

Распределение размеров штатов, округов и городов США: новая информация о форме неравенства

*Г. А. Грачёв*¹

Аннотация: В статье предлагается новый подход к исследованию закономерностей распределения размеров в системах расселения, основанный на анализе формы кривой Парето (РС). Для изучения формы РС использованы коэффициент Джини, асимметрии и, по аналогии с физикой фазовых переходов, показатель степени РС в окрестности нуля. Выполнен эмпирический анализ РС различных уровней агрегации системы расселения США. Форму распределения размеров штатов исследовали по десятилетиям с 1790 по 2010 год. Пространственный анализ формы РС округов и городов выполнен для 2010 года. Результаты проведенного эмпирического исследования показали, что РС штатов на протяжении 220 лет имела левостороннюю асимметрию. РС округов и городов имела как правостороннюю, так и левостороннюю асимметрию. Полученные результаты объясняют, в каких случаях распределение Парето, имеющее РС с правосторонней асимметрией, и логнормальное распределение с симметричной РС объективно могут не соответствовать реальным системам расселения. В качестве альтернативы степенному и логнормальному распределению рассмотрена аналитически простая двухпараметрическая модель с широким диапазоном асимметрии РС, объединяющая в себе свойства комбинации степенного и логнормального распределений. Верификация модели показала, что она адекватно описывает размеры поселений однородных систем расселения.

Ключевые слова: кривая Парето, коэффициент Джини, коэффициент асимметрии, критический показатель степени, модель.

¹ Южный Федеральный Университет, Ростов-на-Дону, Россия,

ORCID: 0000-0003-4532-7886, e-mail: grachev@sfedu.ru

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

*Gennady A. Grachev*²

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, 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 till 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 curve, Gini coefficient, asymmetry coefficient, critical exponent, model.

² Southern Federal University, Rostov-On-Don, Russia,

ORCID: 0000-0003-4532-7886, e-mail: grachev@sfedu.ru

Оглавление

1.	Введение	3
2.	Математические символы и определения	4
2.1.	Кривая Парето	4
2.2.	Измерение неравенства	5
2.2.1.	Коэффициент Джини	5
2.2.2.	Коэффициент асимметрии РС	6
2.2.3.	Критический показатель степени	8
2.3.	Модель кривой Парето	9
3.	Эмпирические данные и методика обработки	10
3.1.	Данные	10
3.2.	Методика обработки	10
4.	Результаты	11
4.1.	Система расселения по штатам	11
4.2.	Система расселения по округам	13
4.3.	Городская система расселения	17
5.	Обсуждение и выводы	19
	Литература	21

1. Введение

Актуальность задачи изучения закономерностей упорядоченности систем расселения обусловлена поиском оптимальных путей развития производства, распределения и потребления. В первой половине XX века эмпирические исследования данной проблемы были сосредоточены на изучении распределения городов по размерам (численности населения) в развитых странах, особенно в США. Основополагающие принципы в распределении размеров были сформулированы Феликсом Ауэрбахом (Auerbach, 1913) и Робертом Гибратом (Gibrat, 1931). Ауэрбах, сопоставив ранжированные по убыванию размера списки городов в пяти европейских странах и США, пришел к выводу, что у крупных городов произведение размера города на его ранг приблизительно равно размеру крупнейшего города страны (правило «ранг — размер»), что формально соответствует распределению Парето с показателем степени, равным единице. Гибрат предположил правило пропорционального роста, которое применительно к городским системам расселения говорит о том, что крупные города в среднем не должны расти быстрее или медленнее, чем более мелкие города. Правило пропорционального роста приводит к логарифмически нормальному распределению. Проблема заключается в том, что если истинное распределение является логнормальным, то весь дистрибутив никогда не сможет одновременно соответствовать распределению Парето. Все последующие исследования систем расселения кратко можно охарактеризовать как анализ ситуаций, когда одно из этих двух распределений предпочтительнее другого (см., например, (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)). Вместе с тем во многих случаях ни степенное, ни логнормальное распределения не описывают системы расселения удовлетворительным образом (см., например, (Soo, 2012)). Учитывая данное обстоятельство, еще в 1982 году Гленн Кэрролл поставил задачу извлечения из существующих эмпирических данных новой информации (основы), позволяющей объяснить причину этого несоответствия (Carroll, 1982, p. 37).

В работе сделано предположение, что недостающую информацию можно получить из анализа формы кривой Парето (РС).

Целью работы является выбор инструментария для изучения формы неравенства РС и эмпирический анализ формы РС разных уровней агрегации системы расселения США.

Работа организована следующим образом. В Разделе 2 кратко очерчены теоретические вопросы, лежащие в основе дальнейшего эмпирического анализа. Описание эмпирических данных и методики их обработки дано в Разделе 3. Результаты эмпирического анализа формы РС трехуровневой системы расселения США представлены в Разделе 4. Обсуждению полученных результатов и выводам посвящен Раздел 5.

2. Математические символы и определения

2.1. Кривая Парето

Пусть система поселений состоит из n элементов (штатов, округов, городов), имеющих доли численности населения w_i . Расположим элементы системы в порядке убывания величины w . В результате получаем упорядоченный (ранжированный) список элементов, в котором порядковый номер каждого элемента называется его рангом и обозначается буквой r . Населенный пункт с наибольшей численностью населения будет иметь ранг 1, второй по численности населения — 2 и т. д. Функция $w_r = w(r)$ называется ранговым распределением системы поселений. Доля населения (S_k), проживающего в k населенных пунктах с рангами 1, 2, ..., k , равна $S_k = \sum_{r=1}^k w_r$. По определению, $\sum_{r=1}^n w_r = 1$. График зависимости S_k от доли населенных пунктов ($p_k = k/n$) называют дискретной функцией Парето.

Кривой Парето (РС) называют неотрицательную неубывающую дифференцируемую функцию $S(p)$, интерполирующую дискретную функцию S_k . Из определения РС следует неравенство $p \leq S(p) \leq 1$ для $p \in [0,1]$ и граничные условия: $S(0) = 0$, $S(1) = 1$.

Форма РС характеризует степень неравномерности распределения размеров населенных пунктов. Чем круче изгиб кривой и чем дальше она отстоит от линии абсолютного равенства, тем больше неравенство в распределении размеров, и наоборот.

