

Düben, Christian; Krause, Melanie

Article — Published Version

Population, light, and the size distribution of cities

Journal of Regional Science

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Düben, Christian; Krause, Melanie (2021) : Population, light, and the size distribution of cities, Journal of Regional Science, ISSN 1467-9787, Wiley, Hoboken, NJ, Vol. 61, Iss. 1, pp. 189-211,
<https://doi.org/10.1111/jors.12507>

This Version is available at:

<https://hdl.handle.net/10419/230241>

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/4.0/>

Population, light, and the size distribution of cities

Christian Düben  | Melanie Krause 

Department of Economics, Hamburg University,
Von-Melle-Park 5, Hamburg, Germany

Correspondence

Christian Düben, Hamburg University,
Department of Economics, Von-Melle-Park 5,
20146 Hamburg, Germany.
Email: christian.dueben@uni-hamburg.de

Funding information

Deutsche Forschungsgemeinschaft,
Grant/Award Number: KR 4936/1-1

Abstract

We provide new insights on the city size distribution of countries around the world. Using more than 10,000 cities delineated via geospatial data and a globally consistent city identification scheme, we investigate distributional shapes in all countries. In terms of population, we find that Zipf's law holds for many, but not all, countries. Contrasting the distribution of population with the distribution of economic activity, measured by nighttime lights, across cities we shed light on the globally variant magnitude of agglomeration economies. Deviations from Zipf's law are to a large extent driven by an undue concentration in the largest cities. They benefit from agglomeration effects which seem to work through area rather than through density. Examining the cross-country heterogeneity in the city size distribution, our model selection approach suggests that historical factors play an important role, in line with the time of development hypothesis.

KEYWORDS

agglomeration economies, cities, geospatial data, nighttime lights, population, urban concentration, Zipf's law

JEL CLASSIFICATION

C18; O18; R11; R12

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Author. *Journal of Regional Science* published by Wiley Periodicals LLC

1 | INTRODUCTION

The rapid urbanization process around the world has brought renewed interest to the question of the size distribution of cities. Many countries are urbanizing at relatively early stages of development (Glaeser, 2014; Jedwab & Vollrath, 2019), and the total urban population of the world is set to grow by 2.7 bn people from 2015 to 2050 (United Nations, 2018). Whether this growth is distributed equally among cities of all sizes, or whether the largest metropolises are further increasing their share, is an important question for policymakers. Bigger cities can reap agglomeration benefits by pooling physical and human capital as well as by exploiting spillovers across and within industries, but they also suffer from plights, such as congestion (Desmet & Rossi-Hansberg, 2013; Rosenthal & Strange, 2004).

It has long been argued that Zipf's law should be an appropriate description of the size distribution of big cities within a country, implying that a city's rank is approximately inversely proportional to its size (Zipf, 1949). There are strong theoretical arguments for Zipf's law as well as empirical evidence with population data, in particular for the United States (Gabaix, 1999; Rozenfeld, Rybski, Gabaix, & Makse, 2011). But cross-country studies have so far yielded mixed results (Rosen & Resnick, 1980; Soo, 2005). Unfortunately, they are often limited to countries with a high statistical quality and miss out on large parts of the rapidly urbanizing developing world. Population data on a country's cities ultimately rely on national statistical agencies. This implies not only large cross-country variations of data quality and availability across time, but also in terms of definitions of what constitutes a city. This is exemplified by a look at the database <https://www.citypopulation.de> (Brinkhoff, 2017), a regularly updated compilation of national census statistics which has been used by many cross-country studies (e.g., Henderson & Wang, 2007; Soo, 2005): The listed cities start at 50,000 inhabitants in the UK, 150,000 in Japan, and 15,000 in Egypt. By contrast, the global threshold of 300,000 inhabitants imposed by the database of the UN World Urbanization Prospects misses out on locally important cities in less populous countries. In fact, the World Bank (2009) argues that a threshold of 50,000 people is reasonable for a sizable settlement in developed and developing countries.

In this paper, we systematically examine the size distribution of cities with a truly global data set. We exploit geospatial data as well as recent theoretical advances on the origin of the city size distribution. In particular, we use a consistent city identification scheme across countries based on the European Commission's Global Human Settlement Layers (GHSL) and include all cities with more than 50,000 inhabitants in every country of the globe. In this way, we contour the issues of limited data availability and comparability plaguing earlier studies. Another advantage of working with such geospatial data is that we can define cities based on their de facto geographical extent rather than working with administrative city boundaries. This can capture the economic and social reality of cities more appropriately. With our approach, we arrive at a data set of 13,844 cities in 194 countries and territories.¹ Aspiring to be the *Zipf paper for the geospatial age*, our paper focuses on the following three investigations: (a) We analyze for each country whether Zipf's law holds in terms of population and examine deviations. (b) In addition to population, our geospatial data allow us to measure city size by nighttime lights proxying for economic activity. By comparing how the distribution of nighttime lights deviates from the distribution of population, we shed light on agglomeration economies. To our knowledge, we are the first to investigate the size distribution of economic activity in countries around the world in comparison to that of population. (c) We also investigate the geographical, institutional, and historical determinants of the cross-country variation we observe in the city size distribution, and we look at changes over time.

More broadly, with our data analysis we extend the Zipf question into a general discussion of how to measure cities' economic and social importance. Satellite data of nighttime lights have been shown to be an appropriate proxy for local economic activity (Donaldson & Storeygard, 2016; Henderson, Storeygard, & Weil, 2012) and are increasingly used for the study of cities in developing countries where up-to-date population data are lacking

¹We will use the term *city* for all observations in our data set, while keeping in mind that there are some agglomerations that do not reflect one but several cities in the administrative sense, such as the Pearl River Delta in China.

(Bluhm & Krause, 2018; Fetzter, Henderson, Nigmatulina, & Shanghavi, 2016; Storeygard, 2016). But it is an open question whether light per capita is the same for all cities in a country and how light output in cities responds to changes in population. We address these points by comparing population and light across the whole city size distribution. We examine the particular role played by primary cities, which have been linked to autocratic structures (Ades & Glaeser, 1995), political centralization (Davis & Henderson, 2003), and low levels of development in general (Henderson, 2003). If luminosity in the largest cities is more responsive to population changes, it suggests that positive agglomeration effects are being reaped. But if primary cities are systematically brighter than the rest and relatively inelastic to population changes, other factors might play a role, such as wasteful political spending (Lipton, 1977).

Going further, our paper exploits the geospatial structure of the data and decomposes city size, in terms of both light and population, into the product of area and density. While larger cities are typically both more extended and denser than smaller ones, we examine which factor dominates. This carries some insights for the types of agglomeration effects at play: Economies of scale and the wider market access of bigger cities (Fujita, Krugman, & Venables, 1999; Krugman, 1991) mainly operate through a larger area, while more frequent human capital interactions, as highlighted by Moretti (2004) and Bettencourt (2013), can be expected to work through a higher density.

The main results of our paper are the following: (a) For many countries, the city size distribution in terms of population can be characterized by Zipf's law, given the appropriate threshold is applied. (b) The city size distribution of light is more unequal than that of population and Zipf's law in light can be rejected for most countries. The distributional difference between light and population is particularly pronounced in African countries. (c) Deviations from Zipf's law can mostly be explained by the top end of the distribution, with primary cities being disproportionately populous and bright. In the biggest cities, economic activity is particularly concentrated and light is less sensitive to population changes. (d) Decomposing city size into density and area, our results are mostly driven by the area effect of the largest cities. This points towards a particular importance of agglomeration effects of scale and market access. (e) The observed cross-country variation in the city size distribution can mostly be accounted for by historical factors, which lends weight to the time of development hypothesis by Henderson, Squires, Storeygard, and Weil (2018). Overall, the city size distribution has exhibited little variation throughout the last decades. Still, the higher growth rates in primary cities are noteworthy and might lead to a more unequal city size distribution for several countries in the future.

The remainder of this paper is structured as follows: In Section 2 we link our paper to the literature on the size distribution of cities and urban primacy. In Section 3, we describe our city identification scheme, the resulting data set of city size in terms of light and population, as well as the econometric estimation approach. Section 4 contains the empirical estimation results on Zipf's law and their implications, while Section 5 investigates the determinants behind the cross-country variation in the city size distribution and gives an outlook. Section 6 concludes. The supporting information appendix contains the accompanying material, such as supplementary information on the data set, a simulation exercise, detailed results for each country as well as numerous robustness checks.

2 | RELATED LITERATURE

As focal points of economic and social activity, cities and their size have attracted researchers' interest for a long time. It was first suggested by Auerbach (1913) that the size distribution of cities in terms of population follows a Pareto distribution

$$N_y = A \cdot y^{-\alpha} \quad (1)$$

with N_y as the number of cities larger than population size y and shape parameter α . The special case of $\alpha = 1$, and constant A equal to the size of the largest city, is referred to as Zipf's law (Zipf, 1949).

