

Measurement of the Bipolarization Events

Stefan V. Stefanescu

Abstract—We intend to point out the differences which exist between the classical Gini concentration coefficient and a proposed bipolarization index defined for an arbitrary random variable which have a finite support.

In fact Gini's index measures only the "poverty degree" for the individuals from a given population taking into consideration their wages. The Gini coefficient is not so sensitive to the significant income variations in the "rich people class".

In practice there are multiple interdependent relations between the pauperization and the socio-economical polarization phenomena. The presence of a strong pauperization aspect inside the population induces often a polarization effect in this society. But the pauperization and the polarization phenomena are not identical. For this reason it isn't always adequate to use a Gini type coefficient, based on the Lorenz order, to estimate the bipolarization level of the individuals from the studied population.

The present paper emphasizes these ideas by considering two families of random variables which have a linear or a triangular type distributions. In addition, the continuous variation, depending on the parameter "time" of the chosen distributions, could simulate a real dynamical evolution of the population.

Keywords—Bipolarization phenomenon, Gini coefficient, income distribution, poverty measure.

I. INTRODUCTION

TO take the best decisions in the social protection work it is necessary to estimate correctly the "pauperization level" which exist inside a population P ([23], [27]-[29], [40]).

In practice there are studied several interdependent relations between the pauperization and the socio-economical polarization phenomena ([2], [5], [19], [23]). The presence of a strong pauperization aspect inside the population P induces often a polarization effect in this society ([2], [5], [13], [17], [19], [23], [41]).

In the literature there are proposed a lot of indices for estimating the level of the social or economical inequalities ([3]-[11], [15]-[18], [27]-[42]). The present paper will use the Gini index as a measure concerning the "poverty" phenomenon ([25], [42]).

In fact, the classical Gini coefficient $\gamma_*(W)$ determines the "concentration degree" of the "small" values produced by the variable W , $W > 0$ ([25]-[27], [31]). In this study the random variable W defines the income distribution of the individuals of P . We mention here many other possible generalizations of Gini's indicator ([26], [42]).

Stefan V. Stefanescu is with the Faculty of Mathematics and Computer Science, University of Bucharest, Romania (e-mail: stefanst@fmi.unibuc.ro).

The literature presents a lot of papers which analyze the socio-economical polarization effect ([11]-[13], [19]). We proposed in [32]-[33] the index $\Delta_*(W)$ for measuring the bipolarization aspect of the random variable W . But we must not neglect other possible polarization indicators ([11], [13], [19]).

Based on the ordinary connection between the poverty and the bipolarization aspects, frequently is accredited in practice the wrong idea that Gini's concentration coefficient $\gamma_*(W)$ could also be used to evaluate the intensity of the income polarization phenomenon of the persons from the population P ([23], [35], [41]). We point out here that the pauperization and the polarization phenomena are not identical. In addition, the $\gamma_*(W)$ coefficient is not so sensitive to the significant income variations within the "rich people" group ([6]-[10], [25], [26], [29]).

For this reason is not adequate to use a Gini type coefficient, based on the Lorenz order, to estimate the bipolarization level of the individuals from P . Operating in a complex research with the single index $\gamma_*(W)$ could result inadequate conclusions and unsuitable decisions.

More details are given in the next section.

II. GINI'S MEASURE

The Gini coefficient is surely the most known index used to measure the inequality of income distribution ([4], [22], [27]). Gini's indicator takes values between 0 and 1. The zero value corresponds to perfect wage equality, that is all the individuals in the population P have the same income. Contrary, the value one for Gini index is obtained only for the strongest inequality case in P where a single person has all the population income while everyone else hasn't any wage ([9], [22], [25], [27]).

The Gini coefficient satisfies four important principles, that is : anonymity, scale independence, population independence and the transfer principle ([6]-[9], [22], [30]).

Let $W \geq 0$ be the variable which defines the variability of the income for the individuals from the population P . The distribution of the random variable (r.v.) W is characterizes by the probability density function (p.d.f.) $f(w)$ or equivalently by the cumulative distribution function (c.d.f.) $F(w)$, $w \geq 0$.

Then the true proportion q , $0 \leq q \leq 1$, of the people having an income less than or equal to w is given by the expression

$$q = \Pr(W \leq w) = \int_0^w f(t) dt = F(w)$$

If the random variable W varies in the interval $[0, b]$ then the proportion z of the sum of all incomes t which are less or equal to w is given by

$$z = \left(\int_0^w t f(t) dt \right) / \left(\int_0^b t f(t) dt \right)$$

Taking in consideration the proportions q and z we define the classical Lorenz curve $z = L(q; f)$, $0 \leq q = F(w) \leq 1$, that is ([22], [42])

$$L(F(w); f) = \left(\int_0^w t f(t) dt \right) / \left(\int_0^b t f(t) dt \right), \quad 0 \leq w \leq b$$

Since $F(y)$ is a nondecreasing function we have for every $0 \leq q \leq 1$ the generalized inverse $Q(q) = F^{-1}(q)$ of the c.d.f. $F(w)$,

$$Q(q) = \inf \{ w \mid F(w) \geq q \}$$

From the equality $F(Q(q)) = q$ which is true for every $0 \leq q \leq 1$, we deduce another form of Lorenz curve $z = L(q; f)$, $0 \leq q \leq 1$,

$$L(q; f) = \left(\int_0^{Q(q)} t f(t) dt \right) / \left(\int_0^b t f(t) dt \right) \quad (1)$$

We can prove that $L(q; f) \leq q$, $\forall 0 \leq q \leq 1$ (see for example [30]). More, the perfect wage equity in the population P is realized when $z = q$ for every $0 \leq w \leq b$. The Gini's index $\gamma_*(W)$ is based on these two properties. More exactly, the inequality of the individuals income in the population P can be evaluated by the value of $\gamma_*(W)$ where ([22])

$$\gamma_*(W) = 2 \int_0^1 (q - L(q; f)) dq \quad (2)$$

The Gini coefficient $\gamma_*(W)$ takes values between 0 and 1. In addition, in the situation of a perfect wage equality in P we have $L(q; f) = q$ for every $0 \leq q \leq 1$ and therefore $\gamma_*(W) = 0$.