Несмотря на то что РС не зависит от числа населенных пунктов, при заданном n она позволяет вычислить доли населения по формуле:

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

Формально для $n \gg 1$ справедливо приближенное равенство:

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

где $\mu = 1/n$.

В то же время соотношение (2) применимо только к населенным пунктам с небольшими размерами, у которых относительная разница между двумя последовательными элементами очень мала. В разы, различающиеся по размеру первые по рангу элементы системы не могут быть выражены через производную РС. Следовательно, пропорции первых по рангу элементов системы не могут быть описаны правилом «ранг — размер», что подтверждают многочисленные эмпирические исследования (см., например, обзор (Arshad, Hu and Ashraf, 2018)).

Из (2) следует, что в точке p_μ , удовлетворяющей уравнению $S'(p_\mu) = 1$, выполняется равенство $w(p_\mu) = \mu$. Другими словами, средний по размеру элемент системы имеет координату на оси долей рангов в точке p_μ . Точка $(p_\mu, S(p_\mu))$, в которой производная РС равна единице, находится на максимальном удалении от эгалитарной линии $y = x$ (Kakwani, 1980).

2.2. Измерение неравенства

2.2.1. Коэффициент Джини

Наиболее распространенной интегральной мерой неравенства размеров является коэффициент Джини (Gini, 1912):

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

Интегрируя (2) по частям, получаем:

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

где $R = \int_0^1 pw(p) dp$.

Из соотношений (3), (4) видно, что G можно интерпретировать как среднее расстояние от РС до эгалитарной линии, удвоенную площадь между РС и эгалитарной линией (Xu, 2005), а также как нормированную координату центра масс системы.

Коэффициент Джини варьируется от минимального значения ноль, когда все элементы равны, до теоретического максимума, равного единице, когда все элементы, кроме одного, имеют размер ноль.

Главным недостатком G является то, что он одинаковым образом суммирует РС разной формы (Voitchovsky, 2005; Osberg, 2017).

2.2.2. Коэффициент асимметрии РС

Обозначим $(p_a, S(p_a))$ точку, в которой РС пересекает альтернативную диагональ $y = 1 - x$ (Kakwani, 1980). По определению:

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

Кривые Парето симметричны относительно альтернативной диагонали, когда выполняется равенство $p_\mu = p_a$, из которого следует условие симметрии РС (Kakwani, 1980):

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

РС имеют правостороннюю асимметрию, когда $p_\mu < p_a$, и левостороннюю асимметрию, когда $p_\mu > p_a$.

В **Таблица 1** представлены три однопараметрические РС. В первой строке — степенная РС с правосторонней асимметрией, во-второй — степенная РС с левосторонней асимметрией, в третьей — симметричная РС распределения Burr III (Burr, 1942).

Таблица 1. Однопараметрические РС

Параметры	РС	G	p_μ	$S(p_\mu)$	$p_\mu - p_a$
$0 < \beta < 1$	p^β	$(1-\beta)/(1+\beta)$	$\beta^{\frac{1}{1-\beta}}$	$\beta^{\frac{\beta}{1-\beta}}$	< 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
$\alpha \geq 1$	$\left(1 - (1-p)^\alpha\right)^{\frac{1}{\alpha}}$	$\frac{\Gamma(1/\alpha)^2}{\alpha \Gamma(2/\alpha)} - 1$	$1 - 0.5^{\frac{1}{\alpha}}$	$0.5^{\frac{1}{\alpha}}$	0

Зависимости p_μ от G для этих РС показаны на **Figure 1**.

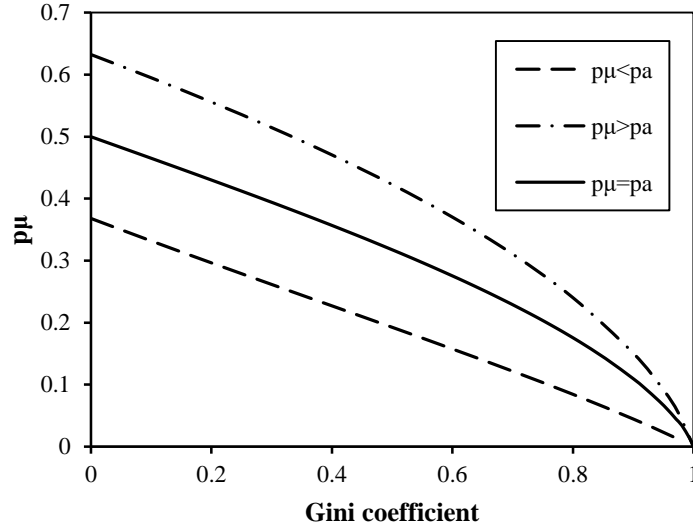


Figure 1. Зависимости p_μ от G

Из представленных на **Figure 1** кривых видно, что $\max |p_a - p_\mu|$ приходится на $G = 0$, а $\min |p_a - p_\mu| = 0$ — на $G = 1$. Таким образом, увеличение дифференциации размеров в системе приводит к уменьшению асимметрии РС.

Классическое определение коэффициента асимметрии (A) (Dagum, 1980; Kakwani, 1980; Damgaard and Weiner, 2000) имеет вид:

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

У симметричных РС $A = 1$. Если $A > 1$, то РС имеет левостороннюю асимметрию, и правостороннюю асимметрию — когда $A < 1$. Вычисляя нижнюю и верхнюю границы A , находим:

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

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

Введем в рассмотрение нормированный коэффициент асимметрии (γ), имеющий значения в конечном промежутке $\gamma \in [-1, 1]$:

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

У симметричных РС $\gamma = 0$, у РС с левосторонней асимметрией — $\gamma > 0$, с правосторонней — $\gamma < 0$.

Отметим, что из представленной на **Figure 1** зависимости p_a от G следует, что полученная Вильфредо Парето при изучении распределения земель в Италии пропорция 80/20 (Pareto, 1916, 1935) имеет место только при $G \approx 0.787$, что объясняет, почему принцип Парето действует далеко не во всех случаях (Грачёв, 2009, 2010, 2011a, 2011b, 2011c, 2012).

