

PRELIMINARY NOTES ON PALMA'S MEASURE OF INEQUALITY

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 18(24)

Katarzyna Ostasiewicz

Wroclaw University of Economics and Business, Poland

e-mail: katarzyna.ostasiewicz@ue.wroc.pl

ORCID: 0000-0002-0115-3696

Achille Vernizzi

University of Milano, Italy

e-mail: achille.vernizzi@unimi.it

ORCID: 0000-0002-1641-5003

ISSN 1644-6739
e-ISSN 2449-9765

© 2020 Katarzyna Ostasiewicz, Achille Vernizzi

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Quote as: Ostasiewicz, K., and Vernizzi, A. (2020). Preliminary notes on Palma's measure of inequality. *Śląski Przegląd Statystyczny*, 18(24).

DOI: 10.15611/sps.2020.18.10

Abstract: Based on Palma's observation that for empirical distributions within the range of the 4th and 9th deciles there are – with a good approximation – 50% of all goods, a new measure was introduced that is in a strong relationship to the Gini index, thus – supposedly – including all the information that it provides. In this paper this relationship is investigated more deeply for cases of some theoretical distributions as well as the relationship of a lower bound of a slightly differently defined 'middle class' on the Gini index.

Keywords: Gini index, Palma measure, inequality, tails.

1. Introduction

Palma (2006) noted that in most of the empirical cases of income/wealth distributions, there exist stable 'middle classes' that start from the 4th decile to the 9th decile, owning approximately half of all goods, which would mean that $L(0.9) - L(0.4) \approx 0.5$. Based on that observation, Palma concluded that "it is all about tails" and pointed out that the measure based on those tails, i.e. the sum of goods owned by the highest 10% as compared to the sum of goods owned by the lowest 40%, brings all the information that the Gini index provides. Palma indicated the strict (or strong) correspondence of the new measure to the Gini index.

Palma's conclusions were further corroborated by other authors who, based on other data, drew still more attention to Palma's observation (e.g. Cobham & Sumner, 2013; Cobham et al., 2013; Palma & Stiglitz, 2016; Cobham, Schlögl, & Sumner, 2016; Pan et al., 2019).

Obviously, showing some examples that corroborate a certain idea is not a proof. As it is not possible to decisively prove this almost complete correspondence of Palma's measure and the Gini index, it is not the aim of the paper to achieve such an ambitious goal. Instead, the author examined some theoretical distributions that are widely used to model the empirical distribution of real data concerning income/spending/wealth, to investigate the relationship between Palma's measure and the Gini index in the wide range of values of inequality, not being restricted by the accessibility of the real data and by the range of inequality revealed by the accessible sets.

After investigating the relationship between Palma's measure and the Gini index, the author examined some other concepts of 'tail measure', which differ slightly from Palma's original one.

The paper is organized as follows. In the next section, four theoretical distributions are presented and their properties with respect to the Gini index. Section 3 compares Palma's measure and the Gini index for these distributions and presents a slightly different approach to defining 'middle class'. Section 4 investigates the dependence of the lower bound of this modified 'middle class' on the values of the Gini index. The final section contains the summary, conclusions and some directions on further investigations.

2. Theoretical distributions

Four theoretical distributions are examined here. All of them are considered as appropriate for modelling the income/wealth distributions and are characterized by one shape-parameter. Thus by changing the value of this parameter, one can cover the whole possible range of the value of the Gini index for the given theoretical distribution.

2.1. Log-normal distribution

The probability density function of this distribution is defined for $x > 0$ and is given by:

$$f_{LN}(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right],$$

where μ is the scale parameter, and $\sigma > 0$ the shape parameter.

The Lorenz curve for log-normal distribution is known to have the analytical form:

$$L_{LN}(p) = \Phi(\Phi^{-1}(p) - \sigma),$$

and the Gini index:

$$G_{LN} = 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1,$$

where Φ denotes the cumulative distribution function of standard normal distribution.

