M. Ibragimov, **C. Necula**, **U. Valetka** (Tashkent, Tashkent State University of Economics, Bucharest, Bucharest Academy of Economic Studies, Minsk?, Belarusian State Technological University and Center for Social and Economic Research Belarus). **City Size Distribution in CIS and Eastern Europe countries.**

This work focuses on a comparative study of city size distributions in Russia, Belarus, Poland, Hungary and the countries of the CIS Central Asia (Kazakhstan, Uzbekistan, Kyrgyzstan, Tajikistan and Turkmenistan) and the CIS Caucasus (Armenia, Azerbaijan and Georgia) in the 20th century. Our empirical results confirm that these distributions have heavy tails that follow Zipf's law with the tail index approximately to 1. On the other hand, the whole distributions of all cities in the countries dealt with satisfy the so-called *Weber–Fechner law* that is considered in this context for the first time in this research. We further introduce models based on hierarchy of logarithms and show that they have very good statistical characteristics in fitting the city size distributions under study.

In the last four decades, we have witnessed a rapid expansion of the study of heavy-tailedness and the extreme outliers phenomena in economics and finance, including the analysis of key variables in financial markets as well as the size of cities and firms (see the discussion and reviews in [1, 3, 5] and references therein). A number of studies (see, among others, [2]) have indicated that the distributions of city sizes Z in many developed countries follow so-called *Zipf's law* with tails that satisfy the power law (Pareto-like behavior)

$$\mathbf{P}\{Z > z\} \sim C/z^{\zeta} \qquad (1)$$

with the tail index $\zeta = 1$. This implies that the city size distributions are dominated by a few megapolises. Furthermore, in particular, the first moments of the variables Z are infinite: $\mathbf{E}\, Z = \infty$. Gabaix [2] explains Zipf's Law for city sizes in developed countries by the properties of the migration process that follows the log-normal distribution.

Several approaches to the inference about the tail index ζ of heavy-tailed power law distributions (1) are available in the literature (see, among others, the review in [1]). The two most commonly used ones are Hill's estimator and the OLS approach using the log-log rank-size regression.

Let $Z_1, Z_2, \ldots, Z_N$ be a sample from a city size population satisfying power law (1). Further, let $Z_{(1)} \geqslant Z_{(2)} \geqslant \cdots \geqslant Z_{(n)}$ be decreasingly ordered largest absolute values of observations in the sample. Despite the availability of more sophisticated methods, a popular way to estimate the tail index ζ is still to run the following OLS log-log rank-size regression with $\gamma = 0$:

$$\log(k - \gamma) = a - b \log(Z_{(k)}), \qquad k = 1, 2, \ldots, n, \qquad (2)$$

or, in other words, calling k the rank of an observation, and $Z(k)$ its size: $\log(\mathrm{Rank}_{\,it} - \gamma) = a - b\log(\mathrm{Size}_{\,it})$, with the subscripts city i and t denoting city and year. Unfortunately, tail index estimation procedures based on OLS log-log rank-size regressions (2) with $\gamma = 0$ are strongly biased in small samples. The recent study by Gabaix and Ibragimov [4] provides a simple practical remedy for this bias, and argues that, if one wants to use an OLS regression, one should use the Rank $= 1/2$, and run $\log(\mathrm{Rank}_{\,it} - 1/2) = a - b\log(\mathrm{Size}_{\,it})$. The shift of $\gamma = 1/2$ is optimal, and reduces the bias to a leading order. The standard error on the Pareto exponent ζ is not the OLS standard error, but is asymptotically $(2/n)^{1/2}\zeta$. Numerical results in [4] further demonstrate the advantage of the proposed approach over the standard OLS estimation procedures (2) with $\gamma = 0$ and indicate that it performs well under deviations from power laws and dependent heavy-tailed processes, including GARCH models. This work conducts estimation of parameters in OLS log-log regression (2) with $\gamma = 1/2$ for 20, 10 and 5 percentage tails of city size distributions in the countries considered. This further allows us to determine the tails of the distributions where Zipf's Law begins to hold.

Zipf's Law with $\gamma = 1$ typically holds only for tails of city size distributions that include only very large cities. Therefore, we also consider fitting the whole city size distributions using alternative parametric models such as the Weber–Fechner Law, whose parameters can be estimated using the regression $\log(\mathrm{Size}_{it}) = a - b\log(\mathrm{Rank}_{it})$, where the coefficient γ is the so-called *Weber's constant*. This constant shows how the size changes with the change in the rank. In case of the Weber–Fechner law, the rank of the city changes in arithmetic progression with the change of city size in geometric progression, while in case of the Zipf's law both rank and the city size change in arithmetic progression. Dependence that follows Weber–Fechner law is typical for the response of perception on stimuli in living organisms. This work presents its first application in urban economics. Our results show that, usually, the Weber–Fechner law describes the whole distribution of all cities and other populated areas a country better than power laws (1) and corresponding log-log rank-size regressions (2).

We further introduce extensions of the Weber–Fechnerlaw given by logarithmic hierarchy models $\log(\mathrm{Rank}_{it})c - a_1\log(\mathrm{Size}_{it}) + a_2\log_2(\mathrm{Size}_{it}) + \cdots + \log_m(\mathrm{Size}_{it})$, where $\ln_k y$ denotes the kth iteration of logarithm (i. e. $\ln_k y = \underbrace{\log\log\cdots\log}_{k} y$, $k \geqslant 1$). The empirical results obtained demonstrate very good performance of these models in fitting the whole distributions of city sizes in the countries considered, with the best statistical characteristics among the models dealt with.

СПИСОК ЛИТЕРАТУРЫ

1. *Embrechts P.*, *Klüppelberg C.*, *Mikosch T.* Modelling extremal events for insurance and finance. New York: Springer, 1997.
2. *Gabaix X.* Zipf's law for cities: an explanation. — Quarterly Journal of Economics, 1999, v. 114, p. 739–767.
3. *Gabaix X.* Power laws in economics and finance. — Annual Review of Economics 2009, v. 1, p. 255–293.
4. *Gabaix X.*, *Ibragimov R.* Rank-1/2: A simple way to improve the OLS estimation of tail exponents. — Journal of Business Economics and Statistics, 2011, v. 29, p. 24–39.
5. *Ibragimov R.* Heavy-tailed densities. — In: S. N. Durlauf and L. E. Blume, eds, The New Palgrave Dictionary of Economics Online. Palgrave Macmillan, http://www.dictionaryofeconomics.com/article?id=pde2009_H000191