2.2.3. Критический показатель степени

Уменьшение симметрии системы при переходе из состояния отсутствия упорядоченности ($S(p) = p$) к состоянию упорядоченности ($S(p) > p$) можно интерпретировать как фазовый переход второго рода. Исходя из этой аналогии, представим $S(p)$ в окрестности нуля ($p \ll 1$) в виде степенной функции (Stanley, 1971):

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

где A, D — константы, особенные для каждой системы расселения.

Параметр D в (11) интерпретируют как фрактальную размерность системы (Allen, 1996), а обратную ему величину ($1/D$) называют критическим показателем степени (Stanley, 1971).

Критический показатель дает объективную информацию о форме РС в окрестности нуля и может быть использован для сравнения размеров первых по рангу элементов системы. Так, например, индекс первенства — отношение размеров двух первых по рангу населенных пунктов — можно вычислить по формуле:

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

Из правила «ранг — размер» ($w_1/w_2 = 2$) находим: $1/D = 0.585$. Таким образом, город-примат, имеющий размер $w_1 > 2w_2$ (Jefferson, 1939), появляется в системах расселения, когда их фрактальная размерность $D \geq 1.71$; соответственно, критический показатель степени меньше или равен 0.585. Появление приматов можно рассматривать как результат асимметричной конкуренции по размеру, при которой более крупные

элементы системы подавляют рост своих более мелких соседей. Модели такой конкуренции можно найти в экологии (см., например, (Weiner and Damgaard, 2006)).

Отметим, что пропорции 80/20 соответствует размерность системы $D = 3$.

2.3. Модель кривой Парето

Поиск распределений, заменяющих логнормальное распределение и объединяющих в себе свойства степенного и логнормального распределений, ведется на протяжении нескольких десятилетий (см., например, (Champernowne, 1956; Ghosh and Basu, 2019; Sánchez, 2019)). Аналитически простую двухпараметрическую модель РС, дающую широкий диапазон асимметрий, можно получить путем обобщения представленных в **Таблица 1** однопараметрических РС (Antoniou *et al.*, 2004):

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

Очевидными достоинствами модели (13) являются простая формула для вычисления РС и явное аналитическое выражение для коэффициента Джини:

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

Известно, что модель (13) хорошо зарекомендовала себя при изучении неравенства доходов населения (Sarabia, Jordá and Trueba, 2013).

Для оценки критического показателя степени модели (13) используем предел (Stanley, 1971):

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

Применяя правило Лопиталя, находим:

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

Соотношение (16) говорит о том, что в окрестности нуля двухпараметрическая РС (13) не зависит от параметра α и может быть аппроксимирована степенной функцией с критическим показателем степени β .

Согласно эмпирическим исследованиям (Ghosh, Chattopadhyay and Chakrabarti, 2014), (Inoue *et al.*, 2015), (Chatterjee, Ghosh and Chakrabarti, 2017), (Banerjee *et al.*, 2020), при выполнении неравенства $G \leq 0.6$ все РС, имеющие одинаковое значение

коэффициента Джини, пересекают альтернативную диагональ в одной точке. Подтверждением этому служит **Figure 2**, на которой показаны симметричная и асимметричные РС, имеющие одинаковое значение $G = 0.5$.

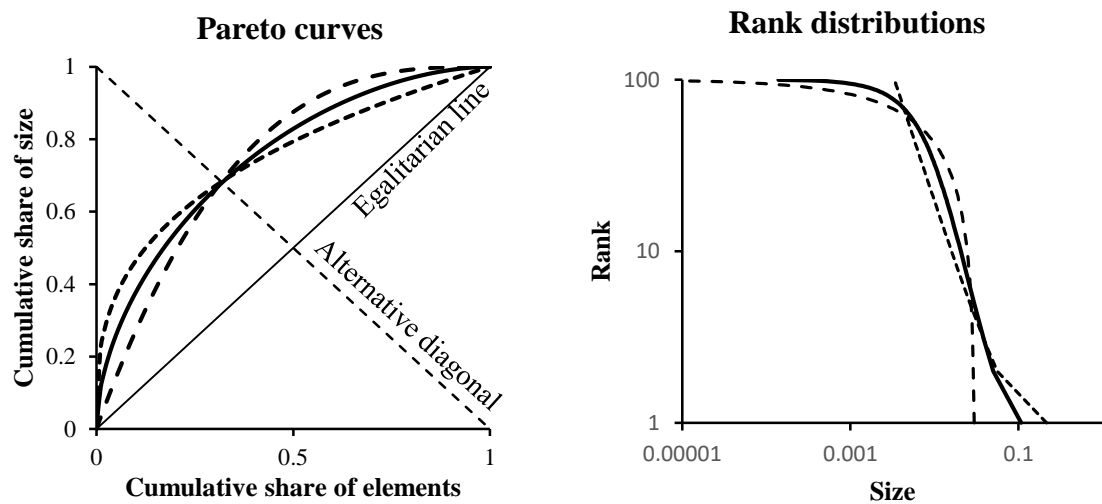


Figure 2. PC и ранговые распределения двухпараметрической модели

Представленные на **Figure 2** ранговые распределения подтверждают, что во всех трех случаях говорить о правиле «ранг — размер» не приходится.

3. Эмпирические данные и методика обработки

3.1. Данные

Основным источником эмпирических данных в настоящей работе является перепись США (Historic US Census data). Эта перепись проводится каждое десятилетие с 1790 года. В первой переписи в 1790 году было только 18 штатов. С 1959 года в состав США входит 50 штатов. Однако данные по всем 51 штату, включая округ Колумбия, присутствуют в переписи начиная с 1900 года.

В качестве базы данных размеров округов и городов США использовали сайт <http://www.citypopulation.de/en/usa/>.

3.2. Методика обработки

Для вычисления коэффициента Джини использовано соотношение (4), которое для дискретного набора данных записано в виде:

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

Оценка параметров α и β была проведена с помощью процедуры «Поиск решения» программы Excel. В качестве целевой функции использована наиболее распространенная для ненормальных данных статистика Колмогорова — Смирнова, которая равна максимуму различий между эмпирической и теоретической РС:

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

где $S(x_i; \hat{\alpha}, \hat{\beta})$ представляет оценочную кривую Парето.

4. Результаты

4.1. Система расселения по штатам

Первым и наиболее агрегированным уровнем системы расселения США является расселение по штатам. Динамика численности населения и количество штатов в переписи приведены в **Таблица 2**.

