

# Size Distribution of States, Counties, and Cities in the USA: New Inequality Form Information

**Gennady A. Grachev**

Southern Federal University, Rostov-On-Don, Russia, E-mail: [grachev@sfedu.ru](mailto:grachev@sfedu.ru)

## Abstract

In this article, we propose a new approach for studying the patterns of size distribution in settlement systems, based on the analysis of the shape of the Pareto curve (PC). To study the shape of the PC, we used the Gini coefficient, the asymmetry coefficient, and, by analogy with the physics of phase transitions, critical exponent — the index of the PC degree in the neighborhood of zero. An empirical analysis of the PC of various levels of aggregation in the US settlement system has been performed. The form of size distribution of states was studied by decades from 1790 to 2010. The spatial analysis of the PC shape for counties and cities was performed for 2010. The results of an empirical study showed that the PC of the states had left-hand asymmetry over 220 years. The PC of districts and cities had both right-hand and left-hand asymmetries. The obtained results explain in which cases the Pareto distribution having a PC with right-hand asymmetry, and the lognormal distribution with a symmetric PC may not correspond objectively to real settlement systems. As an alternative to power-series distribution and lognormal distribution, we considered an analytically simple two-parameter model with a wide range of PC asymmetry that combines the properties of power-series distribution and lognormal distribution. Verification of the model showed that it adequately described the size of settlements in homogeneous settlement systems.

**Keywords:** Pareto (Lorentz) curve, Saez inequality Shape, Gini coefficient, asymmetry coefficient, critical exponent, Pareto curve model.

## 1. Introduction

The relevance of studying the order regularities for settlement systems is determined by the search for optimal ways to develop production, distribution and consumption. In the first half of the 20th century, empirical studies of this problem focused on the distribution of cities by size (population) in developed countries, especially in the United States. The fundamental principles in size distribution were formulated by Felix Auerbach (Auerbach, 1913) and Robert Gibrat (Gibrat, 1931). Having compared lists of cities ranked in descending order of size in five European countries and the United States, Auerbach concluded that for large cities, the product of city size by city rank is approximately equal to the size of the largest city (the rank-size rule) in the country, which formally corresponds to the Pareto distribution with an exponent equal to 1. Gibrat proposed a proportional growth rule, which states, if applied to urban settlement systems, that large cities would not, on average, grow faster or slower than smaller cities. The proportional growth rule results in a logarithmically normal distribution. The problem is that if true distribution is lognormal, then the entire distribution can never simultaneously correspond to the Pareto distribution. All subsequent studies of settlement systems can be briefly described as an analysis of situations where one of these two distributions is preferable to the other (for example, see (Rosen and Resnick, 1980; Carroll, 1982; Clauset, Shalizi and Newman, 2009; Berry and Okulicz-Kozaryn, 2012; Soo, 2012; Bee *et al.*, 2013; Arshad, Hu and Ashraf, 2018; Bee, Riccaboni and Schiavo, 2019)). However, in many cases, neither power-series distribution nor lognormal distribution describe settlement systems in a satisfactory way (for example, see (Soo, 2012)). Given this fact, as early as in 1982, Glenn Carroll aimed at extracting new information (the basis) from the existing empirical data to explain the reason for this discrepancy (Carroll, 1982).

This paper suggests that the missing information can be obtained from the analysis of the shape of the Pareto curve (PC).

The objective of this paper is to select tools for studying the form of inequality of empirical PC and to analyze the shape of the PC at various levels of aggregation in the US settlement system.

The paper is structured as follows. Section 2 briefly outlines the theoretical issues underlying subsequent empirical analysis. For a description of the empirical data and its processing methods, see Section 3. The results of the analysis of the PC shape for three-level settlement system in the US are presented in Section 4. Section 5 deals with the discussion of the results obtained.

## 2. Mathematical symbols and definitions

### 2.1. Pareto Curve

Let the settlement system consist of  $n$  elements (states, counties, cities) that have  $w_i$  shares of population. Let us arrange the system elements in descending order of  $w$ . As a result, we have a sorted (ranked) list of elements, in which the ordinal number of each element is called its rank and is denoted by the letter  $r$ . The settlement with the largest population will have rank 1, the second most populous settlement will have rank 2, and so on. The function  $w_r = w(r)$  is called discrete rank distribution of the settlement system. The share of population ( $S_k$ ) living in  $k$  settlements with ranks 1, 2, ...,  $k$  is equal to  $S_k = \sum_{r=1}^k w_r$ . By convention,  $\sum_{r=1}^n w_r = 1$ . The graph of dependency  $S_k$  on the percentage of settlements ( $p_k = k/n$ ) is called a discrete Pareto function.

A Pareto curve (PC) is a non-negative non-decreasing differentiable function  $S(p)$  that interpolates the discrete function  $S_k$ . From the definition of PC, the inequality  $p \leq S(p) \leq 1$  for  $p \in [0,1]$  follows, as also boundary conditions:  $S(0) = 0$ ,  $S(1) = 1$ .

The PC shape characterizes the degree of uneven size distribution of settlements. The steeper the curve and the further it is from the line of absolute equality, the greater the inequality in the size distribution, and vice versa.

Despite the fact that PC does not depend on the number of settlements, for a given  $n$ , it enables to calculate the shares of population using the formula below:

$$w(p_r) = S(p_r) - S(p_{r-1}). \quad (1)$$

Formally, for  $n \gg 1$ , the approximate equality is valid:

$$w(p) \approx \mu S'(p), \quad p \in (0,1), \quad (2)$$

where  $\mu = 1/n$ .

At the same time, the ratio (2) applies only to small-sized settlements, where relative difference between two sequential elements is very small. The first ranking elements of the system that differ significantly in size cannot be expressed in terms of a PC derivative (Grachev, 2009, 2010). Consequently, even formally, the proportions of the first ranking elements of the system cannot be described by the rank-size rule, which is confirmed by numerous empirical studies (for example, see the review (Arshad, Hu and Ashraf, 2018)).

