

## Zipf's Law for cities: estimation of regression function parameters based on the weight of American urban areas and Polish towns

Dariusz Sokołowski<sup>1, CDFMR</sup>, Iwona Jażdżewska<sup>2, DFMR</sup>

<sup>1</sup>*Nicolaus Copernicus University in Toruń*, Faculty of Earth Sciences and Spatial Management, Department of Spatial Management and Tourism, Lwowska 1, 87-100 Toruń, Poland, e-mail: [sokol@umk.pl](mailto:sokol@umk.pl); <sup>2</sup>*University of Łódź*, Faculty of Geographical Sciences, Institute of Urban Geography, Tourism and Geoinformation, Kopcinskiego 31, 90-142, Łódź, Poland, e-mail: [iwona.jazdzewska@geo.uni.lodz.pl](mailto:iwona.jazdzewska@geo.uni.lodz.pl) (corresponding author)

### How to cite:

Sokołowski, D. Jażdżewska, I. (2021). Zipf's Law for cities: estimation of regression function parameters based on the weight of American urban areas and Polish towns. *Bulletin of Geography. Socio-economic Series*, 53(53): 147-156. DOI: <http://doi.org/10.2478/bog-2021-0028>

**Abstract.** The paper aims at presentation of a methodology where the classical linear regression model is modified to guarantee more realistic estimations and lower parameter oscillations for a specific urban system. That can be achieved by means of the weighted regression model which is based on weights ascribed to individual cities. The major shortcoming of the methods used so far – especially the classical simple linear regression – is the treatment of individual cities as points carrying the same weight, in consequence of which the linear regression poorly matches the empirical distribution of cities. The aim is reached in a several-stage process: demonstration of the drawbacks of the linear parameter estimation methods traditionally used for the purposes of urban system analyses; introduction of the weighted regression which to a large extent diminishes specific drawbacks; and empirical verification of the method with the use of the input data for the USA and Poland

### Article details:

Received: 23 March 2021  
 Revised: 15 May 2021  
 Accepted: 21 September 2021

### Key words:

city size distribution,  
 weight of cities,  
 Zipf's Law,  
 rank-size rule,  
 weighted regression,  
 USA,  
 Poland

### Contents:

1. Introduction	148
2. City rank and size in human settlement systems	149
3. Methods of trend line approximation applied to rank-size rule	150
4. Weighted regression as solution to the problem	151
5. Results and discussion	152
6. Conclusion	153
References	154

## 1. Introduction

Zipf's Law describes the relation between the frequency of a set of items or events and their size, and it is often considered as an example of Power Law (Newman, 2005; Clauset, Shalizi & Newman, 2009; Cristelli, Batty & Pietronero, 2012). With regard to the urban distribution by size, it is commonly known as the rank-size rule (e.g. Berry & Garrison, 1958; Dziewoński, 1972; Gabaix, 1999) or even as Zipf's rank-size rule (Sokołowski, 2001; Berry & Okulicz-Kozaryn, 2012; Jażdżewska, 2006, 2017).

Zipf's Law is an empirical method, where the real size patterns are described by means of statistical and mathematical ones. It can be well applied to some highly-asymmetrical distributions which are often subject of studies in natural, social and exact sciences. Being a linguist, Zipf used the rank-size rule mostly in the empirical studies in his field (Zipf, 1949), but he also suggested it could be applied to socio-economic patterns (Zipf, 1941).

Zipf's Law comes as a very popular instrument in studies of human settlement systems where it facilitates the description and analysis of distributions of population in cities (e.g. Stewart, 1947; Berry, 1964; Rosen & Resnick, 1980; Knudsen, 2001; Ioannides & Overman, 2003; Soo, 2005; Batty, 2008; Berry & Okulicz-Kozaryn, 2012; Jażdżewska, 2017). Comparison was made for: urban settlement systems of the USA, Australia, Sweden (Garner, 1967) or USA, Spain and Italy (González-Val et al., 2015), regions (Xu & Harriss, 2010), (Pumain et al., 2015), OECD countries (Veneri, 2016), individual countries, e.g. France (Pumain, 1989), USA (Fujita, Krugman & Venables, 1999), Denmark (Knudsen, 2001), China (Song & Zhang, 2002; Anderson & Ge, 2005; Chen & Zhou, 2008), Malaysia (Soo, 2005), Poland (Sokołowski, 2001; Jażdżewska, 2006, 2017) and in a global setting (Jiang, Yin & Liu, 2015).

For a few past years the theoretical underpinnings of Zipf's Law have been broadly discussed (for instance: Robson, 2012; Gabaix, 1999; Chen & Zhou, 2008; Gabaix & Ioannides, 2004; Rossi-Hansberg & Wright, 2007; Chen, 2016). They have been under examination, for instance, for the sake of the best explanation of urban growth in the light of Gibrat's or Pareto Law (Eeckhout, 2004;

Garmestani, Allen & Gallagher, 2008; Córdoba, 2008; Giesen, Zimmermann & Suedekum, 2010; Cuberes, 2011). The calculation of the contrast ratio and its juxtaposition with the 'ideal', following the rank-size rule, was one of the elements of the empirical analysis of the settlement system (Garner, 1967; Sokołowski, 2001) as well as the cartographic visualization of the changes in city ranks in the settlement system (Jażdżewska, 2017).

