

PEDTR 2019**18th International Scientific Conference “Problems of Enterprise Development:
Theory and Practice”****THE INFLUENCE OF THE POPULATION DISTRIBUTION
SYSTEM ON REGIONAL DEVELOPMENT**

Y. V. Pavlov (a)*

*Corresponding author

(a) Samara State University of Economics, 443090, Soviet Army Str., 141, Samara, Russia, Pavlov-mlad@mail.ru

Abstract

The population distribution system of a particular territory can be uniform or concentrated. Concentrated settlement leads to an increase in the agglomeration effect and accelerate the development of the region. However, excessive concentration can cause a disruption effect, which is the opposite of the agglomeration effect. This article assesses the balance of the population distribution system in accordance with Zipf's law. The research object selected 37 municipalities in one of the regions of Russia – Samara's region. These municipalities cover 100% of the region's population. It was revealed that for the period 2010-2018 the region's population distribution system has become more unbalanced: the share of high-ranking municipalities has increased, the deviation of the actual share from the ideal has increased in 30 of 37 municipalities. The hypothesis about the relationship between the pace of development of the region and the balance of the population distribution system was tested. The annual growth rate of Gross Regional Product per capita at the prices of the previous year was used as an estimated indicator of the region's development rates. It was revealed that there is a negative functional and statistical relationship: the greater the imbalance of the population distribution system relative to the Zipf's law, the lower the pace of development of the region. It is possible to explain this dependence through the effect of agglomeration diseconomies: the development of the largest municipalities in the Samara's region has reached its limit, therefore, conditions must be created for the development of medium, small municipalities.

2357-1330 © 2020 Published by European Publisher.

Keywords: Zipf's law, population distribution system, territorial distribution, rank size distribution, agglomeration effect, agglomeration diseconomies.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The population of a certain territory is settled in the settlements that form the population distribution system. There are two main groups of models of the population distribution system:

- A concentrated model in which the population is unevenly distributed across settlements;
- A smoothed model in which the population is distributed relatively evenly across settlements.

Typically, population distribution systems develop according to a concentrated model. Various economic theories are known that describe the optimal distribution of the population by population centers, for example, the theories of Christaller, Losch, Zipf, etc. The actual distribution of the population may differ from the ideal values calculated according to one or another theory. There is an imbalance in the population distribution system, which can have a positive or negative impact on the socio-economic development of the region.

This article assesses the degree of balance of the region's population distribution system according to Zipf's law. One of the regions of the European part of Russia, the Samara Region, was taken as an object of study. The following is an assessment of the impact of the degree of balance of the population distribution systems according to Zipf's law on the pace of development of the Samara region.

1.1. Review of previous studies

Zipf proposed a formalized theory using the ratio "rank-size", which shows the optimal distribution of the population by settlements (Zipf, 1949). His formula found application in the analysis of population distribution systems at the global, national, regional and even intracity levels (Friesen, Taubenböck, Wurm, & Pelz, 2018). Using Zipf's law, an analysis was made of the population distribution system of many states, for example, Russia (Rusanovskiy, Markov, & Brovkova, 2018), Poland (Cieslik & Teresiński, 2017), USA (Devadoss, & Luckstead, 2016), Brazil (Chauvin, Glaeser, Ma, & Tobio, 2017), Mexico (Pérez-Campuzano, Guzmán-Vargas, & Angulo-Brown, 2015), China (Farrell & Nijkamp, 2019), India (Luckstead & Devadoss, 2014), Pakistan (Arshad, Hu, & Ashraf, 2019) and others.

In some works, it was found that Zipf's law works best when studying objects distinguished not by administrative, but by functional boundaries (Veneri, 2016). The "rank-size" pattern is used when analyzing the distribution of not only the population, but also other indicators, for example, employment (Tsekeris, 2019), firms (Bee, Riccaboni, & Schiavo, 2017), etc.

1.2. Alternative approaches to assessing the degree of balance of the population distribution system

As an approach to revealing the spatial balance of the population distribution system, not only Zipf's law can be applied, but also the Lorentz curve, Gini coefficient (Farrell & Nijkamp, 2019), the Herfindahl-Hirschman index, the calculations of Christaller, Losch, etc. Zipf's law can be used in combination with other spatial patterns, for example, the distance between cities (González-Val, 2019).