But the computation of the coefficient $\gamma_*(W)$ is hardly for an arbitrary p.d.f. $f(w)$. In this case, for estimating the real $\gamma_*(W)$ value we suggest to use the stochastic Monte Carlo simulation or to apply the following equivalent formula ([22])

$$\gamma_*(W) = \frac{1}{\mu} \int_a^b F(x)(1 - F(x)) dx \quad (3)$$

where $\mu = \text{Mean}(W)$.

For any r.v. W the following inequalities are true ([22])

$$0 \leq \gamma_*(W) \leq 1 \quad (4)$$

We mention here that the formula (3) for computing Gini's value is true when the random variable W is defined on a bounded domain and more if exists a finite average value μ .

III. DEFINING A BIPOLARIZATION COEFFICIENT

Often, the complex socio-economic analysis takes into consideration the study of the bipolarization aspects inside the population P ([2], [17], [19], [41]).

In the last years a lot of papers proposed different statistical indices for measuring the polarization phenomenon (see for example [11], [13], [19], [32], [40]-[41]).

In the papers [32]-[33] we suggested the indicator $\Delta_*(W)$ to estimate the bipolarization degree for an arbitrary r.v. W which has a bounded support $[a, b]$. The variable W could be associated to the wage of the persons from P .

The bipolarization phenomenon assumes implicitly the existence of two disjoint sets I_a and I_b , for example the "poor", respectively the "rich" people.

So, the first problem is to identify a "suitable partition" $\{I_a, I_b\}$ of the domain $[a, b]$.

For our concrete case, since the r.v. W characterizes the income of the individuals from the population P , the subdomains I_a and I_b are convex sets. Therefore we will consider ([32]-[33])

$$\begin{aligned} I_a &= \{x \mid a \leq x \leq \mu_*\} \\ I_b &= \{x \mid \mu_* < x \leq b\} \end{aligned} \quad (5)$$

the subdomains I_a, I_b being separated by the threshold value μ_* .

A main problem arising now is to find an adequate separation threshold μ_* between the two disjoint groups I_a and I_b .

A good cut point μ_* is just the average μ of the r.v. W , that is $\mu_* = \mu = \text{Mean}(W)$ ([32], [33]).

The next step is to establish how opposite are the classes I_a and I_b .

Having in mind this idea, the indicator $\Delta_*(W)$ will evaluate the distance between the "poles" μ_a and μ_b of the groups I_a and I_b , taking also into consideration the "weights" p , respectively $1 - p$, of these groups.

More exactly, if $f(w)$, $a \leq w \leq b$, is the p.d.f. of the r.v. W and

$$\mu_a = \text{Mean}(W | I_a) = \int_a^\mu \frac{wf(w)}{p} dw$$

$$\mu_b = \text{Mean}(W | I_b) = \int_\mu^b \frac{wf(w)}{1-p} dw$$

$$p = \text{Pr}(W \leq \mu) = \int_a^\mu f(w) dw$$

$$1-p = \text{Pr}(W > \mu) = \int_\mu^b f(w) dw$$

then we define ([32], [33])

$$\Delta_*(W) = \frac{4p(1-p)(\mu_b - \mu_a)}{b-a} \quad (7)$$

Since $a \leq \mu_a \leq \mu_b \leq b$ and $0 \leq 4p(1-p) \leq 1$ we obtain the inequalities ([32]-[33])

$$0 \leq \Delta_*(W) \leq 1 \quad (8)$$

IV. THE LINEAR DISTRIBUTION

Definition 1. The random variable (r.v.) X is λ linear distributed, $X \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$, if its p.d.f. $f_1(x; \lambda)$ has the expression

$$f_1(x; \lambda) = 2\lambda x + 1 - \lambda, \quad 0 \leq x \leq 1 \quad (9)$$

The shape of the p.d.f. $f_1(x; \lambda)$ for different values of the parameter $-1 \leq \lambda \leq 1$ is illustrated by the graphic G1.

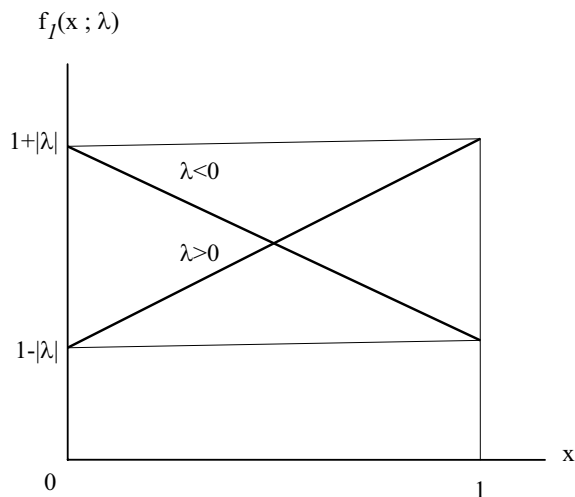


Fig. 1. The p.d.f. $f_1(x; \lambda)$, $0 \leq x \leq 1$, $-1 \leq \lambda \leq 1$.

Remark 1. The cumulative distribution function $F_1(x; \lambda)$ $0 \leq x \leq 1$, of $X \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$, has the form

$$F_1(x; \lambda) = \lambda x^2 + (1 - \lambda)x \quad (10)$$

Remark 2. The r.v. $X \sim \text{Lin}(0)$ is uniformly distributed

on the interval $[0, 1]$, that is $X \sim \text{Uni}([0, 1])$.