First and foremost, this paper concerns empirical issues, in particular the problem of adjusting the approximating line of the urban distribution by size. The classical estimation methods of line parameters often fail, because points deviate too much from the approximation line. It can result from:

- 'outliers' representing the population of the most populated 'primate cities',
- large amount of small towns,
- data availability, which – in the case of many settlement systems – becomes smaller together with the decreasing population.

In comparative analyses other issues emerge: definition of the city in different countries (Sokołowski, 2014; Szymańska, 2007); method of statistical data collection in statistical units (e.g. for the city or urban area, urban cluster). Whereas in dynamic analyses there is a problem of changing population numbers in the studied set of cities.

The methodology proposed below renders it possible to minimize these shortcomings. Owing to that the regression line better matches the layout of points and facilitates a more accurate estimation of a parameter and the theoretical number of city inhabitants.

The method will be introduced on the example of the urban settlement systems of Poland and the USA. Those are two different settlement systems. The former applies to a European country with less than 40 mln inhabitants and approx. 1,000 towns and cities; the second is found in the country with the population of over 300 mln and over 3.5 thousand Urbanized Areas and Urban Clusters. If the conclusions are valid for both of the systems, the method will be recommendable to studies of other settlement systems. In the case of Poland, the Central Statistical Office [GUS] was the source of data; in the case of the USA, the resources of the US Census Bureau were consulted (following:

Thomas Brinkhoff, <https://www.citypopulation.de/USA.html>).

## 2. City rank and size in human settlement systems

At the beginning of the 20<sup>th</sup> century (Auerbach, 1913) pointed to the relation of the city size to the city rank in the human settlement system and suggested that the approximation of city distribution by size be supported with the Pareto distribution. This regularity can be presented in form of the following equation:

$$L_j \times j = \text{const.} \quad (1)$$

where:  $L_j$  – city size expressed by population,  $j$  – city position (rank) in the descending city order.

The above formula was upgraded by the addition of the rank exponent (Lotka 1925), which made it more congruent with the empirical data. Thus, it is possible to determine the theoretical size of the  $j^{\text{th}}$  city in the system basing on the following formula:

$$L_j = L_1 \times j^{-a}, \quad (2)$$

where:  $L_1$  – population of the largest city,  $a$  – contrast exponent of the studied human settlement system. The exponent value may be different for every system and it changes in the course of time reflecting the evolution of the system.

George K. Zipf (1941, 1949) popularized the method and, thus, lent his name to it. The first method had one parameter; the second, revised method had two parameters. The two-parameter method was in the past the most common method used in the studies of settlement systems. The article presents the two-parameter method for which the best method of point approximation was to be found. The two-parameter formula (formula 2) allows for the definition of the theoretical size of  $j$  city in the system. For this calculation two parameters:  $L_1$  and  $a$  are necessary.

According to Zipf (1941), the contrast exponent is the resultant of the impact of two kinds of forces shaping the human settlement system: concentrating force and diffusing force. The 'principle of least effort' posits that the economic system should be

in the state of balance if the costs of transportation are to be minimized (at present the cost of transportation plays a lesser role, but it is beyond the scope of interest of this discussion). Depending which force – differentiating or concentrating – is stronger, a human settlement system may feature a significant number of smaller towns or just a few very big cities. In polycentric systems  $|a| < 1$ , and in the case of population being concentrated in the largest units  $|a| > 1$ .

When a system is in the state of balance, the contrast exponent  $|a|=1$ . Then, the size of every city is conditioned by the size of the largest city, according to the 'law of the primate city' (Jefferson, 1939; Zipf, 1941) and has the following values (cf. formula 2):

$$L_2 = \frac{L_1}{2}, L_3 = \frac{L_1}{3}, \dots, L_n = \frac{L_1}{n}$$

What is questionable in such an interpretation of Zipf's Law is the fact that the size of the cities in the system is dependent on the largest city.

The rank-size rule based on the arithmetic scale does not present itself well graphically, as the points are arranged approximately along an isosceles hyperbola. If a logarithm (log-log plot) is used, the curve turns into a straight line (Figure 1) and the arrangement of points is then approximated by the line described as below:

$$\log L_j = b + a \log j. \quad (3)$$

The length of the line on the graph depends on the number of cities taken for research, only for Poland they were all towns and cities.

The use of linear approximation increases the clarity of the graph and enables its better interpretation, which applies, among other things, to the position of particular points in relation to the approximating line. The contrast exponent  $a$  equates with the gradient angle between the line and x-axis ( $|a| = \text{tg } \alpha$ , thus for  $|a|=1$ :  $\alpha=45^\circ$  or  $\alpha=180^\circ-45^\circ$ ).