Таблица 2. Численность населения США и количество штатов в переписи

Year	US population	States	Year	US population	States
1790	3 929 214	18	1910	92 228 496	51
1800	5 308 483	25	1920	106 021 537	51
1810	7 239 881	28	1930	123 202 624	51
1820	9 638 453	30	1940	132 164 569	51
1830	12 860 702	30	1950	151 325 798	51
1840	17 063 353	31	1960	179 323 175	51
1850	23 191 876	38	1970	203 211 926	51
1860	31 443 321	44	1980	226 545 805	51
1870	38 558 371	48	1990	248 709 873	51
1880	50 189 209	49	2000	281 421 906	51
1890	62 979 766	50	2010	308 745 538	51

Year	US population	States	Year	US population	States
1900	76 212 168	51	2018	327 167 434	51

Ранговые распределения размеров штатов в 1800, 1850, 1900, 1950 и 2000 годах в билогарифмической системе координат показаны на **Figure 3** маркерами, модель (13) — пунктирной линией.

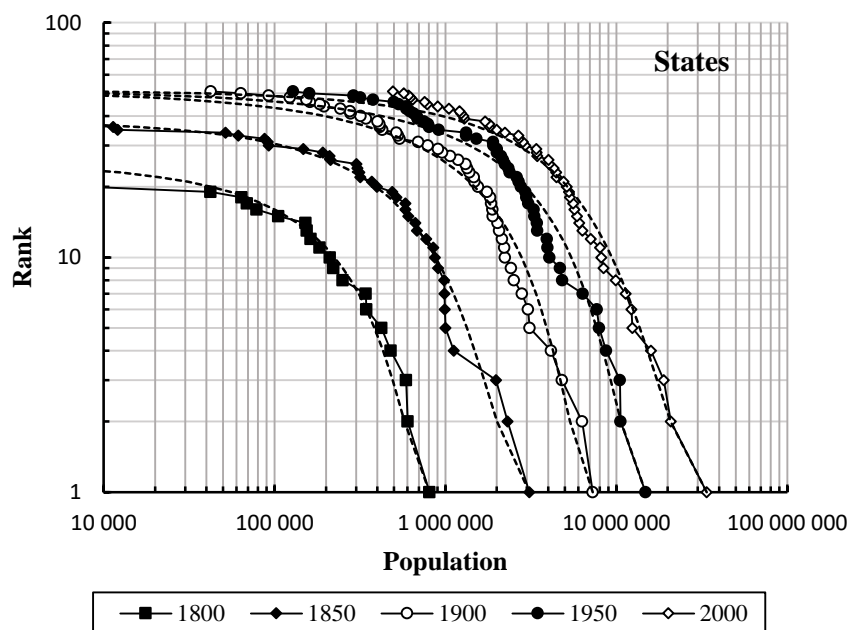


Figure 3. Ранговые распределения размеров штатов

Из визуальной оценки представленных на **Figure 3**. Ранговые распределения размеров штатов видно, что правило «ранг — размер» не является хорошим приближением для описания размеров штатов США. Исследование (Soo, 2012) показало, что использование логнормального распределения несущественно улучшает аппроксимацию представленных эмпирических данных. В то же время из рисунка видно, что модель (13) адекватно описывает размеры штатов во всем диапазоне рангов.

Динамика параметров G , β , p_a и p_μ показана на **Figure 4** маркерами, линейные тренды — штриховой линией. Стандартная ошибка показателя степени 0.005, коэффициента асимметрии 0.01.

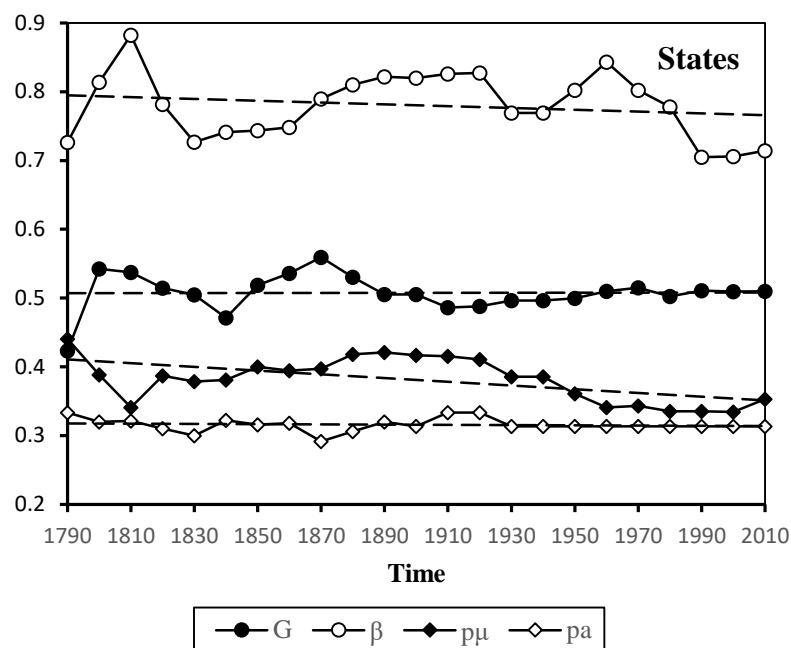


Figure 4. Динамика упорядоченности системы расселения по штатам США

Из представленных на **Figure 4** графиков видно, что на протяжении 220 лет РС системы расселения по штатам имела левостороннюю асимметрию ($p_{\mu} > p_a$), говорящую о том, что логнормальное и Парето распределения, имеющие РС симметричную и правосторонней асимметрией соответственно, не могут адекватно описывать систему расселения по штатам. Критический показатель степени системы расселения по штатам все это время был больше 0.585, что объясняет отсутствие примата у ранговых распределений. Несмотря на то что коэффициент Джини начиная с 1890 года изменялся незначительно, асимметрия РС постоянно уменьшалась, и это еще раз подчеркивает, что РС с разной асимметрией могут иметь одинаковое значение коэффициента Джини.

4.2. Система расселения по округам

Система расселения по округам представляет второй уровень системы расселения США. Численность населения и количество округов в штатах в 2010 году представлены в **Таблица 3**.

Таблица 3. Численность населения и количество округов в штатах в 2010 г.

