Chapter 1
City Size Distributions

Clémentine Cottineau-Mugadza

Abstract Despite high levels of diversity and contingency in the creation and historical development of cities, the distribution of their sizes in a given area presents a surprisingly robust and simple pattern, known as Zipf's law or the rank-size curve. This chapter offers an overview of this regularity in city size distribution. Starting from a short history of early findings, it goes on to present Zipf's contribution, the plethora of empirical material which can be interpreted as a confirmation of the law and some models generating such a regular distribution. The deviations from the law highlighted in the literature are presented along with the theories to explain them. The last section of the chapter reviews recent propositions to advance the study of urbanisation "beyond Zipf's law", namely by using more sophisticated statistical tools or by reintegrating space in the analysis of urban population.

1.1 Introduction

Cities are complex organisations which have emerged independently at several points in human history. Where there was sufficient surplus in agriculture to fund the activity of non-farmers, the concentration of people has allowed a positive feedback loop between the specialisation and complementarity of social groups to produce larger populations and a growing array of functions performed at city level. Depending on their access to resources (from food to tax payers), cities have been

C. Cottineau-Mugadza (✉)
Centre National de la Recherche Scientifique (CNRS), Paris, France
e-mail: c.cottineau@tudelft.nl

Technische Universiteit Delft (TU-Delft), Faculty of Architecture and the Built Environment, Department of Urbanism, Delft, The Netherlands

D. Rybski (ed.), *Compendium of Urban Complexity*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-031-82666-5_1

able to reach significant sizes[1] quite early on, between half a million and a million inhabitants in Ancient Rome [2], over a million in Baghdad in the 8th century and between 150,000 and 200,000 inhabitants in Potosí (South America) around 1560 [3]. Nowadays, depending on city definitions, the largest cities on Earth can sustain thirty to forty million inhabitants, as in Tokyo, Delhi and the Pearl River Delta (Guangzhou/Shenzhen) [4, 5].

Simultaneously, systems of cities have emerged as a higher level of organisation of human settlement, from the interactions between interdependent cities of the territorial organisation [6]. While larger cities have concentrated social, political and economic functions of command and representation, the entire system relies on the presence and interaction with cities of smaller sizes, spatially distributed across territories and serving the basic local functions of socializing, trade, spiritual gatherings, service and production. Systems of cities thus rely on a diversity of city sizes. This is the case even for city-states, which were and are parts of wider networks of cities, as exemplified by Venice and Genoa then (whose trade networks extended to Northern Africa, Central and Eastern Asia) to Dubai and Singapore now (whose airport departure boards reflect the diversity of connexions they have with other cities worldwide).

More than a mere diversity of city sizes, multidisciplinary research over the past hundred years has pointed to the fact that city sizes tend to be distributed very regularly, with a few large cities and many small cities (i.e. fat-tailed distributions). Furthermore, the main parameters of these distributions have been claimed to be quite regular and universal since Auerbach [7]'s contribution. His discovery, the subsequent work done by Lotka [8] and Zipf [9] and the reactions it generated have since fueled a plethora of literature aimed at "confronting the mystery of urban hierarchy" [10]. The Auerbach-Lotka-Zipf (ALZ) law (wrongly attributed to Zipf alone in most cases) is taught in mainstream urban studies degrees and is used in many planning and policy reports. Yet, there is still no consensus as to why most urban systems verify the ALZ law (also known as the "rank-size rule"), why some deviate from it, and what this knowledge can be useful for. In this chapter, I try to provide a critical presentation of the recent multidisciplinary debates over city size distributions, by tracing the origins of the ALZ law to earlier models of urban hierarchy (Sect. 1.2), presenting the empirical, methodological and interdisciplinary discussions it sparks (Sect. 1.3), the attempts to deviate from this model (Sect. 1.4) and the aporia of Zipf's law as a policy tool (Sect. 1.5).

[1] In this chapter, we use population, rather than surface or wealth, as the measuring unit of city size. Indeed, population is the most widely comparable quantity across countries and time periods, and it summarizes a significant number of other urban quantities, such as wealth produced, number of dwellings, jobs etc. [1].

1.2 Early Findings About Regularities in City Size Distribution

G.K. Zipf has become the reference name for describing the regularity of city size distributions in general, and using a power law of exponent of -1 in the distribution of city rank with city size in particular. However, these regularities had already been observed, analysed and modelled by mathematicians and physicists since 1913 at least [7]. Additionally, the contributions of Reynaud [11], Gibrat [12] and Christaller [13] on cities' geographical locations and growth had concluded that systems of cities produced regular (spatial) distributions of cities following a fat-tailed distribution of sizes. Finally, contributions to urban primacy and optimal city size contributed to the scientific understanding of urban hierarchy long before Zipf's seminal publications in the 1940s. In this section, I therefore aim to contextualize Zipf's contribution within the study of city size distribution and to recall the main theoretical conclusions on city size regularity.

1.2.1 Formalisation of a Regular Relationship Between City Sizes and Ranks

The earliest scholarly publication known to address the rank-size distribution of cities is Auerbach [7]'s article: *Das Gesetz der Bevölkerungskonzentration*.[2] This short study (translated recently from German to English [14]) starts from the observation that Germany in 1910 counted about 50 big cities, around 200 small cities, roughly 2,000 towns and nearly 100,000 villages, thus a distribution approaching a geometric progression of the number of settlements with decreasing size. Auerbach noted that if you multiplied the size s of cities by their rank r in the distribution, you obtained a quantity which was a "characteristic product" because it remained more or less constant for German cities order by decreasing size (after some variations for the largest cities)

$$p_i \cdot r_i = AK\,. \tag{1.1}$$

He called this constant AK for Absolute Concentration and used it to construct another measure of population concentration: the specific concentration index (SpK), which corresponds to AK divided by the total population P of the country. He argues with international examples that this index can complement the analysis of population density by giving an idea of the hierarchical organisation of settlements according to their size. Auerbach's formulation Eq. (1.1) was later rewritten by Singer [15][3] as

$$r_i \cdot p_i^{\beta} = A, \quad \beta = 1\,.$$

[2] "The law of population concentration".

[3] According to Guérin-Pace [16] (I was never able to access this reference myself).