### 3. Methods of trend line approximation applied to rank-size rule

The layout of points in Zipf’s graph in the logarithmic scale is approximated by the line which enables:

1. characterization of a human settlement system (gradient of the line as a resultant of the impact of various forces and processes);
2. identification of the position of individual points (cities) relative to the line and analysis of their deviations;
3. calculation of the theoretical population of cities.

Practically, it is possible to juxtapose different human settlement systems and examine changes in a particular system occurring over a period of time; also, upon the analysis of the deviation of points from the line, it is feasible to study the differences between the real and the theoretical number of city inhabitants.

In literature one can encounter various methods of defining line parameters. One of the method of approximation is represented by the popular method of the simple linear regression. Parameters

a and b can be estimated upon solving the system of two linear equations presented in general form as:

$$\sum y_j = bN + a\sum x_j, \tag{4}$$

$$\sum x_j y_j = b\sum x_j + a\sum x_j^2; \tag{5}$$

or upon using the following formulas:

$$a = \frac{N\sum x_j y_j - \sum x_j \sum y_j}{N\sum x_j^2 - (\sum x_j)^2}, \tag{6}$$

$$b = \frac{\sum y_j - a\sum x_j}{N}. \tag{7}$$

With logarithm variables the system of equations will look as below:

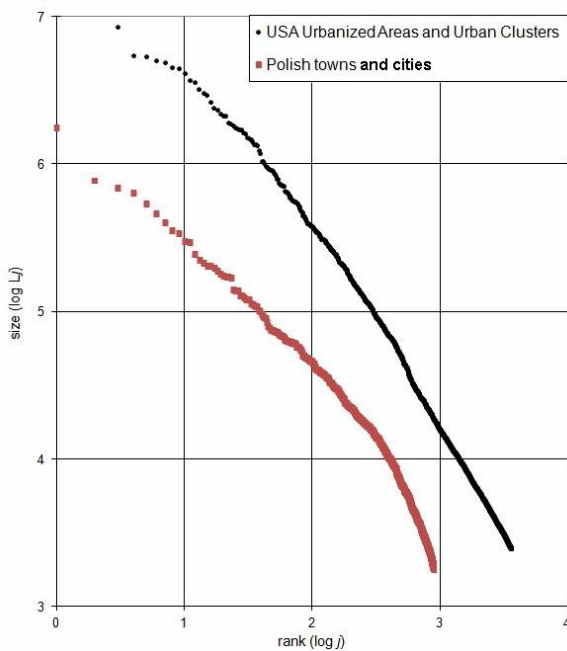
$$\sum \log L_j = bN + a\sum \log j, \tag{8}$$

$$\sum \log j \log L_j = b\sum \log j + a\sum (\log j)^2, \tag{9}$$

and the respective formulas will take the following form:

$$a = \frac{N\sum \log j \log L_j - \sum \log j \sum \log L_j}{N\sum (\log j)^2 - (\sum \log j)^2}, \tag{10}$$

$$b = \frac{\sum \log L_j - a\sum \log j}{N}. \tag{11}$$



**Fig. 1.** Rank-size graph regarding Urban Areas in USA (2010) vs Polish towns and cities (2017)  
Source: own calculations

After substituting the parameter values into the formula (3) the gradient angle between the line and the axis can be defined, where the line reflects the structure of the human settlement system as a whole and enables the estimation of the theoretical size of any city ( $L_j$ ).

Using the classical regression method for the analysis of urban size patterns causes at least two problems:

1. in general, a significant departure of the regression line from the line defined by the set of large cities, which leads to an incorrect estimation of contrast exponent a and to a significant overestimation of theoretical city sizes (calculated on the basis of regression);
2. large discrepancies in line parameter estimations based on the number of units included in the analysis.

The former is well depicted in Figure 2, where the classical linear regression deflects strongly upwards, at the same time ‘losing touch’ with the largest cities. The line approximating the points is not well

matched as it causes substantial overestimation of the theoretical population of cities. For Polish towns and cities the theoretical sum of inhabitants is higher than the real sum by 71%, for American cities it is higher by as much as 176.8%. Noteworthy, the value of  $|\alpha|$  ratio exceeds 1 for Poland and USA.

