Journal of Statistical and Econometric Methods, Vol.10, No.1, 2021, 55-66
ISSN: 2241-0384 (print), 2241-0376 (online)
https://doi.org/10.47260/jsem/1014
Scientific Press International Limited

Aspects of Pareto distributions

Johan Fellman[1]

Abstract

Different skew models, such as the lognormal and the Pareto functions, have been proposed as suitable descriptions of income distribution. Specific distributions are usually applied in empirical investigations. It is a common opinion that the Pareto curve often provides an adequate description of higher incomes. Recently, double Pareto distributions that obey the power law in both the upper and lower tails have been suggested to reflect a general distribution of personal income. In this study, the literature concerning double Pareto models is presented and the model is applied to Finnish income data.

JEL classification numbers: I32.
Keywords: Maximum likelihood estimate, Method of moments, Bayesian method, Mean Squared Error, Lognormal, double Pareto, Coefficient of determination, survival function, Geometric Brownian motion.

[1] Hanken School of Economics

Article Info: Received: March 5, 2021. Revised: March 23, 2021.
Published online: March 29, 2021.

1. Introduction

Different skew models, such as the lognormal and the Pareto functions, have been proposed as suitable descriptions of income distribution when empirical investigations are performed. It is a common opinion that the Pareto curve often provides an adequate description of higher incomes. It is interesting to recall that when Pareto (1897) first presented his law, he emphasized its empirical basis, whereas the process of reasoning by Gibrat (1931) advanced from theory to observations. Already Quensel (1944) stated that the lognormal curve agrees fairly well with the actual distribution of the lower incomes, although the Pareto curve often provides a more adequate description of the higher incomes (Fellman, 2015). Recently, the distribution of personal income has been proposed to be double Pareto, a distribution that obeys the power law in both the upper and lower tails (Reed (2001); Mitzenmacher (2004); Al-Athari (2011); Toda (2012); Shi-Yong et al., (2019)).

2. Methods

Harrison (1981) noted that a number of observed earning distributions were described well by the Pareto distribution defined as

$$F(y) = \begin{cases} 0 & y \leq 1 \\ 1 - y^{-\alpha} & y > 1 \end{cases}\text{'} \tag{1}$$

where $\alpha > 0$ and $y = \frac{Y}{Y_L}$, Y_L being the minimum income. For $\alpha > 1$, the mean is $E(Y) = \frac{\alpha}{\alpha-1}$. Furthermore, the Lorenz curve is $L(p) = 1 - (1-p)^{\frac{\alpha-1}{\alpha}}$ and the Gini coefficient is $G = \frac{1}{2\alpha-1}$. It is convenient to remark here that for commonly occurring values of the parameter α a second moment of the Pareto distribution does not exist unless $\alpha > 2$.

A common technique for estimating the Pareto constant α, is to linearize the survival function by taking logarithms and to then apply ordinary least squares. The survival function is

$$S(y) = 1 - F(y) = \left(\frac{Y}{Y_L}\right)^{-\alpha} \tag{2}$$

After taking natural logarithms, one obtains the linear model

$$ln(S(y)) = -\alpha\, ln(Y) + \alpha\, ln(Y_L) = C - \alpha\, ln(Y). \tag{3}$$

This model indicates a linear, decreasing association between $ln(S(y))$ and $ln(Y)$. A regression analysis gives an estimate of α and the coefficient of determination, R^2, measures the linearity in the model and the goodness of fit of the Pareto model.

In an earlier study, Fellman (2015) applied this analysis on annual taxable incomes in Finland for the year 2009. The data are presented in a grouped table (Table 1). He assumed that the Pareto model may start from ca. $Y = 25000€$. For values equal to or greater than this value he obtained the parameter estimate $\hat{\alpha} = 2.637$ and a coefficient of determination is $R^2 = 0.99241$, indicating good fit. For income distribution for incomes greater than 25000 the Gini coefficient was $G = \frac{1}{2\alpha - 1} = 0.234$.