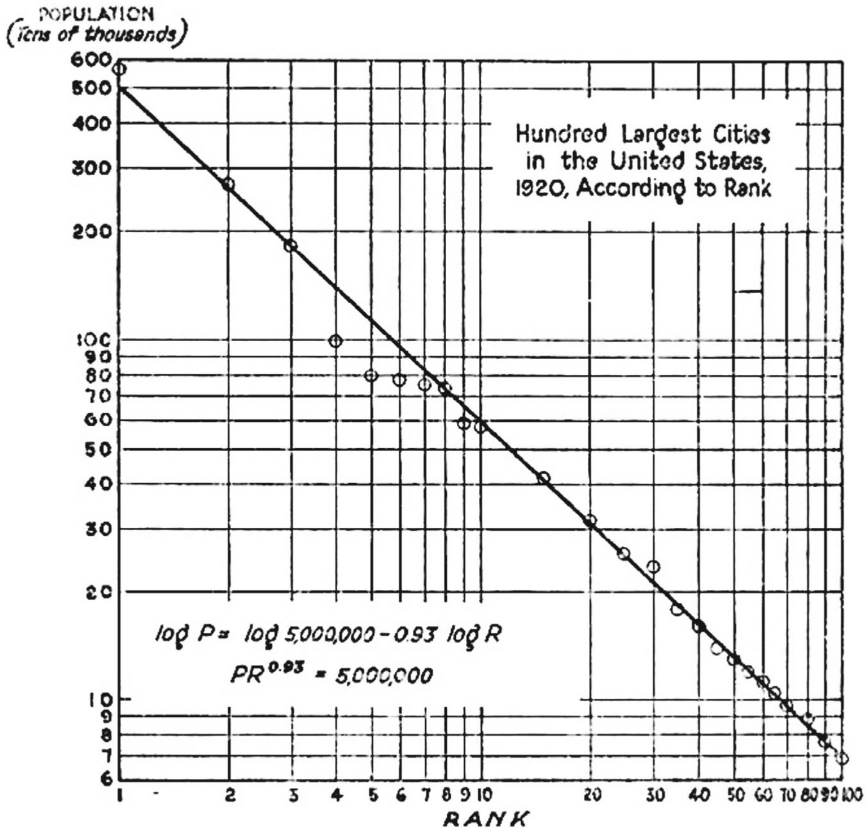

Fig. 1.1 Law of Urban Concentration. *Source*: [8], cited by Rybski and Ciccone [17]

This formalisation is equivalent (assuming $\alpha \simeq 1/\beta$) to the law which we currently know as Zipf's

$$p_i \sim r_i^{\alpha}, \quad \alpha \simeq 1\,. \tag{1.2}$$

The earliest scholarly publication known to visualise the rank-size distribution of cities on a log-log plot and to use the power law formulation of Eq. (1.2) is Lotka [8]'s article [17]. In this figure (cf. Fig. 1.1), Lotka [8] uses the visual representation of cities to estimate the value of the associated exponent α (≈ 0.9 using the 100 largest cities of the United States in 1920).

The earliest scholarly publication known to connect the power law formulation of Eq. (1.2) with the work of Pareto on income distributions is Saibante [18]. In his political economics course at the University of Lausanne, Pareto [19] famously suggested that income distributions follow a pattern best described by

$$\overline{F}(x) = \Pr(X > x) = \begin{cases} \left(\frac{x_\mathrm{m}}{x}\right)^\alpha & x \geq x_\mathrm{m} \\ 1 & x < x_\mathrm{m} \end{cases},$$

where the probability for a person picked at random to earn more than x is an inverse function of x to the power $-\alpha$, when x is over the minimum income x_m.

Building on Lotka's work, Saibante [18] "focused on the exponent [...] and thereby laid the foundation of modern empirical investigations" [17, p. 603].

The mathematical formulation chosen by Zipf relates to other disciplinary precedents such as [20] and [21]. The bibliometric works of mathematician Lotka [20] and librarian Bradford [21] had produced similar expressions to describe the frequency distribution of publishing authors, and were popularised and generalised by the linguist and philologist Zipf [22, 23]. Although he did not come up first with these results, Zipf became one of the first to establish a science of language, working extensively on empirical as well as theoretical matters and trying to explain the regularities observed systematically with quantitative methods [24]. We can explain the reason why his work is remembered today (rather than other scholars' contributions that pre-date his) by the fact that it is accessible (being written in English in a book format rather than more scattered publications) and because it spanned across fields to the social sciences.[4]

After analysing the regularity of word frequency in texts with power laws in earlier publications, G.K. Zipf turned his attention, among other topics in the social sciences, to city sizes, in particular with the books 'National Unity and Disunity' [9] and 'Human Behavior and the Principle of Least Effort' [22]. In these publications, Zipf offers both a formulation of the rank-size regularity, visual representation of empirical data, but also suggests an explanation of this relation with the idea of the "least effort" principle. In a model of economic geography [17], he shows that the optimal spatial distribution of labour in a system of settlement is affected by two opposite forces: the "force of diversification" which brings people to flock to production centers to reduce transport costs, and the "force of unification" which sees goods being transported to smaller towns where labour costs are minimised. In a situation of high transport costs, the "force of diversification" takes precedence and the populations are distributed very unevenly ($\alpha > 1$). In a situation of low transport costs, the "force of unification" takes precedence and the populations are distributed very evenly ($\alpha < 1$). In empirical situations, the two forces are counter-balanced and we find that $\alpha \simeq 1$. This explanation has clear similarities to earlier spatial models of urban settlements, although Christaller [13] is not properly referenced in [17, 22].

[4] For an analysis of the circulation of analytical models in urban studies, see [25].

1.2.2 Formalisation of a Regular Relationship Between City Sizes, Functions and Locations

The most famous spatial model for the distribution of cities is Christaller [13]'s Central Place Theory, later refined by Lösch [26]; although, as shown by Robic [27], an earlier version of the theory was written 100 years earlier in a French encyclopedia by Reynaud [11]. Central Place Theory explains the number, location, function and size of cities of a given region by the combined principles of economic rationality and spatial constraints (such as distance). It predicts that a first level of central places will service population locally by offering frequently needed goods and services. The spatial optimisation of circular catchment areas for these first level cities in an idealised isotropic plain gives way to a hexagon mesh. For goods and services of higher value and less frequent use, a smaller group of higher level cities are picked from lower level central places, whose catchment areas create an additional layer of larger nested hexagons. The resulting distribution of city sizes is therefore fat tailed in that the model creates classes of cities of identical sizes. Its rank-size graphical representation, as in Harris [28], shows a decreasing suite of steps or plateaus where Zipf's law expects a straight line.

Despite the difference in size predictions between Zipf's law and Central Place Theory, Berry and Garrison [29] find two areas of convergence. First, their predictions match in terms of classes of size. Second, the guiding principle behind Zipf's law (unification and diversification) is related to Christaller's catchment areas of goods in a Weberian localization scheme. Zipf's law represents an equilibrium between these two forces, in a similar way the hierarchy of nested hexagons in Central Place Theory represents a spatial equilibrium between servicing population with a diversity of services (in the larger cities) while keeping transport costs low (for frequently used services in smaller centres). An important difference between the models is that Zipf's law is a more generic and simple pattern to generate. Power laws are present in multiple fields of science. Indeed, there exists a wide diversity of generative models which produce a power law, from preferential attachment (e.g. [30] and multiplicative effects (e.g. [12]) to random walks and Yule [31]'s process [32, 33].