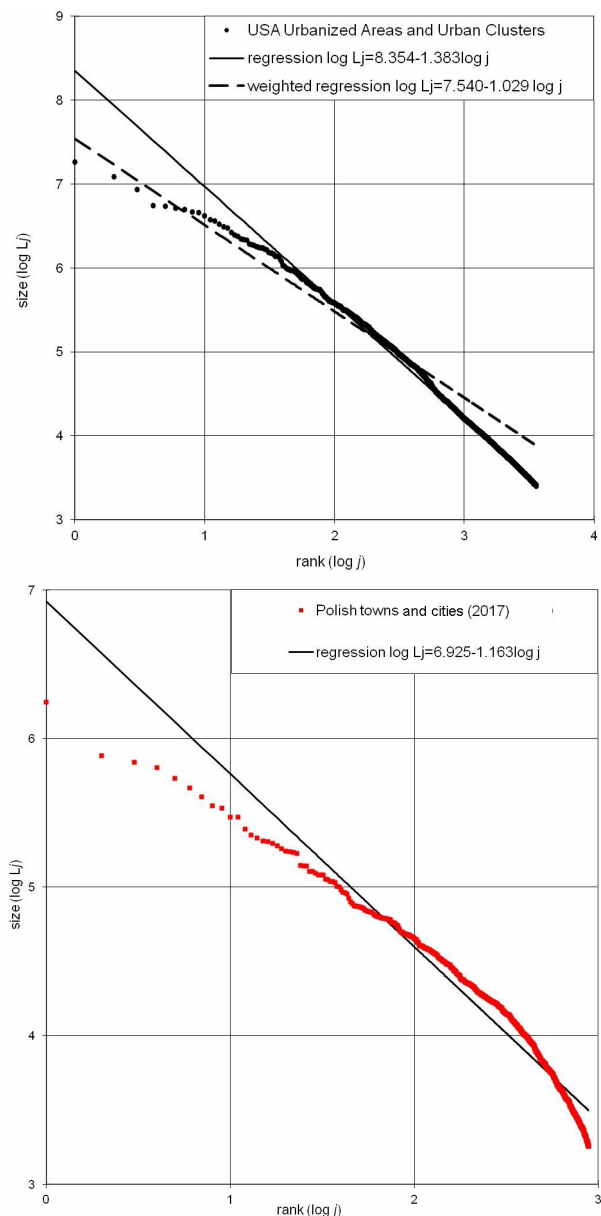
These imperfections of the regression method used for the rank-size rule are especially conspicuous in the comparative analysis of human settlement systems in different countries (or regions), or possibly in the same country (region) in different time frames, if every time another number of settlement units is taken into account. The cardinality of the studied set depends on the pre-defined minimum size limit. In practice, if the classical regression method is used for a large number of units (for example all Polish towns and cities):

1. the size of the major city and sizes of other large cities are estimated unrealistically;
2. the regression gradient oscillates widely in comparative analyses;
3. the parameters are highly unstable and different spatial units have poor comparability (e.g. in comparative studies of different countries).

Weighted regression as solution to the problem

#### 4. Weighted regression as solution to the problem

The simple linear regression is most often not useful for the estimation of the theoretical population of the largest cities, because it yields too high theoretical values. A numerous group of small units has a strong impact on the gradient of the line, despite the fact that their total mass (population) may be lower than the mass of a single large city. In the equations (4-11) used for the calculations of  $a$  and  $b$ ,  $N$  is one of the most important parameters defining the number of studied cities rather than the number of inhabitants in those cities. It is ungrounded to acknowledge that, for instance, the importance of a city with several million inhabitants does not differ from that of a unit which is a thousand times smaller. That is the point of dissimilarity between the empirical situation – typical of urban distributions



**Fig. 2.** Rank-size graph regarding Urban Areas in: USA (2010), towns and cities in Poland (2017) and approximating models of simple line regression

Source: own calculations

by size – and cases studied within the scope of exact sciences or linguistics where no distinction is made between weights of individual points.

Minimization of the said drawbacks is the principal aim of this paper. It can be achieved through an original modification of the method of least squares, which entails accounting for appropriate weight indexes when regression parameters are estimated. Weights should be proportional to the mass (size) of particular settlement units, thus they

will not have the same impact on the line. It will restore the right proportions between settlement units in the rank-size model. The essence of the modification suggested here is best described as the weighted regression. Its use does not directly refer to the uncertain value of one variable (Strutz, 2011), but it relates to the problem of choosing for the weight the number of inhabitants ( $w_j$ ) and their sum ( $\sum w_j$ ) instead of N parameter defining the number of studied cities.

The parameters of the line equation in the weighted regression model can be found upon solving a system of equations, where numerical values  $y_j$  and  $x_j$  are substituted respectively by the product of these values and weight coefficients ( $w_j$ ), and the number of units (N) is replaced with the sum of their weights ( $\sum w_j$ ). In general, the system of equations is as follows:

$$\sum y_j w_j = b \sum w_j + a \sum x_j w_j, \quad (12)$$

$$\sum x_j y_j w_j = b \sum x_j w_j + a \sum x_j^2 w_j, \quad (13)$$

and the direct calculation of constants a and b is possible with these formulas:

$$a = \frac{\sum w_j \sum x_j y_j w_j - \sum x_j w_j \sum y_j w_j}{\sum w_j \sum x_j^2 w_j - (\sum x_j w_j)^2}, \quad (14)$$