Table 1: Taxable income receivers in Finland 2009 (Fellman, 2015)

Classes of annual income (€)	Number of income recipients
- 1000	182281
1000 - 2000	96836
2000 - 3000	80056
3000 - 4000	65800
4000 - 5000	59595
5000 - 6000	62171
6000 - 7000	107558
7000 - 8000	146526
8000 - 9000	114602
9000 - 10000	121555
10000 - 12500	319042
12500 - 15000	329083
15000 - 17500	259979
17500 - 20000	243284
20000 - 25000	481753
25000 - 30000	487376
30000 - 35000	385672
35000 - 40000	266075
40000 - 50000	307810
50000 - 60000	152714
60000 - 80000	120327
80000 -	88488
All	**4478583**

We illustrate his result in Figure 1.

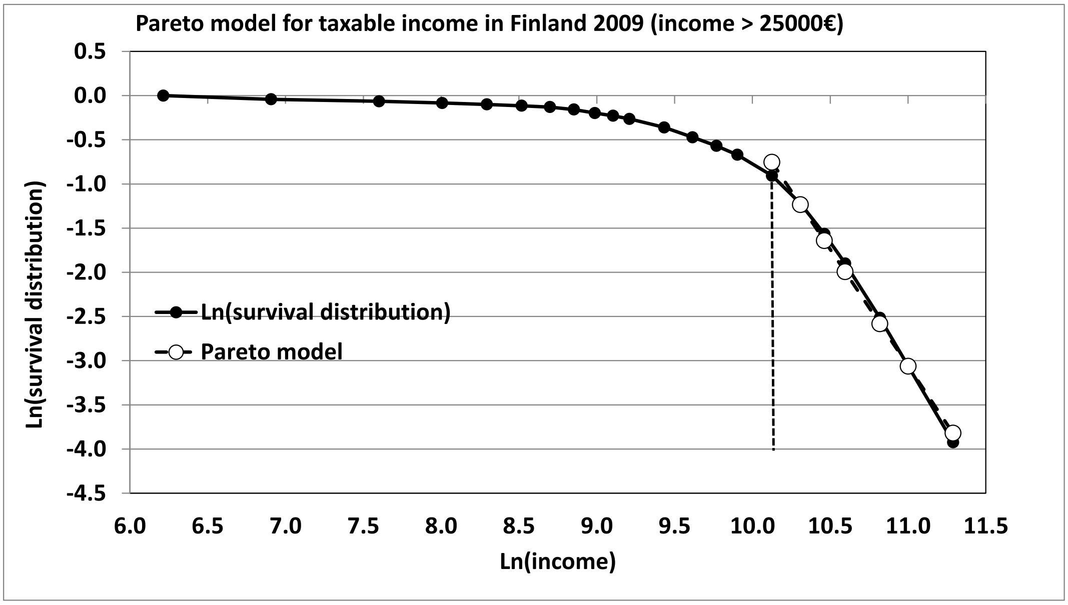

Figure 1: Graphical illustration of the distribution of taxable income in Finland (2009) and a Pareto model for annual incomes greater than $Y = 25000€$ (Fellman, 2015).

Recently, the distribution of personal income has been proposed to be double Pareto, a distribution that obeys the power law in both the upper and lower tails. Toda (2012) proposed a model of income dynamics with a stationary distribution consistent with this law.

3. Results

In this study, we apply the double Pareto model to the income data from Finland (2009) and use the value 25000 € to demarcate the boundary between low and high incomes. The income distribution for Finland (2009) is presented in Figure 2. This figure is a modified version of the original figure in Fellman (2019).