To summarize, the Auerbach-Lotka-Zipf law (subsequently ALZ) originates from a multidisciplinary background and finds precedents in the statistical work of earlier polymaths. Part of the explanation behind this regularity in urban studies stems from early models of spatial distribution of cities. Interestingly Rybski and Ciccone [17] found a disconnect between two strands of early research on city size, based on the citation patterns of early authors: the works cited by Zipf [22] which include [7, 8] and [12] and the works cited by Lösch [26] in an economics tradition (i.e. [12, 19] and [13]). Following these authors, although "Works citing Auerbach (1913) to a large extent also cite Zipf [...] typically 20 % of the works citing Zipf [...] also cite Auerbach (1913)" [17, p. 606]. Additionally, in recent years, Zipf's law has become so much part of the common knowledge of urban studies that a review of 66 English-speaking journal articles reporting an empirical estimation of the exponent of the ALZ law for cities showed that only 5 of them cited [9], 37 cited [22], and a significant proportion did not cite Zipf at all [34].

1.3 Scientific Cacophony Over Zipf's Law for Cities

Despite the wide recognition of the ALZ law in geography, planning, sociophysics and urban studies, there is no consensus regarding the adequacy and opportunity of this law in describing the distribution of city sizes. Three areas of debate persist to this day. The first debate revolves around the empirical observations gathered and their contribution to corroborating or rejecting the "law". The second ongoing debate is methodological and regards the definition of the observations and the fitting techniques used to evaluate the law. Finally an interdisciplinary debate is still ongoing about the universality of the ALZ law for cities, opposing those who find more value in the model to those who find its deviations more interesting.

1.3.1 Empirical Discussions

Despite final claims of empirical validity (such as [35] and [36] based on roughly 150 US cities), seventy years of rank-size rule estimation have produced a strong disparity of results on empirical distributions of city size, and a large pool of quantitative material to compare and analyse. A first array of studies have compared results across the world, identifying countries where the city-size distribution conforms to the ALZ law and countries where it does not. A second strand of research has focused on temporal evolution, identifying when urban systems conform to the law and when they do not. A third way of analysing the diversity of empirical estimations has been through a meta-analysis.

Comparison of Countries

The first available systematic international comparison of the ALZ is Rosen and Resnick [37]'s article in the 1980s.[5] Until the development of systematic recording of urban statistics, the international comparison of rank-size distribution was restricted to a handful of countries only: the United States [15, 20, 39], the USSR [39, 40], Japan [15, 41], France and the UK [15, 42], Germany and Hungary [15] and The Netherlands [7, 43].

In their 1980 paper, economists Rosen and Resnick [37] analyse city size distributions in 44 countries (including "developing" nations at the time such as Brazil,

[5] With the exception of Allen [38]'s study of 58 countries, which is cited in the literature (e.g. in [16]) but which I have never been able to access.

Indonesia, Zaire and Mexico). Their contribution lies in the estimation of the power law exponent of the rank size distribution of cities over 50,000 inhabitants in 1970 for these urban systems. Additionally, they tested the impact of using alternative city definitions for 6 countries and provided some explanations of the deviation. They use the Pareto form of estimation, Eq. (1.3), where the Pareto exponent corresponds to $\beta \simeq 1/\alpha$ in Eq. (1.2)

$$r_i \sim p_i^{-\beta} . \tag{1.3}$$

Empirically, the parameters of this equation are estimated using its equivalent

$$\log(r_i) = B - \beta \cdot \log(p_i) . \tag{1.4}$$

They find that the estimated value of β varies between 0.809 for Morocco and 1.963 for Australia, with an average of 1.136 and 75 % of countries having urban populations distributed more evenly than predicted by the ALZ law ($\beta > 1$). According to them, this meant that "the rank-size rule should be reconsidered" [37, p. 168].

Providing updated evidence on a larger sample of countries (75), Soo [44] relies on the citypopulation.de website[6] to estimate Zipf's law on cities and agglomerations (although the precise delineation used in the website is unknown) of at least 10,000 inhabitants between 1990 and 2001. He finds that, using the Ordinal Least Square (OLS) estimation methods, such empirical data leads to rejecting the ALZ law in most countries, firstly because the distribution does not follow a Pareto/power law distribution, and secondly because "when it does, the Pareto exponent is frequently statistically different from 1, with over half the countries exhibiting values of the Pareto exponent significantly greater than 1." [44, p. 240]. An innovative aspect of his work is to compare two methods of estimation: the OLS method with the more robust Hill estimator. Using both methods as empirical evidence, Soo [44] shows that the ALZ law is rejected in most cases, in terms of significance of the exponent estimated and its value with relation to 1.

Pumain et al. [45] compared the size distribution of urban agglomerations of more than 10,000 inhabitants in BRICS countries between the 1960s and 2010s using harmonised databases built for comparative purposes. They find a large diversity of values estimated for Zipf's law in these countries around 2010, using Eq. (1.2), from $\alpha = 0.8$ in China, i.e. city sizes more evenly distributed than predicted by the ALZ law, to 1.23 in the United States, i.e. city sizes more unevenly distributed than predicted. According to these authors [45, 46], such variations are better explained when looked at dynamically (i.e. over long time periods of time).

[6] www.citypopulation.de, last accessed: 27 November 2023.

Evolution Over Time

One argument used to explain the discrepancy between the ALZ law and the kind of empirical results reported below is that the law represents an equilibrium state for a mature system of cities [35]. Until maturity is reached, a system would not exhibit a power law fit with a unity exponent. By contrast, following the predictions from Evolutionary Urban Geography [47, 48] and the New Economic Geography [49], higher economies of scale and lower transportation costs should benefit larger cities and thus increase city size unevenness (i.e. increase α or reduce β), see also Chap. 3. Given recent economic trends (drastic lowering of transportation costs and higher economies of scale in innovative goods and services due to first-mover advantages in larger cities), one would therefore expect a monotonic evolution of the ALZ exponent towards greater unevenness.[7]

Using municipality data (rather than more systematically defined morphological or functional city definitions), Parr [50] tracked the evolution of Pareto's exponent β over several decades in a dozen of countries at different stages of economic development. He identified several types of evolution: a monotonic decrease of β in Turkey, India or Egypt; a decrease followed by a "slight upturn" in Brazil, Japan, Spain, and the USSR; a clear U-shape in Sweden, France, Austria and the USA. This suggested a U-shape evolution of the Pareto exponent (or an inverse U-shape evolution of the Zipf exponent) over time within a country throughout its economic development.