Various seminal papers have focused on the underlying theoretical processes. From the homogeneity of cities' growth processes with the same rate and variance independent of their size—so-called Gibrat's (1931) Law—one can derive that the entire distribution of all cities and towns should be lognormal (Eeckhout, 2004). But augmenting this homogeneous growth process with a lower bound of city size and some shocks yields a distribution that is Pareto at the top (Gabaix, 1999). A Pareto distribution in city size also emerges from other theoretical consideration, for example, the combination of cities' agglomeration and congestion effects (Rossi-Hansberg & Wright, 2007), the interplay between industry-specific shocks and firms' location decisions (Duranton, 2007), or a combination of sorting of individuals, productivity of firms as well as agglomeration effects (Behrens, Duranton, & Robert-Nicoud, 2014).²

Most of the empirical papers on the topic focus on the United States and provide supportive evidence of a Zipf distribution for the biggest US cities (see, e.g., Fazio & Modica, 2015; Rozenfeld et al., 2011); for example, Gabaix (2016) obtains an alpha coefficient estimate of 1.03. The implication of Zipf's law that a city's rank is approximately inversely proportional to its size is neatly illustrated in the United States, where the biggest city (New York) has twice the population of the second-ranked city (Los Angeles) and three times the population of the third-ranked city (Chicago).³

This paper sets itself apart from these works by its global focus, investigating Zipf's law in countries around the world. It follows earlier cross-country studies which were limited in their data availability and comparability of city definitions. For example, Rosen and Resnick (1980) use census data from 44 countries from the 1970s; Soo (2005) works with the <https://www.citypopulation.de> data based on national statistical offices from 73 countries. But how cities are defined in terms of minimum size and administrative borders varies considerably across countries. It is therefore not clear to what extent the variability of alpha estimates across the sample—Rosen and Resnick's (1980) Pareto alphas range from 0.809 (Morocco) to 1.963 (Australia)—is due to systematic deviations from Zipf's law or how much may be due to measurement issues. With our globally consistent city identification scheme and our threshold discussion of the Pareto tail, we provide a thorough treatment of these topics which was lacking until now. Also, Brakman, Garretsen, Marrewijk, and van den Berg (1999) have already remarked that there is more evidence in favor of Zipf's law when agglomerations rather than core cities are considered as these tend to extend around the largest metropolises. With geospatial data, we are able to investigate this argument thoroughly. Although geospatial data are now widely used in regional and development economics, there only seems to be one other Zipf-related paper which measures the urban extent based on geospatial data: Small, Elvidge, Balk, and Montgomery (2011) look at the size distribution of the world's largest metropolises independent of countries.⁴

To our knowledge, our paper is also the first to look at the size distribution of cities in terms of economic activity proxied for by light, in addition to population. With this, we add to the core discussion in urban economics on how population and economic output are related (Ciccone & Hall, 1996; Glaeser & Gottlieb, 2009). Productivity is known to rise with population density thanks to the presence of agglomeration effects, which would suggest that output scales supralinearly with respect to city size (Bettencourt, 2013; Bettencourt & West, 2010; Lobo, Bettencourt, Strumsky, & West, 2013). While we cannot measure agglomeration effects directly, our data set puts us in a novel position to compare the size distributions of light and population. Nighttime lights have become established as a proxy for local economic activity and are increasingly used in economic research (see, e.g., Henderson et al., 2018; Lessmann & Seidel, 2017); however, its precise scaling relation to population is

²There is also a strand of the literature which focuses on the size distribution of all urban settlements within a country, from small villages over towns to the largest cities. Giesen, Zimmermann, and Suedekum (2010) show that Pareto tails at both ends combined with a lognormal body provide a good fit for the entire settlement distribution.

³Note that even if Zipf's law holds exactly, this rank-size association remains an approximation. It is typically imprecise for the lower-ranked cities and provides a more adequate representation of the many higher-ranked cities (Gabaix & Ioannides, 2004).

⁴Their supportive evidence of Zipf's law in a sample of the world's largest cities irrespective of countries is intriguing. Yet, for policymakers, the city size distribution at the national level is arguably more relevant, and in the absence of a friction-free movement of people and capital across borders, the theoretical processes underlying city growth are more likely to hold at the national than at the global level (Cristelli, Batty, & Pietronero, 2012).

underexplored. Nighttime lights capture consumption, investment, government spending, in particular in public infrastructure (Elvidge, Hsu, Baugh, & Ghosh, 2014; Henderson et al., 2012), but the relation between light and population may be different in cities of different size. For example, primary cities might own their outsized role to both agglomeration effects (Moretti, 2004; Rosenthal & Strange, 2004) and disproportional public resources (Ades & Glaeser, 1995). One reason why nighttime lights have so far not been widely used for the study of largest cities is also the top-coding problem: The classical Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) satellites fail to capture the brightness of big cities due to sensor saturation, thereby underestimating their economic size. In addition to the original data, we therefore work with the top-coding corrected lights data by Bluhm and Krause (2018) to contour this problem and measure the city size distribution of lights in an unbiased way. Under the assumption of supralinear scaling, our a priori hypothesis is that lights should be distributed more unequally than population. We are going to investigate whether a Zipf distribution can therefore be rejected in terms of lights.

By investigating Zipf's law in each country and its patterns of deviations and determinants in different world regions, our work is also related to continent-specific studies of cities. In particular Africa with its rapidly increasing urbanization rates, high primacy ratios, and often insufficient urban infrastructure is the topic of numerous papers (such as Barrios, Bertinelli, & Strobl, 2006; Castells-Quintana, 2017; Christiaensen & Todo, 2014; Jedwab & Vollrath, 2019). It has been argued that the optimal city size distribution might depend on a country's level of development: Urban concentration is thought to be helpful for poor countries with weak physical and human capital resources as well as high transport costs, while a more balanced city size distribution should be beneficial for more advanced economies (Henderson, 2003; Krugman, 1991). Our worldwide data set allows us to investigate to what extent deviations from Zipf's law follow this argument, as well as to give an outlook of the future city size distribution and its consequences.

3 | DATA AND METHODOLOGY

3.1 | City identification and city size measurement

The key contribution of this paper is to conduct a global cross-country study of the size distribution of cities using a consistent city identification scheme.⁵ We employ the European Commission's GHSL to identify urban settlements hosting more than 50,000 residents in 2015.⁶ The lower bound of 50,000 inhabitants has been found to be reasonable for agglomeration effects to take place in countries around the world (World Bank, 2009). To determine the de facto spatial extent of a city, the GHSL data combine Landsat satellite imagery on built-up area with census information (Pesaresi & Freire, 2016). Unlike other possible identification methods derived from the nighttime lights data itself (Small et al., 2011), this method does not suffer from overflow and identification problems at the bottom of the distribution. The global data grid is of a 1-km resolution. Figure 1a illustrates for the case of Paris that the identified urban boundaries (in blue) match the economic and social reality of a city much better than administrative borders (in red). Technical details on the city identification scheme, including the treatment of border cities and a robustness check with shape files from other years, are given in Supporting Information Appendix A. In this way, we identify 13,844 cities in 194 countries and territories. Note that this baseline data set still contains a number of small entities, such as Gibraltar, Réunion or St. Vincent, and the Grenadines, which are

⁵For the reasons outlined above, we use a geospatial data set of cities based on a comparable definition for our cross-country analysis. But as a direct comparison with the existing literature, we repeat our analysis with the <https://www.citypopulation.de> data (Brinkhoff, 2017), which relies on definitions from national statistical agencies. This data set differs from ours in the number of cities included, the available years, and the urban extents of the cities (administrative borders rather than geodata-based extent). The results are discussed in Supporting Information Appendix E.

⁶Since their introduction a few years ago the GHSL strongly gained in popularity and are now widely used in the literature (Stokes & Seto, 2019; Weiss et al., 2018).

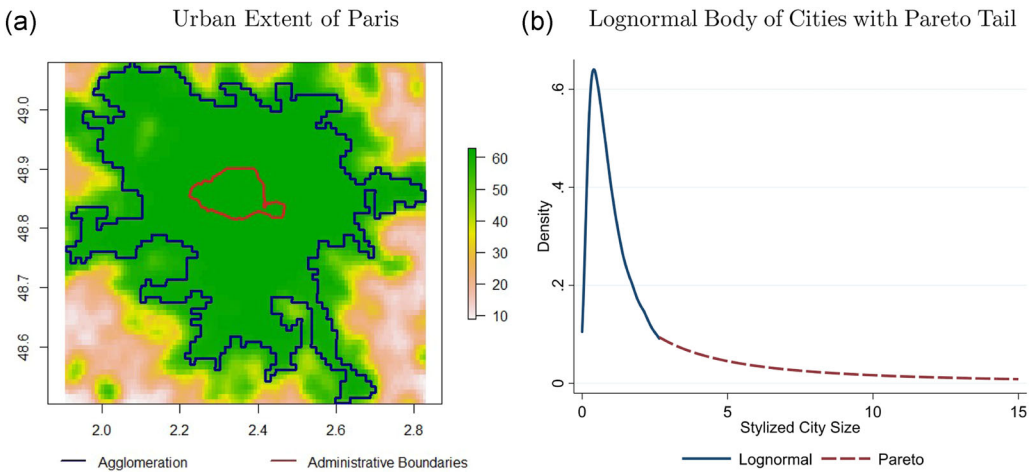


FIGURE 1 Measuring big cities. (a) Urban extent of Paris and (b) lognormal body of cities with Pareto tail. Note: (a) shows the administrative boundaries of Paris as well as the boundaries identified by GHSL. Lit areas (stable lights) are shown in the background. (b) is a stylized city size distribution in line with the literature, showing a lognormal body of towns and smaller cities and a Pareto tail of big cities. [Color figure can be viewed at wileyonlinelibrary.com]

either not independent countries or too small for a sensible estimation for the size distribution of cities. We later on restrict our sample for individual regressions.

For all the identified cities, we use two measures of size, namely, georeferenced population and the sum of lights within their urban extents.

We take the population data from the same data source as the city shapes, the GHSL. They combine Landsat information on built-up area with NASA's Socioeconomic Data and Applications Center's fourth version of the Gridded Population of the World (GPW) to derive a population grid at the same resolution as the city shapes. Hence, we obtain population data for each of our identified cities as a panel for the years 1975, 1990, 2000, and 2015.

In contrast to this, our nighttime light data form a yearly panel from 1992 to 2013. We use the “stable night light images” collected by the DMSP-OLS operated by the National Oceanic Administration Agency (NOAA). The satellites monitor light emissions between 8.30 and 10.00 p.m. local time on a daily basis, and the published pictures are yearly averages of light emissions on cloud-free days. Corrected for glare, auroral lights, forest fire, and gas flares, the resulting lights are assumed to be exclusively man-made. Light emissions at the pixel level are measured by a Digital Number (DN) ranging from 0 (dark) to 63 (fully illuminated). This “stable lights” data set has been used extensively in applications in development and regional economics, see Donaldson and Storeygard (2016) for an overview. However, Bluhm and Krause (2018) point out that the “stable lights” suffer from top-coding. Due to sensor saturation, the satellite cannot capture the full brightness of the biggest cities in which many pixels reach the end of the scale at 63 DN. As big cities form the focus of our analysis, we use the top-coding corrected nighttime lights provided by Bluhm and Krause (2018) as an additional data source. For both the “stable” and corrected lights separately, we add up the DN's for all the pixels within the city boundaries to obtain the total luminosity of the city, which is a standard procedure in the nighttime light literature (Henderson et al., 2012).