State	State Population	Counties	State	State Population	Counties
Alabama	4.78	67	Montana	0.99	56
Alaska	0.71	29	Nebraska	1.83	93

State	State Population	Counties	State	State Population	Counties
Arizona	6.39	15	Nevada	2.70	17
Arkansas	2.92	75	New Hampshire	1.32	10
California	37.25	58	New Jersey	8.79	21
Colorado	5.03	64	New Mexico	2.06	33
Connecticut	3.57	8	New York	19.38	62
Delaware	0.90	3	North Carolina	9.54	100
Florida	18.80	67	North Dakota	0.67	53
Georgia	9.69	159	Ohio	11.54	88
Hawaii	1.36	5	Oklahoma	3.75	77
Idaho	1.57	44	Oregon	3.83	36
Illinois	12.83	102	Pennsylvania	12.70	67
Indiana	6.48	102	Rhode Island	1.05	5
Iowa	3.05	99	South Carolina	4.63	46
Kansas	2.85	105	South Dakota	0.81	66
Kentucky	4.34	120	Tennessee	6.35	95
Louisiana	4.53	64	Texas	25.15	254
Maine	1.33	16	Utah	2.76	29
Maryland	5.77	24	Vermont	0.63	133
Massachusetts	6.55	14	Virginia	8.00	39
Michigan	9.88	83	Washington	6.72	55
Minnesota	5.30	87	West Virginia	1.85	72
Mississippi	2.97	82	Wisconsin	5.69	23
Missouri	5.99	115	Wyoming	0.56	14

Из представленных в **Таблица 1** данных видно, что количество округов в штатах не зависит от численности проживающего в них населения. Так, например, наибольший по численности населения (37.3 млн человек) штат Калифорния разделен на 68 округов, а штат Техас, имеющий численность населения на 12 млн человек меньше, разделен на 254 округа.

Обработка данных выполнена для штатов, у которых количество округов не меньше 14. Графический анализ асимметрии РС округов показан на **Figure 5**. Пунктирной линией на **Figure 5** показаны зависимости верхней и нижней границ коэффициента асимметрии от G . Графический анализ зависимости критического показателя степени РС округов от

коэффициента Джини показан на **Figure 6**. Пунктирной линией на **Figure 6** представлена зависимость нижней границы β от G , вычисленная по формуле $\beta = (1 - G)/(1 + G)$, сплошной линией — порог, ограничивающий появление примата.

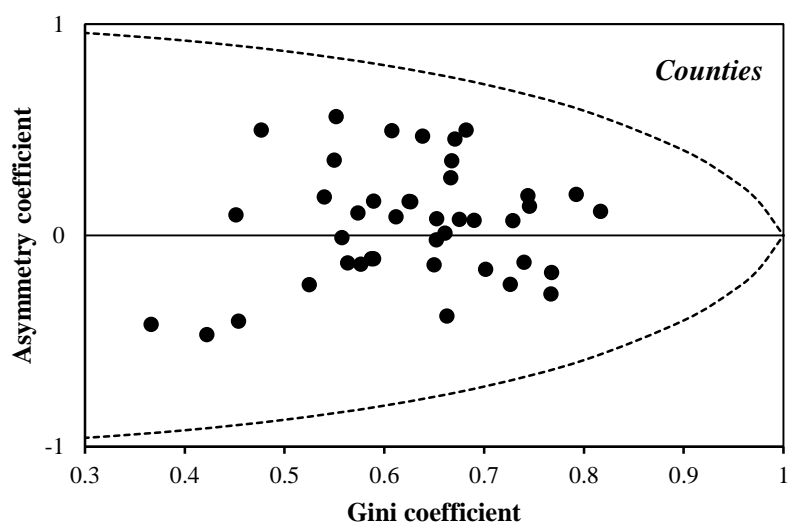


Figure 5. Эмпирическая зависимость асимметрии РС округов от G

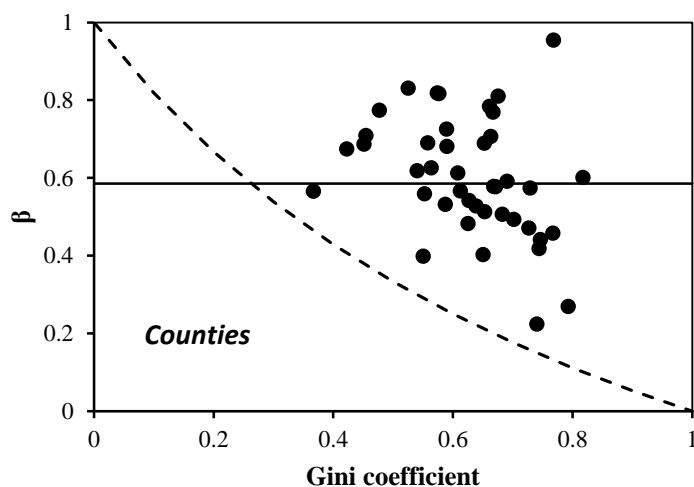


Figure 6. Зависимость критического показателя степени РС округов от G

Из представленных на **Figure 5** и **Figure 6** эмпирических зависимостей видно, что оба параметра упорядоченности систем — коэффициент асимметрии и критический показатель степени — находятся в допустимых для модели (13) пределах.

На **Figure 7** в качестве примера показано распределение населения по округам в шести штатах: Калифорния, Нью-Йорк, Коннектикут, Флорида, Индиана и Вермонт.