Using French "urban units" with a consistent threshold over 161 years, Guérin-Pace [16] highlighted a different pattern: "when the entire system is used [threshold of 2,000 inhabitants], inequality in city size has been accentuated over time. Inversely, when only the top of the hierarchy is considered [threshold of 100,000 inhabitants], the city size inequality has been reduced." [16, p. 557]. Fifteen years later, González-Val [51]'s systematic estimation of the Pareto exponent on US "incorporated places" points to a similar conclusion. Inequality in city sizes increases between 1900 and 2000 when the 19,000+ places are taken into account, yet there is convergence in size among the 1,000 biggest places. This suggests that the ALZ model of city size distribution does not fit empirical systems' evolution over time.

Using large harmonized databases on BRICS cities over the past 50 years, Cura et al. [46] observed two remarkable patterns: 1/ the strong persistence of the peculiarities of the rank-size curve (e.g. its shape or primacy) and 2/ no single trend regarding the evolution of city size inequalities. Indeed, between 1960 and 2010, they find a monotonic increase of α for Brazil and India, a monotonic decrease for China and no clear trend for South Africa and the Former Soviet Union.

Any regular temporal evolution of Zipf's exponent is thus still undemonstrated in empirical studies. One way of resolving discrepancies between national case studies is by pulling all studies together in a meta-analysis.

[7] The same conclusion was apparently reached by Allen [38], according to Guérin-Pace [16].

Meta-Analyses

A meta-analysis is an analytical tool aimed at summarizing the quantitative effect of systematically recorded factors on a value of interest. Developed originally to analyse randomized clinical trials in medical research, meta-analyses are now used to pool results from various empirical studies also in social sciences, in order to go beyond the limits of particular studies and samples and generate more general conclusions. To this end, it contextualises the production of empirical results reported in the literature by taking into account the different choices and constraints affecting the production and reporting of the quantitative value of interest. In the case of the ALZ law for cities, context refers to three broad elements: the territorial context (in which country was the law estimated? At which date?), the statistical context (how was the regression estimated? What is the definition of cities used? Which threshold of population applies?) and the study context (was the estimation reported on its own, as part of a sensitivity analysis or a comparative study?). Theoretical authors have provided plenty of explanations as to why the exponent of the ALZ law might vary over time and across nations (cf. Sects. 1.3.1 and 1.3.1), but meta-analyses tend to show that the statistical context is more important in determining the values of the exponent reported in the literature.

Nitsch [52] conducted the first meta-analysis of the ALZ law for cities. Selecting 29 references from the EconLit database, he was able to quantify the effect of estimation specifications on reported 515 values of Pareto exponents. He concluded that "the estimated exponent in a Zipf regression is on average not 1.0." [52, 97]. Furthermore, he noted that the value of the estimated exponent tended to vary with the type of city definition used in the study (i.e. metropolitan areas vs. built-up areas or municipalities), the year of the estimation and the number of other ALZ exponents estimated in the study. Rather than a picture of universal law, the work of Nitsch [52] pointed to a broad regularity of size distribution whose exponent is very sensitive to the technical specifications of its estimation.

More recently, Cottineau [53] extended the meta-analysis to include 1962 estimations from 86 studies and added more variables to test the effect of the territorial context and to control for the regression method (OLS, Maximum Likelihood, etc.). She finds that "no iron Zipf's law exists for cities" [53, p. 20] and estimates the role of technical specifications in the variation of reported exponent values to about 40 %, i.e. much more than the variation attributable to territorial, temporal and urban differences between countries. Technical specifications pertains to methodological choices made during data collection and statistical implementation of the regression. They represent the bulk of the methodological discussions on-going around the ALZ law for cities.

1.3.2 Methodological Discussions

We highlight four methodological elements to which recent debates attribute an effect on the estimation of the ALZ exponent: the scale of the system of cities analysed, the city definition chosen, the population cutoffs for including cities and the estimation methods selected.

What Scale for the System?

Identifying systems of cities is not straightforward. It means following multi-faceted inter-city interactions and tracing a limit where reciprocal links become stronger within the limit than with cities outside the limit [47]. Because national boundaries facilitate interurban circulations and data collection, most authors default to that scale to estimate the ALZ law. In the MetaZipf database [53],[8] estimations of the rank-size curve with cities selected at the country level represent 80 % of all 2188 estimations. The remaining estimations represent city systems at the sub-national (18 %) or continental (2 %) level. The sub-national scale can be justified in continental-sized countries, for example in Brazil [54]. It seems less appropriate when it results from data collection constraints, such as England and Wales–that is, without Scotland [15]. On the other hand, the continental scale can be justified when interurban interactions are found to exceed national borders, as in Europe or the former Soviet Union [46]. In any case, an appropriate delineation of the city system is important because only an internally consistent system should lead to a strict adherence to the ALZ law for large elements, as demonstrated by Cristelli, Batty, and Pietronero [55]. A correlated issue to that of the system scale is the issue of city delineation.

What is a City?

Delineating cities is not straightforward either. It means favouring one or more features of cities (political power, morphology, functional integration–depending on whether one focuses on people or the built environment) and tracing a limit where the concentration of urbanity seems to drop according to the criterion used.

Along with Rosen and Resnick [37], I would argue that the "entire metropolitan area is the most desirable choice for an urban unit as it represents an integrated economic unit." [37, p. 170]. However, many national statistical bodies do not provide such delineations and offer only municipal level data (the improperly named 'city proper'). Therefore many authors default to that delineation to estimate the ALZ law. In the previously cited MetaZipf database, estimations of the rank-size curve with cities delineated as municipalities rather than morphological or functional agglomerations represent 58 % of all estimations. The remaining estimations represent cities

[8] https://github.com/ClementineCttn/MetaZipf.

defined as morphological agglomerations (24 %), as functionally integrated units–metropolitan areas for instance–(6 %) or a mix of those (12 %). The municipal delineation could be justified for historical periods when municipalities still contained most of the urban footprint and activity. It seems less appropriate nowadays, especially when other definitions are available (or possible to construct, by aggregating municipal data based on morphological and/or functional criteria). In any case, an appropriate delineation of the city system is important because it affects the value of the power law exponent measured [44, 53, 56].