For log-normal distribution, when changing the value of the shape parameter, one can obtain the whole range of the values of the Gini index, from $G \rightarrow 0$ (for $\sigma \rightarrow 0$) and $G \rightarrow 1$ (for $\sigma \rightarrow \infty$).

2.2. Log-logistic distribution

The probability density function of this distribution is defined for $x > 0$ and is given by:

$$f_{LL}(x) = \frac{(a/b)(x/b)^{a-1}}{[1 + (x/b)^a]^2},$$

with b being the scale parameter, while a the shape parameter.

The mean value and the standard deviation for log-logistic distribution exist if $a > 2$, and the Gini index is expressed in terms of this parameter in a very simple way, as:

$$G_{LL} = \frac{1}{a}.$$

By restricting the values of a to the range within which the mean value and the standard deviation are defined, one can obtain the values of the Gini index ranging from 0 ($a \rightarrow \infty$) to 0.5 ($a \rightarrow 2$).

2.3. Weibull distribution

The probability density function of this distribution is defined for $x > 0$ and is given by:

$$f_W(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k},$$

with $\lambda > 0$ being the scale parameter, while $k > 0$ the shape parameter.

The Lorenz curve for the Weibull distribution is known to have the analytical form:

$$L_W(p) = 1 - \frac{\Gamma(-\ln(1-p), 1 + \frac{1}{k})}{\Gamma(1 + \frac{1}{k})},$$

and the Gini index:

$$G_W = 1 - 2^{-\frac{1}{k}}.$$

While the mode exists for $k > 1$, within this restriction the values of the Gini index range from 0 ($k \rightarrow \infty$) to 0.5 ($k \rightarrow 1$).

2.4. Gamma distribution

The probability density function of this distribution is defined for $x > 0$ and is given by:

$$f_G(x) = \frac{1}{\beta^\alpha} x^{\alpha-1} \frac{\exp[-x/\beta]}{\Gamma(\alpha)},$$

with Γ denoting gamma function, $\beta > 0$ being the scale parameter, while $\alpha > 0$ the shape parameter.

The value of the Gini index is given in the analytical form as:

$$G_G = \frac{\Gamma\left(\alpha + \frac{1}{2}\right)}{\Gamma(\alpha + 1)\sqrt{\pi}}.$$

As the mode for Gamma distribution exists for $\alpha > 1$, it is reasonable to be restricted to such values of this parameter. Thus the values of the Gini index range from 0 ($\alpha \rightarrow \infty$) to 0.5 ($\alpha \rightarrow 1$).

3. Palma's measure vs the Gini index

For each aforementioned distribution the author plotted two relationships. Palma's observation is that between the 4th and the 9th decile ('middle class') of virtually each empirical distribution approximately 50% of all goods are included. The left-hand side part of the figure for each theoretical distribution presents this fraction of all goods possessed by the so defined 'middle class', versus the shape of the distribution (and thus the value of the Gini index). The right-hand side of each figure presents Palma's measure vs the Gini index. Each curve is fitted by an exponential curve, following Cobham and Sumner (2013).

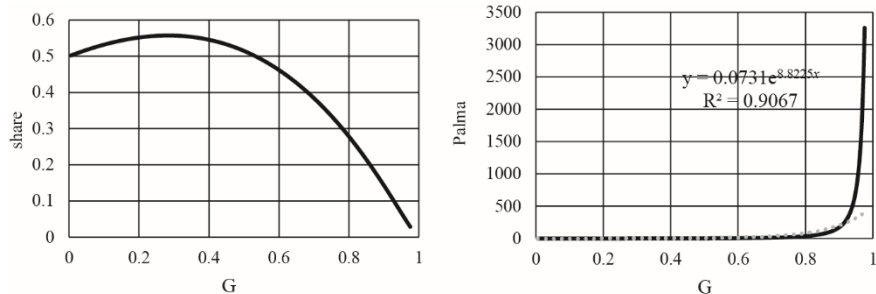


Fig. 1. Share of all goods included between the 4th and the 9th decile, and Palma vs Gini for log-normal distribution

Source: own calculations.