It follows from (2) that at the point  $p_\mu$  that satisfies the equation  $S'(p_\mu) = 1$ , the equality  $w(p_\mu) = \mu$  is valid. In other words, the average-sized element of the system has a coordinate on the axis of shares of ranks at the point  $p_\mu$ . The point  $(p_\mu, S(p_\mu))$  at which the PC derivative is equal to 1 is at the maximum distance from the egalitarian line  $y = x$  (Kakwani, 1980).

## 2.2. Measuring the form of inequality

### 2.2.1. Gini coefficient

The most common integral measure of size inequality is the Gini coefficient (Gini, 1921):

$$G = 2 \int_0^1 (S(p) - p) dp = 2 \int_0^1 S(p) dp - 1. \quad (3)$$

When integrating (2) by parts, we have:

$$G = 2 \left( pS(p) \Big|_0^1 - \int_0^1 pS'(p) dp \right) - 1 = 1 - 2R, \quad (4)$$

where  $R = \int_0^1 pw(p) dp$ .

From relations (3) and (4), it can be seen that  $G$  can be interpreted as the average distance from the PC to the egalitarian line, as the double area between the PC and the egalitarian line (Xu, 2005), and as the normalized coordinate of the center of mass of the system.

The Gini coefficient ranges from a minimum value of zero, where all elements are equal, to a theoretical maximum of 1, where all elements except one have a size of zero.

The main disadvantage of  $G$  is that it sums up PCs of different shapes in the same way (Voitchovsky, 2005; Osberg, 2017; Clementi *et al.*, 2019).

### 2.2.2. PC asymmetry coefficient

Let us specify  $(p_a, S(p_a))$  a point where the PC intersects the alternative diagonal  $y = 1 - x$  (Kakwani, 1980). By convention:

$$p_a + S(p_a) = 1. \quad (5)$$

Pareto curves are symmetric with respect to the alternative diagonal when the equality  $p_\mu = p_a$  from which the PC symmetry condition (Kakwani, 1980) follows is satisfied:

$$p_\mu + S(p_\mu) = 1. \quad (6)$$

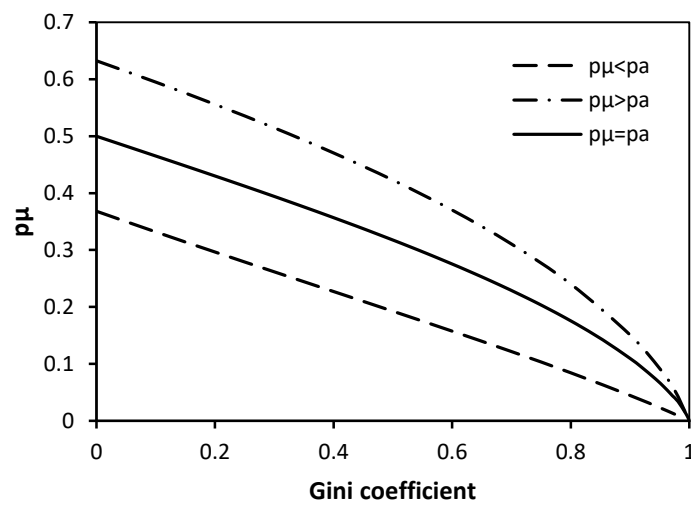
PCs have a right-hand asymmetry when  $p_\mu < p_a$ , and a left-hand asymmetry when  $p_\mu > p_a$ . Thus, the difference  $p_\mu - p_a$  can serve as an indicator of PC asymmetry. The larger the difference modulus, the greater the asymmetry.

There are three single-parameter PCs with different asymmetries in **Table 1**. In the first line, there is a power-series PC with a right-hand asymmetry, in the second line — a power-series PC with a left-hand asymmetry, in the third line — a symmetric PC of the Burr III distribution (Burr, 1942).

**Table 1.** Single-parameter PCs

Parameters	PC	$G$	$p_\mu$	$S(p_\mu)$	$p_\mu - p_a$
$0 < \beta < 1$	$p^\beta$	$(1 - \beta)/(1 + \beta)$	$\beta^{\frac{1}{1-\beta}}$	$\beta^{\frac{\beta}{1-\beta}}$	$< 0$
$\alpha \geq 1$	$1 - (1 - p)^\alpha$	$(\alpha - 1)/(\alpha + 1)$	$1 - \alpha^{\frac{1}{1-\alpha}}$	$1 - \alpha^{\frac{\alpha}{1-\alpha}}$	$> 0$
$\alpha \geq 1$	$(1 - (1 - p)^\alpha)^{\frac{1}{\alpha}}$	$\frac{\Gamma(1/\alpha)^2}{\alpha \Gamma(2/\alpha)} - 1$	$1 - 0.5^{\frac{1}{\alpha}}$	$0.5^{\frac{1}{\alpha}}$	$0$

The dependencies  $p_\mu$  on  $G$  for these PCs are shown in **Fig. 1**.



**Fig. 1.** Dependencies  $p_\mu$  on  $G$

From the curves presented in **Fig. 1**, it can be seen that the maximum of the asymmetry is at  $G = 0$ , and the minimum is at  $G = 1$ . Consequently, an increase in size differentiation in the system leads to a decrease in the PC asymmetry.

Most often, the classical coefficient of asymmetry ( $A$ ) is used to measure the PC asymmetry (Dagum, 1980; Kakwani, 1980; Damgaard and Weiner, 2000):

$$A = p_{\mu} + S(p_{\mu}). \quad (7)$$

For symmetric PCs,  $A = 1$ . If  $A > 1$ , then the PC has a left-hand asymmetry, and a right-hand asymmetry — when  $A < 1$ . When calculating the lower and upper boundaries of  $A$ , we will find:

$$\min A = \lim_{\beta \rightarrow 1} \left( \beta^{\frac{\beta}{1-\beta}} + \beta^{\frac{1}{1-\beta}} \right) = \frac{2}{e} \approx 0.736; \quad (8)$$

$$\max A = 2 - \lim_{\alpha \rightarrow 1} \left( \alpha^{\frac{\alpha}{1-\alpha}} + \alpha^{\frac{1}{1-\alpha}} \right) \approx 1.264. \quad (9)$$

We introduce the normalized asymmetry coefficient ( $\gamma$ ), which has values in a finite interval  $\gamma \in [-1, 1]$ :

$$\gamma = \frac{A - 1}{0.264} = \frac{p_{\mu} + S(p_{\mu}) - 1}{0.264}. \quad (10)$$

For symmetric PCs  $\gamma = 0$ , for PCs with a left-hand asymmetry —  $\gamma > 0$ , for PCs with a right-hand asymmetry —  $\gamma < 0$ .