Table 1 gives an overview of our data set in the year 2000 by presenting summary statistics for all the cities of the world (13,844) as well as those in some selected countries. While the median city has a population of 85,530 inhabitants, the world's largest city has more than 32 m inhabitants (the Pearl River Delta agglomeration in China covering inter alia Shenzhen and Guangzhou). Comparing, for example, the 323 cities in the US and 514 in Bangladesh, we see that the median cities in both countries are of similar size in terms of population but not in terms of sum of lights

**TABLE 1** Summary statistics of the data set in the year 2000

	USA	DEU	CHN	NGA	BGD	World
Number of cities	323	86	2,266	428	514	13,844
Median city size						
Population	100,909	105,124	99,455	72,434	98,226	85,530
Stable light	7,295	4,153	474	53	34	284
Corrected light	16,485	4,833	474	53	34	284
Maximum city size						
Population	14,853,624	7,477,014	32,343,639	7,789,496	15,452,476	32,343,639
Stable light	474,339	299,255	502,338	43,631	58,281	502,338
Corrected light	1,983,732	379,765	845,757	45,687	62,815	1,983,732

Note: The sums of light within the city boundary are measured in DN.

(7,295 DN vs. 34 DN). This underlines the importance of analyzing the city size distribution in each country separately and thus accounting for different development levels and country-specific heterogeneity. The table also shows that the differences between the “stable” and the top-coding corrected lights are larger in richer countries—and, for all countries, they are larger in the biggest cities than at the median. Note that our definition of cities based on the actual urban extent leads to some agglomerations topping the list of the largest cities: The Ruhrgebiet (Essen, Duisburg, etc.) is the most populous “city” in Germany rather than Berlin, Germany's largest city measured by administrative boundaries. The case of the Pearl River Delta in China is similar. By contrast, for the United States (New York), Nigeria (Lagos), and Bangladesh (Dhaka), the name of the most populous city coincides when using the GHSL's urban extent or administrative data.

3.2 | Estimation approach

To test for Zipf's law, we have to limit the analysis to those 103 countries from our data set with a sufficient number of cities, set to 10 here.⁷ We follow the literature in assuming that the size distribution within each country consists of a lognormal body of towns and smaller cities on the one hand and the Pareto tail of big cities on the other hand (Figure 1b). Yet, it is an open question where the threshold is located. While our data set only contains cities with at least 50,000 inhabitants, a Chinese city of that size might still belong to the lognormal part of its country's distribution. Ioannides and Skouras (2013) argue that in the United States the switch between the two portions of the distribution occurs between 30,000 and 60,000 inhabitants, but a cross-country discussion of this issue is missing so far. In Supporting Information Appendix B, we conduct a Monte Carlo simulation which shows the potential distortion of the coefficient estimate when the Pareto estimation is carried out using too low a threshold. We also implement a numeric identification algorithm to identify the threshold between the lognormal body and Pareto tail for each country. We find that across our data set, setting a country's median city size as the threshold solves the tradeoff between ensuring a pure Pareto tail with a rather high threshold and keeping as many observations as possible with a rather low threshold. We therefore follow a twofold strategy in our empirical estimation: In each country, we use (a) all cities in our data set and (b) only those cities above the median.⁸

⁷In Supporting Information Appendix B, we motivate why we set a minimum requirement in the quantity of given cities, showing that Pareto tails cannot be established empirically for tiny countries' fewer cities.

⁸We also test other thresholds. In Supporting Information Appendix B we discuss a linear combination of a cutoff that is first horizontal and then increasing in the number of cities per country, analogous to the literature on international poverty lines (see, e.g., Ravallion & Chen, 2011).

For our actual estimation, we use log-rank regressions, as is usually done in the literature. If city size y is Pareto distributed with shape parameter α above the threshold y_c as determined above, we have $\text{rank}(y) \approx Ny_c^\alpha y^{-\alpha}$, or, in logarithms, $\log \text{rank}(y) - \log N \approx \alpha \log y_c - \alpha \log y$. OLS estimation of this equation underestimates the true coefficients and standard errors in small samples due to the ranking procedure (Gabaix & Ioannides, 2004). However, subtraction of one half from the rank has been shown to improve the estimation of the Pareto alpha (Gabaix & Ibragimov, 2011) so that we will estimate the following log-rank regression by OLS:

$$\log\left(\text{rank}(y) - \frac{1}{2}\right) - \log(N) = \text{cons.} - \alpha \cdot \log(y) + \epsilon \quad (2)$$

with the corrected OLS standard errors given as $\sqrt{2/N} \cdot \hat{\alpha}$ by Gabaix and Ibragimov (2011).⁹

4 | RESULTS ON ZIPF'S LAW AROUND THE WORLD

4.1 | Results in terms of population and light

Running (2) individually for all countries with at least 10 cities, and measuring city size by population and light, we determine whether or not Zipf's law holds.¹⁰

Figure 2 shows the densities of the obtained Pareto alpha coefficients in the year 2000. The main insight is that Zipf's law is an appropriate characterization of the city size distribution for many countries in terms of population but not in terms of economic activity as proxied for by light.

The green curve of the population alpha coefficients is centered around 1, and for 77% of countries the coefficients lie within 95% confidence bands around $\alpha = 1$.¹¹ When using the full distribution of cities above 50,000 inhabitants in each country (Figure 2a), the alpha density is slightly more left-skewed than when only including cities above the median (Figure 2b). As the former potentially features observations from the lognormal body, inequality in the distribution is slightly overstated. Nevertheless, the mode at $\alpha = 1$ is a robust characteristic. Our result that an approximately equal share of countries has coefficient estimates slightly smaller and larger than 1 differs from the previous literature, which typically finds a dominating share of countries with larger Pareto alphas (Rosen & Resnick, 1980; Soo, 2005). Apart from the larger number of countries in our data set and the consistent city identification scheme, this may be due to the different urban extents we measure. Administrative boundaries fail to capture the economic and social extent of larger cities and suggest a more egalitarian city size distribution than in our results.

When comparing our results from the different samples, we note that hypotheses of alphas equal to or smaller than 1 are more difficult to reject in short than they are in long city size distributions. The restriction to above-median cities overcomes the bias at the cost of a larger variance. The resulting bigger confidence intervals therefore include a larger set of values apart from $\alpha = 1$, for example, $\alpha = 0.8$.¹²

Turning towards light, Figure 2 clearly shows that the Pareto alphas are more unequally distributed than they are for population. This confirms our prior hypothesis. The distributions of the light Pareto alphas are centered around coefficient estimates smaller than 1. Only 52% of "stable light" alpha estimates, and 39% of "corrected light" estimates, lie within a 95% confidence interval around $\alpha = 1$. In line with our expectations, top-coding corrected lights are even more inegalitarian than the "stable lights" as they capture the full brightness of the largest cities.

⁹An alternative to OLS is the Hill (1975) estimator, which is the maximum likelihood estimator if the data are Pareto distributed but which is less robust to deviations. In Supporting Information Appendix D we repeat our analysis with the Hill estimator as a robustness check for our results. They are very similar for most countries; for example, the correlation between the OLS and Hill alpha estimates is 0.813 for "stable lights" in the above-median sample.

¹⁰An alternative to our countrywise approach would be to merge the city size distributions of strongly interacting countries (Brakman, Garretsen, & van Marrewijk, 2019).

¹¹In Supporting Information Appendix C we list Pareto alpha coefficients for all countries with their standard errors.

¹²Point estimates and standard errors are listed in Supporting Information Appendix C.

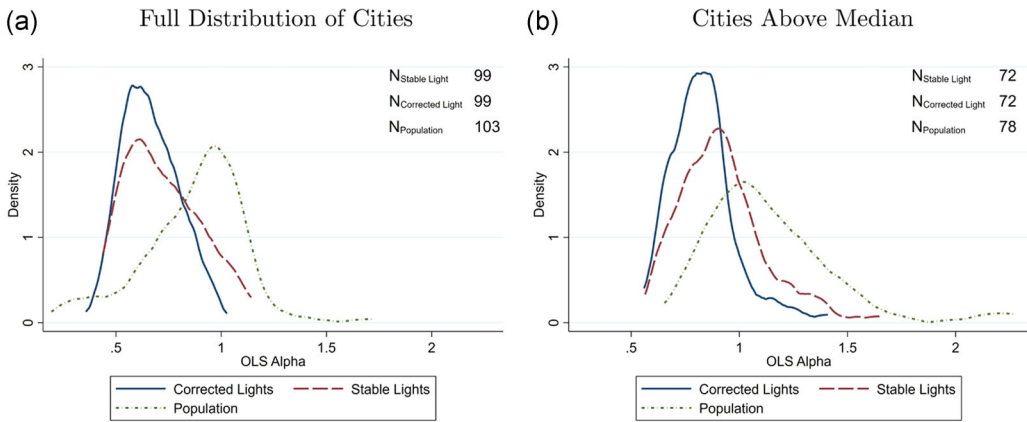


FIGURE 2 Density plots of countries' estimated Pareto alphas in the year 2000. (a) Full distribution of cities (b) cities above median. Note: The estimated Pareto alphas of all countries with more than 10 cities. Figure 2a estimates (2) for the whole distribution of countries and Figure 2b uses only cities above the median. Light is based on satellite F15. [Color figure can be viewed at wileyonlinelibrary.com]

This is particularly evident when only cities above the median are considered, of which large metropolises command a larger share (Figure 2b).¹³

There is, however, evidence that the size distribution of cities in terms of light follows a Pareto distribution; it is just not the special Zipf case of $\alpha = 1$ for many countries. Figure 3 uses the example of the two otherwise very different countries of Germany and South Africa, which both display remarkably linear Zipf plots in terms of population, “stable” and corrected light. It is the slopes of the OLS best-fit lines that differ and deviate slightly more strongly from the green Zipf line ($\alpha = 1$) for light than for population.