$$b = \frac{\sum y_j w_j - a \sum x_j w_j}{\sum w_j}, \quad (15)$$

where:  $w_j$  is the point weight of the object with  $j$  rank; in the case of the city, as a rule, it will be the value expressed by the number of inhabitants. With logarithm variables entered into the above-presented system of equations and formulas, the results will look as follows:

$$\sum \log L_j w_j = b \sum w_j + a \sum \log j w_j \quad (16)$$

$$\sum \log j \log L_j w_j = b \sum \log j w_j + a \sum (\log j)^2 w_j, \quad (17)$$

and:

$$a = \frac{\sum w_j \sum \log j \log L_j w_j - \sum \log j w_j \sum \log L_j w_j}{\sum w_j \sum (\log j)^2 w_j - (\sum \log j w_j)^2} \quad (18)$$

$$b = \frac{\sum \log L_j w_j - a \sum \log j w_j}{\sum w_j}. \quad (19)$$

## 5. Results and discussion

The comparison of the two methods of regression parameter estimation was based on the data on Urban Areas (Urbanized Areas and Urban Clusters) in the USA for 2010 and all town and cities for Poland (2017).

The analysis concerned: deviation of the theoretical sizes of particular units from their real sizes (Figure 3); sums of the rests of regression show sum of deviations (regression residual) as an index towns of showing how well the regression matches the pattern in reality (Table 1) and interpretation of a value.

What can be concluded from the graph (Figure 3) is that the classical regression highly overestimates the size of large units, while the weighted regression overestimates the size of the smallest units.

In order to measure how well-matched regression is, it was necessary to establish the deviations of the estimated sizes from the real sizes, i.e. the regression residual (Table 1). As a result of an inaccurate regression estimation, mistakes cumulate in the total population in the country. In the case of American cities the overestimation for the classical regression is 176.8%, for the weighted regression it is 8 % and for Polish towns and cities 79,8% versus 3%.

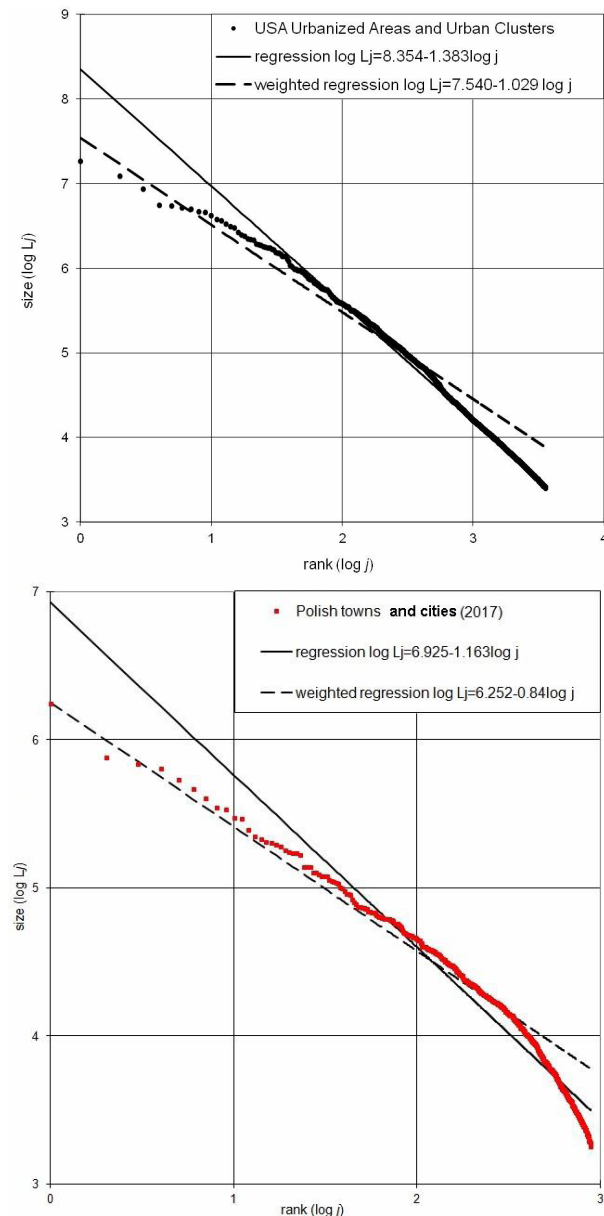
As mentioned before, regression lines lead to incongruences in the estimation of the theoretical population of cities. The three largest cities or urbanized areas in Poland and the USA are the case in point (Table 2). The theoretical number of inhabitants calculated by means of the classical regression for NY is more than ten times higher than the real number, 5 times higher in the case of Warsaw. As far as the weighted regression is concerned, the results are different, although much closer to the real values (Table 2). For NY they are almost twice as high as the real number of population and for Warsaw they are 1.2 times higher. The second and the third largest cities have even better match of theoretical and real values.