Population Cutoff and Sample Size

Another aspect of city definition pertains to the minimum population size for an urban aggregate to be considered a city. This minimum population (i.e. cutoff point of the distribution) has a mechanistic impact on the number of cities considered to examine the ALZ law: with low population cutoffs (from a few hundred to a few thousand people), many small settlements are considered as cities and the number of observations is large; when high population cutoffs are applied (hundred of thousands or million people), only a handful of cities are considered in the regression for most countries. The issue with this cutoff point was already mentioned by Auerbach [7] but remains relevant because there is no absolute justification for a given cutoff (no magic number distinguishes cities from non-urban settlements everywhere and at all times), yet the value of the cutoff chosen has a strong impact on regressed value of the ALZ exponent [37, 51, 53]. Using Monte Carlo simulations of US cities with a random rolling sample method, Nota and Song [57] found that "in 1980, 1990, and 2000, an increase of approximately one percent in the number of urban areas used in regression would cause a 0.16 percent reduction in the value of the estimated exponent. […] In other words, the rank-size rule is not necessarily an economic regularity and may in fact be a statistical phenomenon" [57, p. 30].

As noted by Rosen and Resnick [37], the population cutoff and sample size used to estimate the ALZ law would not be important if the power law fit was perfect. "But since deviations from the Pareto curve occur, our choice of sample size becomes important" [37, p. 171]. The question of deviations and their meaning is analysed in Sect. 1.4.

Fitting Technique and Estimation Methods

Finally, the fitting technique used to examine the ALZ law empirically has been found to play a role in the value of the exponent estimated. Soo [44] compares the results of the regression of the ALZ law for a variety of countries with the Ordinary Least Squares (OLS) and the Hill estimator. In the first case, the best fitting line is found heuristically, whereas in the second case, a Pareto distribution is assumed in the first place. The author concludes that "the results we obtain depend on the estimation method used, and in turn, the preferred estimation method would depend on our sample size and on our theoretical priors–whether or not we believe that Zipf's Law holds" [44, p. 261]. In a meta-analysis of empirical estimates of the ALZ law for cities, Cottineau [53] considers the four dominant estimation methods present in the literature: OLS, the likelihood method (which includes the Hill estimator), Gabaix and Ibragimov [58]'s modification of the traditional OLS regression (which simply replaces R_i by $R_i - 0.5$) and Markov chains. She finds that the likelihood method is indeed significantly associated with a lower Pareto estimate[9] (OLS meta-analysis) and that Gabaix and Ibragimov [58]'s estimation method leads to estimate systematically higher Pareto exponents on average (fixed-effect meta-analysis). These effects add significant, unnecessary (and usually overlooked) noise to the interpretation of the ALZ law for cities and its deviation.

1.3.3 Disciplinary Angles

An important aspect of the multidisciplinary interest in the ALZ law (including biology, sociology, physics and economics [59]) is that it reflects different attitudes towards modelling, residuals and deviations (cf. Sect. 1.4). In fields such as economics and physics, the model takes the centre stage, and its status is either validated or rejected. By contrast, geographers will use the model as a "generality" filter and will focus on understanding the local deviations to the model. In a "meta-"meta-analysis, Cottineau [34] tested the influence of this potential disciplinary bias on the reporting of estimation results in empirical studies of Zipf's law by authors from various disciplines. She hypothesised that the usual publication bias, whereby "positive" results (i.e. validating an hypothesis) get more easily published than "negative" results (i.e. the absence of significant evidence), play differently for different disciplines, since positive results in physics and economics consist in validating the ALZ law whereas positive results in geography can mean rejecting the model in favour of its deviations. She computed similarity matrices between 66 empirical articles estimating a coefficient for Zipf's law, in terms of their wording, the references cited, and the disciplines of the journals cited (cf. Fig. 1.2). She shows 1/ that similar wording between pairs of articles is significantly associated with similar values of estimated coefficient (average and dispersion) and 2/ that similar citation patterns between pairs of articles is significantly associated with similar values of estimated coefficient (average only). "These results point towards the existence of a combined publication and reporting bias [...] Indeed, despite the fact that city sizes could be observed identically by all researchers using the same specifications, I find that different authors tend to delineate cities and report Zipf estimations differently depending on the field they come from, which translates into the words they use and the citations they mobilise" [34, p. 1459].

[9] In the original article, the coefficient of this regression is positive since the regression of Zipf's law uses the Lotka equation form, with the population expressed as a function of the rank, unlike Eq. (1.4).

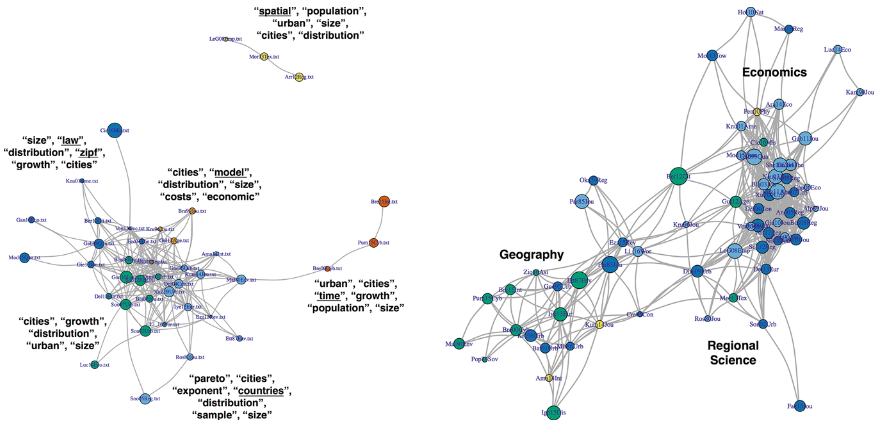

Fig. 1.2 Wording and citation similarity between articles reporting an empirical estimation of the ALZ law for cities. *Source*: [34, Figs. 4–6] distributed under the terms of the Creative Commons CC BY license

1.4 Deviating from the ALZ Law

Beyond the variation of value of the β coefficient discussed above, some analyse the poor fit to a power law model in terms of meaningful curve deviations. This section reviews the theoretical propositions to explain both the deviation from the value of 1 for the β coefficient and the deviation from a power law distribution.

There are two main sets of explanation of empirical deviation to the ALZ law for cities. The first one relates to theories of socioeconomic development and their consequence in terms of population settlement (Sect. 1.4.1). This type of top-down explanations is in line with Zipf's original take on why city sizes are so regularly distributed. The second set of explanations (Sect. 1.4.2) corresponds more directly to urbanisation processes. This type of bottom-up explanation emanates from the study of cities themselves and their evolution.