2. Problem Statement

2.1. Research hypothesis

The hypothesis of the study is the connection between the pace of development of the region and the balance of the population distribution system. The annual growth rate of GRP (Gross Regional Product) per capita at the prices of the previous year is used as an indicator characterizing the economy of the region. The Zipf's coefficient is used to characterize the degree of balance of the population distribution system.

The concentration of the population can have both positive and negative effects. For example, with concentration, an agglomeration effect occurs, which accelerates development compared to a uniform distribution of the population. However, excessive concentration can lead to the emergence of a agglomeration diseconomies, which is the opposite of the agglomeration effect, which will ultimately undermine the competitive advantages over the even distribution of the population. Of course, the balance of the population distribution system does not determine the development of the region by 100%, because there are many other factors. In addition, the degree of influence of the population distribution system is due to the geographical and economic connectivity of the settlements within this system.

2.2. Study Assumptions

The following assumptions were made as part of the study:

- 1) Cities or another settlements are usually accepted as elements of a population distribution system. In this study, instead of settlements, local administrative unit (which called municipalities) of the first level (they are not included in other municipalities) were investigated: urban districts and municipal areas. Within municipal areas there are several lower-level municipalities (e.g. urban settlements, rural settlements). In this study, such a high level of detail of the population distribution system was not carried out, confining itself to first-level municipalities.
- 2) Municipalities were considered within their administrative boundaries. However, in reality, some administrative units form a single socio-economic system, their built-up territories are combined, as a result of which an agglomeration is formed. Therefore, it would be more logical to consider some integrated municipalities as a single municipality (as an agglomeration). However, in Russia there are no statistics on agglomerations.
- 3) As an indicator characterizing the development of the region, the annual growth rate of GRP per capita at the prices of the previous year was taken. An alternative indicator could be labor productivity, average monthly wages, etc.

3. Research Questions

The study posed the following questions:

- 1) Does the degree of balance of the population distribution system affect the pace of development of the region?
- 2) If the degree of balance of the population distribution system has an effect, is it positive or negative?

4. Purpose of the Study

In accordance with the research questions, the following goals were set:

- 1) Find the deviation of the region's population distribution system from the ideal value according to Zipf's law;
- 2) To study the relationship between the indicator of the balance of the population distribution system and the annual growth rate of GRP per capita at the prices of the previous year;
- 3) Draw conclusions.

The practical significance of this research is expressed in the fact that the revealed dependencies will show the advisability of choosing one or another policy in the formation of a population distribution system (population concentration; the policy of "unloading" of large municipalities).

5. Research Methods

5.1. Zipf's law, other methods and data sources

Zipf's law is described by the formula (1):

$$P_n = \frac{P_1}{n} (1)$$

where P_n – optimal population of a city with a rank "n"; P_1 – population of the largest city of this population distribution system.

Note that there are more complex interpretations of Zipf's law, for example, in the form of constructing a linear regression dependence between the decimal logarithm of the actual population and the logarithm of rank (Zipf, 1949; Andreev, Lukyanova, & Kadyshev, 2018):

$$\lg N = A + a * \lg R \quad (2)$$

where N – population of a city with a rank "R"; R – rank; a – Zipf's coefficient; A – constant.

If $|a| > 1$, then there is a concentration of the population in large cities. If $|a| < 1$, then there is a concentration of the population in medium and small cities. The constant "A" shows the target population of the largest city to approximate the distribution according to Zipf's law (when "a" equals "-1").

The study uses regression, correlation and comparative analyzes. The source of statistical information was the website of the Federal State Statistic Service of Russia (Federal State Statistic Servi..., 2018).

5.2. Object of study

In Russia, it is possible to distinguish the following population distribution systems: at the local level, subregional (part of the region), regional, at the level of the federal district (several regions), at the state. However, within the state and the federal district, the institutional conditions are uneven: regional policy, the volume of transfers from the federal center, and so on are different. Therefore, it would be logical to conduct an assessment at the regional level. The object of research is the Samara region. The analyzed period: 2010-2018.