Proposition 1. If $X \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$, then

$$\mu_1 = \text{Mean}(X) = \frac{\lambda + 3}{6} \quad (11)$$

Proof. We have

$$\begin{aligned} \mu_1 &= \text{Mean}(X) = \int_0^1 x f_1(x; \lambda) dx = \\ &= \int_0^1 (2\lambda x^2 + (1 - \lambda)x) dx = \frac{\lambda + 3}{6} \end{aligned}$$

Proposition 2. For $X \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$, the Gini index $\gamma_*(X)$ has the expression

$$\gamma_1(\lambda) = \gamma_*(X) = \frac{5 - \lambda^2}{5(3 + \lambda)} \quad (12)$$

Proof. If J_1 is the definite integral

$$J_1 = \int_0^1 F_1(x; \lambda)(1 - F_1(x; \lambda)) dx$$

then it results

$$\begin{aligned} J_1 &= \int_0^1 (\lambda x^2 + (1 - \lambda)x)(1 - \lambda x^2 + (1 - \lambda)x) dx = \\ &= \frac{5 - \lambda^2}{30} \end{aligned}$$

Applying formula (3) and Proposition 1 we deduce that the Gini's index $\gamma_*(X)$ is given by the expression

$$\gamma_1(\lambda) = \gamma_*(X) = \frac{J_1}{\mu_1} = \frac{(5 - \lambda^2)/30}{(\lambda + 3)/6} = \frac{5 - \lambda^2}{5(3 + \lambda)}$$

Corollary 1. The function $\gamma_1(\lambda)$ decreases for any $-1 \leq \lambda \leq 1$.

Proof. Since $\lambda^2 + 6\lambda + 5 = (\lambda + 1)(\lambda + 5) \geq 0$ for any $-1 \leq \lambda \leq 1$ we conclude that the derivative $\gamma_1'(\lambda)$ of the function $\gamma_1(\lambda)$ is not positive on the domain $[-1, 1]$,

$$\gamma_1'(\lambda) = -\frac{\lambda^2 + 6\lambda + 5}{5(3 + \lambda)^2} \leq 0$$

that is the Gini's index $\gamma_1(\lambda)$ decreases on the interval $[-1, 1]$.

Corollary 2. For an arbitrary $-1 \leq \lambda \leq 1$ we have the following inequalities

$$\frac{1}{5} \leq \gamma_1(\lambda) \leq \frac{2}{5} \quad (13)$$

Proof. Taking into consideration the monotony of the $\gamma_1(\lambda)$ function defined by the formula (12) on $[-1, 1]$ interval we get

$$\frac{1}{5} = \gamma_1(1) \leq \gamma_1(\lambda) \leq \gamma_1(-1) = \frac{2}{5}$$

Corollary 3. For any $-1 \leq \lambda \leq 1$ the equality $\gamma_1(\lambda) = \gamma_1(-\lambda)$ is true if and only if $\lambda = 0$.

Proof. Obviously for $\lambda = 0$ we have

$$\gamma_1(\lambda) = \gamma_1(0) = \gamma_1(-\lambda).$$

Reciprocally, from $\gamma_1(\lambda) = \gamma_1(-\lambda)$ we get

$$\frac{5 - \lambda^2}{5(3 + \lambda)} = \frac{5 - \lambda^2}{5(3 - \lambda)}, \quad -1 \leq \lambda \leq 1, \text{ that is } \lambda = 0.$$

Proposition 3. The polarization index $\Delta_*(X)$ of the r.v. $X \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$, is given by the formula

$$\Delta_1(\lambda) = \Delta_*(X) = \frac{(9 - \lambda^2)^2}{162} \quad (14)$$

Proof. Indeed, interpreting the formula (7), the polarization index $\Delta_1(\lambda) = \Delta_*(X)$ of the r.v. $X \sim \text{Lin}(\lambda)$ has the expression

$$\Delta_1(\lambda) = \Delta_*(X) = \frac{4p_1(1-p_1)(\mu_{1b} - \mu_{1a})}{b-a}$$

where

$$a = 0 \quad b = 1$$

$$p_1 = Pr(X \leq \mu_1) = \int_0^{\mu_1} f_1(x; \lambda) dx = \int_0^{(\lambda+3)/6} (2\lambda x + 1 - \lambda) dx = \frac{(\lambda+3)(\lambda^2 - 3\lambda + 6)}{36} = \frac{\lambda^3 - 3\lambda + 18}{36}$$

$$1 - p_1 = Pr(X > \mu_1) = 1 - \frac{\lambda^3 - 3\lambda + 18}{36} = \frac{18 + 3\lambda - \lambda^3}{36} = \frac{(3 - \lambda)(\lambda^2 + 3\lambda + 6)}{36}$$

$$\mu_{1a} = \int_0^{\mu_1} \frac{x f_1(x; \lambda)}{p_1} dx = \frac{1}{p_1} \int_0^{(\lambda+3)/6} (2\lambda x^2 + (1 - \lambda)x) dx = \frac{(\lambda+3)(2\lambda^2 - 3\lambda + 9)}{18(\lambda^2 - 3\lambda + 6)}$$

$$\mu_{1b} = \int_{\mu_1}^1 \frac{x f_1(x; \lambda)}{1 - p_1} dx = \frac{1}{1 - p_1} \int_{(\lambda+3)/6}^1 (2\lambda x^2 + (1 - \lambda)x) dx = \frac{(\lambda+3)(2\lambda^2 + 9\lambda + 27)}{18(\lambda^2 + 3\lambda + 6)}$$

Therefore, the polarization index $\Delta_*(X)$ associated to the r.v. $X \sim \text{Lin}(\lambda)$ has the expression

$$\Delta_1(\lambda) = \Delta_*(X) = \frac{4p_1(1-p_1)(\mu_{1b} - \mu_{1a})}{b-a} = \frac{(9 - \lambda^2)(3 + \lambda)(108 - 36\lambda)}{2^3 3^6} = \frac{(9 - \lambda^2)^2}{162}$$