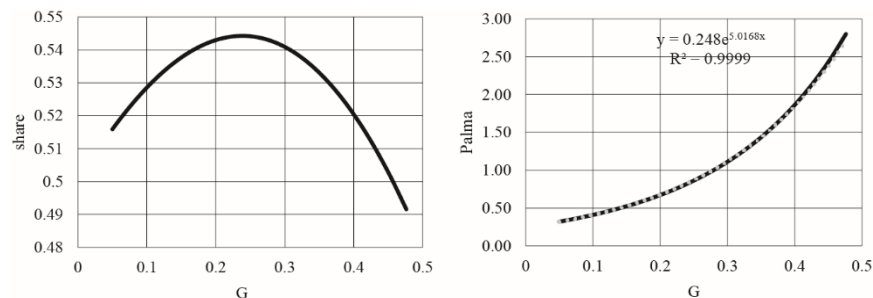


Fig. 2. Share of all goods included between the 4th and the 9th decile, and Palma vs Gini for log-logistic distribution

Source: own calculations.

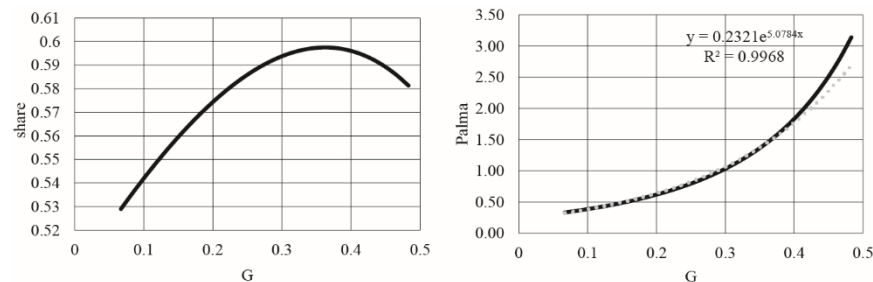


Fig. 3. Share of all goods included between the 4th and the 9th decile, and Palma vs Gini for Weibull distribution

Source: own calculations.

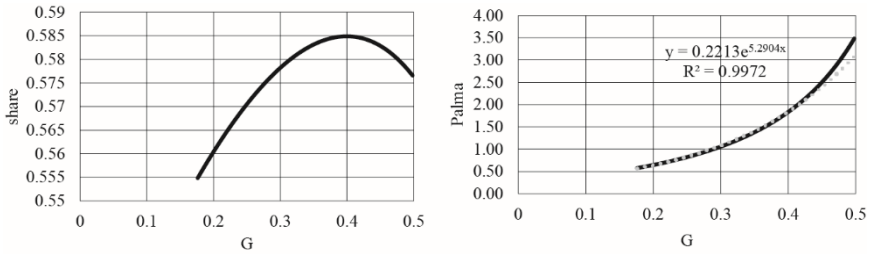


Fig. 4. Share of all goods included between the 4th and the 9th decile, and Palma vs Gini for Gamma distribution

Source: own calculations.

Two general observation can be made.

First, the robustness of approximately 50% belonging to the defined ‘middle class’ (between the 4th and the 9th decile) holds for the theoretical distributions for low and intermediate values of the Gini index, more or less, for $G < 0.5$. Second, Palma vs Gini can be approximated by the exponential curve whose parameters depend on the theoretical distribution (see Table 1). Based on empirical data for various countries and various years, Cobhan and Sumner (2013) fitted the Palma vs Gini relationship with the exponential curve (with parameters: 0.1969 and 5.6739) with a high degree of accuracy ($R^2 > 0.98$). The empirical data from Cobhan and Sumner covered Gini indexes from about 0.2 to 0.65. If we restrict our theoretical distributions to the range of parameters for which $0.2 \leq G \leq 0.65$ (or $0.2 \leq G \leq 0.50$ for distributions for which the maximum possible value of Gini is 0.5), there may be found the highly consistent values of these parameters for all distributions (see Table 2).