It is important to answer the question about the sample of settlements within the region. In paragraph 2.2. it has already been said that first level municipalities were taken. When applying the Zipf's law, there are different approaches to sampling:

- 1) A mixed sample of first-level municipalities (urban districts) and second-level municipalities (urban settlements). Those, those municipalities that are conditionally referred to as "cities" are selected. Sometimes a selection is added according to the criterion "population" (for example, cities with 100 thousand people are selected);
- 2) A mixed sample of first-level municipalities (urban districts) and second-level municipalities (urban settlements, rural settlements);
- 3) In this study, an alternative approach is used: only the municipalities of the first level (urban districts and municipal areas) are analyzed. Such a sample will cover 100% of the region's population, and it is not necessary to collect statistics on small rural settlements in municipal areas.

In the Samara's region there are 37 first-level municipalities (which are not included in any other municipalities): 10 urban districts, one of which is an urban district with intracity division, 27 municipal districts. Samara as an urban district with intracity division includes 9 intracity districts (municipalities of the second level). Within 27 municipal districts there are 296 municipalities of the second level: 12 urban settlements and 284 rural settlements. I.e. total 305 municipalities of the second level (Government of the Samara's region, 2018).

6. Findings

6.1. Description of the research object

The population of the Samara's region at the beginning of 2018 is 3193514 people who are accommodated in 37 first-level municipalities (Table 01).

Table 01. Distribution of the population of the Samara region by municipalities (2018)

Municipality	Population	Rank	Perfect share	Actual share
Samara	1163440	1	23,80%	36,43%
Togliatti	707408	2	11,90%	22,15%
Syzran	172070	3	7,93%	5,39%
Novokuybyshevsk	104279	4	5,95%	3,27%
Volzhsky municipal district	99500	5	4,76%	3,12%
Stavropol municipal district	73794	6	3,97%	2,31%
Chapaevsk	72778	7	3,40%	2,28%
Kinel	58239	8	2,98%	1,82%
Zhigulevsk	57687	9	2,64%	1,81%
Krasnoyarsk municipal district	57226	10	2,38%	1,79%
Otradny	47180	11	2,16%	1,48%
Sergievsky municipal district	45193	12	1,98%	1,42%
Kinel-Cherkasy municipal district	44266	13	1,83%	1,39%
Bezenguijsky municipal district	39774	14	1,70%	1,25%
Neftegorsky municipal district	33138	15	1,59%	1,04%
Kinelsky municipal district	32552	16	1,49%	1,02%
Pokhvistnevo	29203	17	1,40%	0,91%
Pokhvistnevsky municipal district	27317	18	1,32%	0,86%

Oktyabrsk	26438	19	1,25%	0,83%
Syzran municipal district	24604	20	1,19%	0,77%
Borsky municipal district	23717	21	1,13%	0,74%
Privolga municipal district	23321	22	1,08%	0,73%
Koshkinsky municipal district	22081	23	1,03%	0,69%
Shigonsky municipal district	19430	24	0,99%	0,61%
Bolsheglushitsky municipal district	18503	25	0,95%	0,58%
Bolshevvernigovsky municipal district	17499	26	0,92%	0,55%
Krasnoarmeysky municipal district	17052	27	0,88%	0,53%
Pestrovka municipal district	16575	28	0,85%	0,52%
Hvorostyansky municipal district	16165	29	0,82%	0,51%
Shentalinsky municipal district	15278	30	0,79%	0,48%
Chelno-Vershinsky municipal district	14755	31	0,77%	0,46%
Klyavlinsky municipal district	14452	32	0,74%	0,45%
Bogatovsky municipal district	14355	33	0,72%	0,45%
Isaklinsky municipal district	12363	34	0,70%	0,39%
Alekseevsky municipal district	11799	35	0,68%	0,37%
Kamyshlinsky municipal district	10638	36	0,66%	0,33%
Elkhovsky municipal district	9445	37	0,64%	0,30%

Source: author based on (Federal State Statistic Service of Russia, 2018).

6.2. Perfect share according to Zipf's law, actual share of municipalities in the region's population

It is necessary to compare the actual distribution with the perfect. The perfect share was found as follows (Table 02 & Table 03):

- 1) It is suggested that a city with a rank "1" already has a perfect population (in step 5, this assumption is reviewed);
- 2) Perfect population can be found for all other municipalities: the rank is known for all municipalities, so formula (1) can be used;
- 3) The total perfect population of all municipalities is found (through summation);
- 4) The perfect share of each municipality from the total perfect population is found;
- 5) Based on the actual total population of all municipalities and the known perfect share of each municipality, we can find the perfect population of each municipality (taking into account of the actual total population).