Corollary 4. For any $\lambda^2 \leq 1$ the following inequalities are true

$$0.395 \approx \frac{32}{81} = \Delta_1(1) \leq \Delta_1(\lambda) \leq \Delta_1(0) = \frac{1}{2} = 0.500 \quad (15)$$

V. THE TRIANGULAR DISTRIBUTION

Definition 2. For any $0 \leq \lambda \leq 2$ the r.v. Y has a λ triangular distribution, $Y \sim \text{Tri}(\lambda)$, if its p.d.f. $f_2(y; \lambda)$ has

the form

$$f_2(y; \lambda) = \begin{cases} 2 - \lambda + 4(\lambda - 1)y & \text{for } 0 \leq y \leq 1/2 \\ 3\lambda - 2 + 4(1 - \lambda)y & \text{for } 1/2 < y \leq 1 \end{cases} \quad (16)$$

The shape variation of the p.d.f. $f_2(y; \lambda)$, $0 \leq y \leq 1$, depending on the parameter $0 \leq \lambda \leq 2$, is suggested by the graphic G2.

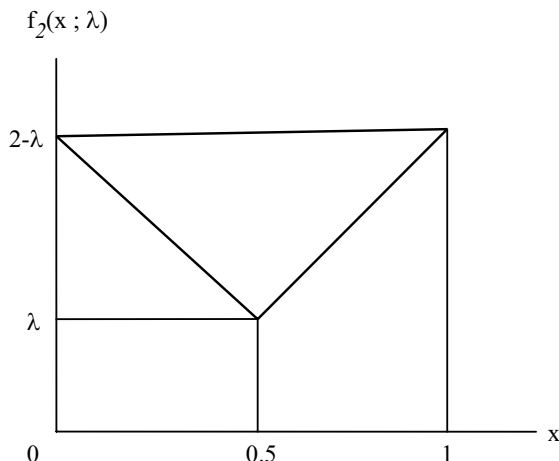


Fig. 2. The p.d.f. $f_2(x; \lambda)$, $0 \leq x \leq 1$, $0 \leq \lambda \leq 2$.

Remark 3. The c.d.f. $F_2(y; \lambda)$ of the r.v. $Y \sim \text{Tri}(\lambda)$ is given by the expression

$$F_2(y; \lambda) = \begin{cases} (2 - \lambda)y + 2(\lambda - 1)y^2 & \text{for } 0 \leq y \leq 1/2 \\ 1 - \lambda + (3\lambda - 2)y + 2(1 - \lambda)y^2 & \text{for } 1/2 < y \leq 1 \end{cases} \quad (17)$$

Remark 4. The r.v. $Y \sim \text{Tri}(1)$ is uniformly distributed on the interval $[0, 1]$, $Y \sim \text{Uni}([0, 1])$.

Proposition 4. If $Y \sim \text{Tri}(\lambda)$, $0 \leq \lambda \leq 2$, then we get

$$\mu_2 = \text{Mean}(Y) = \frac{1}{2} \quad (18)$$

Proof. Obviously we obtain

$$\begin{aligned} \mu_2 = \text{Mean}(Y) &= \int_{-\infty}^{+\infty} y f_2(y; \lambda) dy = \\ &= \int_0^{1/2} y f_2(y; \lambda) dy + \int_{1/2}^1 y f_2(y; \lambda) dy = \\ &= \int_0^{1/2} ((2 - \lambda)y + 4(\lambda - 1)y^2) dy + \\ &+ \int_{1/2}^1 ((3\lambda - 2)y + 4(1 - \lambda)y^2) dy = \\ &= \frac{2 + \lambda}{24} + \left(\frac{\lambda + 2}{6} - \frac{5\lambda - 2}{24} \right) = \frac{1}{2} \end{aligned}$$

Proposition 5. For any $0 \leq \lambda \leq 2$ the Gini index $\gamma_*(Y)$ of the r.v. Y , $Y \sim \text{Tri}(\lambda)$, is given by the formula

$$\gamma_2(\lambda) = \gamma_*(Y) = \frac{J_{21}}{\mu_2} = \frac{24 - 3\lambda - \lambda^2}{60} \quad (19)$$

Proof. Let J_{2a} and J_{2b} be the integrals

$$J_{2a} = \int_0^{1/2} F_2(y; \lambda) (1 - F_2(y; \lambda)) dy = \int_0^{1/2} ((2 - \lambda)y + 2(\lambda - 1)y^2) (1 - ((2 - \lambda)y + 2(\lambda - 1)y^2)) dy$$

$$J_{2b} = \int_{1/2}^1 F_2(y; \lambda) (1 - F_2(y; \lambda)) dy = \int_{1/2}^1 (1 - \lambda + (3\lambda - 2)y + 2(1 - \lambda)y^2) (\lambda - (3\lambda - 2)y - 2(1 - \lambda)y^2) dy$$

After a direct calculus we obtain

$$J_{2a} = \frac{24 - 3\lambda - \lambda^2}{240}$$

More, applying to the integral J_{2b} the transform $y = 1 - t$ we get just the value J_{2b} .

Hence

$$J_2 = \int_{-\infty}^{+\infty} F_2(y; \lambda) (1 - F_2(y; \lambda)) dy = J_{2a} + J_{2b} = 2J_{2a} = \frac{24 - 3\lambda - \lambda^2}{120}$$

Therefore, from (3) and Proposition 4, the Gini's index $\gamma_*(Y)$ of the r.v. Y has the expression (19).

Corollary 5. The Gini coefficient $\gamma_2(\lambda)$ is a strict decreasing function on the interval $[0, 2]$.

Proof. Indeed, the derivative $\gamma_2'(\lambda) = -(3 + 2\lambda) / 60$ of the function $\gamma_2(\lambda)$ which has the expression (19) is strict negative for any $0 \leq \lambda \leq 2$.