Table 1. Values of parameters of fitted exponential curve in the form: Ce^{ax} , for theoretical distributions within the range of G for which the theoretical distribution is unimodal and has a defined mean and standard deviation

Distribution	C	a	R^2
log-normal	0.0731	8.8225	0.9067
log-logistic	0.248	5.0168	0.9999
Weibull	0.2321	5.0784	0.9968
Gamma	0.2213	5.2904	0.9972

Source: own calculations.

Table 2. Values of parameters of fitted exponential curve for theoretical distributions within the range of G : $0.2 \leq G \leq 0.65$ and for which the theoretical distribution is unimodal and has a defined mean and standard deviation

Distribution	C	a	R^2
log-normal	0.1929	5.7479	0.9970
log-logistic	0.2388	5.1385	0.9999
Weibull	0.1881	5.8770	0.9975
Gamma	0.2139	5.3906	0.9972

Source: own calculations.

Moreover, if one collects values for all four theoretical distributions (for $0.2 \leq G \leq 0.65$) and presents them on the single plot, the data may be approximated by the exponential curve with parameters 0.2046 and 5.5582 (and $R^2 = 0.9967$), see Figure 5.

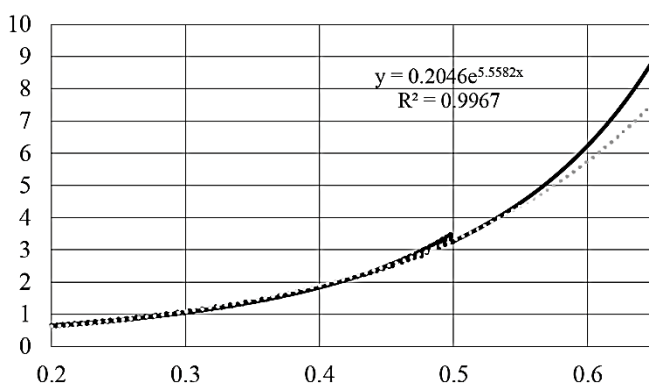


Fig. 5. Palma vs Gini for all four distributions put together, within the range of $0.2 \leq G \leq 0.65$, fitted by the exponential curve

Source: own calculations.

The values of these parameters are strikingly close to the values from the empirical data (0.2046 vs 0.1969 and 5.5582 vs 5.6739) the more so that – as the authors of the empirical study observed – the empirical data was not always consistent as some sources gave Ginis for incomes and some for expenditures, etc.

4. Modification of Palma's 'middle class'

It is obvious that for equal distribution, any 50% of the population own 50% of the total wealth. It is also obvious that this statement does not hold

in the case of unequal distribution. From the very definition of inequality, the poorest 50% of population own less than 50% of the total wealth, while the richest 50% – more than 50% of the total wealth. The question may be posed: is it possible – or under which conditions it is possible – to determine such a half of the population owning exactly half of the total wealth? The additional requirement is that this half of the population has to be a continuous subset of all observations ordered in a non-decreasing way, i.e. exactly as for the construction of the Lorenz curve.

The answer to that question in the case of continuous distribution (i.e. virtually for all the cases when the number of observations is large) is easily illustrated (Ostasiewicz, 2020).