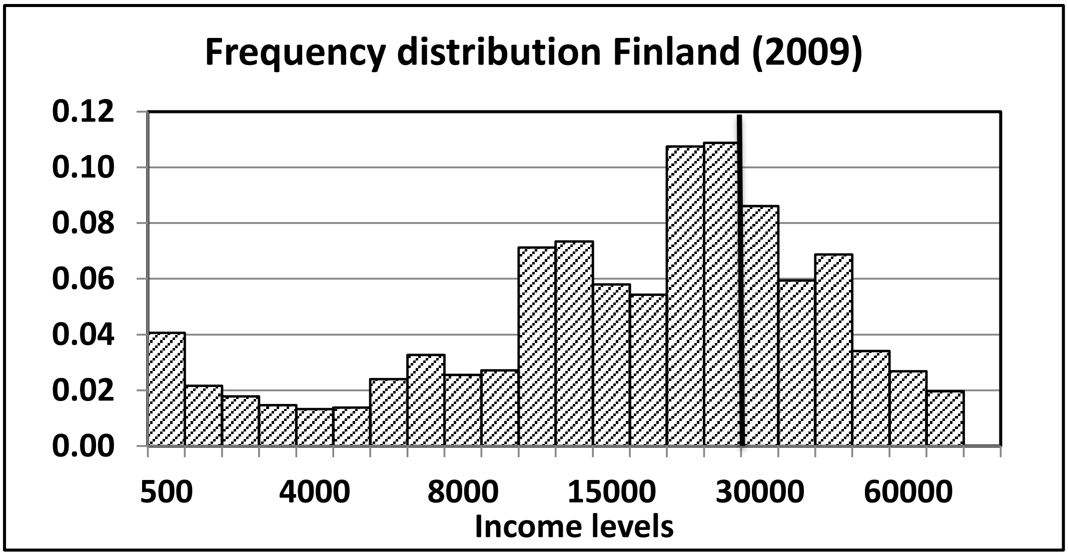

Figure 2: Distribution of income in Finland (2009). In this figure, we have indicated the boundary (25000 €) between low and high incomes and slightly modified the figure presented earlier (Fellman, 2019)

If one applies the double Pareto model on the Finland (2009) data, the power law in the lower tail results in a bad fit. For the low incomes, we introduce the logarithmic model

$$ln(f(x)) = C + \alpha\, ln(x) \text{ for } x < 25000, \qquad (4)$$

where x is the income level and $f(x)$ is the number of individuals receiving the amount *x*. We obtain the estimated results

$$ln(f(x)) = 8.861067 + 0.346812\, ln(x), \qquad (5)$$

where $\bar{R}^2 = 0.240$ and the low $\bar{R}^2$ value indicates a rather poor fit.
For the upper part of the incomes, we apply the model

$$ln(f(x)) = C - \beta\, ln(x) \text{ for } x > 25000, \qquad (6)$$

where *x* is the income level and $f(x)$ is the number of individuals r0eceiving the amount *x*. We obtain the estimated results

$$ln(f(x)) = 28.0283 - 1.46125\, ln(x) \qquad (7)$$

and $\bar{R}^2 = 0.943$ The high $\bar{R}^2$ value indicates a good fit and strong agreement with the results in (Fellman, 2015). In Table 2, we present the estimated models of the low and high income data and the corresponding test results for Finland (2009).

Table 2: Estimate results for Finland (2009)

	Low	**incomes**	**High**	**incomes**
	Estimate	**SE**	**Estimate**	**SE**
Intercept	8.861067	1.304	28.0283	1.566
Slope	0.346812	0.149	−1.46125	0.145
$\bar{R}^2$	0.240	0.581	0.943	0.151
F	5.427	0.037	101.044	<0.001

The statistical results are presented in Figure 3.