Using the monotony property of the application $\gamma_2(\lambda)$ we deduce

Corollary 6. For an arbitrary $0 \leq \lambda \leq 2$ there are true the relations

$$0.233 \approx \frac{7}{30} = \gamma_2(2) \leq \gamma_2(\lambda) \leq \gamma_2(0) = \frac{2}{5} = 0.4 \quad (20)$$

Corollary 7. For any $\lambda_1 \neq \lambda_2$ values belonging to the interval $[0, 2]$ we have

$$\gamma_2(\lambda_1) \neq \gamma_2(\lambda_2) \quad (21)$$

Proof. This relation results by using the strict monotony of the function $\gamma_2(\lambda)$ which is defined on the interval $[0, 2]$ (see formula (19)).

Proposition 6. The polarization index $\Delta_*(Y)$ of the r.v. $Y \sim \text{Tri}(\lambda)$ has the expression

$$\Delta_2(\lambda) = \Delta_*(Y) = \frac{4 - \lambda}{6}, \quad 0 \leq \lambda \leq 2 \quad (22)$$

Proof. Interpreting the formulas (6) and (7) the polarization index $\Delta_2(\lambda) = \Delta_*(Y)$ of the r.v. $Y \sim \text{Tri}(\lambda)$ has the form

$$\Delta_2(\lambda) = \Delta_*(Y) = \frac{4p_2(1 - p_2)(\mu_{2b} - \mu_{2a})}{b - a}$$

where

$$a = 0 \quad b = 1$$

$$p_2 = \Pr(Y \leq \mu_2) = F_2(\mu_2; \lambda) = F_2(1/2; \lambda) = \frac{1}{2}$$

$$\mu_{2a} = \int_a^{\mu_2} \frac{y f_2(y; \lambda)}{p_2} dy = \int_0^{1/2} ((2 - \lambda)y + 4(\lambda - 1)y^2) dy = \frac{\lambda + 2}{12}$$

$$\mu_{2b} = \int_{\mu_2}^b \frac{y f_2(y; \lambda)}{1 - p_2} dy = \int_{1/2}^1 ((3\lambda - 2)y + 4(1 - \lambda)y^2) dy = \frac{10 - \lambda}{12}$$

Therefore it results

$$\Delta_2(\lambda) = \Delta_*(Y) = \mu_{2b} - \mu_{2a} = \frac{4 - \lambda}{6}$$

Corollary 8. The function $\Delta_2(\lambda)$ decreases on the interval $[0, 2]$.

Corollary 9. For any $0 \leq \lambda \leq 2$ the following inequalities are true

$$\frac{1}{3} = \Delta_2(2) \leq \Delta_2(\lambda) \leq \Delta_2(0) = \frac{2}{3} \quad (23)$$

VI. COMPARING THE TWO INDICATORS

In this section we intend to compare the values of the indicators $\gamma_*(W)$ and $\Delta_*(W)$ for different r.v.-s W .

We have in mind that the variable W defines the income distribution of the individuals from the population P .

Remark 5. When the r.v. W varies in the interval $[0, 1]$ then the variable $V = 1 - W$ reverses significance of the "poor" and "rich" people in P . So, a rich individual having the wage W is transformed in a poor person with the income $V = 1 - W$. In the same manner, by applying the income transformation $v = 1 - w$ for all individuals of P , the poor people becomes wealthy.

For the subsequent we'll consider $W \equiv X_\lambda$ with $X_\lambda \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$ or $W \equiv Y_\lambda$ where $Y_\lambda \sim \text{Tri}(\lambda)$, $0 \leq \lambda \leq 2$, the population P being denoted in these cases by P_X , respectively P_Y .

It is easy to prove that

Proposition 7. For any fixed $0 \leq x, y \leq 1$ the functions $f_1(x; \lambda)$, $F_1(x; \lambda)$, $\gamma_1(\lambda)$, $\Delta_1(\lambda)$, respectively $f_2(y; \lambda)$,

$F_2(y; \lambda)$, $\gamma_2(\lambda)$, $\Delta_2(\lambda)$ are continuous depending on the λ variable ($-1 \leq \lambda \leq 1$ or $0 \leq \lambda \leq 2$).

Remark 6. The continuous variation, depending on the "time" λ , of the distributions of the random variables X_λ or Y_λ could be interpreted as a dynamical income evolution of the individuals from the population P_X , respectively P_Y .

Remark 7. When the values of the parameter λ increases, from the graphics G1 and G2 we deduce that the percentage of the poor people is diminished in the populations P_X or P_Y (the probability to have individuals in P with "small income" decreases). Taking into consideration the practical significance of Gini's coefficient $\gamma_*(W)$ which is strongly sensitive especially to the proportion of the "poor class", we conclude that the functions $\gamma_1(\lambda) = \gamma_*(X_\lambda)$ and $\gamma_2(\lambda) = \gamma_*(Y_\lambda)$ must decrease. This intuitive observation is validated by Corollary 1, respectively Corollary 5.

We will establish new others properties of the indices $\gamma_*(W)$ and $\Delta_*(W)$ when $W = X_\lambda$ or $W = Y_\lambda$.