Exploiting the global nature of our data set, we assemble the population and “stable light” alpha coefficients for available countries in the scatter plot in Figure 4. Two lines representing Zipf's law dissect the plot into four quadrants. The clustering of countries in the lower left and upper right quadrants shows that those whose city size distribution is slightly more (less) egalitarian in terms of population is also more (less) egalitarian in terms of light. Yet, the sizable number of countries in the upper-left quadrant of more equality in the population size distribution and more inequality in light suggests a more nuanced relation. Clearly, population increases do not always translate one to one into increases in economic activity as proxied for by light, but the effects may differ across the city size distribution, as we will examine in the following.

Furthermore, we note patterns of cross-country heterogeneity in the estimated coefficients: Table 2 shows that on average, population and light are distributed more equally in Europe than in Africa, Asia, and the Americas. In terms of population, this confirms the cross-country pattern observed by Soo (2005). Going beyond that, our results show that (a) this holds for light as well, even as light is more unequally distributed than population, (b) working with the full distribution of cities rather than those above the median yields coefficients of a larger magnitude but with similar cross-continent patterns, and (c) the distributional differences between light and population are not as pronounced in Europe as on other continents. Across African countries, the mean alpha estimate is 0.746 for the above-median corrected light distribution and 0.609 for the full distribution, indicating strong deviations from Zipf's law.

¹³The rightward shift in all three distributions when the sample is restricted from the full sample to above-median cities is in line with the threshold discussion in Section B.1. The inclusion of the lognormal body induces a downward bias in the OLS Pareto alpha estimates.

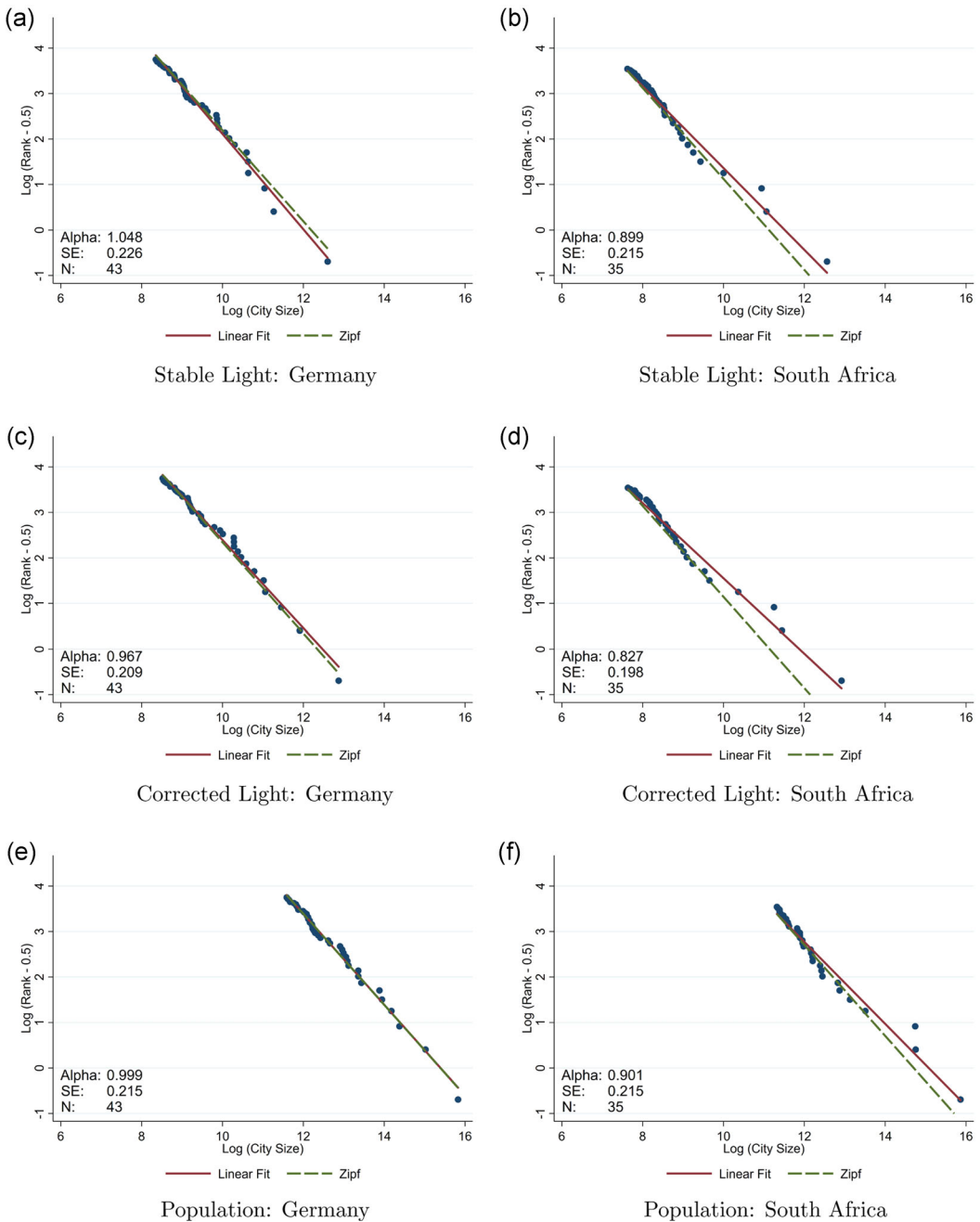


FIGURE 3 Zipf plots for Germany and South Africa (above-median setting, year 2000, satellite F15). (a) Stable light, Germany; (b) stable light, South Africa; (c) corrected light, Germany; (d) corrected light, South Africa; (e) population, Germany; and (f) population, South Africa [Color figure can be viewed at wileyonlinelibrary.com]

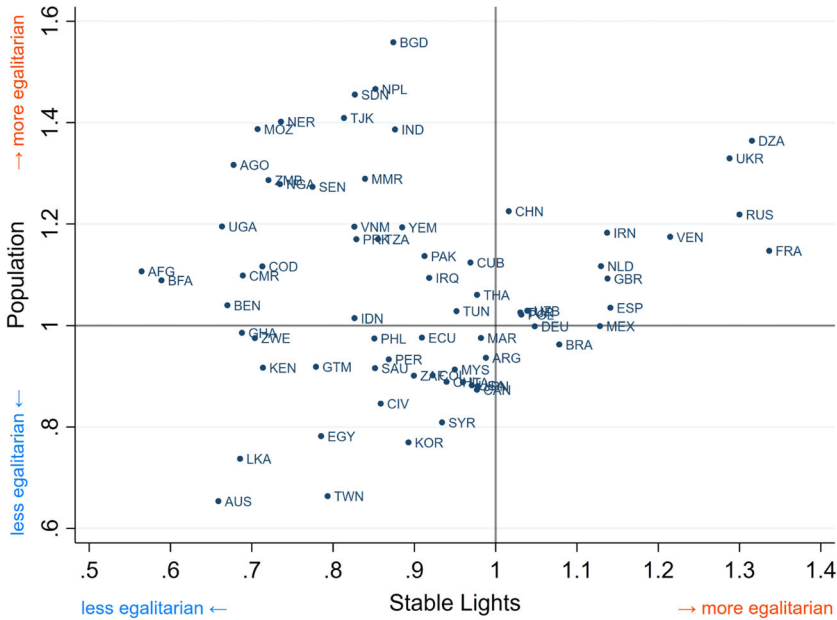


FIGURE 4 Scatter plot of population and stable light Pareto alpha coefficients (above-median setting, year 2000, satellite F15, correlation: 0.06, omits outliers: Kazakhstan, Romania, and Ethiopia) [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Summary statistics for Pareto alphas (year 2000)

	Africa	Americas	Asia	Europe	World
<i>Stable light</i>					
Full dist.					
Mean	0.626	0.706	0.643	0.955	0.709
(SD)	(0.126)	(0.153)	(0.125)	(0.130)	(0.178)
N	29	17	32	17	99
Above median					
Mean	0.776	0.964	0.886	1.145	0.904
(SD)	(0.156)	(0.118)	(0.161)	(0.129)	(0.204)
N	22	12	26	9	72
<i>Corrected light</i>					
Full dist.					
Mean	0.609	0.624	0.607	0.829	0.654
(SD)	(0.102)	(0.096)	(0.104)	(0.106)	(0.135)
N	29	17	32	17	99
Above median					
Mean	0.746	0.810	0.816	0.979	0.817
(SD)	(0.113)	(0.082)	(0.128)	(0.133)	(0.152)
N	22	12	26	9	72

(Continues)

TABLE 2 (Continued)

	Africa	Americas	Asia	Europe	World
<i>Population</i>					
Full dist.					
Mean	0.811	0.891	0.875	0.989	0.868
(SD)	(0.263)	(0.117)	(0.271)	(0.085)	(0.244)
N	31	17	32	17	103
Above median					
Mean	1.195	0.955	1.152	1.094	1.137
(SD)	(0.294)	(0.098)	(0.331)	(0.130)	(0.308)
N	25	13	27	9	78

Note: The table presents the summary statistics of OLS alpha estimates by continent. Luminosity in the year 2000 is here defined as the average of the two DN values obtained from the two satellites that were active in that year (F14 and F15). Asia includes Oceania.

4.2 | Explaining patterns along the city size distribution in population and light

Let us now examine which part of the distribution is driving the deviations from Zipf's law. We compute primary shares as the proportion of the distribution's total population or light that is accounted for by the largest city.

Figure 5 shows that higher primary shares are strongly correlated with smaller Pareto alphas, both in terms of population and light.¹⁴ The best-fit line for light is flatter, so that a smaller Pareto alpha is associated with an even larger primacy share for light. We see that, overall, countries with a more unequal city size distribution exhibit an undue concentration of population and economic activity in primary cities.