1.4.1 Socioeconomic Developments

A first strand of explanations for the deviation from the ALZ law suggests that if a system of cities does not fit a power law distribution of their sizes, or if the exponent of this power law differs from 1, it means that the system has not reached maturity and is therefore on its way there. This type of explanation encompasses both theories of economic development (whereby "wealthier countries [...] have more evenly distributed populations [... because] high income allows a country to support a network of intermediate-sized cities as demand thresholds are passed" [37, p. 182]) as well as theories of territorial integration (whereby newly integrated federal states tend to duplicate small to medium-sized cities which were the heads of previously independent systems of cities, as 19th century India and China [60] or the Soviet Union [28]). The lack of integration is supposed to produce rank-size curves which are convex instead of linear in a log-log plot, whereas the lag in socioeconomic development can lead to convex curves (when small cities merge with one another through spatial growth [61]), to concave curves (when urban growth rates are autocorrelated over time [62]) when the primate cities are larger than expected), or simply to different exponent values.

Finally, Pumain [47] suggests that systems of cities tend to be more uneven over time because larger cities grow faster on average, due to first-mover advantages in innovations. This is another illustration of the autocorrelation of growth rates which would lead to a deviation from the ALZ law in terms of the value of its coefficient.

1.4.2 Urbanisation Patterns

The explanations related to urbanisation processes are the most powerful in predicting the deviation of the empirical exponent to its theoretical value of 1 [53]. One of them relates to the coevolution of urban systems and transportation systems. Therefore, recently urbanised systems are expected to be more unequal than ancient ones because the transportation networks available at the time of urbanisation were slower in old systems, a larger amount of small cities were necessary. In areas urbanised with railroads and highways, these small cities tend to have been short-circuited [47, 63]. According to Morrill [64], small territories have more uneven distribution of city sizes than large countries because the concentration of power in the largest cities is not balanced by a sufficiently large set of secondary cities, which is seconded by Rosen and Resnick [37]'s empirical inquiry. Finally, the spatial dependence of urban growth (or the positive autocorrelation of growth rates), which contrary to Gibrat's model assumptions,[10] can produce empirical deviations from the ALZ law [65, 66].

1.5 Beyond Zipf's Law

The doubt formulated around the ALZ law and the reduction of city systems to a single value lead many authors to consider new paths to better describe city size distributions and the regularities of urban systems more generally. The first approach is to consider other distributions than power laws. The second is to reintegrate geographical space in the analysis of city sizes.

[10] Gibrat's model being the model most commonly accepted to generate Zipf's regularity [35].

1.5.1 From Power Laws to Other Distributions

The most frequent distribution preferred to the power law is the lognormal distribution [67–69]. There are two main reasons for this. Firstly, these distributions belong to the same type of mathematical distributions, i.e. skewed distributions [70]. Empirical data such as city populations are usually not sufficient to detect and distinguish significantly between these distributions [71]. Secondly, the lognormal distributions is usually preferred to a power law when analysing city sizes because the generative mechanisms suggested to explain them tend to produce lognormal distributions rather than power laws. This is the case of [30]'s preferential attachment model or Gibrat's stochastic model [35].

Other distributions have been suggested, like the Beta or Lavalette distributions [72], or hybrid forms of distributions [73–75], for instance a Pareto upper tail combined with a lognormal lower tail [76]. This strand of research trades the simplicity of the ALZ model for a more accurate representation of city size distribution. It must be noted that this quest can only achieve significant results if it relies of carefully delineated and harmonised urban data, since the quality of any statistical fit depends heavily on the quality of the underlying data [46].

1.5.2 From Power Laws to Spatial Models

A more radical improvement of the model involves the consideration of geographical space and spatial relations in the study of city size distributions. For instance, Gallo and Chasco [65] consider spatial dependence in their estimation of Zipf's law, using a spatial autoregressive model with spatially autocorrelated errors. They find statistical evidence that considering spatial dependence in urban growth impacts the estimated value of the ALZ law for cities. Bergs [75] demonstrates the necessity for economic approaches to analyse the ALZ law spatially, both with theoretical and empirical datasets.

The consideration of spatial dependence in empirical estimations goes hand in hand with the focus on spatial models to generate and potentially explain the emergence and stability of city size distributions. Such models range from social physics models (for instance [77]) with their call for simplicity, to more complex agent-based models involving non-deterministic interactions (for instance [78]).

1.6 Conclusion

> Despite the early discovery, the quest for a robust theoretical model to explain such an empirical distribution of cities remains elusive; starting with Christaller (1933) who described Auerbach's finding as 'a most incredible law,' yet essentially was 'nothing more than just

playing with numbers'. The view that the Zipf's Law is 'atheoretical' carries on to the present date. [79, p. 2439]

More than a century since the first attempts at summarizing city size distributions with an elegant and simple mathematical expression, urban studies scholars can be left with a demoralising impression regarding the rank-size regularity and the ALZ law. Past the initial sense of mystery and pleasing simplicity, many question the future of this model. Is the law pointless? What should we do with this statistical relationship? Since much empirical evidence rejects the law and numerous self-reinforcing process can generate a power law distribution, can we really use the ALZ law to validate our theories on how cities grow and why urbanisation differ from country to country? Does it have any value as a policy tool?

This short review of the recent literature on the ALZ law suggests that the rank-size curve is still a popular model to analyse urbanisation patterns and to explain national differences. However, the cacophony of interpretations derives from the wide variations in urban definitions and estimation techniques. Until these differences are accounted for across disciplines and academic traditions, theoretical models will be challenging to validate empirically. Until theoretical models are validated empirically, it seems adventurous to try to use the ALZ law as a policy tool, let alone as a policy objective.

Acknowledgements The author thanks Denise Pumain, Marie-Claire Robic, Michael Batty, Robin Morphet and John Parr for their insight and many discussions about Zipf, the ALZ law and the urban hierarchy over the past decade.

References

1. D. Pumain, Une théorie géographique des villes. Bulletin de la Société Géographique de Liége **55**(2), 5–15 (2010)
2. G.R. Storey, The population of ancient Rome. Antiquity **71**(274), 966–978 (1997). https://doi.org/10.1017/S0003598X00085859
3. A. Aruska de Souza Santos, S. AS Iamamoto, The Difficult Legacy of Mining in Past and Contemporary Potosí and Ouro Preto, in *Journal of Latin American Geography* 18.3 (2019), pp. 110–131. https://www.jstor.org/stable/48618853
4. United Nations, *The world's cities in 2018* (Tech. rep. New York, United, 2018)
5. O.E.C.D. European Commission, *A new perspective on urbanisation* (OECD, Tech. rep, 2020)
6. D. Pumain, *Hierarchy in natural and social sciences*. (Springer, 2006)
7. F. Auerbach, Das Gesetz der Bevölkerungskonzentration, in *Petermanns Geogr. Mitteilungen* 59.74 (1913), pp. 74–76. https://zs.thulb.uni-jena.de/receive/jportal_jparticle_00586390
8. A.J. Lotka, *Elements of Physical Biology*. (Williams and Wilkins, 1925)
9. G. Kingsley Zipf, *National unity and disunity; the nation as a bio-social organism.* (Principia Press, 1941)
10. P. Krugman, Confronting the mystery of urban hierarchy. Journal of the Japanese and International economies **10**(4), 399–418 (1996). https://doi.org/10.1006/jjie.1996.0023
11. J.R. Villes, Encyclop é die nouvelle **8**, 670–687 (1841)