Proposition 8. If $X \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$ and $Z = 1 - X$ then $Z \sim \text{Lin}(-\lambda)$

Proof. Let $f_1(x; \lambda)$ and $g_1(z; \lambda)$ be the probability density functions of the random variables X , respectively Z . So, applying the one to one transform $z = 1 - x$ we get

$$\begin{aligned} g_1(z; \lambda) &= f_1(x; \lambda) \left| \frac{\partial x}{\partial z} \right| = f(x; \lambda) = f(1 - z; \lambda) = \\ &= 2\lambda(1 - z) + 1 - \lambda = -2\lambda z + 1 + \lambda = f(z; -\lambda) \end{aligned}$$

Corollary 10. If $X \sim \text{Lin}(\lambda)$ and $Z = 1 - X$ then for any $0 < |\lambda| \leq 1$ we have

$$\gamma_*(Z) \neq \gamma_*(X) \quad (24)$$

Proof. When $\lambda \neq 0$ from formula (12) it results

$$\gamma_*(Z) = \gamma_1(-\lambda) = \frac{5 - \lambda^2}{5(3 + \lambda)} \neq \frac{5 - \lambda^2}{5(3 - \lambda)} = \gamma_1(\lambda) = \gamma_*(X)$$

Corollary 11. For any $|\lambda| \leq 1$, if $X \sim \text{Lin}(\lambda)$ and $Z = 1 - X$ then it is true the equality

$$\Delta_*(Z) = \Delta_*(X) \quad (25)$$

Proof. Indeed from (14) we get

$$\Delta_*(Z) = \Delta_1(-\lambda) = \frac{(9 - \lambda^2)^2}{162} = \Delta_1(\lambda) = \Delta_*(X)$$

Proposition 9. If $Y \sim \text{Tri}(\lambda)$, $0 \leq \lambda \leq 2$ and $V = 1 - Y$ then $V \sim \text{Tri}(\lambda)$.

Proof. Indeed, if $f_2(y; \lambda)$ and $g_2(v; \lambda)$, $0 \leq y, v \leq 1$, are the p.d.f.-s of the r.v.-s Y and V then we have

$$\begin{aligned} g_2(v; \lambda) &= f_2(y; \lambda) \left| \frac{\partial y}{\partial v} \right| = f_2(y; \lambda) = f_2(1 - v; \lambda) = \\ &= \begin{cases} 2 - \lambda + 4(\lambda - 1)(1 - v) & \text{for } 0 \leq 1 - v \leq 1/2 \\ 3\lambda - 2 + 4(1 - \lambda)(1 - v) & \text{for } 1/2 \leq 1 - v \leq 1 \end{cases} \end{aligned}$$

$$= \begin{cases} 3\lambda - 2 + 4(1 - \lambda)v & \text{for } 1/2 \leq v \leq 1 \\ 2 - \lambda + 4(\lambda - 1)v & \text{for } 0 \leq v \leq 1/2 \end{cases} = f_2(v; \lambda)$$

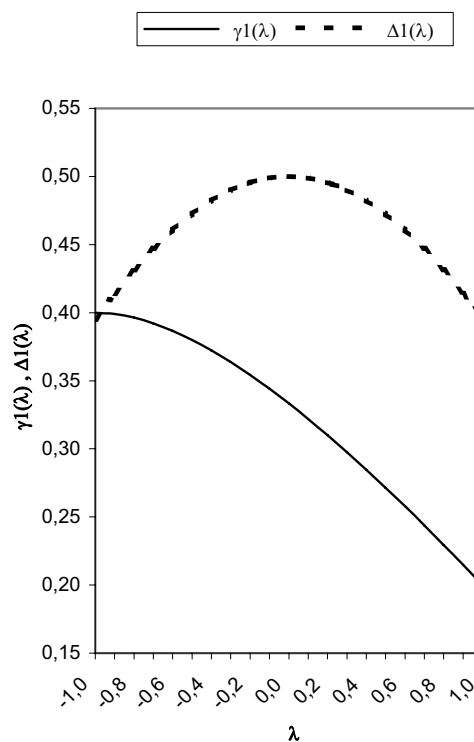


Fig. 3. The variation of the indices $\gamma_1(\lambda)$, $\Delta_1(\lambda)$.

Remark 8. The polarization coefficient $\Delta_*(W)$ measures the differences between the "poor" and the "wealthy" classes. Applying the transform $T = 1 - W$ the status of the individuals in P is changed, a "rich" individual becoming a "poor" man and viceversa. In fact, after this income transformation we get the same groups of persons but with another "names". For this reason the bipolarization index must remain invariant to the antithetic transformation $T = 1 - W$ ([32]-[33]).

More explicitly

Property AT ([32], [33]). For any r.v. $0 \leq W \leq 1$, we always have

$$\Delta_*(W) = \Delta_*(1 - W) \quad (26)$$

In particular the property **AT** must be true for the r.v.-s X_λ and Y_λ (see the proofs from Corollary 11 and Proposition 9).

Remark 9. As a rule, the Gini coefficient has not the property **AT** (see Corollary 10). This result makes the main distinction between the Gini coefficient $\gamma_*(W)$ and the bipolarization index $\Delta_*(W)$.

Remark 10. Having in mind the last remarks we conclude that usually the behavior of the indicators $\gamma_*(W)$ and $\Delta_*(W)$

is not the same. So, for the population P_X , in the graphic G3 are compared the evolution of the indices values $\gamma_1(\lambda) = \gamma_*(X_\lambda)$ and $\Delta_1(\lambda) = \Delta_*(X_\lambda)$ with $X_\lambda \sim \text{Lin}(\lambda)$, $-1 \leq \lambda \leq 1$. We mention here that the functions $\gamma_1(\lambda)$ and $\Delta_1(\lambda)$ do not have the same intervals of monotony (graphic G3, Corollary 1, formula (14)).

Remark 11. Contrary, in the case of the population P_Y , both functions $\gamma_2(\lambda) = \gamma_*(Y_\lambda)$ and $\Delta_2(\lambda) = \Delta_*(Y_\lambda)$, decrease for $Y_\lambda \sim \text{Tri}(\lambda)$, $0 \leq \lambda \leq 2$ (graphic G4, Corollary 5, Corollary 8). The monotony of the indicator $\Delta_*(Y_\lambda)$ is imposed by the symmetry of the p.d.f. $f_2(y; \lambda)$ (formula (16), graphic G2). So, the reduction of the number of persons which have "very small" income in the population P_Y is equivalent to the diminishing, by the same proportion, of the number of "very rich" individuals in P_Y (to interpret the graphic G2). For this reason, the "poor" and "rich" classes become more distinct when the percentage of poor in P_Y increases. Therefore the both indicators $\gamma_*(Y_\lambda)$ and $\Delta_*(Y_\lambda)$ will fluctuates in the same manner (see also graphic G4).