The Lorenz curve is concave. If we plot the Lorenz curve and the line of equal distribution, the length of this line which is limited by the intersection points with the Lorenz curve is equal to $\sqrt{2}$. Let us move the line downward, without changing its slope. There are always two intersection points of this line with the Lorenz curve, until the line reaches the ‘bottom’ of the curve, i.e. there will be the only one intersection (junction) point where $L'(p) = 1$. As $L'(p) = F^{-1}(p)/\mu$ (F – cumulative distribution function), $L'(p)$ will be equal to 1 for: $p = F(\mu)$. One may say that in the sequence of pairs of points being the intersections of $y = p$ line with $L(p)$, both points approach each other, merging into one. Simultaneously, the length of the line between these two points changes from $\sqrt{2}$ to 0. As the process is continuous (for continuous distribution function), there will be the moment when the length of the line will be equal to $\sqrt{2}/2$. However, as the line moves downward without changing its slope, the edges of the right-angled triangle whose hypotenuse is this fragment of $y = p$ line (see Figure 6), are equal to 1/2 and 1/2. Yet as the length of these edges correspond to the fraction of the population (horizontal edge) and the fraction of total goods (vertical edge), this means that these two points delimit half of the population owning half of all the goods.

For any unequal distribution it is obvious that the range of this ‘50% class’ will start from $p > 0$ – from the very definition of inequality, the lowest half of the population, $0 \leq p \leq 0.5$ does not own half of all goods. It is also intuitively obvious that the higher the inequality, the further the lower bound of this class will be shifted toward $p \rightarrow 0.5$. Naturally, this lower bound cannot be greater than 0.5 as the class we are looking for is defined as the class covering half of the whole population, and if started from $p > 0.5$ there would be always be less than 50% of the population inside it.

Like in the previous section, the paper investigated the four theoretical distributions within the possible range of the value of the Gini index (the limits of this value are described in Section 2).

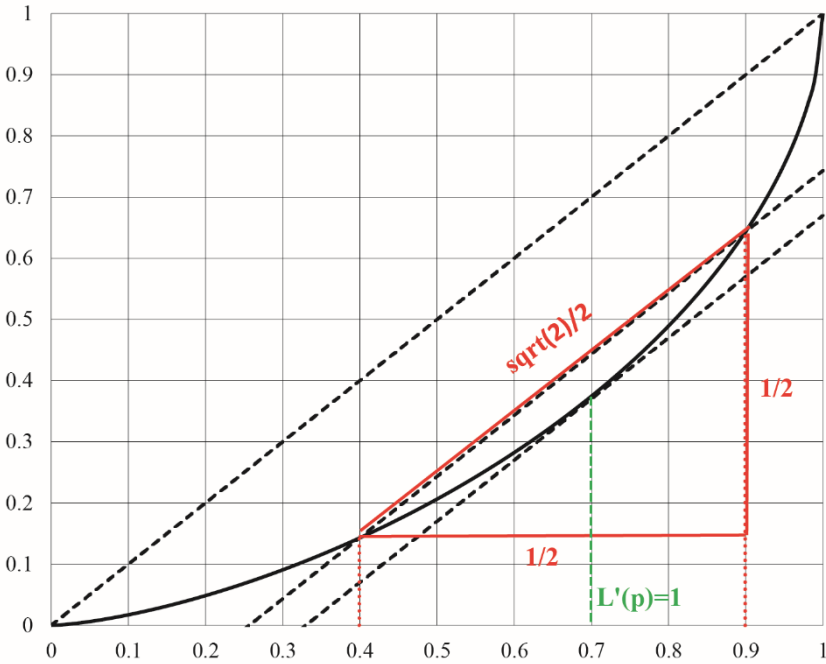


Fig. 6. Definition of a ‘middle class’ as the range of 50% of observations, owning exactly 50% of all goods

Source: own elaboration.

The plots for these distributions – the lower bound of the 50% class vs the value of the Gini index – are presented in Figures 7 to 10, alongside the linear fit.

The fit of the straight line for each particular distribution is almost perfect. However, the slope in each case is a bit different, and each distribution forms a distinct line. Nevertheless, the gathered results – while being fitted by the straight line with much lower R^2 – still suggest some interesting dependence which might be observed within the empirical data.

Hence an interesting question to be answered in the future arises: do the empirical data fit the linear dependence, and if so, what is the slope of the line, and does it correspond to any of the slopes for the theoretical distributions?