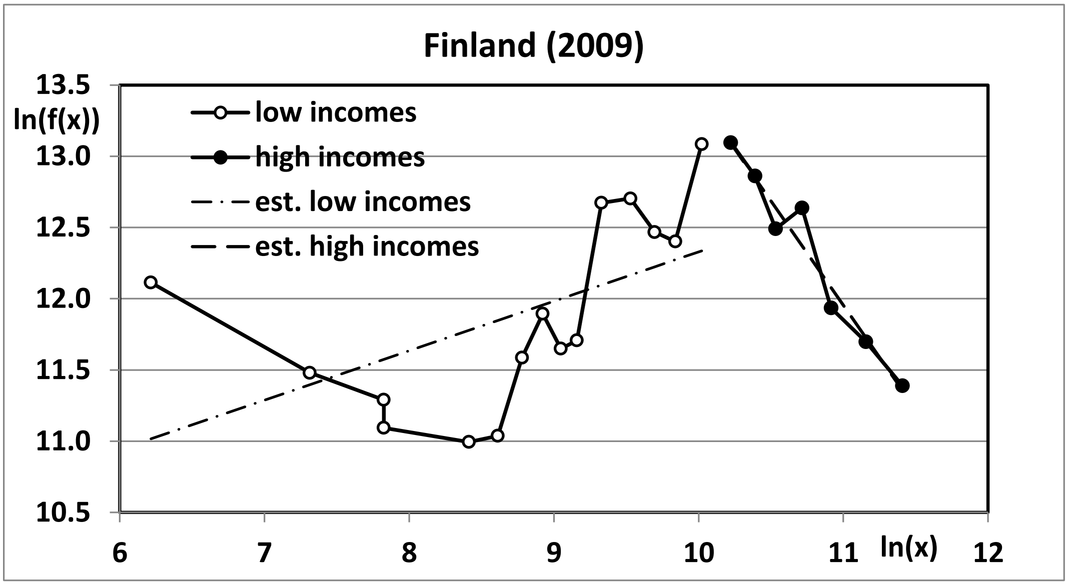

Figure 3: Estimated results of low and high income data for Finland (2009). Note the poor fit for low incomes and the good fit for high incomes. Estimated regression models are included in the figure

4. Discussion

Many empirical distributions encountered in economics and other realms of inquiry exhibit power-law behaviour in the upper tail. Reed (2001) presented a simple explanation for this. In economics, he also noted lower-tail power-law behaviour, which is verified empirically for income and city size data as well as for standardized price returns on individual stocks or stock indices. This widespread observed regularity has been explained in many ways. It continues to fascinate both natural scientists, who have recently proposed explanations based on such current ideas as self-organized criticality and highly optimized tolerance (Newman, 2000), and economists, as recent papers by Gabaix (1999) and Brakman et al. (1999) testify. While it seems unlikely that there is a single general theory that could explain all instances of power-law behaviour, there is, as Reed (2001) claimed, a simple, plausible explanation that has apparently been overlooked and can explain many examples in economics (including the Pareto and Zipf laws) and other areas. The temporal evolution of many phenomena exhibiting power-law behaviour is often considered to involve a varying, but size-independent proportional growth rate, which mathematically can be modelled by geometric Brownian motion (GBM).

Allowing for variation in initial sizes will modify this somewhat, but one would still expect power-law behaviour in the upper tail. Thus, provided all income earners had the same starting income, the current distribution of incomes should be that of a GBM observed after an exponentially distributed time T. This distribution is what Reed (2001) called a *double Pareto distribution*. He provided a simple explanation for this and also predicted lower-tail power-law behaviour, which was verified empirically for income and city size data. For example, if new stock issues occurred in a Poisson process and individual stock prices evolved following GBM, one might expect that the distribution of the ratio of current price to issue price over all such stocks would follow a power law in each tail. Reed (2001) stressed that phenomena frequently modelled by GBM include the evolution of stock prices, firm sizes, city sizes and individual incomes. It is well known that the state of a GBM after a fixed time T follows a lognormal distribution, which does not exhibit power-law behaviour. Why then should power-law behaviour occur for phenomena evolving as GBM? Reed (2001) claimed that the solution lies in the fact that the time of observation, T, should itself be regarded as a random variable, often with a distribution close to an exponential distribution. He considered, for instance, a census or sample survey of incomes. Even though each individual income may follow GBM, the time during which it has been so evolving will vary from individual to individual. If recruitment to the workforce has been growing at a more or less constant rate, the distribution of time in the workforce of any individual will follow an exponential distribution. Thus, provided all income earners had the same starting income, the current distribution of incomes should be that of a GBM observed after an exponentially distributed time T. This distribution, the double Pareto distribution, has a density proportional to $x^{-\alpha-1}$ for $x > 1$ and proportional to $x^{\beta-1}$ for $x < 1$. Thus, not only does this simple model offer a plausible explanation of the Pareto Law of Incomes (upper tail), it also predicts power-law behaviour in the lower tail. In fact, lower-tail power-law behaviour has been identified before (Champernowne, 1953).