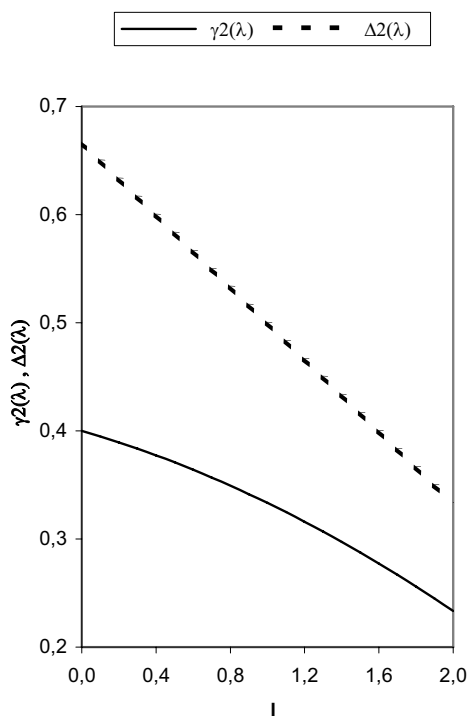


Fig. 4. The variation of the indices $\gamma_2(\lambda)$ and $\Delta_2(\lambda)$.

VII. CONCLUDING REMARKS

Between the pauperization and the bipolarization socio-economical phenomena strong interconnections are present. Often in practice we associate a bipolarization effect to a

powerful pauperization event which appeared inside the population P . We remark here that the mentioned phenomena are not identical.

Taking into consideration the distribution W of the income for the individuals from P we have chosen the classical Gini's coefficient $\gamma_*(W)$ to measure the intensity of the pauperization event. Using some examples we proved that this measure correspond to our intuition (see also Remark 7 and graphic G3).

A bipolarization effect presumes implicitly the presence of two distinct groups. We chose the measure $\Delta_*(W)$ which is based essentially on this idea. The present paper validated indirectly, on some intuitive examples, the good behavior of the indicator $\Delta_*(W)$ (Remark 8 and **AT** property, interpret the graphic G3).

So, the Gini coefficient $\gamma_*(W)$ establishes the poverty "concentration" level in the population P and the index $\Delta_*(W)$ indicates a possible bipolarization property of W values. The obtained results confirmed without any doubt that both coefficients, $\gamma_*(W)$ and $\Delta_*(W)$, are suitable measures for these two distinct aspects of the social reality (reflect on the signification of the p.d.f.-s $f_1(x; \lambda)$ and $f_2(y; \lambda)$, explain intuitively the shape of the indices $\gamma_k(\lambda)$, $\Delta_k(\lambda)$, $k \in \{1, 2\}$, and analyze the monotony intervals for these indicators, eventually).

In general, if we reduce substantially the percentage of the "poorest" individuals in P we do not always affect significantly a strong bipolarization phenomenon. This perception is also confirmed practically by the graphic G3.

Two families of distributions were chosen to shape the income variation of the individuals from the population P . So, we selected the linear $\{X_\lambda\}_\lambda$ and the triangular $\{Y_\lambda\}_\lambda$ families of random variables taking into consideration their intuitive interpretations. More, we proved that the indicators $\gamma_k(\lambda)$, $\Delta_k(\lambda)$ vary in a continuous manner (Proposition 13). This property permits us to have a good real imagine about a dynamic evolution of the income.

In the last years, is detectable the presence of a hard economical bipolarization tendency especially for East European countries. This special effect couldn't be practically pointed out by using only a general income inequality measure as Gini's coefficient. For a complex reality is insufficient to analyze the socio-economical events, in time and also in space, based exclusively on the poverty concentration coefficient $\gamma_*(W)$. A complementary study concerning the variation of the bipolarization indicator $\Delta_*(W)$ or using any other polarization indices is absolutely necessary. We mention here that, as a rule, is inadequate to use the Gini coefficient $\gamma_*(W)$ for measuring the bipolarization degree inside a given population P . See, for example, the graphic G3.

But sometimes, depending on the income distribution in P (for example when $W \sim \text{Tri}(\lambda)$) the indicators $\gamma_*(W)$ and $\Delta_*(W)$ could have a similar behavior (see Remark 11 and interpret graphic G4).

Concluding, we strongly recommend to use, in a complementary way, the $\Delta_*(W)$ bipolarization index to emphasize a grouping phenomenon and to apply Gini's coefficient $\gamma_*(W)$ only for estimating the intensity of the "poverty level" inside the population P .

ACKNOWLEDGMENT

The author thanks to Professor Sant Sharan Mishra and also to the referees for improving the initial version of the manuscript.