Table 02. Hypothetical example of calculating the perfect population for a fictitious population distribution system of 4 municipalities

Rank	Actual population, thousand people	Perfect population, thousand people (step 2)	Perfect share from the total perfect population (step 4)	Perfect population (taking into account of the actual total population), thousand people (step 5)
1	1200 (step 1)	1200	48,00%	1046
2	700	600	24,00%	523
3	170	400	16,00%	349
4	110	300	12,00%	262
-		Bcero		
-	2180	2500 (step 3)	100%	2180

Notes. The parentheses indicate the stages of identifying the perfect population size.

Source: author.

All 37 municipalities of the Samara's region are divided into 3 approximately equal groups (12, 12, 13 municipalities). Each group can be conditionally classified taking into account the population as relatively large, relatively medium, relatively small. For a real example of the Samara's region, the perfect shares are shown in Table 03 (Table 03). These values, when they are rounded to the nearest hundredth, are relevant for each year from the period 2010-2018.

Table 03. The perfect share of each group of municipalities in the total population in the Samara's region (relevant for each year from the period 2010-2018)

Year	Share of 12 municipalities with a rank of 1-12	Share of 12 municipalities with a rank of 13-24	Share of 13 municipalities with a rank of 25-37
Perfect distribution	73,86%	16,01%	10,13%

Source: author.

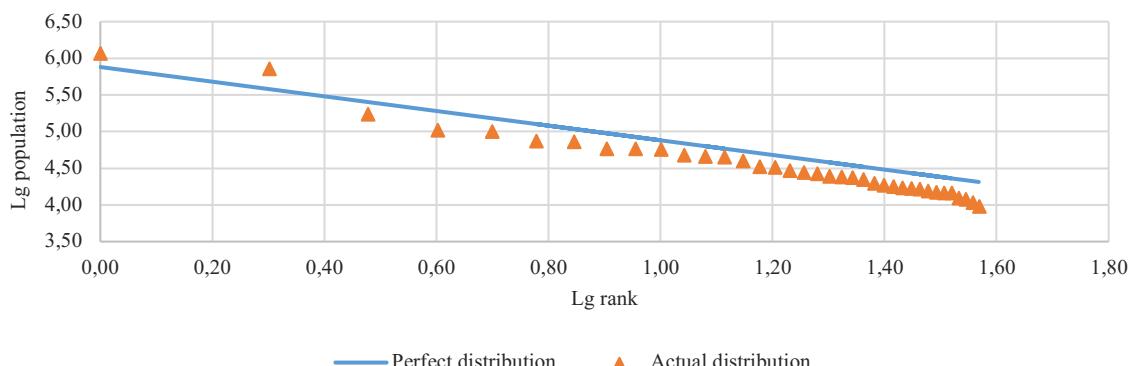
Table 04 shows the actual shares of each group (Table 04).

Table 04. Actual share of each municipal groups in the population of the Samara's region

Year	Share of 12 municipalities with a rank of 1-12	Share of 12 municipalities with a rank of 13-24	Share of 13 municipalities with a rank of 25-37
2010	82,35%	11,31%	6,34%
2011	82,48%	11,24%	6,28%
2012	82,63%	11,17%	6,20%
2013	82,75%	11,11%	6,15%
2014	82,85%	11,04%	6,11%
2015	82,99%	10,97%	6,04%
2016	83,05%	10,94%	6,00%
2017	83,15%	10,89%	5,96%
2018	83,26%	10,83%	5,91%

Source: author.

Thus, the actual share of large municipalities in 2018 is 83.26% (Table 04), which is almost 10% higher than the perfect share of 73.86%. Based on the fact that, according to Zipf's law, the perfect shares of each municipality are known, it is possible to construct a graph showing the actual and perfect distribution of the population (Figure 01).

