

| Katarzyna Ostasiewicz

INCOME AND EXPENDITURE INEQUALITIES AMONG HOUSEHOLDS IN POLAND

Abstract

Stratification and its inverse concept, overlapping, are strictly connected with inequality, playing an essential role in one kind of decomposition of the Gini index. Properties of overlapping indexes are investigated in this paper. Decomposition of inequalities in Poland (of incomes, expenditures, and their difference), which includes the overlapping-due term, is performed and analysed.

Keywords: stratification, overlapping, Gini, decomposition, between-group inequality.

JEL Classification: D14, D63, I32.

1. Introduction

Inequalities of wealth, or, in general, in quality of life and standards of living, are an urgent problem nowadays, gaining much attention among the public and policy-makers alike. To better understand the situation, it is frequently useful to know how overall inequality is distributed among various geographical regions and/or social groups. Thus, the question arises of the proper decomposition of inequality measures.

As for the Gini index, the most popular and most frequently used measure of inequality, its decomposition with respect both to different groups and to different sources of the total quantity in question (e.g. different sources of income), is known and used. Within the standard approach to decomposition of the Gini index with respect to different groups, three components are identified (see, for instance, Bhattacharya & Mahalanobis 1967; Pyatt 1976; and Dagum 1997). First, it is the within-group component that is a weighted

sum of the Gini indexes for each of the subgroups, while the weighting factor is the share of the group in the total population times the share of the total good belonging to the particular group. The second component is the between-group one, which is calculated as the Gini index while neglecting all diversity within groups – that is, treating all individuals within a particular group as if they possessed an amount of good equal to the average for this group. These two components in general do not constitute the overall Gini index, and there remains the so-called residual term, which results from the overlapping of the distributions of different groups. This means that if the group with the lower average did not involve any individual which had more than any individual from the group with the higher average value, the overlap component would be equal to zero. However, this is rarely the case for any set of natural groups. Of course, if one defined groups as quintile groups, for instance, they would not overlap by their very definition. However, once we are willing to investigate any geographical or socio-economic groups, the overlapping will occur and thus the third term of decomposition will have non-zero value. This residual term takes into account all pairs of individuals from the whole populations whose order (with respect to the investigated quantity, e.g. income) is inverse to the order of the averages of the groups the individuals belong to.

Another approach to decomposition of the Gini index has been proposed by Yitzhaki (1994), who focussed on the concept of stratification.

Stratification is in fact one of the many faces of inequality, along with others such as segregation and polarisation. Under the name of stratification, four types of strata are considered: caste, class, estate, and slavery. An example of perfect caste stratification is the caste system in India. In the case of class stratification, the essential problem concerns the determination of the three “basic” classes: working class, middle class, and upper, or capitalist class. Problems concerning economic polarisation, which means polarisation according to economic status, are investigated at length in M. Kot’s monograph (2008). The scope of this paper is rather to shed some light on the problem of stratification of inequalities when strata are overlapping.

Assuming that definitions of social classes are dependent not merely on income but on the type of occupation, lifestyle, social capital, etc., some members of the middle class may be richer than some members of the aristocracy. Thus, stratification of social classes is not perfect. Again, the groups or classes might be defined in such a way as to ensure perfect stratification. However, when dealing with natural socio-economic groups or

geographical subpopulations, one may expect no stratification will be perfect and that some overlapping will occur between each pair of groups.

Within the scope of Yitzhaki and Lerman's decomposition (1991), the within-group part is identical as in the standard decomposition. The between-group part is expressed in terms of stratification and assumes, in general, a different value from the standard between-group Gini. The residual term in the latter decomposition differs from the standard one with respect to establishing the position of the particular group within the whole population. While in the standard method it is taken as the rank of the average of the group in the set of the groups' averages, the latter decomposition makes use of more information. The rank of the group is the average rank of all individuals' ranks (all individuals belonging to the group). Thus, it may be different from the former. For example, if only one individual makes the average of his group very high (and accordingly gives it a high rank), while all of the remaining individuals are placed very low in the whole population, the average of individuals' ranks may be much lower than the rank of the overall average for this group.