Furthermore, Reed (2001) gave other examples, outside of economics, for which a similar explanation might hold such as the body-size distribution of animal species (May, 1988). Here it would be assumed that the body mass of any individual species evolved through natural selection following GBM, while speciations occurred in a Yule process. Power-law behaviour in the lower tail of particle-size distributions and in the upper tail of forest-fire size distributions could also be explained in a similar way. In the former case, repeated random fractures indicate a form of random geometric decay, while in the latter case the area burnt might follow random proportional growth through time until stopped at random (*e.g.* by the onset of suitably heavy rainfall).

Later Reed & Jorgensen (2004) introduced a family of probability densities, and stressed that it has proved useful in modelling the size distributions of various phenomena, including incomes and earnings, human settlement sizes, oil-field volumes and particle sizes. The distribution, named the *double Pareto-lognormal* (dPLN) distribution, arises similarly to the state of a GBM, with a lognormally distributed initial state, after an exponentially distributed length of time (Reed, 2001). Reed & Jorgensen (2004) stressed that a number of phenomena can be viewed as resulting from such a process (e.g. incomes, settlement sizes), which explains the good fit. Furthermore, they derived properties of the distribution and discussed the estimation methods. They found that the distribution exhibits Paretian (power-law) behaviour in both tails, and when plotted on logarithmic axes, its density shows hyperbolic-type behaviour (Reed & Jorgensen, 2004).

Mitzenmacher (2004) started from the double Pareto distributions recently suggested to describe income distributions and other power-law phenomena (Reed & Jorgensen, 2004). As he shows, such distributions have a lognormal body and a Pareto tail, which matches some previous studies of empirical data for file sizes. He believed that such distributions may be useful for modelling other power-law phenomena in computer systems, and that his generative model might prove useful for other applications. Mitzenmacher provided a detailed analysis of his Recursive Forest File model which is interesting in its own right. In particular, he found several connections to the theory of random graphs that he expected would provide a useful framework for future work. Furthermore, he showed how to cope with the effects of correlation that are implicit in a file system model where new files are derived from existing files, using a martingale analysis.

Reed and Wu (2008) introduced two parametric models for income distributions. The models fitted to log(income) are the 4-parameter *normal-Laplace* (NL) and the 5-parameter *generalized normal-Laplace* (GNL) distributions. The NL model for log(income) is equivalent to the double Pareto-lognormal (dPLN) distribution applied to income itself. Definitions and properties are presented along with methods for maximum likelihood estimation of parameters. Both models along with 4- and 5-parameter beta distributions were fitted to nine empirical distributions of family income. In all cases, the 4-parameter NL distribution fits better than the 5-parameter generalized beta distribution. The 5-parameter GNL distribution provides an even better fit. They found that fitting of the GNL is numerically slow since there are no closed-form expressions for its density or cumulative distribution functions. They found that 5-parameter beta distribution is the best fitting, and the results suggest that the NL should be seriously considered as *a parametric model for income distributions.*

Al-Athari (2011) stressed that the double Pareto distribution appears most often as a model for a variety of fields, including archaeology, biology, economics, environmental science, finance and physics. Furthermore, the family of double Pareto distributions has recently been proposed for modelling growth rates such as annual gross domestic product, stock prices, foreign currency exchange rates and company sizes.

Todorova and Vogta (2011) stated that power-law distributions are very common in studies of natural sciences. They analysed high-frequency financial data using maximum likelihood estimation and the Kolmogorov–Smirnov statistic to test whether the power-law hypothesis holds also for these data. They found that the universality and scale invariance properties of the power law are violated. Furthermore, the returns of some shares traded simultaneously on both exchanges follow a power law at one exchange, but not at the other. These results raise some questions about the no-arbitrage condition. Finally, they found that an exponential function provides a better fit for the tails of the sample distributions than a power-law function. They did not associate the power law with the double Pareto distribution.