12. R. Gibrat, *Les inégalités économiques: applications aux inégalités des richesses, à la concentration des entreprises... d'une loi nouvelle, la loi de l'effet proportionnel*. Libr. du Recueil Sirey, 1931
13. W. Christaller, *Die Zentrale Orte in S ü ddeutschland* (Fischer, Jena, 1933)
14. F. Auerbach, A. Ciccone, The law of population concentration. Environment and Planning B: Urban Analytics and City Science **50**(2), 290–298 (2023). https://doi.org/10.1177/23998083221147139
15. H.W. Singer, The "courbe des populations.âŁž A parallel to Pareto's Law, in The Economic Journal **46**(182), 254–263 (1936). https://doi.org/10.2307/2225228
16. F. Guérin-Pace, Rank-size distribution and the process of urban growth. Urban studies **32**(3), 551–562 (1995). https://doi.org/10.1080/00420989550012960
17. D. Rybski, A. Ciccone, Auerbach, Lotka, and Zipf: pioneers of power-law city-size distributions. Archive for History of Exact Sciences **77**(6), 601–613 (2023). https://doi.org/10.1007/s00407-023-00314-0
18. M. Saibante, La concentrazione della popolazione, in *Metron–Rivista Internszionale di Statistica* 7.2 (1928), pp. 53–99. https://ebiblio.istat.it/digibib/Metron/MetronV7N2_1928.pdf
19. V. Pareto, *Cours d' é conomie politique*. (Rouge, 1896)
20. A.J. Lotka, The frequency distribution of scientific productivity. Journal of the Washington academy of sciences **16**(12), 317–323 (1926)
21. S.C. Bradford, Sources of information on specific subjects. Engineering **137**, 85–86 (1934)
22. G. Kingsley Zipf, *Human behavior and the principle of least effort: an introd. to human ecology*. (Addison-Wesley Press, 1949)
23. Y.-S. Chen, F.F. Leimkuhler, A relationship between Lotka's law, Bradford's law, and Zipf's law. Journal of the American Society for Information Science **37**(5), 307–314 (1986). https://doi.org/10.1002/(SICI)1097-4571(198609)37:5<307::AID-ASI5>3.0.CO;2-8
24. C. Prün, R. Zipf, Biographical notes on G. K. Zipf, in *Glottometrics* 3 (2002), pp. 1–10. https://www.ram-verlag.eu/wp-content/uploads/2018/08/g3zeit.pdf#page=5
25. C. Cottineau et al., The role of analytical models and their circulation in urban studies and policy, in *Urban Studies* (2024), p. 00420980241237410. https://doi.org/10.1177/00420980241237410
26. A. Losch, *Economics of location*. (Yale University Press, 1954)
27. M.-C. Robic, Cent ans avant Christaller... une théorie des lieux centraux, in *L'Espace g é ographique* (1982), pp. 5–12
28. C. Dennison Harris, *Cities of the Soviet Union: studies in their functions, size, density, and growth* (Published for Association of American Geographers by Rand McNally, Chicago, 1970)
29. B.J.L. Berry, W.L. Garrison, Alternate explanations of urban rank-size relationships. Annals of the Association of American Geographers **48**(1), 83–90 (1958). https://doi.org/10.1111/j.1467-8306.1958.tb01559.x
30. H.A. Simon, On a class of skew distribution functions. Biometrika **42**(3/4), 425–440 (1955). https://doi.org/10.2307/2333389
31. U.G. Yule, A mathematical theory of evolution, based on the conclusions of Dr. JC Willis, FRS, in Philosophical Transactions of the Royal Society of London Series B **213**, 21–87 (1925). https://doi.org/10.1098/rstb.1925.0002
32. M. Mitzenmacher, A brief history of generative models for power law and lognormal distributions. Internet mathematics **1**(2), 226–251 (2004). https://doi.org/10.1080/15427951.2004.10129088
33. M.E.J. Newman, Power laws. Pareto distributions and Zipf's law, in Contemporary physics **46**(5), 323–351 (2005). https://doi.org/10.1080/00107510500052444
34. C. Cottineau, What do analyses of city size distributions have in common?" In: *Scientometrics* (2022), pp. 1–25. https://doi.org/10.1007/s11192-021-04256-8
35. X. Gabaix, Zipf's law for cities: an explanation. The Quarterly journal of economics **114**(3), 739–767 (1999). https://doi.org/10.1162/003355399556133
36. B.J.L. Berry, A. Okulicz-Kozaryn, The city size distribution debate: Resolution for US urban regions and megalopolitan areas. Cities **29**, S17–S23 (2012). https://doi.org/10.1016/j.cities.2011.11.007