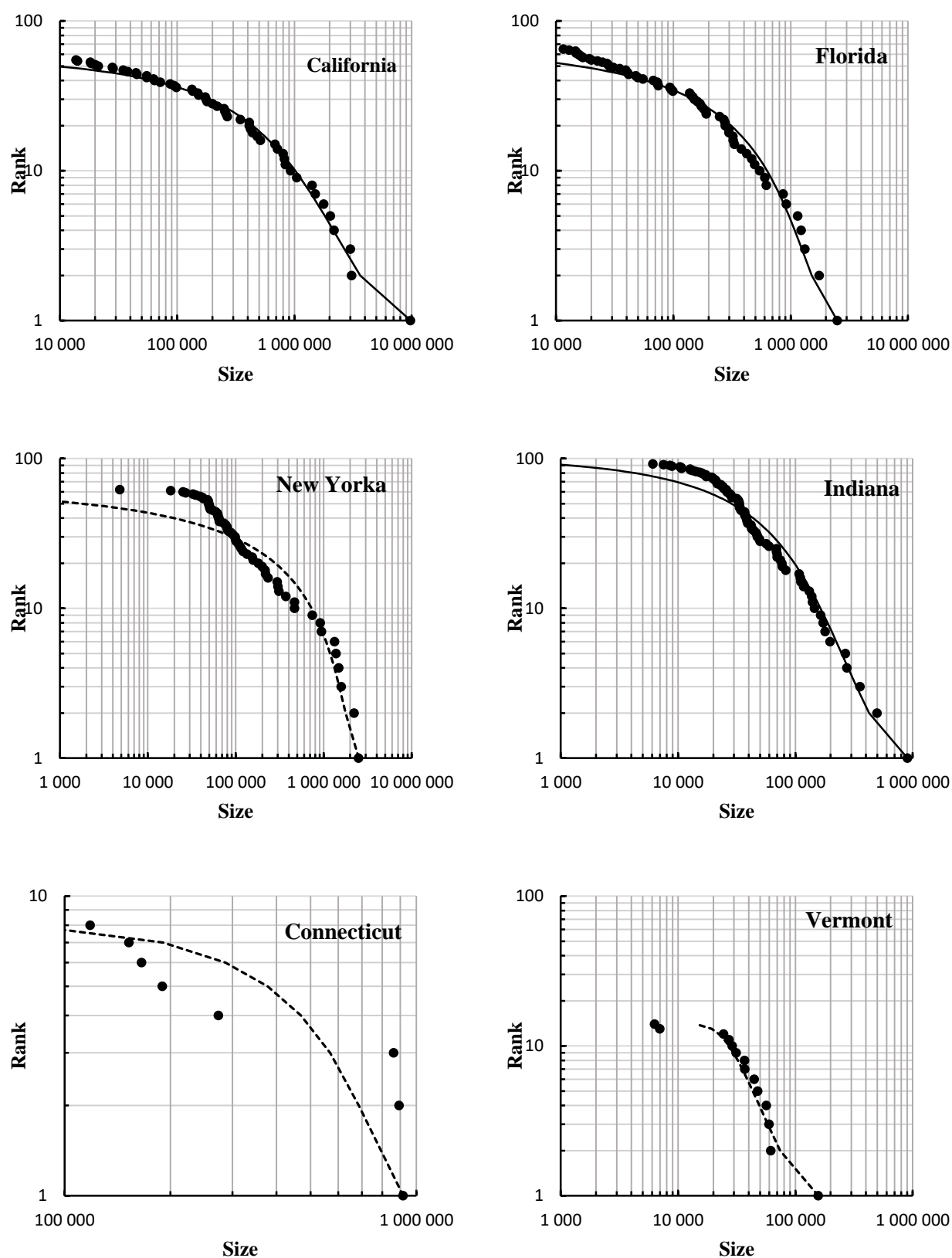


Figure 7. Окружные системы расселения штатов США в 2010 г.

Из представленных на **Figure 7** ранговых распределений видно, что, за исключением штатов Калифорния и Вермонт, они не могут быть описаны единственным выражением,

что указывает на их неоднородность, которая может быть следствием некорректного объединения поселений в округа.

4.3. Городская система расселения

Третьим уровнем административно-территориального деления США являются городские муниципалитеты, имеющие наименьшую агрегацию по сравнению со штатами и округами. Поскольку в США нет однозначно определенного понятия «город», то ниже термин «город» используется для обозначения любого поселения с численностью населения не меньше тысячи человек.

Графический анализ асимметрии РС городов в штатах показан на **Figure 8** жирными точками, пунктирной линией — зависимости верхней и нижней границ коэффициента асимметрии от G . Графический анализ зависимости критического показателя степени РС городов от коэффициента Джини показан на **Figure 9**. На **Figure 10** представлены ранговые распределения городских систем расселения тех же штатов, что и на **Figure 7**.