This method can be recommended for calculations of theoretical population; however, it is still far from perfect. Noteworthy, in the weighted regression the value of a is lower than in the classical regression. For Polish towns and cities it is below 1 in the weighted regression (0.84) and exceeds 1 in

**Table 1.** Adjusting of regression to empirical population of Urbanized Areas and Urban Clusters in USA vs Poland (towns and cities) – sums of the rests of regression

Country	Sums of the rests of regression (thousands inhabitants)			
	classic regression		weighted regression	
USA	446,794	(176.8%)	20,284	(8.0%)
Poland	18,626	(79.8%)	0,720	(3.0%)

Source: own calculations



**Fig. 3.** Rank-size graph regarding Urban Areas in USA (2010) vs towns and cities in Poland (2011) and approximating models of simple line regression and weighted regression  
Source: own calculations

the classical regression (1.163). As mentioned above, the value of  $|\alpha| < 1$  is characteristic of the systems where smaller towns dominate – in other words: the urban settlement system in Poland in 2017. For American cities a value was above 1, which means that population is concentrated in the largest cities, but in the weighted regression it approximates 1 (1.029) – thus, the system is balanced in terms of the number of cities and population. As illustrated by the above examples, the differences between a are considerable and necessitate various interpretations of the ratio according to Zipf’s Law.

This conclusion requires further, more extensive discussion and comparative studies.

## 6. Conclusion

Among the practical problems connected to the application of the rank-size rule to human settlement systems one can mention such estimation of approximating line parameters which would optimally correspond to the empirical urban distribution by size. The paper presents the originally modified regression model which significantly reduces the inconveniences related to the use of the classical regression. The merits of the suggested method were demonstrated on the basis of the USA and Poland.

The weighted regression yields the approximating line better matched to the empirical values (see Figure 3), which translates into more realistic estimations of size of individual cities (Table 2) and a lower sums of the rests of regression residuals (Table 1), as well as a higher stability of regression parameters (Table 2). Therefore, the weighted regression facilitates comparative analyses of different human settlement systems and reduces

**Table 2.** The theoretical population (determined on the basis of regression) and regression parameters of the three largest USA units (2010) vs Polish towns and cities (2017)

Country	Units <sup>*</sup>	Theoretical size of the main units (thousands of inhabitants) determined on the basis of:		Real population (thousands of inhabitants)	Difference between real and theoretical values population determined on the basis of:	
		classic regression	weighted regression		classic regression	weighted regression
		USA	NY		225,722	34,669
LA	86,517		16,993	12,151	405%	26%
Ch	49,372		11,198	8,608	222%	14%
Poland	W	8,410	1,787	1,765	377%	1%
	C	3,756	999	767	169%	13%
	L	2,344	710	690	94%	1%

\* Urbanized Areas: NY – New York–Newark, NY-NJ-CT; LA – Los Angeles–Long Beach–Anaheim, CA, Ch – Chicago, IL-IN, Poland W – Warsaw, C – Cracow, L – Lodz

Source: *own calculations*

the sensitivity of the model to the cardinality of a studied set. The method presented may be treated as an alternative, or a complementary method, to the classical regression. It should be tested on other sets of cities.

Discussions which have been held for many years on the theoretical and empirical use of Zipf's rank-size theory prove that the search for the best point approximation methods is going to be continued (Richardson, 1973), because the variability and complexity of the world, including settlement systems, is immense. The proposed methods may contribute to progress in:

- comparative studies (settlement systems of several regions, countries),
- dynamic studies (one or several settlement systems),
- studies on the best match of the approximating line to define the theoretical number of population in cities,
- studies on a parameter for the best description of settlement systems.

## References

- Anderson, G. Ge, Y. (2005). The size distribution of Chinese cities. *Regional Science and Urban Economics*, 35(6): 756–776. DOI: <https://doi.org/10.1016/j.regsciurbeco.2005.01.003>
- Auerbach, F. (1913). The law of population concentration (in German). *Petermanns Geographische Mitteilungen*: 74–76.
- Batty, M. (2008). The size, scale, and shape of cities. *Science*, 319(5864): 769–771. DOI: [10.1126/science.1151419](https://doi.org/10.1126/science.1151419)
- Berry, B.J.L. (1964). Cities as systems within systems of cities. *Papers of the Regional Science Association*, 13(1): 146–163. DOI: [10.1007/BF01942566](https://doi.org/10.1007/BF01942566)
- Berry, B.J.L. Garrison, W.L. (1958). Alternate explanations of urban rank-size relationships. *Annals of the Association of American Geographers*, 40(1): 83–90. DOI: [10.1111/j.1467-8306.1958.tb01559.x](https://doi.org/10.1111/j.1467-8306.1958.tb01559.x)
- Berry, B.J.L. Okulicz-Kozaryn, A. (2012). The city size distribution debate: Resolution for US urban regions and megalopolitan areas. *Cities*, 29(Supplement 1): S17–S23. DOI: [10.1016/j.cities.2011.11.007](https://doi.org/10.1016/j.cities.2011.11.007)
- Central Statistical Office [GUS], <https://stat.gov.pl/>
- Chen, Y. (2016). The evolution of Zipf's law indicative of city development. *Physica A: Statistical Mechanics and its Applications*, 443: 555–567. DOI: [10.1016/j.physa.2015.09.083](https://doi.org/10.1016/j.physa.2015.09.083)
- Chen, Y. Zhou, Y. (2008). Scaling laws and indications of self-organized criticality in urban systems. *Chaos, Solitons and Fractals*, 35(1): 85–98. DOI: [10.1016/j.chaos.2006.05.018](https://doi.org/10.1016/j.chaos.2006.05.018)