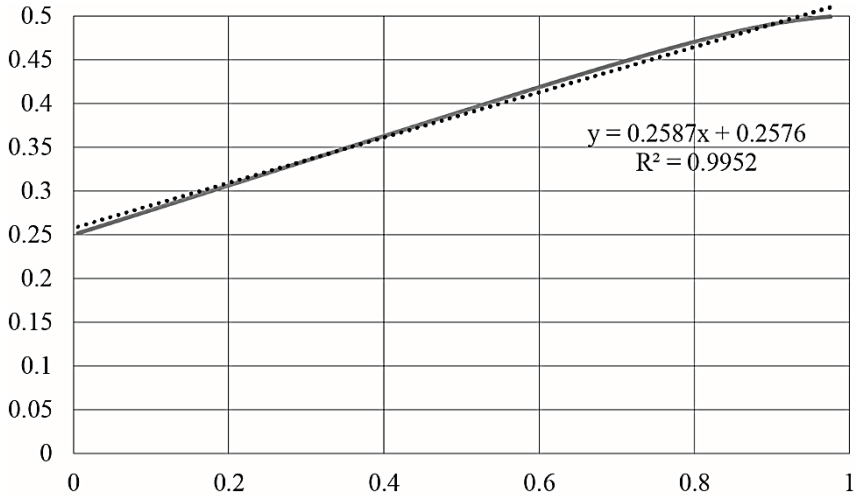


Fig. 7. The lower bound of the 50% class vs the value of the Gini index.
Log-normal distribution

Source: own calculations.

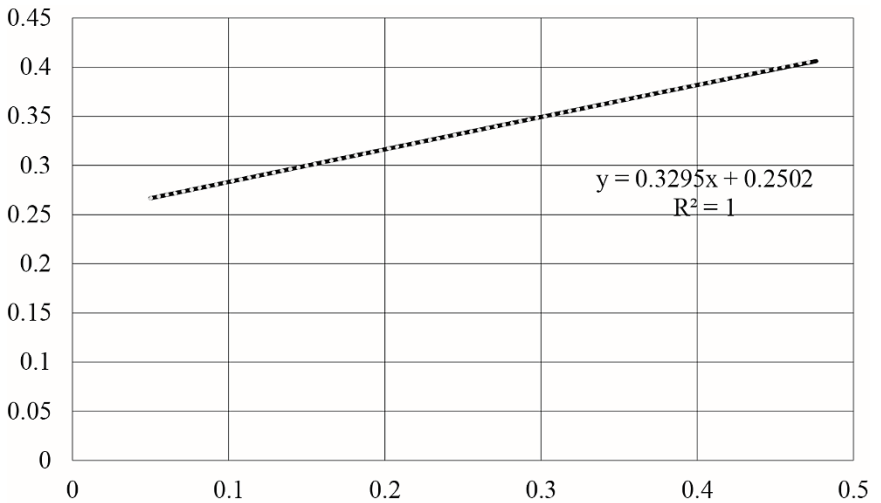


Fig. 8. The lower bound of the 50% class vs the value of the Gini index.
Log-logistic distribution

Source: own calculations.