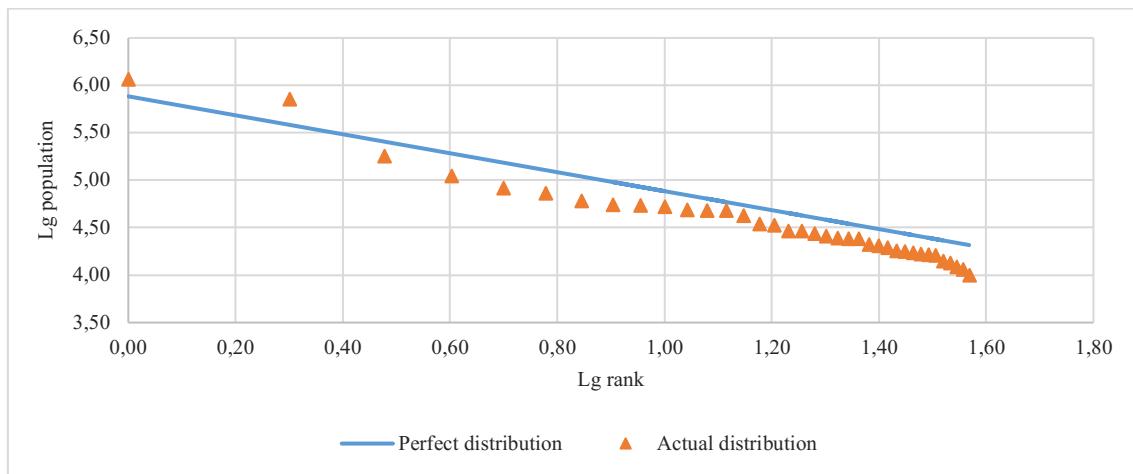


Source: author.

Figure 01. Municipality size distribution according to Zipf's law in the Samara's region for 2018

From Figure 01 it can be seen that only two municipalities in the Samara Region have a population higher than the recommended one; all other municipalities have a shortage of population, which makes the system unbalanced.

Figure 02 shows a graph of the actual and perfect distribution of the population for 2010 (Figure 02).



Source: author.

Figure 02. Municipality size distribution according to Zipf's law in the Samara's region for 2010

We apply formula (2) to find the parameters of the regression equation that reflects the population distribution according to Zipf's law for the Samara region (Table 05). A regression equation is constructed in which "x" are the logarithms of the ranks, "y" are the logarithms of the population.

Table 05. Testing the Zipf's law for municipalities in the Samara's region

Year	Constant "A"	Zipf's coefficient "a"	R ²
2010	5,909	-1,164	0,96
2011	5,911	-1,168	0,96
2012	5,916	-1,174	0,96
2013	5,919	-1,178	0,97
2014	5,922	-1,182	0,97
2015	5,928	-1,188	0,97
2016	5,932	-1,193	0,97
2017	5,936	-1,198	0,97
2018	5,940	-1,203	0,97

Notes. All coefficients are significant by the p-value criterion at the level of 99%. Some indicators are rounded to three decimal places for comparison purposes.

Source: authors.

The regression equation is obtained according to the Zipf's law for 2018 (3):

$$\lg N = 5,94 - 1,203 * \lg R \quad (3)$$

The dynamics of the Zipf's coefficient "a" for the period 2010-2018 years shows its growth, which suggests that relatively large municipalities have increased their share in the population. This is consistent with the data in Table 04 (population growth in the group of the first 12 municipalities). The population distribution system of the Samara's region is becoming increasingly unbalanced with a shift to the dominance of large municipalities.

The visual analysis of Figures 01 and 02 gives the impression that the actual distribution of the population in 2018 compared to 2010 has approached the perfect level. However, calculations show that in 2018, compared with 2010, the deviation of the actual share in the population from the perfect increased in 30 of the 37 municipalities of the Samara's region. So the population distribution system as a whole has become truly less balanced.

6.3. The impact of the population distribution system balance on the pace of development of the region

It is necessary to find out whether there is a functional relationship between the balance of the population distribution system and the annual growth rate of GRP per capita in the prices of the previous year. The Zipf coefficient "a" acts as a factor variable, and the rate of change in GRP as the resulting variable (Table 06).

Table 06. Data on the rate of change in GRP

Year	Annual growth rate of GRP (Gross Regional Product) per capita at the prices of the previous year
2010	109,0%
2011	106,1%
2012	104,6%
2013	105,1%
2014	101,7%
2015	99,1%
2016	97,5%
2017	100,9%
2018	-

Source: author.

The obtained linear regression equation is shown in formula (4):

$$(Annual\ growth\ rate\ of\ GRP\ per\ capita) = 4,49 + 2,93 * a \quad (4)$$

The p-value criterion shows that the equation is significant with a probability of more than 99%. The determination coefficient R^2 is 0.81; 81% of the variation in the rate of development depends on the variation of the population distribution system.