- Clauset, A. Shalizi, C.R. Newman, M.E.J.** (2009). Power-law distributions in empirical data. *SIAM review* 51(4): 661–703. DOI: [10.1137/070710111](https://doi.org/10.1137/070710111)
- Córdoba, J.C.** (2008). On the distribution of city sizes. *Journal of Urban Economics*, 63(1): 177–197. DOI: [10.1016/j.jue.2007.01.005](https://doi.org/10.1016/j.jue.2007.01.005)
- Cristelli, M. Batty, M. Pietronero, L.** (2012). There is more than a power law in Zipf. *Scientific Reports*, 2(812): 1–7. DOI: [10.1038/srep00812](https://doi.org/10.1038/srep00812)
- Cuberes, D.** (2011). Sequential city growth: Empirical evidence. *Journal of Urban Economics*, 69(5): 229–239. DOI: [10.1016/j.jue.2010.10.002](https://doi.org/10.1016/j.jue.2010.10.002)
- Dziewoński, K.** (1972). General theory of rank-size distributions in regional settlement systems: Reappraisal and reformulation of the rank-size rule. *Papers of the Regional Science Association*, 29: 73–86. DOI: [10.1007/BF01962286](https://doi.org/10.1007/BF01962286)
- Eeckhout, J.** (2004). Gibrat's law for (all) cities. *American Economic Review*, 94(5): 1429–1451. DOI: [10.1257/0002828043052303](https://doi.org/10.1257/0002828043052303)
- Fujita, M. Krugman, P. Venables, A.J.** (1999). *The spatial economy*. MIT PRESS.
- Gabaix, X.** (1999). Zipf's Law for Cities: An Explanation. *The Quarterly Journal of Economics*, 114(3): 739–767. DOI: [10.1162/003355399556133](https://doi.org/10.1162/003355399556133)
- Gabaix, X. Ioannides, Y.M.** (2004). The Evolution of City Size Distributions. In: *Handbook of Regional and Urban Economics*: 2341–2378. DOI: [10.1016/S0169-7218\(04\)07053-4](https://doi.org/10.1016/S0169-7218(04)07053-4)
- Garmestani, A.S. Allen, C.R. Gallagher C.M.** (2008). Power laws, discontinuities and regional city size distributions. *Journal of Economic Behavior and Organization*, 68(1): 209–216. DOI: [10.1016/j.jebo.2008.03.011](https://doi.org/10.1016/j.jebo.2008.03.011)
- Garner, B.** (1967). Models of urban geography and settlement location. In: R.J. Chorley, P. Haggett (Eds), *Models in geography*, 303–360. Methuen Co. Ltd.
- 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. DOI: [10.1016/j.jue.2010.03.007](https://doi.org/10.1016/j.jue.2010.03.007)
- González-Val, R. Ramos, A. Sanz-Gracia, F. Vera-Cabello, M.** (2015). Size distributions for all cities: Which one is best? *Papers in Regional Science*, 94(1): 93–115. DOI: [10.1111/pirs.12037](https://doi.org/10.1111/pirs.12037)
- Ioannides, Y.M. Overman, H.G.** (2003). Zipf's law for cities: An empirical examination. *Regional Science and Urban Economics*, 33(2): 127–137. DOI: [10.1016/S0166-0462\(02\)00006-6](https://doi.org/10.1016/S0166-0462(02)00006-6)
- Jażdżewska, I.** (2006). K. Zipf's rank-size rule in search of the Best method of point approximation (in Polish). *Czasopismo Geograficzne*, 77(3): 206–219.
- Jażdżewska, I.** (2017). Spatial and dynamic aspects of the rank-size rule method. Case of an urban settlement in Poland. *Computers, Environment and Urban Systems*, 62: 199–209. DOI: [10.1016/j.compenvurbsys.2016.11.006](https://doi.org/10.1016/j.compenvurbsys.2016.11.006)
- Jefferson, M.** (1939). The Law of the Primate City. *Geographical Review*, 29(2): 226–232. DOI: [10.1177/019263656104526231](https://doi.org/10.1177/019263656104526231)
- Jiang, B. Yin, J. Liu, Q.** (2015). Zipf's law for all the natural cities around the world. *International Journal of Geographical Information Science*, 29(3): 498–522. DOI: [10.1080/13658816.2014.988715](https://doi.org/10.1080/13658816.2014.988715)
- Knudsen, T.** (2001). Zipf's Law for Cities and Beyond: The Case of Denmark. *American Journal of Economics and Sociology*, 60(1): 123–146. DOI: [10.1111/1536-7150.00057](https://doi.org/10.1111/1536-7150.00057)
- Lotka, A.** (1925). *Elements of Physical Biology*, 435. Williams and Wilkins Company. <https://doi.org/10.2105/AJPH.15.9.812-b>
- Newman, M.E.J.** (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5): 323–351. DOI: [10.1080/00107510500052444](https://doi.org/10.1080/00107510500052444)
- Pumain, D.** (1989). Spatial dynamics and urban models. In: J. Hauer, H. Timmermans, N. Wrigley (Eds), *Urban Dynamics and Spatial Choice Behaviour*, 155–173. Springer. Netherlands.
- Pumain, D. Swerts, E. Cottineau, C. Vacchiani-Marcuzzo, C. Ignazzi, A. Bretagnolle, A. Delisle, F. Cura, R. Lizzi, L. Baffi, S.** (2015). Multilevel comparison of large urban systems. *Cybergeo: European Journal of Geography*, 7. DOI: [10.4000/cybergeo.26730](https://doi.org/10.4000/cybergeo.26730)
- Richardson, H.W.** (1973). Theory of the distribution of city sizes: Review and prospects. *Regional Studies*, 7(3): 239–251. DOI: <https://doi.org/10.1080/09595237300185241>
- Robson, B.T.** (2012). *Urban growth: An approach*. DOI: [10.4324/9780203716649](https://doi.org/10.4324/9780203716649)
- Rosen, K.T. Resnick, M.** (1980). The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics*, 8(2): 165–186. DOI: [10.1016/0094-1190\(80\)90043-1](https://doi.org/10.1016/0094-1190(80)90043-1)
- Rossi-Hansberg, E. Wright, M.L.J.** (2007). Urban Structure and Growth. *Review of Economic Studies*, 74(2): 597–624. DOI: [10.1111/j.1467-937X.2007.00432.x](https://doi.org/10.1111/j.1467-937X.2007.00432.x)