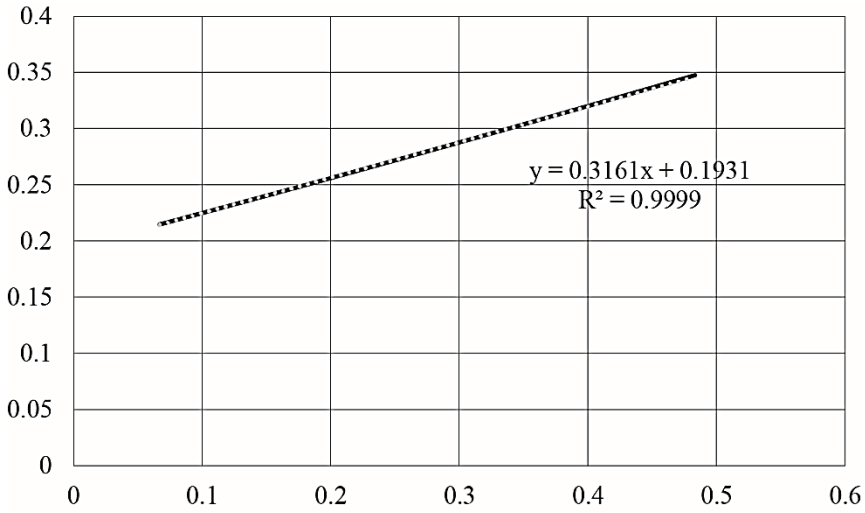


Fig. 9. The lower bound of the 50% class vs the value of the Gini index.
Weibull distribution

Source: own calculations.

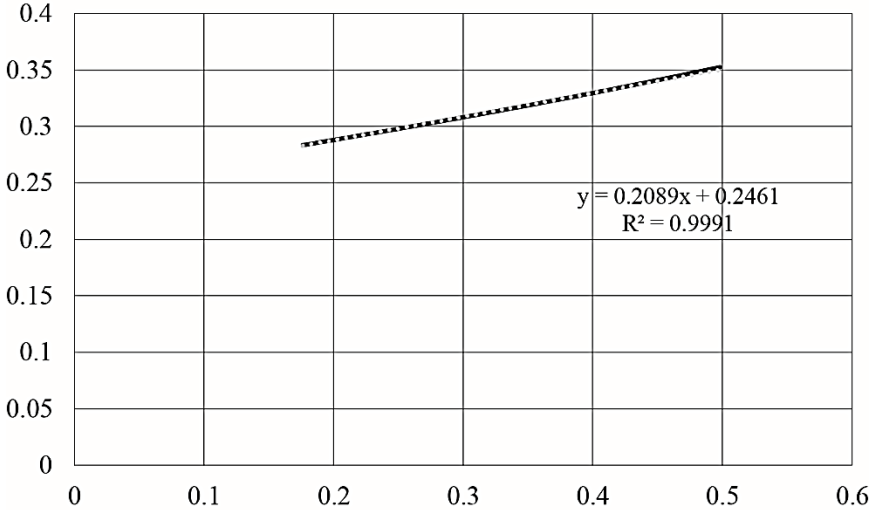


Fig. 10. The lower bound of the 50% class vs the value of the Gini index.
Gamma distribution

Source: own calculations.

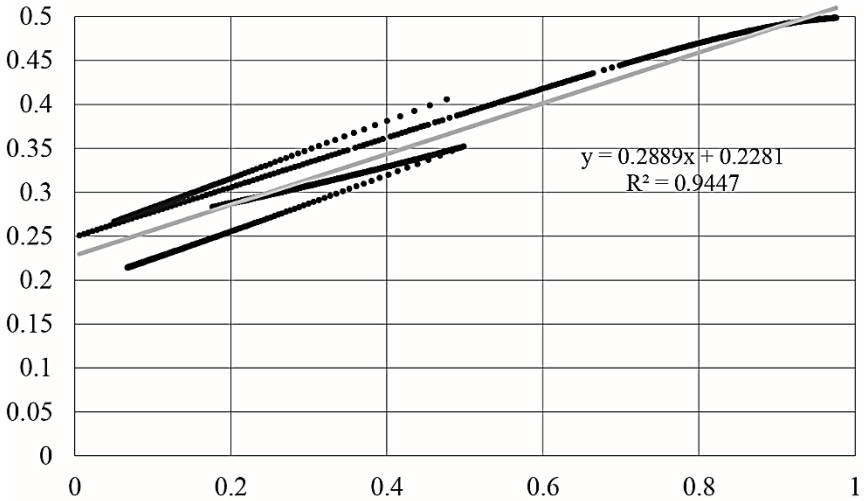


Fig. 11. Lower bound of the 50% class vs the value of the Gini index, all the four theoretical distributions together

Source: own calculations.