How does this concentration in primary cities relate to light being less equally distributed than population? Emitted light per capita is, obviously, not constant throughout the distribution. And the link between population and light, that is, economic activity, varies with city size. We therefore run regressions of the logarithm of total light emissions in city i in country j and year t on the logarithm of population, including country- and year-fixed effects, δ_j and γ_t :

$$\log(\text{Light}_{ijt}) = \beta_0 + \beta_1 \log(\text{Population}_{ijt}) + \delta_j + \gamma_t + \varepsilon_{ijt}. \quad (3)$$

According to the estimates displayed in the first two columns of Table 3 a 1% larger population is associated with a significantly more than a 1% higher light output.¹⁵ This significance is unrelated to the displayed asterisks referring to $H_0: \beta = 0$, but can easily be derived from the given coefficients and standard errors. The stronger impact on corrected lights arises from the fact that the total luminosity of larger cities is more severely underestimated by top-coding than it is the case for smaller cities. To see whether the elasticity is different for primary cities compared with the rest of the distribution we add Primacy_{ijt} , a dummy that equals one for the largest city in terms of population in a country in a certain year, and its interaction with $\log(\text{Population}_{ijt})$ to the model. We can see that primary cities are ceteris paribus on average brighter than other cities, irrespective of population, but their light emissions respond less strongly to population level variation.¹⁶ That the coefficient estimate for population is larger than one can be understood by nighttime lights as a combination of population and activity per

¹⁴The correlation coefficients are -0.55 in Figure 5a and -0.60 in Figure 5b.

¹⁵As robustness checks we repeat the estimation in a Seemingly Unrelated Regression Equations (SURE) framework, which leads to virtually identical results. Regressions with modifications along the fixed effects dimensions suggest that both, inter-temporal and cross-sectional variation, drive the results. Replacing $\log(\text{light})$ with $\log(\text{light per capita})$ intuitively only decreases the β_1 coefficient estimate by 1 and alters the adjusted R^2 , keeping all else equal.

¹⁶When clustering standard errors at the city rather than country level, the coefficient of the interaction term coefficient is strongly significant for both light measures. Given that Primacy rarely moves between cities over time, city fixed effects render the coefficient estimate insignificant.

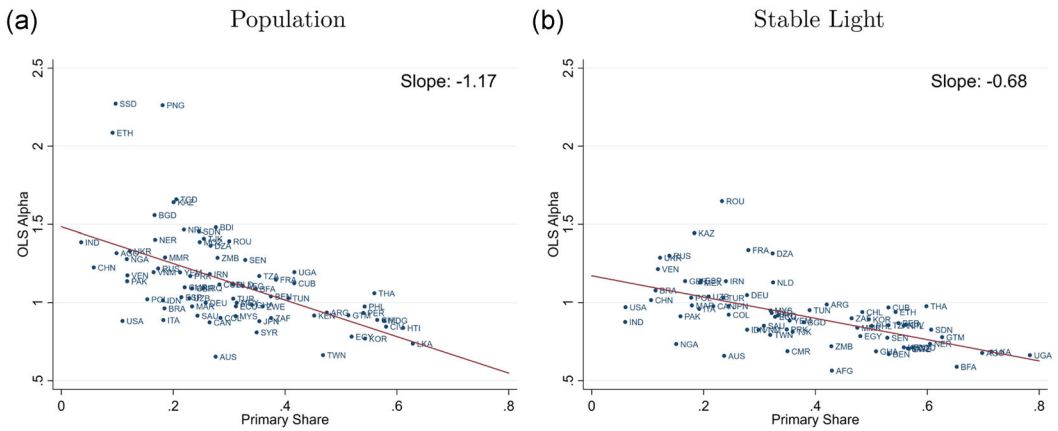


FIGURE 5 Primary share and estimated Pareto alpha in the year 2000. (a) Population and (b) stable light. *Note:* The scatter plots show the relation for the estimated Pareto alphas (above-median setting, all countries with more than 10 cities) with the primary share, defined as the share of the distribution’s population or total luminosity that is due to the largest city. The stable light alphas are based on the satellite F15. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Light–population elasticities

Dependent variable: log(Light)						
	Stable (1)	Corrected (2)	Stable (3)	Corrected (4)	Stable (5)	Corrected (6)
log(Pop.)	1.098*** (0.045)	1.171*** (0.050)	1.094*** (0.050)	1.160*** (0.056)	1.080*** (0.068)	1.127*** (0.076)
Primacy			2.419*** (0.927)	1.874** (0.943)		
log(Pop.) × Primacy			-0.155** (0.068)	-0.106 (0.072)		
TopTen					0.865 (0.926)	-0.122 (1.014)
log(Pop.) × TopTen					-0.048 (0.079)	0.036 (0.087)
Constant	-8.273*** (0.600)	-9.120*** (0.662)	-8.239*** (0.659)	-9.016*** (0.728)	-8.139*** (0.839)	-8.718*** (0.931)
Country F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
N	254,689	254,689	254,689	254,689	254,689	254,689
Adj. R ²	0.730	0.755	0.730	0.755	0.731	0.756

Note: Standard errors clustered at the country level in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Primacy is a dummy that equals one for the largest city in terms of population in the respective year and country. TopTen is a dummy that equals one for the 10 largest cities in terms of population in the respective year and country.

capita, which might be expected to scale superlinearly with city size, in the terms of Bettencourt (2013). But the particular results for primary cities require an additional interpretation. We argue that they are in line with theories and empirical evidence by, *inter alia*, Ades and Glaeser (1995), Henderson and Wang (2007), and Gollin, Kirchberger, and Lagakos (2017). These authors find that disproportionately many resources are pooled into primary cities, and that primary cities play an outsized political, social, and economic role. Potentially because they are already bright due to their special status, their emitted luminosity responds less strongly to population growth than that of smaller cities. Regressions depicted in the last two columns of Table 3 show that this effect is indeed specific to primary cities and not just a big city effect: Replacing the Primacy_{ijt} dummy with TopTen_{ijt} , which equals one for the 10 most populated cities per country and year, yields insignificant results for the dummy and the interaction term.

That primary cities in countries around the world behave so differently from their second- and third-largest peers is worth a closer look. With their outsized position, they benefit from agglomeration effects particularly strongly: They are home to many firms and skilled workers, are often the center of political power, and offer a vibrant social and cultural life. This adds to their productivity, which, in turn makes the city more attractive to newcomers, further propelling its role. The literature on agglomeration effects and positive externalities splits these factors into a scaling and a density effect: More extended cities offer larger markets with more opportunities for both firms and workers (Fujita et al., 1999; Krugman, 1991) thanks to their pure size, while denser cities facilitate the interactions of individuals (Bettencourt, 2013) and make them benefit from human capital externalities (Diamond, 2016; Moretti, 2004).

While our global data set does not lend itself to a micro-analysis of wages, firms, and human capital externalities, we can shed some light on the aggregate level and conduct a decomposition of city size. Let us keep in mind that city size—be it in terms of population or light—of city i at time t is the product of area and density.

$$\text{Size}_{it} = \text{Area}_{it} \cdot \text{Density}_{it}. \quad (4)$$

This means that primary cities can be bigger than secondary cities for two reasons: either they extend to a larger area or they have a higher density of population. Obviously, we often have a combination of both contributors, but they are not equally important. We divide city size, area, and density for the primary and secondary city of each country and summarize the mean values across measures and continents in Table 4.¹⁷

Several observations can be made: (a) Across the world, a country's primary city is on average 4.3 times as populous and 4.6–5.5 times as bright as the city at rank two. These are sizable proportions, given that the Zipf's law predicts a factor of two, albeit with large confidence bands (Gabaix & Ioannides, 2004). We conclude that primary cities are disproportionately large, underlining their special role (Ades & Glaeser, 1995; Fetzer et al., 2016; Storeygard, 2016). (b) In Africa, primary cities are more than 4 times as populous and about 7–8 times as bright as secondary cities, a higher factor than on any other continent. This goes in line with the well-known high primacy share in Africa (Henderson & Wang, 2007; Junius, 1999). Europe, by contrast, has the lowest proportions between primary and secondary cities and they are of a similar magnitude for population and light. (c) The size differences between primary and secondary cities can mostly be attributed to differences in area rather than density. Across all continents it holds that the largest cities are only slightly denser—in terms of both light and population—but much more extended than secondary cities. Hence, area appears to be driving inequality at the top of the city size distribution, which generalizes the observations by Rozenfeld et al. (2011) for the United States and the United Kingdom to a worldwide setting.

¹⁷These results are for the latest available years, 2013 for light and 2015 for population. They include the countries with 10 or more cities for which we calculate the Pareto alphas. Pooling the data across all available years as well as including countries with fewer than 10 cities does not qualitatively change the results, see Supporting Information Appendix F.

TABLE 4 Comparing primary and secondary cities (countries with 10 or more cities), years 2013/15

	World	Africa	Americas	Asia	Europe
<i>Population</i>					
Size	4.322 (4.032)	4.661 (2.924)	5.011 (3.545)	4.192 (5.456)	2.871 (2.216)
Density	1.324 (0.905)	1.118 (0.691)	1.459 (0.544)	1.547 (1.255)	1.082 (0.377)
Area	4.241 (4.021)	5.651 (4.764)	3.849 (3.078)	3.559 (3.944)	2.854 (1.914)
<i>Stable light</i>					
Size	4.666 (4.561)	6.819 (6.195)	4.095 (3.201)	3.555 (3.066)	2.704 (1.665)
Density	1.255 (0.590)	1.333 (0.478)	1.099 (0.261)	1.335 (0.833)	1.021 (0.121)
Area	3.792 (3.119)	4.930 (3.707)	3.604 (2.387)	3.181 (2.994)	2.674 (1.604)
<i>Corrected light</i>					
Size	5.507 (5.453)	7.904 (6.819)	5.462 (4.677)	4.240 (4.363)	2.642 (1.574)
Density	1.530 (0.941)	1.522 (0.526)	1.472 (0.791)	1.703 (1.298)	1.141 (0.789)
Area	3.865 (3.163)	4.950 (3.691)	3.642 (2.465)	3.314 (3.139)	2.768 (1.641)

Note: The values are computed as $\frac{1}{N} \sum_{i=1}^N \frac{\text{PrimaryCitySize}_i}{\text{SecondaryCitySize}_i}$, $\frac{1}{N} \sum_{i=1}^N \frac{\text{PrimaryCityDensity}_i}{\text{SecondaryCityDensity}_i}$, $\frac{1}{N} \sum_{i=1}^N \frac{\text{PrimaryCityArea}_i}{\text{SecondaryCityArea}_i}$ with country i and N as the total number of countries on the respective continent. The respective standard deviations are denoted in parentheses. Asia includes Oceania.