Note that it follows from the dependencies presented in **Fig. 1** that the 80/20 ratio (Pareto, 1897) obtained by Wilfredo Pareto when studying the distribution of land in Italy only occurs when  $G \approx 0.787$  (Grachev, 2009), which explains why the Pareto principle does not apply in all cases (Grachev, 2010, 2011c, 2011d, 2011a, 2011e, 2011b, 2012, 2013a, 2013b; Verbitsky and Grachev, 2013).

### 2.2.3. Critical exponent of PC

The decrease in the symmetry of the system during the transition from the state of a lack of order ( $S(p) = p$ ) to the state of order ( $S(p) > p$ ,  $0 < p < 1$ ) can be interpreted as a second-kind phase transition. Based on this analogy, we represent  $S(p)$  in the neighborhood of zero ( $p \ll 1$ ) as a power function (Stanley, 1971):

$$S(p) \approx Ap^{\frac{1}{D}} \text{ for } p \ll 1, \quad (11)$$

where  $A, D$  are constants specific to each settlement system.

The parameter  $D$  is interpreted as the fractal dimension of the order parameter, and the reciprocal  $(1/D)$  is called critical exponent (Stanley, 1971). Note that the fractal dimension in (11) differs from the fractal dimension of Yanguang Chen (Chen, 2016) not only in size, but also in physical meaning.

The critical exponent provides objective information about the shape of the PC in the neighborhood of zero and can be used to compare the size of the first ranking elements of the system. For example, the primacy index — the ratio of the size of the two first ranking settlements — can be calculated using the formula:

$$\frac{w_1}{w_2} = \frac{S_1}{S_2 - S_1} = \frac{1}{2^{\frac{1}{D}} - 1}. \quad (12)$$

From the rank-size rule,  $(w_1/w_2 = 2)$  we find:  $1/D = 0.585$ . Thus, a primate city that has a size of  $w_1 > 2w_2$  (Jefferson, 1939) appears in settlement systems when their fractal dimension  $D \geq 1.71$ ; accordingly, the critical exponent is less than or equal to 0.585. The emergence of primates can be viewed as the result of asymmetric size competition, in which larger elements of the system suppress the growth of their smaller neighbors. Models of such competition can be found in ecology (for example, see (Weiner and Damgaard, 2006)).

#### 2.2.4. The Pareto curve model

The search for distributions that replace the lognormal distribution and combine the properties of the power-series distribution and lognormal distribution has been going on for several decades (for example, see (Champernowne, 1956; Ghosh and Basu, 2019; Sánchez, 2019)). An analytically simple two-parameter PC model that gives a wide range of asymmetries can be obtained by generalizing the single-parameter PCs presented in **Table 1** (Antoniou *et al.*, 2004):

$$S(p) = \left(1 - (1-p)^\alpha\right)^\beta, \quad 1 \leq \alpha, \quad 0 < \beta \leq 1. \quad (13)$$

The obvious advantages of the model (13) are a simple formula to calculate the PC and an explicit analytical expression for the Gini coefficient:

$$G(\alpha, \beta) = \frac{2\Gamma\left(\frac{1}{\alpha}\right)\Gamma(\beta+1)}{\alpha\Gamma\left(\frac{1}{\alpha} + \beta + 1\right)} - 1. \quad (14)$$

It is known that the model(13) has shown good results in the study of inequality of incomes of the population(Sarabia, Jordá and Trueba, 2013).

To estimate the critical exponent of the model (13), we use the limit (Stanley, 1971):

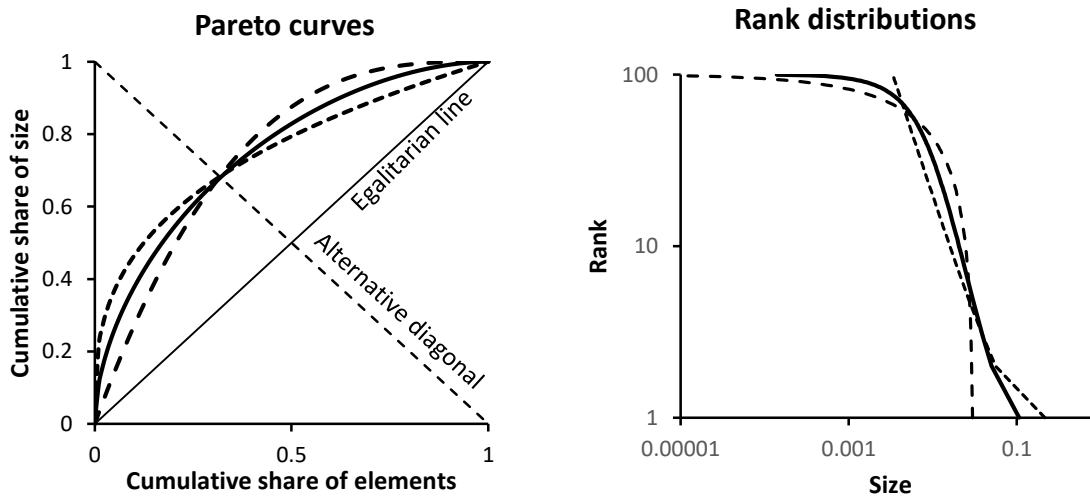
$$\frac{1}{D} = \lim_{p \rightarrow 0} \frac{\ln S(p)}{\ln p}. \quad (15)$$

When applying L'Hôpital's rule, we will find:

$$\frac{1}{D} = \lim_{p \rightarrow 0} \frac{\ln(1 - (1-p)^\alpha)^\beta}{\ln p} = \beta. \quad (16)$$

The relation .(16) suggests that in the neighborhood of zero, the two-parameter PC (13) does not depend on the parameter  $\alpha$  and can be approximated by power function with a critical exponent  $\beta$ .