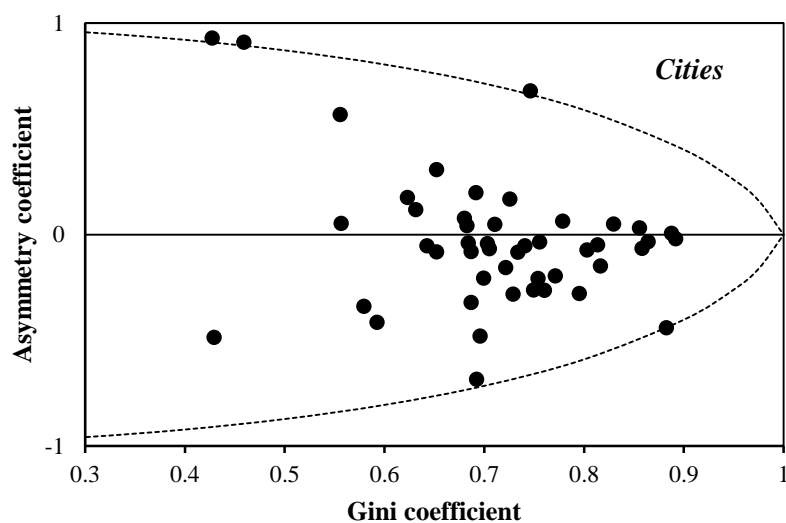


Figure 8. Эмпирическая зависимость асимметрии РС округов от G

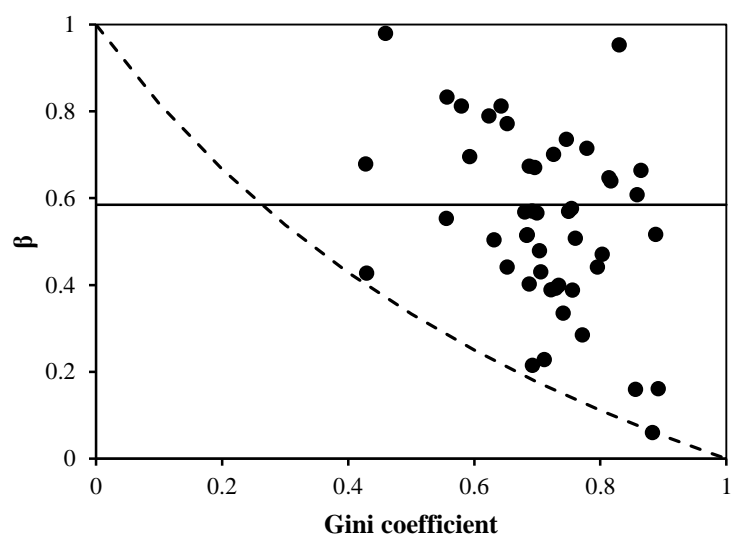
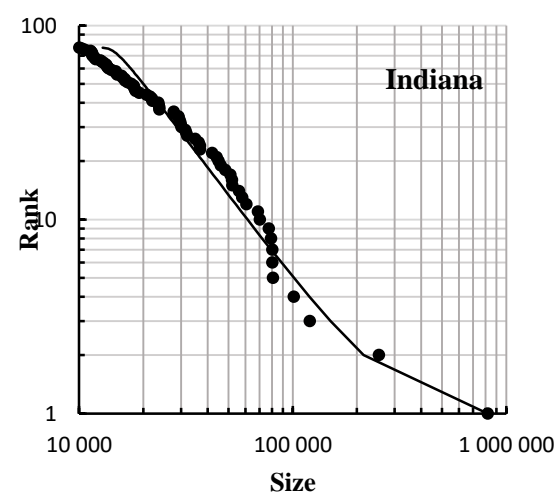
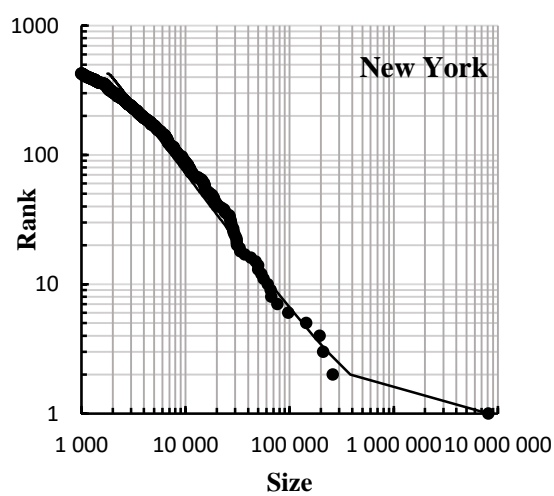
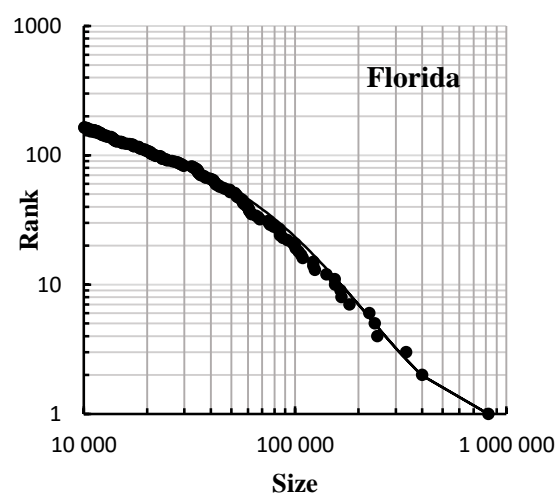
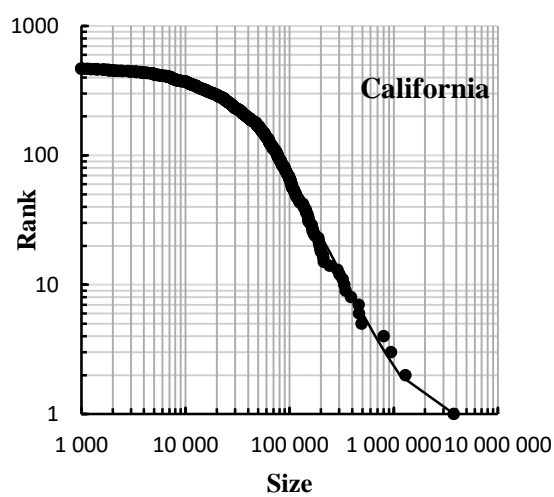


Figure 9. Зависимость критического показателя степени PC городов от G



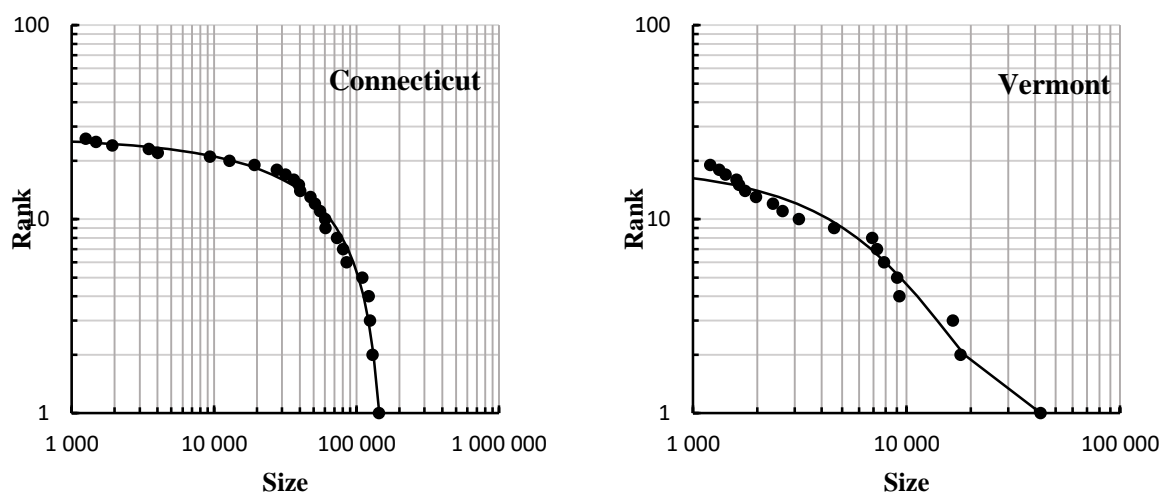


Figure 10. Городские системы расселения штатов США в 2010 г.

Из представленных на **Figure 10** эмпирических зависимостей видно, что все эмпирические значения коэффициента асимметрии и критического показателя степени находятся в допустимых для модели пределах. В то же время по сравнению с системой расселения по округам четыре штата имеют РС с предельной асимметрией, а один штат — с предельным значением критического показателя степени.