One way to interpret our finding is that agglomeration effects in the largest cities work through scaling rather than through density, which holds lessons for policymakers. A more extended city brings economies of scale and larger product and labor markets, and it can also ease congestion away from packed areas (Rosenthal & Strange, 2004). By contrast, a higher density is associated with more frequent interactions of its inhabitants, with pooled resources and larger productivity (Ciccone & Hall, 1996; Kerr & Kominers, 2015). In the model of Bettencourt (2013), the cost of human interactions increases with the traverse dimension, which is the area of the city. In this context, inncity infrastructure plays a key role. Our results show an outsized concentration of economic activity in the largest cities. But the agglomeration benefits will be limited if cities are so extended and fragmented that they fail to connect their inhabitants. For example, the typically insufficient public infrastructure in many African cities, as inter alia remarked by Castells-Quintana (2017), Lall, Henderson, and Venables (2017), and Bluhm and Krause (2018), can be an obstacle on the way of channeling their outsized primary cities into hubs of productivity.

5 | DETERMINANTS OF THE CITY SIZE DISTRIBUTION AND TIME VARIATION

5.1 | Determinants of cross-country variation

What are the underlying factors that engender such a city size distribution that is more unequal in some countries than in others? Earlier papers have linked the city size distribution to various institutional and geographic factors,

TABLE 5 Time of development as determinants of the city size distribution

	Stable lights		Corrected lights		Population	
	Alpha	Pr. Sh.	Alpha	Pr. Sh.	Alpha	Pr. Sh.
Education in 1950	-0.010 (0.025)	0.009 (0.013)	-0.005 (0.018)	0.002 (0.013)	-0.012 (0.021)	-0.001 (0.011)
Urbanization in 1950	0.006*** (0.002)	-0.003* (0.001)	0.004*** (0.001)	-0.001 (0.001)	-0.004* (0.002)	0.003** (0.001)
GDP p.c. in 1950	-0.008 (0.020)	-0.025** (0.012)	-0.015 (0.014)	-0.031** (0.012)	0.013 (0.016)	-0.029*** (0.009)
Constant	0.758*** (0.032)	0.543*** (0.035)	0.740*** (0.024)	0.567*** (0.034)	1.149*** (0.048)	0.341*** (0.029)
Year-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	1,358	1,742	1,358	1,742	255	324
Adj. R ²	0.217	0.198	0.100	0.150	0.142	0.047

Note: Clustered standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. Education, average years of schooling; urbanization, percent urbanized; GDP p.c., 1,000 dollars (in 2005 PPP); alpha, OLS alpha estimate; Pr. Sh., primary share.

such as total area and population (Rosen & Resnick, 1980), trade openness (Moomaw & Shatter, 1996), infrastructure (Junius, 1999), autocracy (Ades & Glaeser, 1995), governmental decentralization (Henderson & Wang, 2007), fiscal expenditure (Davis & Henderson, 2003), and ethnic fractionalization (Mutlu, 1989).

While all of these factors have been shown to play a role, we are here going to test them in connection with an overarching theory on the evolution of the spatial distribution of economic activity put forward by Henderson et al. (2018). Their key concept is the time of development. According to this argument, the larger spatial equality in economic activity in early developed countries can be traced back to cities' formation in agricultural regions at times when transport costs were still high. This is also in line with Motamed, Florax, and Masters (2014) who use historic population data to show that places with good agricultural quality urbanized earlier. As agglomeration patterns exhibit strong persistence, these structures are still visible today. By contrast, in countries which developed later, when transport costs were already low, fewer and larger cities were built, often in strategically important coastal locations. While Henderson et al. (2018) provide evidence for this theory in terms of the spatial variation of nighttime lights across the total area of countries, our data set puts us in a position to test it using countries' actual city size distribution. As variables proxying for early and late development, we use education, urbanization, and GDP per capita of countries in 1950, just as Henderson et al. (2018).

Table 5 shows the result of the panel regression with year-fixed effects of

$$\text{ineqcitysize}_{it} = \beta_0 + \beta_1 \text{educ1950}_i + \beta_2 \text{urban1950}_i + \beta_3 \text{gdp1950}_i + \gamma_t + \varepsilon_{it}, \quad (5)$$

where inequality in the city size distribution of country i at time t is expressed either by the Pareto alpha estimate (based on the above-median distribution) or the primary share.¹⁸ According to Table 5 countries which were highly urbanized in 1950, that is, early developed, tend to exhibit greater Pareto alpha coefficients in their city size

¹⁸Running a cross-sectional regression yields almost the same estimates. To increase the sample size we use the panel regression (5) and cluster standard errors at the country level, while taking into account year-fixed effects.

distributions nowadays, both in terms of light and population. The negative coefficient estimate of early development on primary cities' relative size supports that notion.

We now investigate the connection between this time of development theory of the city size distribution and other variables, including, but not limiting ourselves to, those that have been used in previous studies. We look at a total of 36 possible explanatory variables ranging from population structure (such as total population, fertility, and migration) over physical geography (such as terrain ruggedness, coastal border, and continent), institutions (such as time of independence, political rights, and fiscal centralization), economic structure (such as agricultural share, energy use, and patent applications), and international connectedness (such as exports and interstate war).¹⁹

Due to the collinearity of this huge number of possible determinants, we employ a model selection approach: We regress the country's Pareto alpha coefficient on up to 7 out of the 36 determinants X_{it} at a time, including year-fixed effects:

$$\hat{\alpha}_{it} = \beta_0 + \sum_{i=1}^l \beta_i X_{it} + \delta_t + \varepsilon_{it} \quad \text{with } l \in \{1, \dots, 7\}. \quad (6)$$

This amounts to 10,739,175 regressions to compare, out of which we select the best ones based on akaike information criterion and bayesian information criterion. Table 6 shows the coefficient estimates of the determinants included in the selected models to explain the size distribution of "stable" light, corrected light, and population. Even using such a purely algorithmic approach, we see the importance of historic variables, such as the year of independence as well as population in 1400, which are clearly in line with the time of development framework by Henderson et al. (2018). The continent dummies also play a large role in all three selected models, capturing in particular a more egalitarian city size distribution in (early developed) Europe. Also, current urbanization rates predict a higher Pareto alpha. The effect of trade-related variables on city size is ambiguous, in line with the literature (Duranton, 2008; Fujita & Mori, 1996). However, the association of coastal proximity with a more unequal city size distribution can be seen in light of the recent works by Bonfatti and Poelhekke (2017) and Jedwab and Moradi (2016), who argue that outward-oriented colonial infrastructure in many developing countries still has an effect on their distribution of economic activity. Taken together, our results suggest to a much stronger extent than previous Zipf-related studies that historical determinants can explain a lot of the cross-country variation of the city size distribution.

5.2 | Changes over time and outlook

What do these results mean in a dynamic perspective? If the city size distribution in terms of light and population can be explained to a large degree by historical factors, it will only change very slowly. It has been shown for particular countries, such as the United States (Black & Henderson, 2003; González-Val, 2010), France (Duranton, 2007), and Japan (Eaton & Eckstein, 1997), that the city size distribution has been rather persistent across decades despite structural economic change and city growth. Here we exploit the panel structure of our global data set to verify this hypothesis for countries around the world: Figure 6 shows that across the available years for population (1975–2015) and light (1992–2013), the alpha coefficients and their cross-country distributions exhibit little variation. In terms of population, the range of Pareto alpha coefficients around 1 seems to have narrowed so that some countries are getting a bit closer to Zipf's law from either side. In terms of corrected light,

¹⁹A complete list of the 36 variables including their sources is contained in Supporting Information Appendix G. We do not use transportation variables, such as the road and rail network, which we consider endogenous to the city size distribution even though they have been used in some other studies, such as Rosen and Resnick (1980). For a discussion of the link between communication technology, another potentially endogenous variable, and the size distribution of cities, we refer to Ioannides, Overman, Rossi-Hansberg, and Schmidheiny (2008).

TABLE 6 Coefficient estimates of selected models

	Stable light alpha	Corrected light alpha	Population alpha
Coastal proximity	-0.002*** (0.000)	-0.001*** (0.000)	
Independence before 1914	0.049*** (0.010)	0.070*** (0.011)	
Between 1946 and 1989	0.060*** (0.009)	0.047*** (0.009)	
After 1989	0.233*** (0.012)	0.203*** (0.011)	
Continent dummy Americas	-0.031*** (0.010)	-0.019* (0.010)	0.038 (0.025)
Asia	0.051*** (0.009)	0.084*** (0.008)	0.121*** (0.027)
Europe	0.166*** (0.013)	0.223*** (0.011)	0.272*** (0.033)
Oceania	-0.395*** (0.021)	-0.187*** (0.020)	-0.062 (0.058)
Trade	-0.229*** (0.029)	-0.032*** (0.008)	0.047 (0.031)
Urbanization	0.004*** (0.000)	0.002*** (0.000)	
Fertility	-0.048*** (0.002)		
Exports	0.004*** (0.001)		
Agriculture		-0.001 (0.000)	
Population in 1,400		0.012*** (0.002)	0.043*** (0.005)
Interstate war			0.042 (0.059)
Patent applications			-0.001 (0.001)
GDP p.c.			-0.021*** (0.007)
Ethnic fractionalization			0.255*** (0.038)
Constant	0.894*** (0.024)	0.655*** (0.023)	0.742*** (0.035)
Satellite-year F.E.	Yes	Yes	
Year F.E.			Yes
N	2,225	2,165	123
Adj. R ²	0.705	0.482	0.713

Note: Standard errors in parentheses, * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$. The base categories for the two categorical variables are non-colonies and Africa respectively. The category "independence between 1914 and 1945" drops out due to the lack of observations. Scaling of some variables differs from original sources: trade is here measured in permyriad of GDP, population in 1400 in 10,000,000 people, patent applications in 10,000 and GDP p.c. in 10,000 USD.