According to an empirical study (Ghosh, Chattopadhyay and Chakrabarti, 2014; Inoue *et al.*, 2015; Chatterjee, Ghosh and Chakrabarti, 2017; Banerjee *et al.*, 2020), with  $G \leq 0.6$  all PCs that have the same  $G$  value intersect the alternative diagonal at a single point. From the PCs presented in **Fig. 2**, which have the same values  $G=0.5$ , it can be seen that the model (13) describes this phenomenon adequately.



**Fig. 2.** PC and rank distributions of the two-parameter model



**Fig. 2** is also interesting because it clearly demonstrates the possibility of unambiguously determining the critical exponent and the ambiguity of evaluating the exponent of the Pareto distribution over rank distributions.

### 3. Empirical data and processing methodology

#### 3.1. Data

The main source of empirical data for state sizes (the first level of aggregation) is the US census (Historic US Census data). This census has been conducted every decade since 1790. In the first census in 1790, there were only 18 states. Since 1959, the United States comprise 50 States. However, data for all 51 states, including the district of Columbia, has been present in the census since 1900.

As a database of sizes of counties (the second level of aggregation) and cities (the third level), the United States used website <http://www.citypopulation.de/en/usa/>.

#### 3.2. Data processing methodology

To calculate the Gini coefficient, the ratio (4) was used, which for a discrete data set was calculated using the formula (Sen, 1973):

$$G = \frac{n+1}{n} - \frac{2}{n} \sum_{r=1}^n r w_r. \quad (17)$$

The estimates of the parameters  $\alpha$  and  $\beta$  were obtained using the Solution Search function in Microsoft Excel. As an objective function, we have used the minimum of the Kolmogorov — Smirnov statistics that is most common for abnormal data, which is equal to the maximum of differences between the empirical and theoretical PCs:

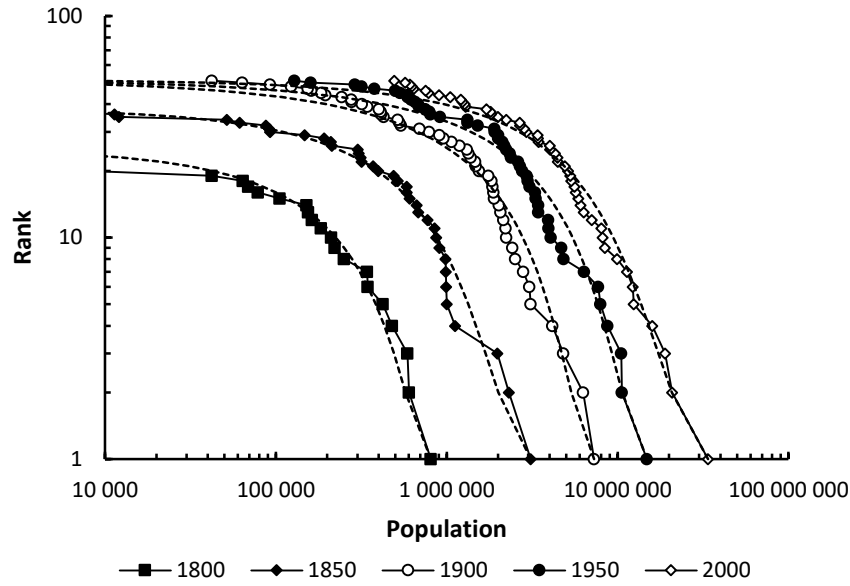
$$KS = \sup_x |S(x_i) - S(x_i; \hat{\alpha}, \hat{\beta})|, \quad (18)$$

where  $S(x_i; \hat{\alpha}, \hat{\beta})$  represents the estimated Pareto curve.

### 4. Results

#### 4.1. The system of settlement in states

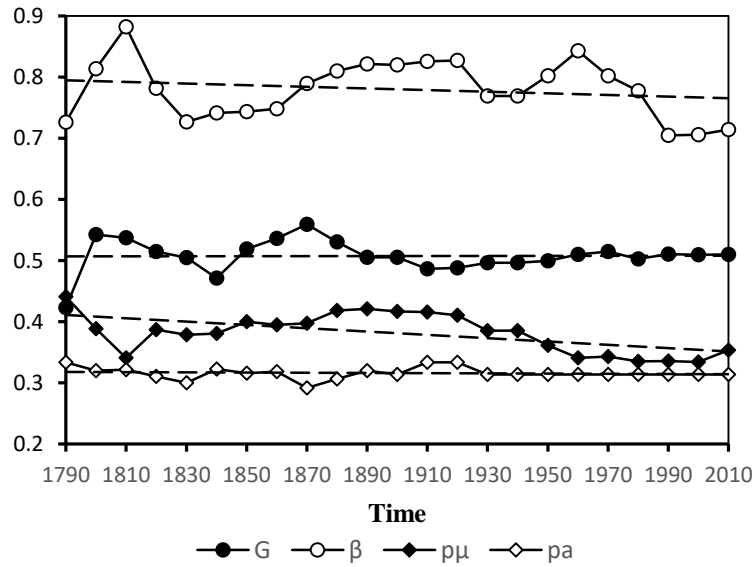
The first and most aggregated level of the US settlement system is the settlement in states. The rank distributions of states in 1800, 1850, 1900, 1950, and 2000 in the bilogarithmic coordinate system are shown by markers in **Fig. 3**. The results of their approximation by the model (13) are indicated with a dotted line.