Из сравнения ранговых распределений, представленных на **Figure 7** и **Figure 10** видно, что однородность распределения размеров у городской системы расселения выше, чем у окружной.

5. Обсуждение и выводы

Исходной посылкой этой статьи является предположение, что неадекватность описания систем расселения распределением Парето, имеющим РС с правосторонней асимметрией, и логнормальным распределением, имеющим симметричную РС, возникает в тех случаях, когда реальные кривые Парето имеют левостороннюю асимметрию.

Для проверки этой гипотезы в работе выполнен анализ динамики формы кривой Парето системы расселения по штатам и пространственный анализ формы РС окружных и городских систем расселения.

Эмпирический анализ формы РС системы расселения по штатам США показал, что в период с 1790 по 2018 год РС имела левостороннюю асимметрию, и это объясняет, почему Парето и логнормальное распределения не описывают систему расселения по штатам удовлетворительным образом (Soo, 2012).

Анализ формы РС систем расселения по округам и городам в 2010 году показал, что системы расселения США второго и третьего уровней имеют приблизительно поровну РС с лево- и правосторонней асимметрией, что объясняет, почему в половине случаев к этим системам неприменимы Парето и логнормальное распределения.

Анализ влияния агрегации системы на интегральное неравенство размеров и критический показатель степени РС показал, что система расселения по штатам имеет наименьшее значение коэффициента Джини ($G = 0.51 \pm 0.02$) и наибольшее значение критического показателя степени ($\beta = 0.78 \pm 0.05$). Средние значения G и β приходятся на систему расселения по округам ($G = 0.63 \pm 0.017$; $\beta = 0.60 \pm 0.026$), а система расселения по городам штатов имеет наибольшее значение коэффициента Джини ($G = 0.69 \pm 0.02$) и наименьшее значение критического показателя степени ($\beta = 0.54 \pm 0.03$). Учитывая, что приматы могут иметь место только в системах с критическим показателем степени $\beta \leq 0.585$, полученный результат объясняет, почему в США они появляются главным образом в городских системах расселения. Однако исключать их из выборки данных при оценке параметров модели, как это рекомендуют делать в работе (Benguigui and Blumenfeld-Lieberthal, 2007), можно только в том случае, когда критический показатель степени ниже установленного в работе минимума.

Верификация простой двухпараметрической модели, объединяющей в себе свойства степенного и логнормального распределений, на эмпирических данных системы расселения США показала, что она адекватно описывает однородные системы расселения и удовлетворяет всем ограничениям для коэффициента асимметрии и критического показателя степени.

Так как государство может оказывать влияние на направленность развития систем расселения (например, различные налоговые льготы для определенных территорий, федеральные субсидии малым городам для привлечения жителей и т. д.), то возникает вопрос: какую асимметрию кривых Парето целесообразно взять за основу для управления развитием систем расселения?

Согласно изопериметрическому неравенству, чем более «правильную» форму имеет плоская фигура заданной площади, тем меньше ее периметр (Pólya and Szegő, 1951). Или, что то же самое, при одном и том же значении коэффициента Джини длина симметричной РС меньше, чем асимметричной. Учитывая, что концептуально изопериметрическое неравенство связано с принципом наименьшего действия, объясняющим, почему масляные пятна на поверхности воды имеют форму окружности, а

капля воды — форму шара, логично предположить, что нормально функционирующая система должна иметь симметричную РС.

Поскольку проблема распределения размеров представляет интерес для систем любой природы, то полученные в работе результаты и методика изучения формы РС могут найти применение в разных областях знания.

Это исследование не получало грантов от финансирующих агентств государственного, коммерческого или некоммерческого секторов.

Литература

Allen, P. M. (1996) *Cities and Regions as Self-Organizing Systems: Models of Complexity*, Gordon and Breach Science Publishers.

Antoniou, 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.

Bee, M. *et al.* (2013) ‘The size distribution of US cities: Not Pareto, even in the tail’, *Economics Letters*, 120, pp. 232–237. Available at: <https://www.sciencedirect.com/science/article/pii/S0165176513002139> (Accessed: 25 February 2020).

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) ‘Beyond the power law—a new approach to analyze city size distributions’, *Computers, Environment and Urban Systems*, 31, pp. 648–666. doi: 10.1016/j.compenvurbsys.2006.11.002.

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.

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.

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.

Dagum, C. (1980) 'Inequality Measures between Income Distributions with Applications', *Econometrica*. Econometric Society, 48(7), pp. 1791–1803. Available at: <https://ideas.repec.org/a/ecm/emetrp/v48y1980i7p1791-1803.html> (Accessed: 10 July 2020).

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. (1912) *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche*. Bologna: Tipografia di Paolo Cuppini.

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. (1916) *Trattato di sociologia generale*, G. Barbera, Firenze. Available at: https://books.google.ru/books/about/Trattato_di_sociologia_generale.html?id=j04uAAAAYAAJ&redir_esc=y (Accessed: 8 July 2020).

Pareto, V. (1935) *The Mind and Society*, New York, Harcourt, Brace and Company. Available at: <https://www.peterharrington.co.uk/the-mind-and-society-131510.html> (Accessed: 8 July 2020).

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.

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.

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*.

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.

Грачёв, Г. А. (2009) 'Системные закономерности неравенства доходов населения', *Экономические науки*, 8(57), pp. 327–331.

Грачёв, Г. А. (2010) 'Модель оптимального состояния системы городского расселения', *Известия Российской академии наук. Серия географическая*, (3), pp. 46–51.

Грачёв, Г. А. (2011a) 'К оценке политической стабильности по результатам голосования на выборах', *Полис. Политические исследования*, (5), pp. 123–127.

Грачёв, Г. А. (2011b) 'К прогнозированию оптимальной структуры российской банковской системы', *Проблемы прогнозирования*, 22(5), pp. 103–109.

Грачёв, Г. А. (2011c) *Моделирование принципа Парето, Ростов-на-Дону: Изд-во Южного федерального ун-та*.

Грачёв, Г. А. (2012) 'Оценка политической стабильности на прошедших выборах парламента и президента России', *Полис. Политические исследования*, (3), pp. 30–35.