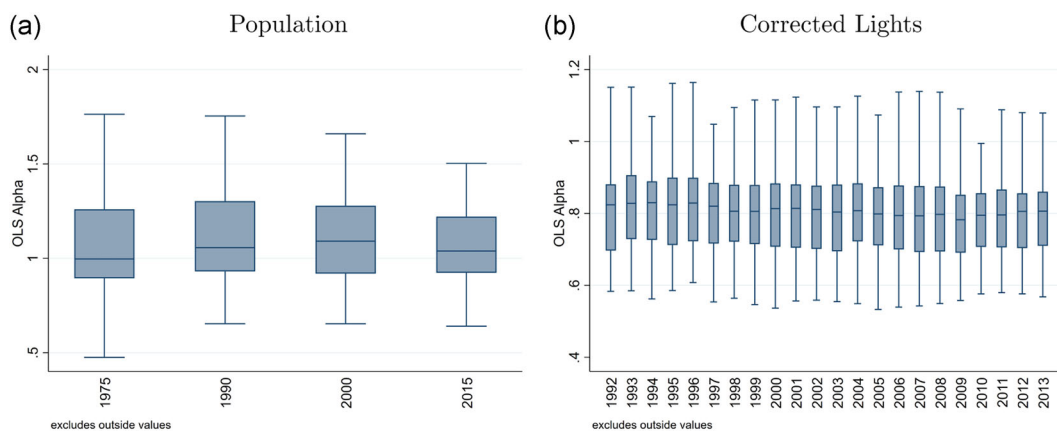


FIGURE 6 Stability of estimated Pareto alpha coefficients over time. (a) Population and (b) corrected lights. *Note:* The estimated Pareto alphas of all countries with more than 10 cities, based on the above-median distribution, are shown in the boxplots. The boxes indicate the 25–75% percentiles with the median in between. The range of the plots go up the adjacent values, omitting outliers. In years where more than one satellite per year is available for the lights data, values are averaged. [Color figure can be viewed at wileyonlinelibrary.com]

the cross-country distribution of Pareto alphas remains centered around a value of 0.8 for the entire period. Tracing the coefficients for each individual country over time shows similarly little movement, underlining the strong persistence in the city size distribution for most countries.

How can we then expect the future size distribution to look like? Extrapolating the results from our analysis over time would suggest that most countries will remain close to Zipf's law in population and keep their more unequal distribution in terms of light. However, it is conceivable that the growth rates of the biggest cities might add a twist here. In the long run, cities of different size have to grow at the same average rate for the distribution to remain stable, according to Gibrat's Law. Figure 7 compares the population growth rates of primary and secondary (2nd to 10th largest) cities in each country from 2000 to 2015. In the scatter plot of growth rates (Figure 7a), we see a stronger tendency of countries to be distributed above the main diagonal than below it. Hence, in more countries, primary cities have outgrown secondary cities, pointing to a relative consolidation rather than a catch-up. This holds for countries with both egalitarian and inegalitarian city size distributions, as the different colors of the dots show.²⁰ Figure 7b compares the cross-country growth rate distributions: The growth rate distribution of primary cities is located a bit to the right of that of secondary cities and has its mode at a slightly higher level; the Kolmogorov–Smirnov test rejects equality of the distributions at the 10% level.

Although this evidence is rather tentative, it fits in with other recent results from the literature about higher population growth rates of primary cities. Bluhm and Krause (2018) find that in sub-Saharan Africa, primary cities are growing significantly faster than secondary cities in terms of light once the top-coding correction is applied to the data. United Nations (2018) forecast that in 2030, there will be 43 megacities with more than 10 m inhabitants, most of them in developing countries. Given our result from Section 4 that large primary shares drive inequality of the total size distribution of cities, if concentration is indeed increasing, it might have wide-ranging effects. In terms of Zipf's law, we can expect existing deviations not only to persist, but they might become larger rather than smaller, if the largest cities draw away from the rest. For policymakers, managing the growth of such evolving megacities is vital, in particular in developing countries. There is a debate whether overall poverty is lower in larger

²⁰Correlation coefficients: 0.48 (blue dots) and 0.38 (green dots).

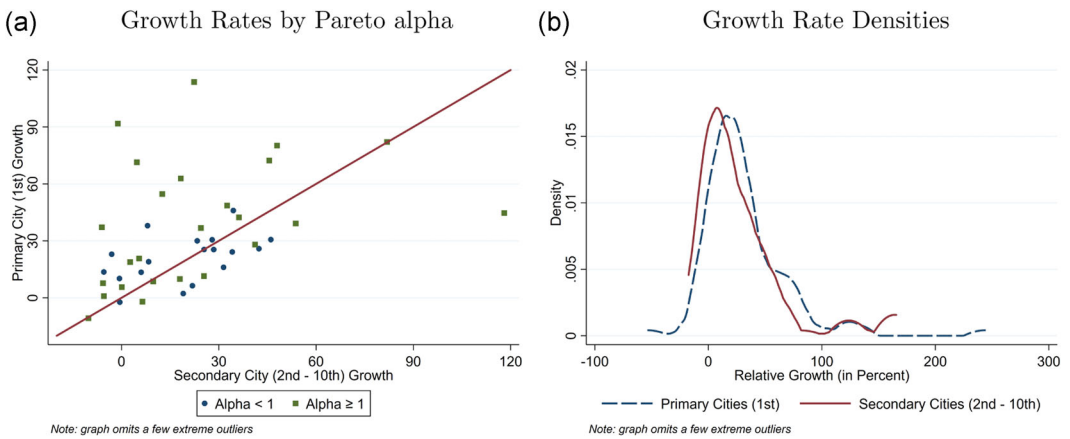


FIGURE 7 Population growth rates of primary and secondary cities (2000–2015). (a) Growth rates by Pareto alpha and (b) growth rate densities [Color figure can be viewed at wileyonlinelibrary.com]

or smaller cities (Christiaensen & Todo, 2014; Ferre, Ferreira, & Lanjouw, 2012). But growing primary cities will struggle to reap the benefits of agglomeration if (a) living conditions remain bad (Castells-Quintana, 2017; Glaeser, 2014) and (b) they are disconnected neighborhoods with poor infrastructure (Bluhm & Krause, 2018; Lall et al., 2017).

6 | CONCLUDING REMARKS

We revisit the discussion about Zipf's law in a cross-country setting by exploiting recent geospatial data. We use a consistent city identification scheme, provide a rigorous treatment of the threshold issue, and compare the city size distribution in each country based on both population and light proxying for economic activity. Our data set contains more than 10,000 cities in countries around the world.

The main insight from our analysis is that Zipf's law is an adequate characterization for the size distribution of cities for many, but not for all, countries. Economic activity, however, is typically distributed more unequally, so that Zipf's law does not hold for most countries in terms of light, in line with the existence of economic agglomeration effects. Such deviations can be explained to a large extent by an undue concentration of resources in the largest cities. We also note that the size effect is mainly driven by area rather than density, underlining the importance of economies of scale and market access.

To explain the cross-country heterogeneity in the size distribution of cities, we make use of recent time of development theories. Factors related to economic history turn out to be robust explanatory factors in our model selection procedure. Despite this persistence of the city size distribution, recent growth rates of the largest cities lead us to suggest that we might see a further move away from rather than towards Zipf's law in several countries.

There are two main avenues for further research to build upon the results of our paper. On the one hand, there is the question of which distribution of city size is optimal for a given country. Despite the theoretical arguments behind Zipf's law, country-specific factors, such as the current development level, might make a different distribution more appropriate (Henderson, 2003). In particular, it has been argued that a stronger concentration of resources at earlier stages of development might be beneficial, while later on a more balanced size distribution should emerge (Davis & Henderson, 2003; Hansen, 1990; Junius, 1999). It would be insightful to revisit this normative debate based on the new results.

Moreover, further research might focus on the connection between the city size distribution, regional inequality, and household-level income inequality. Which role large cities play in the evolution of income inequality is a highly relevant question (Diamond, 2016). Despite recent results on the nonlinear relation between the city size distribution and income inequality (Castells-Quintana, 2018), this issue is only starting to be explored.

ACKNOWLEDGMENTS

This paper has been presented at the Urban Economics Conferences in New York City and Düsseldorf, the European Regional Science Congress in Cork, the Development Economics and Policy Conference in Zurich, the 8th ifo Dresden Workshop on Regional Economics in Dresden, the ERSA-OECD Winter School in Trento, the Spring Meeting of Young Economists in Brussels, and various seminars. We would like to thank Richard Bluhm, David Castells-Quintana, Lewis Dijkstra, Gilles Duranton, Martin Gassebner, Roland Hodler, Remi Jedwab, Christian Lessmann, Adam Storeygard, and David Weil for very helpful comments and suggestions. We gratefully acknowledge financial support from the Germany Science Foundation (DFG). All remaining errors are unfortunately ours. Open access funding enabled and organized by Projekt DEAL.