**Fig. 3.** Rank distributions of state sizes

It can be seen from the rank distributions presented in **Fig. 3** that the model (13) adequately describes the sizes of states in the entire range of ranks.

The dynamics of the parameters  $G$ ,  $\beta$ ,  $p_a$  and  $p_\mu$  is indicated in **Fig. 4** by markers, and the linear trends — by a dashed line. The standard error of the exponent does not exceed 0.005, and the asymmetry coefficient – 0.01.

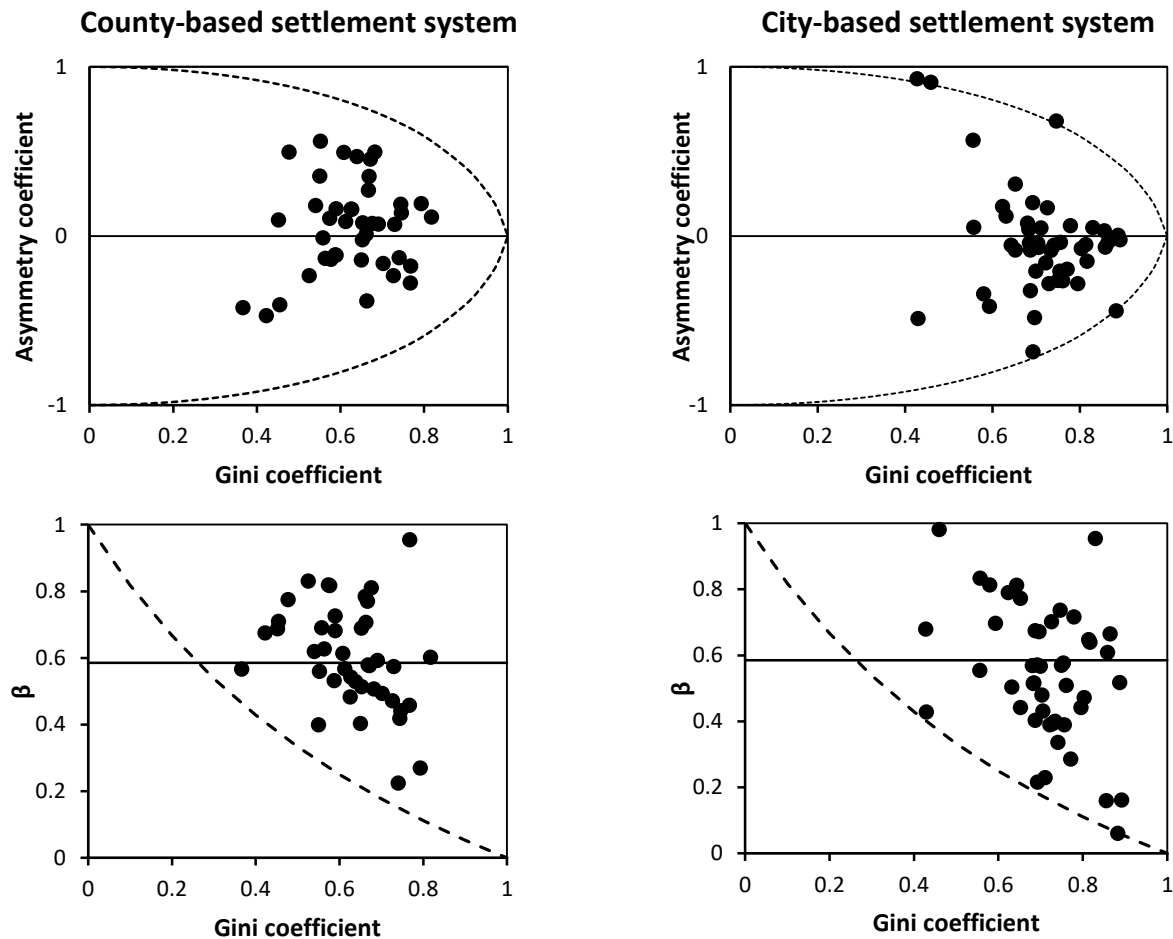


**Fig. 4.** Behavior of the order of settlement in states

The behavior of PC shape parameters presented in **Fig. 4** shows that for 220 years, the PC of the state settlement system had a left-hand asymmetry ( $p_\mu > p_\sigma$ ) decreasing as the settlement system developed.