- Sokołowski, D.** (2001). Application of the weighted regression model to the approximation of rank-size graph (in Polish). In: H. Rogacki, H. (Ed.). *Koncepcje teoretyczne i metody badań geografii społeczno-ekonomicznej i gospodarki przestrzennej*, 169–178. Bogucki Wydawnictwo Naukowe. Poznań.
- Sokołowski, D.** (2014). New towns in Poland. *Bulletin of Geography. Socio-economic Series*, 23: 149-160. DOI: <https://doi.org/10.2478/bog-2014-0010>
- Szymańska, D.** (2007). Urbanizacja na świecie (Urbanization in the world) (wydanie pierwsze 2007; wydanie drugie 2008), Wydawnictwo Naukowe PWN, Warszawa, s. 390.
- Song, S. Zhang, K.H.** (2002). Urbanisation and city size distribution in China. *Urban Studies*, 39(12): 2317–2327. DOI: [10.1080/0042098022000033890](https://doi.org/10.1080/0042098022000033890)
- Soo, K.T.** (2005). Zipf's Law for cities: A cross-country investigation. *Regional Science and Urban Economics*, 35(3): 239–263. DOI: [10.1016/j.regsciurbeco.2004.04.004](https://doi.org/10.1016/j.regsciurbeco.2004.04.004)
- Stewart, J.Q.** (1947). Empirical Mathematical Rules concerning the Distribution and Equilibrium of Population. *Geographical Review*, 37(3): 461-485. DOI: [10.2307/211132](https://doi.org/10.2307/211132)
- Strutz, T.** (2011). *Data Fitting and Uncertainty: A practical introduction to weighted least squares and beyond*. Vieweg + Teubner. DOI: [10.1007/978-3-8348-9813-5](https://doi.org/10.1007/978-3-8348-9813-5)
- US Census Bureau, following: Thomas Brinkhoff,** <https://www.citypopulation.de/USA.html>
- Veneri, P.** (2016). City size distribution across the OECD: Does the definition of cities matter? *Computers, Environment and Urban Systems*, 59: 86–94. DOI: [10.1016/j.compenvurbsys.2016.05.007](https://doi.org/10.1016/j.compenvurbsys.2016.05.007)
- Xu, Z. Harriss, R.** (2010). A spatial and temporal autocorrelated growth model for city rank-size distribution. *Urban Studies*, 43: 14–24. DOI: [10.1177/0042098009348326](https://doi.org/10.1177/0042098009348326)
- Zipf, G.K.** (1941). *National Unity and Disunity, The Nation as a Bio-Social Organism*. The Principia Press Inc.: Blomington, IN.
- Zipf, G.K.** (1949). *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley: Cambridge, MA.