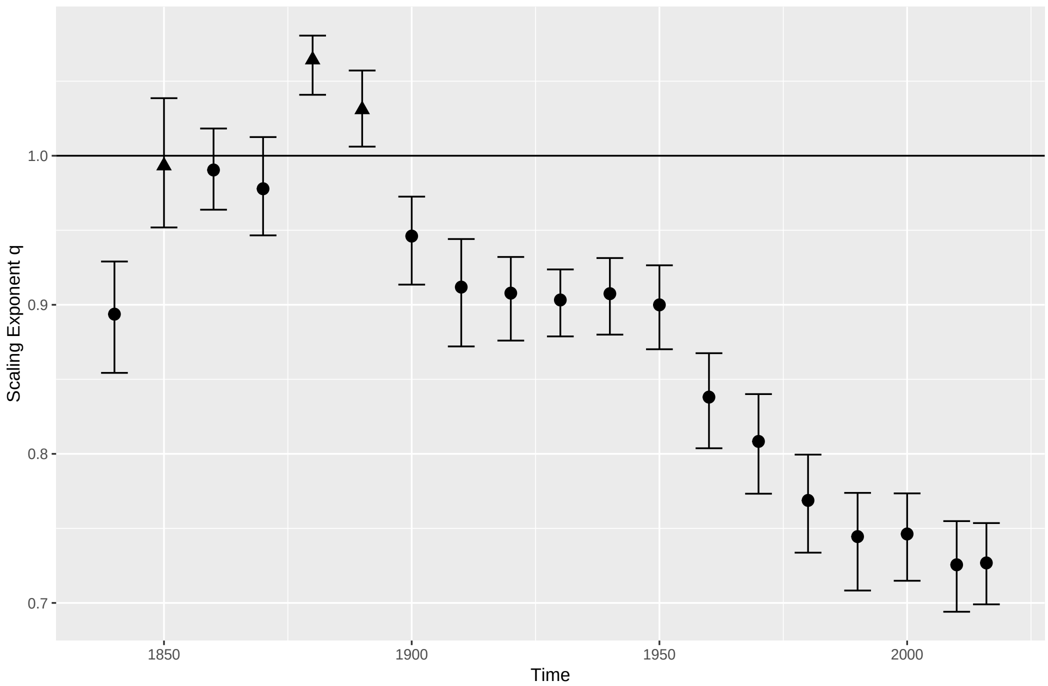


Figure 6.1: Evolution of the scaling exponent q between 1840 and 2016.
Notes: The circles (triangles) show the calculated scaling exponents for $k = 0$ ($k = 1$), when the R^2 reaches its maximum. For each estimated scaling exponent q, the 95%-confidence intervals for the bootstrapped estimate of q with 10000 replications are depicted. *Source:* Own illustration based on data by United States Census Bureau.

three-parameter Zipf model. Further, the scaling exponent q is decreasing over the time horizon (see Figure 6.1). 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.

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 does not reach its maximum for $k = 0$, but for $k = 1$. Hence, the U.S. city size distribution follows the three-parameter Zipf's law 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.

To check the robustness of our findings, we considered larger datasets, which yields rather similar estimation results.[5] 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[6] follows a two-parameter Zipf's law 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.

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 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 (see section 4).

Cities are ranked into 7 levels. If the one- or the two-parameter Zipf's law 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 A.3).

The estimation of a three-parameter or the special form of the two-parameter model 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.[7]

Looking at Table A.4 in the Appendix, we see that the the city structure follows a hierarchical scaling law from 1840 to 1950 as well as for 1990-2016 when larger datasets are used. For smaller datasets, we observe a pronounced divergence from the

[5] For 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.

[6] At a 5% level of significance.

[7] The classification for the years 1880 and 2016, when a three-parameter model and a two-parameter model hold, is presented in Table A.3 in the Appendix.

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 reaming 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 administratively defined 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 reaming largest U.S. cities.
How can we relate our findings to the existing relevant literature? Surprisingly, there is scant literature regarding the long-term evolution of Zipf's law for the U.S.. The great majority of studies uses cross-sectional data to check whether or not Zipf's law holds exactly.[8] For instance, Krugman (1996) and Gabaix (1999) use data for U.S. Metropolitan Statistical Areas (MSAs) and find that the one-parameter Zipf's law 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. Using U.S. census data, Soo (2005) found that the largest administratively defined 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)). Our results 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 law. Black and Henderson (2003) and Dobkins and Ioannides (2001) which are closest to our study found that U.S. MSAs have become more unequally distributed during the twentieth century. Thus, our results can be seen as a confirmation of the fact that we observe an increasing suburbanization in the

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

growth process of the largest U.S. urban areas starting in the 1960s (Soo (2005)).[9]

[9] According 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).