#### 4.2. County-based and city-based systems of settlement in states

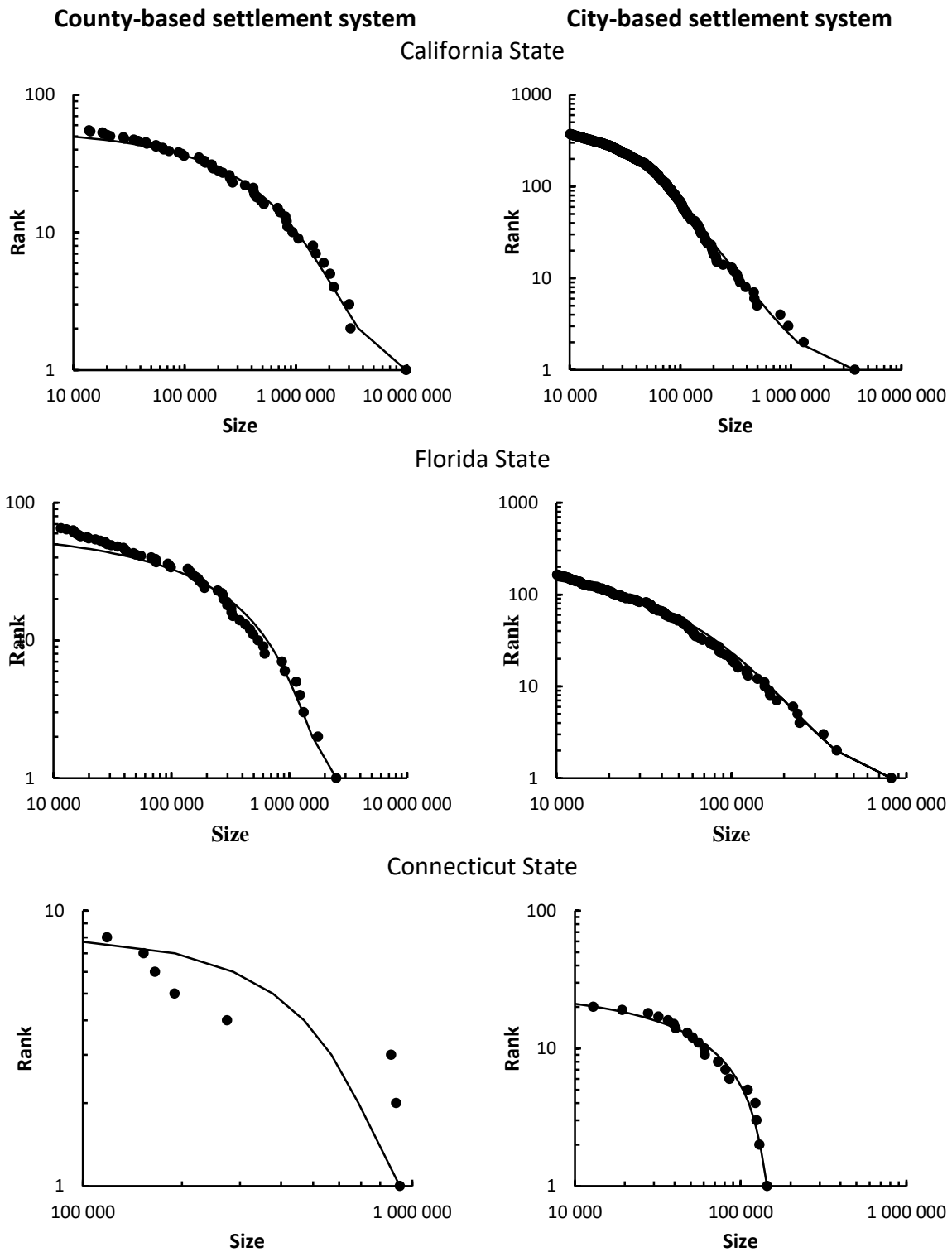
The empirical dependences of  $\beta$  and  $\gamma$  on  $G$  for county-based  $n \geq 14$  and city-based settlement systems are shown in **Fig. 5**.



**Fig. 5.** The dependences of the asymmetry coefficient and the critical exponent of PC for counties and cities on  $G$

The dotted line in **Fig. 5** shows the dependence of the upper and lower boundaries of the asymmetry coefficient on  $G$ . The dashed line shows the dependence of the lower boundary  $\beta$  on  $G$ , as calculated by the formula  $\beta = (1-G)/(1+G)$ , and the solid line is the threshold limiting the occurrence of a primate.

Ranked population distributions by county and city in the states of California, Florida and Connecticut are shown in **Fig. 6**.



**Fig. 6.** County-based and city-based settlement systems of the United States in 2010

It can be seen from the empirical rank distributions presented in **Fig. 6**, that the uniformity of the size distribution in the city-based settlement system is higher than in the county-based settlement system.

## 5. Discussion of the results. Conclusions

The initial premise of this paper is the assumption that the inadequacy of describing settlement systems by a Pareto distribution having a PC with a right-hand asymmetry and by a lognormal distribution having a symmetric PC occurs in cases where real Pareto curves have a left-hand asymmetry.

To test this hypothesis, the paper provides analysis of behavior of the Pareto curve shape in the system of settlement in states and spatial analysis of the PC shape of the county-based and city-based settlement systems of the United States. The Gini coefficient, the asymmetry coefficient and the critical exponent of the PC were used as indicators of the PC shape.

An empirical analysis of the behavior of the PC shape in the system of settlement in states showed that the PC had a left-sided asymmetry between 1790 and 2010, which explains why the Pareto and lognormal distributions do not describe the system of settlement in states in a satisfactory way (Soo, 2012).

Spatial analysis of the systems of settlement in counties and cities in 2010 showed that both systems have approximately equal PC with left-hand and right-hand asymmetries, which explains why Pareto and lognormal distributions are not applicable to these systems in half of the cases.

When analyzing the influence of aggregation of settlement systems on the parameters of the PC shape (see **Table 2**), we can see that the system of settlement in states has the lowest average of the Gini coefficient and the highest average of the critical exponent. The average values of  $G$  and  $\beta$  fall on the system of settlement in counties, while the system of settlement in cities has the highest average Gini coefficient and the lowest average critical exponent. Given that primates can only occur in systems with a critical exponent  $\beta \leq 0.585$ , the result explains why they appear primarily in city-based settlement systems in the United States. However, excluding them from the data sample when evaluating model parameters, as suggested in the paper (Benguigui and Blumenfeld-Lieberthal, 2007), is allowed only in exceptional cases.

**Table 2** Parameters of the PC shape for three levels of the US settlement system

	Average			$G$		$\beta$		$\gamma$	
	$G$	$\beta$	$\gamma$	min	max	min	max	min	max
State	0.51	0.78	0.18	0.42	0.56	0.70	0.88	0.05	0.37
Counties	0.63	0.60	-0.04	0.37	0.82	0.22	0.95	-0.48	0.52
City	0.71	0.54	-0.03	0.43	0.89	0.06	0.98	-0.68	0.96

A verification of a two-parameter model built in (Antoniou *et al.*, 2004) that combines the properties of power-series distribution and lognormal distribution, based on empirical data of the US settlement system, has shown that it adequately describes homogeneous settlement systems.

A comparison of the uniformity of the size of counties and cities in the States showed that the uniformity of the size distribution in the city-based settlement system is higher than that of the district-based settlement system. Most likely, this is due to the fact that the city-based settlement system is self-organized under the influence of objective laws of geographical systems (Batty, 2008), while the aggregation of settlements into counties depends not only on objective historical and local events, but also on subjective views of state leaders on ways to improve the efficiency of the region (Allen, 1996).