37. K.T. Rosen, M. Resnick, The size distribution of cities: an examination of the Pareto law and primacy. Journal of urban economics **8**(2), 165–186 (1980). https://doi.org/10.1016/0094-1190(80)90043-1
38. G.R. Allen, The 'courbe des populations', a Further Analysis. Bulletin of the Oxford University Institute of Statistics **16**, 179–189 (1954)
39. F.T. Moore, A note on city size distributions. Economic Development and Cultural Change **7**(4), 465–466 (1959). https://doi.org/10.1086/449818
40. V.R. Popov, Investigation of the System of Urban Places of Crimea Oblast. Soviet Geography **15**(1), 19–23 (1974). https://doi.org/10.1080/00385417.1974.10770642
41. A. Okabe, An expected rank-size rule: A theoretical relationship between the rank size rule and city size distributions. Regional Science and Urban Economics **9**(1), 21–40 (1979). https://doi.org/10.1016/0166-0462(79)90019-X
42. B. Ward, City structure and interdependence. Papers of the Regional Science Association **10**(1), 207–221 (1962). https://doi.org/10.1007/BF01934688
43. M.C. Deurloo, De wet der urbane concentratie. Tijdschrift voor economische en sociale geografie **63**(5), 306–314 (1972). https://doi.org/10.1111/j.1467-9663.1972.tb00557.x
44. K. Tong Soo, Zipf's Law for cities: a cross-country investigation, in *Regional science and urban Economics* 35.3 (2005), pp. 239–263. https://doi.org/10.1016/j.regsciurbeco.2004.04.004
45. D. Pumain et al., Multilevel comparison of large urban systems, in *Cybergeo: European Journal of Geography* 706 (2015). https://doi.org/10.4000/cybergeo.26730
46. R. Cura et al., The old and the new: qualifying city systems in the world with classical models and new data. Geographical Analysis **49**(4), 363–386 (2017). https://doi.org/10.1111/gean.12129
47. D. Pumain, Pour une théorie évolutive des villes, in *L'Espace géographique* (1997), pp. 119–134
48. D. Pumain, Une théorie géographique pour la loi de Zipf. R é gion et d é veloppement **36**(31–54), 125–150 (2012)
49. M. Fujita, P.R. Krugman, A. Venables, *The Spatial Economy: Cities, Regions, and International trade*. (MIT press, 1999)
50. J.B. Parr, A note on the size distribution of cities over time. Journal of Urban Economics **18**(2), 199–212 (1985). https://doi.org/10.1016/0094-1190(85)90017-8
51. R. González-Val, The Evolution of US City Size Distribution from a Long-Term Perspective (1900–2000). Journal of Regional Science **50**(5), 952–972 (2010). https://doi.org/10.1111/j.1467-9787.2010.00685.x
52. V. Nitsch, Zipf zipped. Journal of Urban Economics **57**(1), 86–100 (2005). https://doi.org/10.1016/j.jue.2004.09.002
53. C.C. MetaZipf, A dynamic meta-analysis of city size distributions, in *PloS one* 12.8 (2017), e0183919. https://doi.org/10.1371/journal.pone.0183919
54. L. Suarez-Villa, Rank size distribution, city size hierarchies and the beckmann model: some empirical results. Journal of Regional Science **20**(1), 91–95 (1980). https://doi.org/10.1111/j.1467-9787.1980.tb00628.x
55. M. Cristelli, M. Batty, L. Pietronero, There is more than a power law in Zipf, in *Scientific reports* 2 (2012), p. 812. https://doi.org/10.1038/srep00812
56. P. Veneri, City size distribution across the OECD: Does the definition of cities matter? Computers, Environment and Urban Systems **59**, 86–94 (2016). https://doi.org/10.1016/j.compenvurbsys.2016.05.007
57. F. Nota, S. Song, Further analysis of the Zipf's law: Does the rank-size rule really exist?" In: *Journal of Urban Management* 1.2 (), pp. 19–31. https://doi.org/10.1016/S2226-5856(18)30058-X
58. X. Gabaix, R. Ibragimov, Rank-1/2: a simple way to improve the OLS estimation of tail exponents. Journal of Business & Economic Statistics **29**(1), 24–39 (2011). https://doi.org/10.1198/jbes.2009.06157
59. M.J. Beckmann, City hierarchies and the distribution of city size. Economic Development and Cultural Change **6**(3), 243–248 (1958). https://doi.org/10.1086/449769

60. G.A. Johnson, Rank-size convexity and system integration: a view from archaeology. Economic geography **56**(3), 234–247 (1980). https://doi.org/10.2307/142715
61. Y. Ijiri, H.A. Simon, Interpretations of departures from the Pareto curve firm-size distributions. Journal of Political Economy **82**(2), 315–331 (1974). https://doi.org/10.1086/260194
62. D.R. Vining Jr., Autocorrelated growth rates and the Pareto law: a further analysis. Journal of Political Economy **84**(2), 369–380 (1976). https://doi.org/10.1086/260440
63. F. Moriconi-Ebrard, *L'urbanisation du monde depuis 1950* (Anthropos, Economica, 1993)
64. R.L. Morrill, *The spatial organization of society*. (Duxbury Press, 1970)
65. J. Le Gallo, C. Chasco, Spatial analysis of urban growth in Spain, 1900–2001. Empirical economics **34**(1), 59–80 (2008). https://doi.org/10.1007/s00181-007-0150-5
66. K.-M. Cheng, Y. Zhuang, Spatial Econometric Analysis of the Rank-size Rule for the Urban System: A Case of Prefectural-level Cities in China's Middle Area. Scientia Geographica Sinica **32**(8), 905–912 (2012)
67. J.B. Parr, K. Suzuki, Settlement populations and the lognormal distribution. Urban Studies **10**(3), 335–352 (1973). https://doi.org/10.1080/00420987320080471
68. B.T. Robson, *Urban growth: an approach*. (Methuen, 1979)
69. J. Eeckhout, Gibrat's law for (all) cities. American Economic Review **94**(5), 1429–1451 (2004). https://doi.org/10.1257/0002828043052303
70. M. Barbut, A family of distributions: from Paretian to contra-Paretian. Applications to the study of urban concentration and its evolution, in *Cybergeo, European Journal of Geography* 266 (2004)
71. A. Clauset, C. Rohilla Shalizi, M.E.J. Newman, Power-law distributions in empirical data, in *SIAM review* 51.4 (2009), pp. 661–703. https://doi.org/10.1137/070710111
72. O. Fontanelli et al., Beyond Zipf's law: the Lavalette rank function and its properties, in *PloS one* 11.9 (2016), e0163241. https://doi.org/10.1371/journal.pone.0163241
73. M. Levy, Gibrat's law for (all) cities: Comment. American Economic Review **99**(4), 1672–75 (2009). https://doi.org/10.1257/aer.99.4.1672
74. Y. Malevergne, V. Pisarenko, D. Sornette, Testing the Pareto against the lognormal distributions with the uniformly most powerful unbiased test applied to the distribution of cities, in *Physical Review E* 83.3 (2011), p. 036111. https://doi.org/10.1103/PhysRevE.83.036111
75. R. Bergs, Spatial dependence in the rank-size distribution of cities–weak but not negligible, in *PloS one* 16.2 (2021), e0246796. https://doi.org/10.1371/journal.pone.0246796
76. Y. Ioannides, S. Skouras, US city size distribution: Robustly Pareto, but only in the tail. Journal of Urban Economics **73**(1), 18–29 (2013). https://doi.org/10.1016/j.jue.2012.06.005
77. D. Rybski, A. García Cantú Ros, J.P. Kropp, Distance-weighted city growth, in *Physical Review E* 87.4 (2013), p. 042114. https://doi.org/10.1103/PhysRevE.87.042114
78. C. Schmitt et al., Half a billion simulations: Evolutionary algorithms and distributed computing for calibrating the SimpopLocal geographical model. Environment and Planning B: Planning and Design **42**(2), 300–315 (2015). https://doi.org/10.1068/b130064p
79. Y. Mansury, L. Gulyas, The emergence of Zipf's Law in a system of cities: An agent-based simulation approach. Journal of Economic Dynamics and Control **31**(7), 2438–2460 (2007). https://doi.org/10.1016/j.jedc.2006.08.002