Fellman (2012) analysed classes of theoretical Lorenz curves with varying Gini coefficients. Especially he compared Gini estimates for the Pareto distributions. If one defines the Pareto distribution as $F(x) = 1 - x^{-\alpha}$, where $x \geq 1$ and $\alpha > 1$, then the frequency function is $f(x) = \alpha x^{-\alpha}$, the mean is $\mu = \frac{\alpha}{\alpha-1}$, the quantiles are $x_p = \left(\frac{1}{1-p}\right)^{\frac{\alpha-1}{\alpha}}$, the Lorenz curve is $L(p) = 1 - (1-p)^{\frac{\alpha-1}{\alpha}}$ and the Gini coefficient is $G = \frac{1}{2\alpha-1}$. Fellman considered $1.5 \leq \alpha \leq 5.0$; the Gini coefficient then satisfies the inequalities $0.111 \leq G \leq 0.500$. This *G* interval corresponds to the most common Gini coefficients. He noted that Simpson´s and Golden´s (2011) rules yield similar accuracy, but the trapezium rule shows the largest errors for all levels of Gini coefficients. This theoretical study indicated that Golden´s rule is not uniformly better than the trapezium rule in performing comparisons between the estimated and theoretical Gini coefficients.

Toda (2012) found double power law behaviour in income distributions. Conditional on education and experience, the distribution of personal labour income appears to be double Pareto. This “double power law” is not rejected by goodness-of-fit tests. He compared two diffusion processes (one mean-reverting, the other unit root) with a stationary double Pareto distribution as a model of income dynamics. The data favour the mean-reverting process over the unit root process for modelling income dynamics.

Al-Athari and Jaber (2013) derived a Bayesian estimator for this distribution with multi-parameter Jeffreys´ prior. The theoretical part of the fore mentioned study reveals that Bayesian estimator for the scale parameter does not exist. The simulation part reveals that Bayesian and maximum likelihood estimators are equally better than the method of moments estimator when the sample size is larger than or equal to 14, otherwise the method of moments and the maximum likelihood estimators are equally good.

Yu et al. (2019) stated that although societal scale and mode of production from foraging to farming correlate with increases in economic inequality there is no consensus regarding the relative importance of those factors or the role of institutions in the variance of inequality across time and space. Furthermore, they specified that to better understand the dynamics of economic inequality it is necessary to expand the analytical horizon beyond the present into the deeper past. However, an analytical protocol especially oriented towards the systematic study of economic inequality with archaeological data is lacking. They proposed the utility of grave size as a reliable proxy for estimating prehistoric social inequality and provided a methodological framework for analysing this type of data. Their case studies using grave-size data from two Neolithic settlements in North and East China suggested that the asymmetric double Pareto distribution can be used as an alternative model to fit the size distribution of grave wealth, which is usually skewed and long-tailed. Based on the analytical connection between the probability density function and the Lorenz curve, they derived a parsimonious algebraic expression of the Gini coefficient. Furthermore, Yu et al (2019) stated that this analytical protocol also can serve as a convenient tool for quantifying economic inequality in prehistoric societies using other types of archaeological data such as land and house areas.

5. Conclusions

Different skew models, such as the lognormal and the Pareto functions, have been proposed as suitable descriptions of income distribution. It is a common opinion that the Pareto curve often provides an adequate description of higher incomes (Harrison, 1981; Fellman 2015). Recently, double Pareto distributions that obey the power law in both the upper and lower tails have been proposed to reflect a general distribution of personal income (Toda, 2012). The double Pareto distribution appears often as a model for a variety of fields, including archaeology, biology, economics, environmental science, finance and physics (Al-Athari, 2011). The distribution exhibits Paretian power-law behaviour in both tails. Many empirical distributions encountered in economics and other realms of inquiry exhibit power-law behaviour in the upper tail. The symmetric double Pareto distribution is frequently used to fit data on income, growth rates, finance and physical and biological problems. The temporal evolution of many phenomena exhibiting power-law behaviour is often considered to involve a varying, but size-independent proportional growth rate, which mathematically can be modelled by GBM (Reed, 2001). No doubt there are many other examples fitting within this paradigm, whose essential elements are random proportional (geometric) change and random stopping or observation. Phenomena evolving according to Gibrat’s law, which are observed after an exponentially distributed period of time, should be expected to exhibit distributions with power-law tail behaviour (Reed, 2001).