References

Arshad S., S. Hu, B. N. Ashraf 2018. Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A: Statistical Mechanics and its Applications*. Vol. 492. pp. 75-92

Auerbach, F. 1913. Das Gesetz der Bevölkerungskonzentration. *Petermann's Geographische Mitteilungen*. Vol. 59/1. pp. 74-76

Bayoh I., E. G. Irwin, T. Haab 2006. Determinants of residential location choice: How important are local public goods in attracting homeowners to central city locations? *Journal of Regional Science*. Vol. 46/1. pp. 97-120

Black D., V. Henderson 2003. Urban evolution in the USA. *Journal of Economic Geography*. Vol. 3/4. pp. 343-372

Boustan L., A. Shertzer 2013. Population Trends as a Counterweight to Central City Decline, 1950-2000. *Demography*. Vol. 50/1. pp. 125-147

Chen Y. 2012a. Zipf's law, 1/f noise, and fractal hierarchy. *Chaos, Solitons and Fractals*. Vol. 45/1. pp. 63-73

Chen Y. 2012b. The mathematical relationship between Zipf's law and the hierarchical scaling law. *Physica A: Statistical Mechanics and its Applications*. Vol. 391/11. pp. 3285-3299

Chen Y. 2015. The distance-decay function of geographical gravity model: Power law or exponential law? *Chaos, Solitons and Fractals*. Vol. 77. pp. 174-189

Chen Y. 2016. The evolution of Zipf's law indicative of city development. *Physica A: Statistical Mechanics and its Applications*. Vol. 443. pp. 555-567

Chen Y., Y. Zhou 2008. Scaling laws and indications of self-organized criticality in urban systems. *Chaos, Solitons and Fractals*. Vol. 35. pp. 85-98

Dobkins L. H., Y. M. Ioannides 2001. Dynamic Evolution of the U.S. City Size Distribution. *In: J. Huriot, J. Thisse: The Economics of Cities*. Cambridge, Cambridge University Press

Gabaix X. 1999. Zipf's Law for cities: An explanation. *The Quarterly Journal of Economics*. Vol. 114. pp. 739-767

Gan L., D. Li, S. Song 2006. Is the Zipf law spurious in explaining city-size distributions? *Economics Letters*. Vol. 92. pp. 256-262

Ioannides Y. M., H. G. Overman 2003. Zipf's law for cities: an empirical examination. *Regional Science and Urban Economics*. Vol. 33. pp. 127-137

Kim S. 2000. Urban Development in the United States, 1690-1990. *Southern Economic Journal*. Vol. 66/4. pp. 855-880

Krugman P. 1996. Confronting the Urban Mystery. *Journal of the Japanese and International Economies*. Vol. 10. pp. 399-418

Mandelbrot B. B. 1982. The fractal nature of geometry. New York, W. H. Freeman and Company

Mieszkowski P., E. S. Mills 1993. The Causes of Metropolitan Suburbanization. *Journal of Economic Perspectives*. Vol. 7/3. pp. 135-147

Schmidheiny K., J. Suedekum 2015. The pan-European population distribution across consistently defined functional urban areas. *Economics Letters*. Vol. 133. pp. 10-13

Soo K. T. 2005. Zipf's Law for cities: a cross-country investigation. *Regional Science and Urban Economics*. Vol. 35. pp. 239-263

Zhang J., T. Yu 2010. Allometric Scaling of Countries. *Physica A: Statistical Mechanics and its Applications*. Vol. 389/21. pp. 4887-4896

Zipf G. 1949. Human Behaviour and the Principle of Least Effort. Cambridge MA, Addison-Wesley

Online-Appendix

A.1 Estimation and Validation Procedure

The detailed approach, how to study the evolution of the U.S. city size distribution relies on Chen (2016). It is as follows:

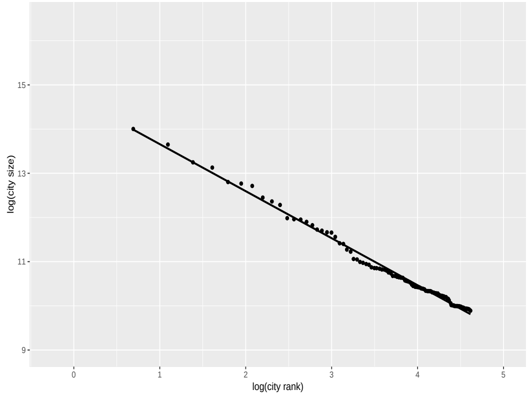
First step: Determine scaling range
As a first step, the scaling range is determined, which is a straight line on the plot with the logarithmized size of the city on the y-axis and the logarithmized rank of the city on the x-axis. Cities beyond this scaling range represent underdeveloped cities and they 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.