Since the government can influence the areas of development of settlement systems (for example, various tax benefits for certain territories, federal subsidies to small cities to attract residents, etc.), the question arises: which asymmetry of the Pareto curves is appropriate to take as a basis for managing the development of settlement systems?

According to the isoperimetric inequality, the more "regular" shape of a flat figure of a given area is, the smaller its perimeter will be (Polya and Szegő, 1951). Or, by the same token, for the same value of the Gini coefficient, the length of the symmetric PC is less than that of asymmetric PC. Given that the isoperimetric inequality is conceptually related to the principle of least action, which explains why oil spots on water surface have the shape of a circle, and a drop of water has the shape of a ball, as well as the decrease in the PC asymmetry of the state-based settlement system over time and the equality of the average asymmetry of counties and state cities to zero, it is logical to assume that a normally functioning system should have a symmetrical PC.

Thus, the study confirmed that in settlement systems, the shape of the Pareto curve has a wide range of asymmetries that cannot be described by Pareto and lognormal distributions.

It has been established that primates in settlement systems occur when the critical exponent is less than 0.585. An increase in the Gini coefficient usually leads to an increase in the probability of a primate occurring.

The well-proven tools for evaluating the PC shape and the two-parameter PC model in this study can be used in other fields of knowledge.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References

- Allen, P. M. (1996) *Cities and Regions as Self-Organizing Systems: Models of Complexity*, Gordon and Breach Science Publishers.
- Antonioni, I. et al. (2004) 'On the efficient resources distribution in economics based on entropy', *Physica A*, 336, pp. 549–562. doi: 10.1016/j.physa.2003.12.051.
- Arshad, S., Hu, S. and Ashraf, N. (2018) 'Zipf's law and city size distribution: A survey of the literature and future research agenda', *Physica A*, 492, pp. 75–92. doi: 10.1016/j.physa.2017.10.005.
- Auerbach, F. (1913) 'Das Gesetz der Bevölkerungskonzentration', *Petermanns Geographische*, pp. 74–76.
- Banerjee, S. et al. (2020) 'On the Kolkata index as a measure of income inequality', *Physica A: Statistical Mechanics and its Applications*. Elsevier B.V., 545, p. 123178. doi: 10.1016/j.physa.2019.123178.
- Batty, M. (2008) 'The size, scale, and shape of cities', *Science*. American Association for the Advancement of Science, pp. 769–771. doi: 10.1126/science.1151419.
- Bee, M. et al. (2013) 'The size distribution of US cities: Not Pareto, even in the tail', *Economics Letters*, 120, pp. 232–237.
- Bee, M., Riccaboni, M. and Schiavo, S. (2019) 'Distribution of city size: Gibrat, pareto, zipf', in *Modeling and Simulation in Science, Engineering and Technology*. Springer Basel, pp. 77–91. doi: 10.1007/978-3-030-12381-9\_4.
- Benguigui, L. and Blumenfeld-Lieberthal, E. (2007) 'A dynamic model for city size distribution beyond Zipf's law', *Physica A*, 384, pp. 613–627. doi: 10.1016/j.physa.2007.05.059.
- Berry, B. J. L. and Okulicz-Kozaryn, A. (2012) 'The city size distribution debate: Resolution for US urban regions and megalopolitan areas', *Cities*. Pergamon, 29(SUPPL. 1), pp. S17–S23. doi: 10.1016/j.cities.2011.11.007.
- Burr, I. W. (1942) 'Cumulative Frequency Functions', *The Annals of Mathematical Statistics*. Institute of Mathematical Statistics, 13(2), pp. 215–232. doi: 10.1214/aoms/1177731607.
- Carroll, G. R. (1982) 'National city-size distributions: what do we know after 67 years of research?', *Progress in Human Geography*, 6(1), pp. 1–43. doi: 10.1177/030913258200600101.
- Champernowne, D. G. (1956) 'Discussion on a paper by Hart and Prais', *Journal of the Royal Statistical Society*, part II(119), pp. 181–183.
- Chatterjee, A., Ghosh, A. and Chakrabarti, B. K. (2017) 'Socio-economic inequality: Relationship between Gini and Kolkata indices', *Physica A: Statistical Mechanics and its Applications*. Elsevier B.V., 466, pp. 583–595. doi: 10.1016/j.physa.2016.09.027.
- Chen, Y. (2016) 'The evolution of Zipf's law indicative of city development', *Physica A: Statistical Mechanics and its Applications*. Elsevier, 443, pp. 555–567. doi: 10.1016/j.physa.2015.09.083.
- Clauset, A., Shalizi, C. R. and Newman, M. E. J. (2009) 'Power-law distributions in empirical data', *SIAM Review*, pp. 661–703. doi: 10.1137/070710111.
- Clementi, F. et al. (2019) 'Mis-measurement of inequality: a critical reflection and new insights', *Journal of Economic Interaction and Coordination*. Springer Verlag. doi: 10.1007/s11403-

Dagum, C. (1980) 'Inequality Measures between Income Distributions with Applications', *Econometrica*. Econometric Society, 48(7), pp. 1791–1803.

Damgaard, C. and Weiner, J. (2000) 'Describing inequality in plant size or fecundity', *Ecology*. John Wiley & Sons, Ltd, 81(4), pp. 1139–1142. doi: 10.1890/0012-9658(2000)081[1139:DIIPSO]2.0.CO;2.

Ghosh, A. and Basu, B. (2019) 'Universal City-size distributions through rank ordering', *Physica A: Statistical Mechanics and its Applications*. Elsevier B.V., 528. doi: 10.1016/j.physa.2019.121094.

Ghosh, A., Chattopadhyay, N. and Chakrabarti, B. K. (2014) 'Inequality in societies, academic institutions and science journals: Gini and k-indices', *Physica A: Statistical Mechanics and its Applications*, 410, pp. 30–34. doi: 10.1016/j.physa.2014.05.026.