References

[1] Al-Athari, F. M. (2011). Parameter Estimation for the Double Pareto Distribution. Journal of Mathematics and Statistics, 7(4):289-294.

[2] Al-Athari, F. M. & Jaber, K, K. (2013) Bayesian estimation for the symmetric double Pareto distribution with multi-parameter Jeffreys´ prior information. International Journal of Academic Research.

[3] Brakman, S, H.Garretsen, C. Van Marrewijk and M. van den Berg, (1999). The return of Zipf: Towards a further understanding of the rank-size distribution. Journal of Regional Science, 39:739-767.

[4] Champernowne, D., (1953). A model of income distribution. Economic Journal, 63:318-351.

[5] Fellman, J. (2012). Estimation of Gini coefficients using Lorenz curves. Journal of Statistical and Econometric Methods. 1(2):31-38.

[6] Fellman, J, (2015). Mathematical analysis of distribution and redistribution of income. Science Publishing Group 166 pp ISBN: 978-1-940366-25-8, (2015).

[7] Fellman, J. (2019) Empirical analyses of income: Finland (2009) and Australia (1967-1968). Journal of Statistical and Econometric Methods submitted

[8] Gabaix, X.,(1999. Zipf's law for cities: an explanation. Quarterly Journal of Economics, 114 739-767.

[9] Gibrat, (1931). Les Inégalités Economiques. Librairie du Recueil Sirey, Paris.

[10] Golden, J. (2008). A simple geometric approach to approximating the Gini coefficient. J. Economic Education 39(1):68-77.

[11] Harrison, A. (1981). Earnings by size: A tale of two distributions. Review of Economic Studies 48:621-631.

[12] May, R. M., (1988). How many species are there on earth? Science, 241 :1441-1449.

[13] Mitzenmacher M (2004) Dynamic models for file sizes and double Pareto distributions. Internet Math 1:305–333.

[14] Newman, N. (2000). The power of design. Nature, 405 :412-413.

[15] Pareto, V. (1897). Cours d'Economie Politique. Lausanne, Suisse,.

[16] Quensel, C. E. (1944). Inkomstfördelning och skattetryck. Sveriges industriförbund, Lund,.

[17] Reed, W. J. (2001). The Pareto, Zipf and other power laws. Economics Letters,74(1), https://doi.org/10.1016/S0165-1765(01)00524-9.reed@math.uvic.ca

[18] Reed, W. J. & Jorgensen M (2004). The double Pareto-lognormal distribution -a new parametric model for size distributions. Commun. Stat. Theory. Methods 33:1733–1753

[19] Reed, W. J. & Wu, F. (2008). New four- and five- parameter models for income distributions. In: Modeling Income Distributions and Lorentz Curves,

Part of the Economic Studies in Equality, Social Exclusion and Well-Being (EIAP, volume 5) D. Chotikapanich (ed.) Springer.

[20] http://www.sciencepublishinggroup.com/book/B-978-1-940366-25-8.aspx

[21] Toda A. A. (2012). The double power law in income distribution: Explanations and evidence. Journal of Economic Behavior & Organization Volume 84, Issue 1, September 2012, Pages 364-381. alexisakira.toda@yale.edu

[22] Todorova, L. & Vogt, B. (2011). Power law distribution in high frequency financial data? An econometric analysis. Physica A: Statistical Mechanics and its Applications, 390 (23–24): 4433-4444.doi:10.1016/j.physa.2011.07.035.

[23] Yu, Shi-Yong, Xue-Xiang Chen & Hui Fang. (2019). Inferring inequality in prehistoric societies from grave sizes: a methodological framework. Archaeological and Anthropological Sciences. Volume 11, Issue 9:4947–4958 . https://doi.org/10.1007/s12520-019-00845-0