REFERENCES

- [1] A. Agresti, *An introduction to categorical data analysis*. New York: Wiley Series in Probability and Statistics, 1996.
- [2] C. D'Ambrosio, "Household characteristics and the distribution of income in Italy - An application of social distance measures," *The Review of Income and Wealth*, vol. 47, no. 1, pp. 43-64, 2001.
- [3] A. B. Atkinson, "On the measurement of inequality," *Journal of Economic Theory*, no. 2, pp. 244-263, 1970.
- [4] D. J. Bartolomew, *The statistical approach to social measurement*. London: Academic Press, 1996.
- [5] N. Bhattacharya, and B. Mahalanobis, "Regional disparities in household consumption in India," *Journal of the American Statistical Association*, vol. 62, pp. 143-161, 1967.
- [6] F. Bourguignon, "Decomposable income inequality measures," *Econometrica*, vol. 47, pp. 901-920, 1979.
- [7] A. Chateauneuf, T. Gajdos, and P. H. Wiltien, "The principle of strong diminishing transfer," *Journal of Economic Theory*, vol. 103, pp. 311-333, 2002.
- [8] F. A. Cowell, "On the structure of additive inequality measures," *Review of Economic Studies*, vol. 47, pp. 521-531, 1980.
- [9] F. A. Cowell, and S. P. Jenkins, "How much inequality can we explain? A methodology and an application to the United States," *The Economic Journal*, vol. 105, pp. 421-430, 1995.
- [10] J. B. Davies, and A. Shorrocks, "Optimal grouping of income and wealth data," *Journal of Econometrics*, vol. 42, pp. 97-108, 1989.
- [11] J. Esteban, and D. Ray, "On the measurement of polarization," *Econometrica*, vol. 62, no. 4, pp. 819-852, 1994.
- [12] J. Esteban, and D. Ray, "Conflict and distribution," *Journal of Economic Theory*, vol. 87, pp. 379-415, 1999.
- [13] J. Esteban, C. Gradin, and D. Ray, "Extension of a measure of polarization with application to the income distribution of five OECD countries," Maxwell School of Citizenship and Public Affairs, Syracuse University, New York, Working paper no. 218, November 1999.
- [14] B. S. Everitt, S. Landau, and M. Leese, *Cluster analysis*. London: Arnold, 2001.
- [15] J. Foster, and A. K. Sen, *On economic inequality*. Oxford: Clarendon Press, 1997.
- [16] J. E. Foster, and A. A. Shneyerov, "Path independent inequality measures," *Journal of Economic Theory*, Oxford, vol. 91, no. 2, pp. 199-222, 2000.
- [17] M. Fournier, "Inequality decomposition by factor component - A new approach illustrated on the Taiwanese case," CERDI, Université d'Auvergne, November 1999.
- [18] J. Gastwirth, "The estimation of a family of measures of economic inequality," *Journal of Econometrics*, no. 3, pp. 61-70, 1975.
- [19] C. Gradin, "Polarization by sub-populations in Spain, 1973-1991," *Review of Income and Wealth*, vol. 46, no. 4, pp. 457-474, December 2000.
- [20] P. E. Hart, "The comparative statistics and dynamics of income distributions," *Journal of the Royal Statistical Society*, vol. 139, pp. 108-125, 1976.
- [21] N. C. Kakwani, "Statistical inference in the measurement of poverty," *Review of Economics and Statistics*, vol. 75, no. 3, pp. 632-639, 1993.
- [22] C. Kleiber, and S. Kotz, *Statistical size distributions in economics and actuarial sciences*. New Jersey: John Wiley & Sons - Interscience, 2003.
- [23] J. O. Lanjouw, and P. Lanjouw, "How to compare apples and oranges - Poverty measurement based on different definitions of consumption," *Review of Income and Wealth*, vol. 47, no. 1, pp. 25-42, 2001.
- [24] A. W. Pedersen, "Inequality as relative deprivation - Theoretical issues and implications for empirical research," preprint ISA RC19 Conference 2001, "Old a new social inequalities", University of Oviedo, September 2001.
- [25] G. Pyatt, "On the interpretation and disaggregation of Gini coefficient," *The Economic Journal*, vol. 86, pp. 243-255, 1976.
- [26] F. Schmid, "A general class of poverty measures," *Statistical Papers*, vol. 34, pp. 189-211, 1993.
- [27] A. K. Sen, *On economic inequality*. Oxford: Oxford University Press, 1973.
- [28] A. F. Shorrocks, "The class additively decomposable inequality measures," *Econometrica*, vol. 48, pp. 613-625, 1980.
- [29] A. F. Shorrocks, "Inequality decomposition by population subgroups," *Econometrica*, vol. 52, pp. 1369-1385, 1984.
- [30] P. Stefanescu, and S. Stefanescu, "Extending the Gini index to measure inequality and poverty," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 35, no. 1-4, pp. 145-154, 2001.
- [31] P. Stefanescu, and S. Stefanescu, "Comparing Gini index with three other concentration measures," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 36, no. 1-4, pp. 173-184, 2002.
- [32] P. Stefanescu, and S. Stefanescu, "The polarization index for bounded exponential distributions," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 40, no. 3-4, pp. 211-218, 2006.
- [33] P. Stefanescu, and S. Stefanescu, "The properties of a polarization index for bounded exponential distributions," *U. P. B. - Scientific Bulletin, Series A : Applied Mathematics and Physics*, vol. 68, no. 4, pp. 9-20, 2006.
- [34] P. Stefanescu, and S. Stefanescu, "The Monte Carlo estimation of a polarization index for an arbitrary distribution," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 41, no. 3-4, pp. 181-192, 2007.
- [35] S. Stefanescu, "Measuring the socio-economic bipolarization phenomenon," *Romanian Journal of Economic Forecasting*, vol. 9, no. 1, pp. 149-161, 2008.
- [36] K. Y. Tsui, "Multidimensional inequality and multidimensional entropy measures - An axiomatic derivation," *Social Choice and Welfare*, vol. 16, no. 1, pp. 145-157, 1999.
- [37] B. Wilfling, "Lorenz ordering of power function order statistics," *Statistics & Probability Letters*, vol. 30, pp. 313-319, 1996.
- [38] M. Wolfson, "When inequalities diverge," *American Economic Review*, vol. 84, no. 2, pp. 353-358, 1994.
- [39] M. Wolfson, "Divergent inequalities - Theory and empirical results," *The Review of Income and Wealth*, vol. 43, no. 4, pp. 401-422, 1997.
- [40] S. Yitzhaki, "Economic distance and overlapping of distributions," *Journal of Econometrics*, vol. 61, pp. 147-159, 1994.
- [41] X. Zhang, and R. Kanbur, "What difference do polarization measures make? - An application to China," *Journal of Development Studies*, vol. 37, pp. 85-98, 2001.
- [42] C. Zoli, "Intersecting generalized Lorenz curves and the Gini index," *Social Choice and Welfare*, vol. 16, pp. 183-196, 1999.