ORCID

Christian Düben  <http://orcid.org/0000-0002-4281-1130>

Melanie Krause  <https://orcid.org/0000-0001-9299-2366>

REFERENCES

- Ades, A. F., & Glaeser, E. L. (1995). Trade and circuses: Explaining urban giants. *Quarterly Journal of Economics*, 110(1), 195–227.
- Auerbach, F. (1913). Das Gesetz der Bevölkerungskonzentration. *Petermanns Geographische Mitteilungen*, 59, 74–76.
- Barrios, S., Bertinelli, L., & Strobl, E. (2006). Climatic change and rural–urban migration: The case of sub-Saharan Africa. *Journal of Urban Economics*, 60(3), 357–371.
- Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3), 507–553.
- Bettencourt, L. M. A. (2013). The origins of scaling in cities. *Science*, 340(6139), 1438–1441.
- Bettencourt, L. M. A., & West, G. (2010). A unified theory of urban living. *Nature*, 467, 912–913.
- Black, D., & Henderson, J. (2003). Urban evolution in the USA. *Journal of Economic Geography*, 3(4), 343–372.
- Bluhm, R., & Krause, M. (2018). *Top lights—Bright cities and their contribution to economic development*. CESifo Working Paper 7411.
- Bonfatti, R., & Poelhekke, S. (2017). From mine to coast: Transport infrastructure and the direction of trade in developing countries. *Journal of Development Economics*, 127, 91–108.
- Brakman, S., Garretsen, H., Marrewijk, C. V., & van den Berg, M. (1999). The return of Zipf: A further understanding of the rank–size distribution. *Journal of Regional Science*, 39(1), 183–213.
- Brakman, S., Garretsen, H., & van Marrewijk, C. (2019). *An introduction to geographical and urban economics: A spiky world*. Cambridge: Cambridge University Press.
- Brinkhoff, T. (2017). *City population*. Retrieved from <https://www.citypopulation.de>
- Castells-Quintana, D. (2017). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics*, 98, 158–173.
- Castells-Quintana, D. (2018). Beyond kuznets: Inequality and the size and distribution of cities. *Journal of Regional Science*, 58(3), 564–580.
- Christiaensen, L., & Todo, Y. (2014). Poverty reduction during the rural–urban transformation—The role of the missing middle. *World Development*, 63, 43–58.
- Ciccone, A., & Hall, R. (1996). Productivity and the density of economic activity. *American Economic Review*, 86(1), 54–70.
- Cristelli, M., Batty, M., & Pietronero, L. (2012). There is more than a power law in Zipf. *Scientific Reports*, 812(2), 1–7.
- Davis, J., & Henderson, J. (2003). Evidence on the political economy of the urbanization process. *Journal of Urban Economics*, 53(1), 98–125.
- Desmet, K., & Rossi-Hansberg, E. (2013). Urban accounting and welfare. *American Economic Review*, 103(6), 2296–2327.
- Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000. *American Economic Review*, 106(3), 479–524.

- Donaldson, D., & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4), 171–198.
- Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *American Economic Review*, 97(1), 197–221.
- Duranton, G. (2008). Viewpoint: From cities to productivity and growth in developing countries. *Canadian Journal of Economics*, 41(3), 689–736.
- Eaton, J., & Eckstein, Z. (1997). Cities and growth: Theory and evidence from France and Japan. *Regional Science and Urban Economics*, 27(4–5), 443–474.
- Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, 94(5), 1429–1451.
- Elvidge, C. D., Hsu, F. C., Baugh, K., & Ghosh, T. (2014). National trends in satellite-observed lighting: 1992–2012. In Q. Weng (Ed.), *Global urban monitoring and assessment through earth observation* (pp. 97–120). Boca Raton: CRC Press.
- Fazio, G., & Modica, M. (2015). Pareto or lognormal? Best fit and truncation in the distribution of all cities. *Journal of Regional Science*, 55(5), 736–756.
- Ferre, C., Ferreira, F. H., & Lanjouw, P. (2012). Is there a metropolitan bias? The relationship between poverty and city size in a selection of developing countries. *The World Bank Economic Review*, 26(3), 351–382.
- Fetzer, T., Henderson, V., Nigmatulina, D., & Shanghavi, A. (2016). *What happens to cities when countries become democratic?* (Unpublished paper).
- Fujita, M., Krugman, P., & Venables, A. (1999). *The spatial economy: Cities, regions and international trade*. Cambridge: MIT Press.
- Fujita, M., & Mori, T. (1996). The role of ports in the making of major cities: Self-agglomeration and hub-effect. *Journal of Development Economics*, 49(1), 93–120.
- Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114(3), 739–767.
- Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives*, 30(1), 185–206.
- Gabaix, X., & Ibragimov, R. (2011). Rank-1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business and Economics Statistics*, 29(1), 24–39.
- Gabaix, X., & Ioannides, Y. (2004). The evolution of city size distribution. In J. Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics* (pp. 2341–2378). Amsterdam: Elsevier.
- Gibrat, R. (1931). *Les Inégalités Économiques: Applications: Aux Inégalités des Richesses, à la Concentration des Entreprises, aux Populations des Villes, aux Statistiques des Familles, etc. D'une Loi Nouvelle, la Loi de l'Effet Proportionnel*. Paris: Librairie du Recueil Sirey.
- Giesen, K., Zimmermann, A., & Suedekum, J. (2010). The size distribution across all cities—Double Pareto lognormal strikes. *Journal of Urban Economics*, 68(2), 129–137.
- Glaeser, E. L. (2014). A world of cities: The causes and consequences of urbanization in poorer countries. *Journal of the European Economic Association*, 12(5), 1154–1199.
- Glaeser, E. L., & Gottlieb, J. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature*, 47(4), 938–1028.
- Gollin, D., Kirchberger, M., & Lagakos, D. (2017). *In search of a spatial equilibrium in the developing world*. Working Paper No. 23916. National Bureau of Economic Research.
- González-Val, R. (2010). The evolution of the U.S. city size distribution from a long-term perspective (1900–2000). *Journal of Regional Science*, 5(50), 952–972.
- Hansen, N. (1990). Impacts of small- and intermediate-sized cities on population distribution: Issues and responses. *Regional Development Dialogue*, 11(1), 60–79.
- Henderson, J. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth*, 8(1), 47–71.
- Henderson, J., Squires, T., Storeygard, A., & Weil, D. (2018). The global distribution of economic activity: Nature, history, and the role of trade. *Quarterly Journal of Economics*, 133(1), 357–406.
- Henderson, J., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2), 994–1028.
- Henderson, J., & Wang, H. (2007). Urbanization and city growth: The role of institutions. *Regional Science and Urban Economics*, 37(3), 283–313.
- Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5), 1163–1174.
- Ioannides, Y., Overman, H., Rossi-Hansberg, E., & Schmidheiny, K. (2008). The effect of information and communication technologies on urban structure. *Economic Policy*, 23, 201–242.
- Ioannides, Y., & Skouras, S. (2013). U.S. city size distribution: Robustly Pareto, but only in the tail. *Journal of Urban Economics*, 73(1), 18–29.
- Jedwab, R., & Moradi, A. (2016). The permanent effects of transportation revolutions in poor countries: Evidence from Africa. *Review of Economics and Statistics*, 98(2), 268–284.

- Jedwab, R., & Vollrath, D. (2019). The urban mortality transition and poor country urbanization. *American Economic Journal: Macroeconomics*, 11(1), 223–275.
- Junius, K. (1999). Primacy and economic development: Bell shaped or parallel growth of cities? *Journal of Economic Development*, 24(1), 1–22.
- Kerr, W., & Kominers, S. (2015). Agglomerative forces and cluster shapes. *Review of Economics and Statistics*, 97(4), 877–899.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3), 483–499.
- Lall, S. V., Henderson, J. V., & Venables, A. J. (2017). *Africa's cities: Opening doors to the world*. Washington, DC: World Bank.
- Lessmann, C., & Seidel, A. (2017). Regional inequality, convergence, and its determinants—A view from outer space. *European Economic Review*, 92, 110–132. <https://doi.org/10.1016/j.euroecorev.2016.11.009>
- Lipton, M. (1977). *Why poor people stay poor: Urban bias in world development*. Cambridge, MA: Harvard University Press.
- Lobo, J., Bettencourt, L., Strumsky, D., & West, G. (2013). Urban scaling and the production function for cities. *PLOS One*, 3, 8.
- Moomaw, R., & Shatter, A. (1996). Urbanization and economic development: A bias toward large cities? *Journal of Urban Economics*, 40(1), 13–37.
- Moretti, E. (2004). Human capital externalities in cities. In J. Henderson & J. Thisse (Eds.), *Handbook of regional and urban economics* (pp. 2243–2291). Amsterdam: Elsevier North Holland.
- Motamed, M., Florax, R., & Masters, W. (2014). Agriculture, transportation and the timing of urbanization: Global analysis at the grid cell level. *Journal of Economic Growth*, 19(3), 339–368.
- Mutlu, S. (1989). Urban concentration and primacy revisited: An analysis and some policy conclusions. *Economic Development and Cultural Change*, 37(3), 611–639.
- Pesaresi, M., & Freire, S. (2016). *GHS settlement grid following the REGIO model 2014 in application to GHSL Landsat and CIESIN GPW v4-multitemporal (1975–1990–2000–2015)* (Tech. Rep.). European Commission, Joint Research Centre (JRC).
- Ravallion, M., & Chen, S. (2011). Weakly relative poverty. *Review of Economics and Statistics*, 93(4), 1251–1261.
- Rosen, K., & Resnick, M. (1980). The size distribution of cities—An examination of the Pareto law and primacy. *Journal of Urban Economics*, 8(2), 165–186.
- Rosenthal, S., & Strange, W. (2004). Evidence on the nature and sources of agglomeration economics. In J. Henderson & J. Thisse (Eds.), *Handbook of regional and urban economics* (pp. 2119–2171). Amsterdam: Elsevier North Holland.
- Rossi-Hansberg, E., & Wright, M. (2007). Urban structure and growth. *Review of Economic Studies*, 74(2), 597–624.
- Rozenfeld, H., Rybski, D., Gabaix, X., & Makse, H. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review*, 101(5), 2205–2225.
- Small, C., Elvidge, C., Balk, D., & Montgomery, M. (2011). Spatial scaling of stable night lights. *Remote Sensing of Environment*, 115(2), 269–280.
- Soo, K. (2005). Zipf's law for cities: A cross-country investigation. *Regional Science and Urban Economics*, 35(3), 239–263.
- Stokes, E. C., & Seto, K. C. (2019). Characterizing urban infrastructural transitions for the sustainable development goals using multi-temporal land, population, and nighttime light data. *Remote Sensing of Environment*, 234(111430), 1–11.
- Storeygard, A. (2016). Farther on down the road: Transport costs, trade and urban growth in sub-Saharan Africa. *Review of Economic Studies*, 83(3), 1263–1295.
- United Nations. (2018). *World urban prospects, the 2018 revision*. New York: United Nations.
- Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., ... Gething, P. W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, 553, 333–336.
- World Bank. (2009). *World development report 2009: Reshaping economic geography*. Washington DC: World Bank.
- Zipf, G. (1949). *Human behavior and the principle of least effort*. Cambridge: Addison-Wesley Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Düben C, Krause M. Population, light, and the size distribution of cities. *J Regional Sci.* 2021;61:189–211. <https://doi.org/10.1111/jors.12507>