5. Summary and conclusions

In this paper the author investigated only four simple theoretical distributions chosen with respect to their one shape-parameter. The merit of this property is that one can uniquely examine the whole (available for the given distribution) range of the Gini index. For more complicated analytical forms of theoretical distributions (e.g. Singh-Maddala or Dagum distribution) there exist two shape parameters, and an additional task is to choose the proper combination of them to produce a given value of the Gini index – or to accept the ambiguity of a combination of these values. Yet investigating more complex theoretical distributions is a task to be addressed in the future.

Palma's observation on the strong relationship between the Gini index and the newly defined Palma index may be interpreted two-fold. First, it may be suggested that there is no need to calculate the Gini index as the much simpler to obtain Palma index involves all the information that the Gini index provides. On the other hand, it may be argued that the Gini index is too sensitive to extreme values (e.g. Monti, 1991) and is overwhelmed by the tails – and in order to investigate changes within the middle class, other approaches would be more useful. Both approaches, namely replacing the Gini index with the much simpler one and introducing some other measures that better reflect changes within the middle class, might be obviously for simultaneous use.

Palma's observations bring more interesting questions. Is the distribution with the middle class similar for all empirical distributions (and the corresponding theoretical distributions) or is it rather the predominant impact of the tails? What is the source of such a strong dependence of the Palma index on the one hand, and the lower bound of the modified middle class on the other, on the value of the Gini index? Seeking answers to these questions may throw some light onto the deeper characteristics of both the empirical and the theoretical distributions of income/wealth, and at the same time help to better understand the changes within them.

Literature

- Cobham, A., & Sumner, A. (2013). Is it all about the tails? The Palma measure of income inequality. *Center for Global Development Working Paper*, (343).
- Cobham, A. et al. (2013). *Putting the Gini back in the bottle? 'The Palma' as a policy-relevant measure of inequality*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.365.4686>
- Cobham, A., Schlögl, L., & Sumner, A. (2016). Inequality and the tails: the Palma proposition and ratio. *Global Policy*, 7(1), 25-36.
- Monti, A. C. (1991). The study of the Gini concentration ratio by means of the influence function. *Statistica*, 51(4), 561-580.
- Ostasiewicz, K., The 21st International Scientific Conference Quantitative Methods in Economics 2020, presentation.
- Palma, J. G. (2006). *Globalizing inequality: 'centrifugal' and 'centripetal' forces at work* (DESA Working Paper No. 35, ST/ESA/2006/DWP/35).
- Palma, J. G., & Stiglitz, J. E. (2016). Do nations just get the inequality they deserve? The "Palma Ratio" re-examined. In *Inequality and growth: Patterns and policy* (pp. 35-97). London: Palgrave Macmillan.
- Pan, X., Wang, H., Wang, Z., Lin, L., Zhang, Q., Zheng, X., & Chen, W. (2019). Carbon Palma Ratio: A new indicator for measuring the distribution inequality of carbon emissions among individuals. *Journal of Cleaner Production*, 241, 118-418.

WSTĘPNE UWAGI NA TEMAT MIARY NIERÓWNOŚCI PALMY

Streszczenie: Nowa miara nierówności, zaproponowana na podstawie spostrzeżeń Palmy, iż dla empirycznych rozkładów dochodów pomiędzy czwartym a dziewiątym decylem zawarte jest – z dobrym przybliżeniem – 50% wszystkich dóbr, okazuje się ściśle związana z indeksem Giniego, niosąc tyle samo co on informacji. W niniejszym artykule zależność ta analizowana jest dla pewnych rozkładów teoretycznych. Badana jest również zależność dolnej granicy „klasy środkowej”, zdefiniowanej nieco odmiennie od analogicznej klasy według Palmy, od współczynnika Giniego.

Słowa kluczowe: współczynnik Giniego, współczynnik Palmy, nierówności, ogony rozkładów.