There is a negative functional relationship: the greater the imbalance in the population distribution system according to Zipf's law, the lower the growth rate of GRP per capita per year.

Evaluation of the statistical relationship gives similar results: the correlation coefficient "r" is 0.90; the statistical relationship is strong (in the range of 0.7-0.9); with $(8-2) = 6$ degrees of freedom, the actual value of the correlation exceeds the critical value of 0.834 with a significance level of 99%. The connection is negative: the stronger the unbalanced system, the lower the pace of development of the region.

7. Conclusion

The analysis showed that the population distribution system of the Samara's region for the period 2010-2018 increased its imbalance in accordance with Zipf's law. The municipalities with high rank have increased their share in the region's population. However, an increase in population concentration did not lead to the expected effect of agglomeration. On the contrary, the pace of development of the region began

to decline. Accordingly, regional policy should take into account that in the context of the intensive development of “growth points” and the achievement of the agglomeration effect, it is important not to upset the balance of the existing population distribution system. The revealed dependence can be explained by the influence of agglomeration diseconomies: pulling an excessive amount of resources into one “growth point” leads to their inefficient use, which can slow down the development of the region. To reduce imbalances in the population distribution system, it is necessary to create comfortable conditions for work and life in medium and small municipalities of the Samara's region. The development of this research is possible by comparing estimates of the balance of the population distribution system according to Zipf's law with estimates by other theories.

References

- Andreev, V. V., Lukyanova, V. Y., & Kadyshev, E. N. (2018). Analysis of people territorial distribution in regions of the Volga Federal District on the base of Zipf and Gibrat laws. *Applied Econometrics*, 48, 97–121. [in Rus.].
- Arshad, S., Hu, S., & Ashraf, B. N. (2019). Zipf's law, the coherence of the urban system and city size distribution: Evidence from Pakistan. *Physica A: Statistical Mechanics and its Applications*, 513, 87-103.
- Bee, M., & Riccaboni, M., & Schiavo, S. (2017). Where Gibrat meets Zipf: Scale and scope of French firms. *Physica A: Statistical Mechanics and its Applications*, 481, 265-275.
- Chauvin, J. P., Glaeser, E., Ma, Y., & Tobio, K. (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics*, 98, 17-49.
- Cieslik, A., & Teresiński, J. (2017). Does Zipf's law hold for Polish cities? *Miscellanea Geographica*, 20 (4), 5-10.
- Devadoss, S., & Luckstead, J. (2016). Size distribution of U.S. lower tail cities. *Physica A: Statistical Mechanics and its Applications*, 444, 158-162.
- Farrell, K., & Nijkamp, P. (2019). The evolution of national urban systems in China, Nigeria and India. *Journal of Urban Management*, 8(3), 408-419.
- Federal State Statistic Service of Russia (2018). Database of indicators of municipalities. As of the beginning of 2018. Retrieved from <https://www.gks.ru/storage/mediabank/munst.htm>
- Friesen, J., Taubenböck, H., Wurm, M., & Pelz, P. F. (2018). The similar size of slums. *Habitat International*, 73, 79-88.
- González-Val, R. (2019). The spatial distribution of US cities. *Cities*, 91, 157-164.
- Government of the Samara's region (2018). Local government. As of the beginning of 2018. Retrieved from https://www.samregion.ru/cities_regions/
- Luckstead, J., & Devadoss, S. (2014). A nonparametric analysis of the growth process of Indian cities. *Economics Letters*, 124(3), 516-519.
- Pérez-Campuzano, E., Guzmán-Vargas, L., & Angulo-Brown, F. (2015). Distributions of city sizes in Mexico during the 20th century. *Chaos, Solitons & Fractals*, 73, 64-70.
- Rusanovskiy, V. A., Markov, V. A., & Brovkova, A. V. (2018). Modeling the effects of spatial localization in urban agglomerations of Russia. *Ekonomicheskaya Politika*, 13(6), 136-163. [in Rus.]. DOI: 10.18288/1994-5124-2018-6-136-163
- Tsekeris, T. (2019). Rank-size distribution of urban employment in labour market areas. *Cities*, 95, 102472.
- Veneri, P. (2016). City size distribution across the OECD: Does the definition of cities matter? *Computers, Environment and Urban Systems*, 59, 86-94.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, M.A.: Addison-Wesley Press.