Gibrat, R. (1931) *Les Inégalités Économiques: Applications, aux Inégalités des Richesses, a la Concentration des Entreprises, aux Populations des Villes, aux Statistiques des Familles, etc.: d'une Loi Nouvelle la Loi de l'Effet Proportionnel*. Recueil Sirey, Paris.

Gini, C. (1921) 'Measurement of Inequality of Incomes', *The Economic Journal*. Oxford University Press (OUP), 31(121), p. 124. doi: 10.2307/2223319.

Grachev, G. A. (2009) 'Sistemnye zakonomernosti neravenstva dokhodov naseleniya', *Ekonomicheskie nauki (In Russ.)*, (8(57)), pp. 327–331.

Grachev, G. A. (2010) 'Model' optimal'nogo sostoyaniya sistemy gorodskogo rasseleniya'', *Izvestiya Rossiiskoi Akademii Nauk. Seriya Geograficheskaya (In Russ.)*, (3), pp. 46–51.

Grachev, G. A. (2011a) 'Forecasting the optimal structure of the Russian banking system', *Studies on Russian Economic Development (in Russ.)*, 22(5), pp. 103–109. doi: 10.1134/S1075700711050042.

Grachev, G. A. (2011b) 'K analizu strukturnoi effektivnosti gosudarstvennykh soyuzov', *Ekonomicheskie nauki (In Russ.)*, (6(79)), pp. 212–216.

Grachev, G. A. (2011c) *Modelirovanie printsipa Pareto*. Rostov-na-Donu: Izd-vo Yuzhnogo federal'nogo un-ta. (in Russ.).

Grachev, G. A. (2011d) 'On the Assessment of Political Stability on the Basis of Results of Voting at the Elections', *POLIS-Politicheskiye Issledovaniya (In Russ.)*, (5), pp. 123–127.

Grachev, G. A. (2011e) 'Otsenka strukturnoi effektivnosti rashodnoi chasti gosudarstvennykh byudzhetrov', *Terra Economicus (in Russ.)*, 9(3), pp. 10–12.

Grachev, G. A. (2012) 'Assessment of the political stability at the latest Russian parliamentary and presidential elections', *POLIS-Politicheskiye Issledovaniya (In Russ.)*, (3), pp. 30–35.

Grachev, G. A. (2013a) 'Sistemologiya rasseleniya i regional'nogo ustroystva Rossii', *Terra Economicus (in Russ.)*, 11(4), pp. 81–89.

Grachev, G. A. (2013b) 'Upravlenie strukturoi denezhnykh dokhodov naseleniya na osnove printsipa Pareto', *Terra Economicus (in Russ.)*, 11(1), pp. 57–63.

Inoue, J. I. et al. (2015) 'Measuring social inequality with quantitative methodology: Analytical estimates and empirical data analysis by Gini and k indices', *Physica A: Statistical Mechanics and its Applications*. Elsevier, 429, pp. 184–204. doi: 10.1016/j.physa.2015.01.082.



- Jefferson, M. (1939) 'The Law of the Primate City', *Geographical Review*. JSTOR, 29(2), p. 226. doi: 10.2307/209944.
- Kakwani, N. (1980) *Income inequality and poverty: methods of estimation and policy applications*. Published for the World Bank [by] Oxford University Press.
- Osberg, L. (2017) 'On the Limitations of Some Current Usages of the Gini Index', *Review of Income and Wealth*. Blackwell Publishing Ltd, 63(3), pp. 574–584. doi: 10.1111/roiw.12256.
- Pareto, V. (1897) 'Cours d'Économie Politique, Vol. I. Pp. 430. 1896. Vol. II. Pp. 426. 1897. Lausanne: F. Rouge', *The ANNALS of the American Academy of Political and Social Science*. SAGE Publications, 9(3), pp. 128–131. doi: 10.1177/000271629700900314.
- Polya, G. and Szegő, G. (1951) *Isoperimetric Inequalities in Mathematical Physics*. Princeton University Press. doi: 10.1515/9781400882663.
- Rosen, K. T. and Resnick, M. (1980) 'The size distribution of cities: An examination of the Pareto law and primacy', *Journal of Urban Economics*, 8(2), pp. 165–186. doi: 10.1016/0094-1190(80)90043-1.
- Sánchez, E. (2019) 'Burr type-XII as a superstatistical stationary distribution', *Physica A: Statistical Mechanics and its Applications*. Elsevier B.V., 516, pp. 443–446. doi: 10.1016/j.physa.2018.10.044.
- Sarabia, J. M., Jordá, V. and Trueba, C. (2013) 'The Lamé class of Lorenz curves', *Communications in Statistics - Theory and Methods*. Taylor and Francis Inc., 46(11), pp. 5311–5326. doi: 10.1080/03610926.2013.775306.
- Sen, A. (1973) *On economic equality*. Oxford: Clarendon Press.
- Soo, K. T. (2012) 'The size and growth of state populations in the United States', *Economics Bulletin*. AccessEcon, 32(2), pp. 1238–1249.
- Stanley, H. (1971) *Introduction to phase transitions and critical phenomena*. New York: Oxford University Press.
- Verbitsky, E. V. and Grachev, G. A. (2013) 'Analysis of Efficiency of Human Sleep by Pareto Principle', *Human Physiology (in Russ.)*, 39(6). doi: 10.1134/S0362119713060121.
- Voitchovsky, S. (2005) 'Does the Profile of Income Inequality Matter for Economic Growth?: Distinguishing Between the Effects of Inequality in Different Parts of the Income Distribution', *Journal of Economic Growth*, 10, pp. 273–296.
- Weiner, J. and Damgaard, C. (2006) 'Size-asymmetric competition and size-asymmetric growth in a spatially explicit zone-of-influence model of plant competition', *Ecol Res*, (21), pp. 707–712. doi: 10.1007/s11284-006-0178-6.
- Xu, K. (2005) 'How has the Literature on Gini's Index Evolved in the Past 80 Years?', *SSRN Electronic Journal*. Elsevier BV. doi: 10.2139/ssrn.423200.