Second step: Estimate Zipf model
Having defined the scaling range, we apply an OLS estimation 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 (2.4). In case that $k = 0$ and $q \neq 1$ or else $k > 0$ and $q = 1$, we obtain the two-parameter Zipf model (2.3) 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 (2.2).

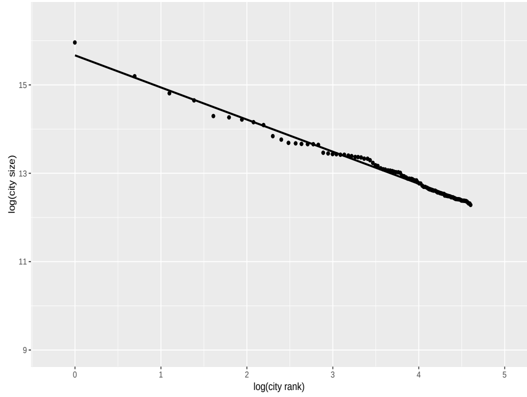
Third step: Validate Zipf model
The ascertained model can be transformed into the hierarchical scaling law (3.1), which is based on the hierarchy constructed by the city number law or city size law. In order to validate our Zipf models, we apply the city number law. Given a number ratio of e.g. $\delta = 2$, then the number of cities in the different levels will be a geometric sequence such as $1, 2, 4, ..., 2^{m-1}$. 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.

A.2 Figures

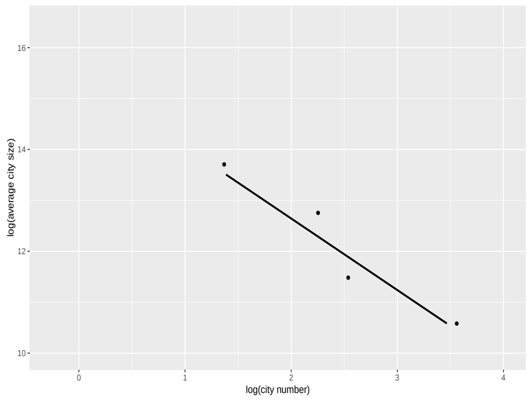
Figure A.1: Estimation Results for the 100 largest U.S. cities in 1880 and 2016



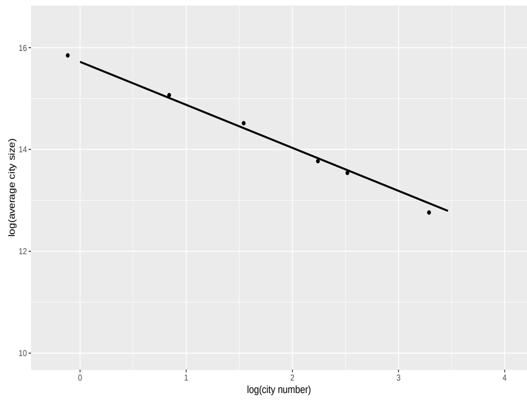
(a) Three-parameter Zipf distribution



(b) Two-parameter Zipf distribution



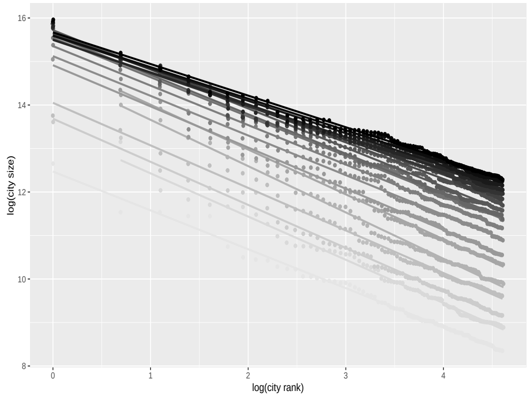
(c) Hierarchical Scaling pattern (without leading cities)



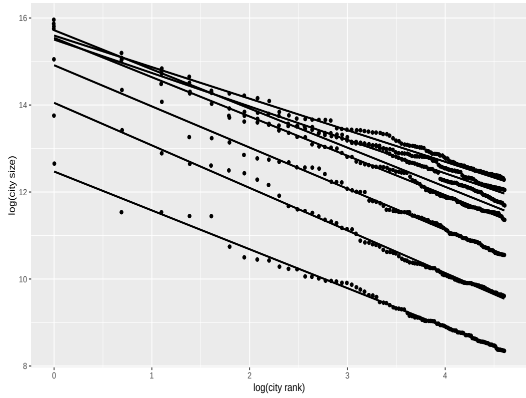
(d) Hierarchical Scaling pattern

Notes: Figure (a) and (c) relate to the year 1880, when a three-parameter Zipf's law holds and Figure (b) and (d) relate to the year 2016, when a two-parameter Zipf's law holds.

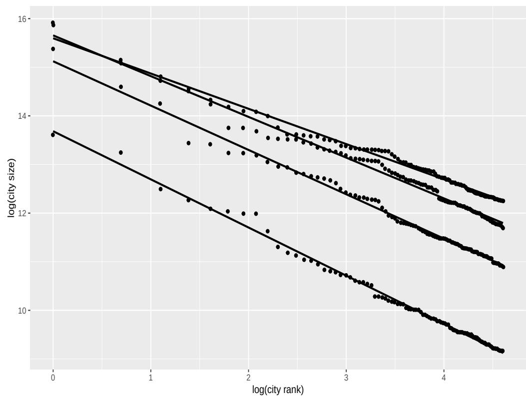
Figure A.2: Estimation Results for the 100 largest U.S. cities in selected years



(a) Regression results for all years (1840 to 2016)



(b) Regression results for the years 1840,1870,1900,1930,1960,1990,2016



(c) Regression results for the years 1860,1910,1960,2010

A.3 Tables

Table A.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	Chicago city	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	New Bedford	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	Phoenix	1,445,632
6	Baltimore	939,024	Philadelphia	1,526,006
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

Table A.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, two^*: special form of the two-parameter Zipf model

Table A.3

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

	1880		2016	
Level	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 A.4

Validation Results

year	range	P_{1-k}	q	s	R^2	D^{OLS}	D^{RMA}	$D^{RMA} - D^{OLS}$
1840	1 to 6	12.4574	1.0032	0.0597	0.9860	0.9829	0.9898	0.0070
1850	3 to 6	14.0423	1.2783	0.0840	0.9914	0.7756	0.7789	0.0034
1860	1 to 6	13.6340	1.0935	0.0155	0.9992	0.9137	0.9141	0.0004
1870	1 to 6	13.8666	1.0238	0.0324	0.9960	0.9728	0.9748	0.0020
1880	3 to 6	15.4551	1.4057	0.0533	0.9971	0.7093	0.7104	0.0010
1890	3 to 6	15.6590	1.3299	0.0951	0.9899	0.7443	0.7481	0.0038
1900	1 to 6	14.9061	1.0670	0.0604	0.9874	0.9254	0.9313	0.0059
1910	1 to 6	15.1616	1.0509	0.0784	0.9782	0.9308	0.9411	0.0103
1920	1 to 6	15.3536	1.0285	0.0647	0.9844	0.9571	0.9647	0.0076
1930	1 to 6	15.5641	1.0299	0.0536	0.9893	0.9606	0.9658	0.0052
1940	1 to 6	15.6100	1.0283	0.0569	0.9879	0.9607	0.9666	0.0059
1950	1 to 6	15.6713	0.9942	0.0555	0.9877	0.9935	0.9996	0.0062
1960	1 to 6	15.6424	0.9422	0.0663	0.9806	1.0408	1.0510	0.0102
1970	1 to 6	15.6347	0.9147	0.0762	0.9730	1.0637	1.0784	0.0147
1980	1 to 6	15.5572	0.8920	0.0680	0.9773	1.0957	1.1083	0.0127
1990	1 to 6	15.5846	0.8800	0.0709	0.9747	1.1076	1.1218	0.0143
2000	1 to 6	15.6673	0.8777	0.0679	0.9766	1.1127	1.1260	0.0133
2010	1 to 6	15.6730	0.8540	0.0681	0.9752	1.1420	1.1564	0.0144
2016	1 to 6	15.7209	0.8448	0.0640	0.9775	1.1571	1.17034	0.0132
				Larger datasets				
1990	1 to 9	15.5061	0.8225	0.0323	0.9893	1.2028	1.2093	0.0065
2000	1 to 9	15.5790	0.8120	0.0322	0.9891	1.2181	1.2248	0.0067
2010	1 to 8	15.6121	0.8052	0.0397	0.9856	1.2240	1.2329	0.0089
2016	1 to 8	15.6674	0.8022	0.0369	0.9875	1.2310	1.2388	0.0078

Notes: P_{1-k}: size of the (1-k)th-city or size of the largest city P_1 for $k = 0$, q: scaling exponent, s: standard error, 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)