The concept of stratification, and its inverse – overlapping, allows us to present the mutual relationship between each pair of groups, not only the overall picture, which might be advantageous in many situations. Moreover, the overlapping index, as will be shown in later sections of this paper, is standardised as it does not exceed the value of 2 (and cannot be negative). This is an advantage in the presence of negative values. It is known that the Gini index is not normalised in the presence of negatives, that is, it is not bounded from above by 1 and it is difficult to compare two different Ginis in such a situation. Of course, the method of decomposition does not overcome the problem for the final Gini result, however: it postpones the moment of incomparability of values until the final step of composing the overall Gini from the group and between group-values.

The problem of stratification of Polish households based on data from the Household Budget Survey has been investigated recently by A. Jędrzejczak (2014). The essential difference of the approach presented in this paper is that it focusses on overlapping rather than on stratification and uses a different definition of the within-group term. Jędrzejczak (2014) adopted the within-group term as defined in the work of Yitzhaki and Lerman (1991), which is thus different from the within-group term as defined in the "standard" decomposition used by, for instance, C. Dagum (1997), G. Pyatt (1976) or A. F. Shorrocks (1984). Using the overlapping index rather than the stratification index enables us to represent the within-group term

in a way consistent with the standard way of decomposition and to compare the remaining parts of inequality. Moreover, while the stratification index describes the degree of stratification of a given group with respect to the whole population, the overlapping index is in a natural way a weighted sum of overlapping indexes between each pair of groups – thus, it is easy to see the mutual relationship of any pair of groups, which may be useful in identifying some social groups.

The aim of this paper is to present the Yitzhaki-Lerman methodology for calculating overlapping between different groups and to use it for Polish households of different types. Before presenting the results for empirical data, some simple theoretical examples are used to support the intuition of circumstances in which the overlapping index is low/high and symmetrical/unsymmetrical. Readers who are not interested in the mathematical details of how the formulae are obtained can go straight to Section 3 for these illustrations.

The paper is organised as follows: in the next section the notation and basic formulae are introduced. In the third section, three simple examples of distributions with a detailed presentation of stratification and overlapping indexes, along with the Gini values, are presented. The following section presents some results for Poland, and the presence of stratification with respect to types of households is examined. The final section offers a summary and conclusions.

2. Basic Formulae and Notation

In what follows, the population consisting of N members is divided into g groups, each of them having n_i members, $\sum_{i=1}^g n_i = N$. The share of the population of group i in the whole population is denoted by p_i : $p_i = \frac{n_i}{N}$, $\sum_{i=1}^g p_i = 1$.

Let us consider some quantity X , e.g. income, which is considered to be a random variable, distributed according to F_X (which will simply be denoted also as F). The probability distribution F is considered to be a mixture of g distributions characterising each of the g groups:

$$F = \sum_{i=1}^g \frac{n_i}{N} F_i. \quad (1)$$

Denoting by μ the overall mean value and by μ_i the mean value of X_i , the share of group i in the whole amount of investigated good is denoted by

$$s_i: s_i = \frac{n_i \mu_i}{N \mu} = p_i \frac{\mu_i}{\mu}, \sum_{i=1}^g s_i = 1.$$

The index of overlapping, O_{ji} , measuring the degree to which distribution j is included in the range of distribution i , is defined as (Yitzhaki & Lerman 1991):

$$O_{ji} = \frac{\text{COV}_i(F_j(X), X)}{\text{COV}_i(F_i(X), X)}, \quad (2)$$

where $\text{COV}_i(Y, X)$ denotes the covariance between random variables Y and X with respect to distribution F_i , i.e.

$$\text{COV}_i(Y, X) = \int (Y - \mu_Y)(X - \mu_X) dF_i. \quad (3)$$

From the definition (2), it follows that $O_{ii} = 1$. Also, when the distributions F_i and F_j are identical, $F_i \equiv F_j$, it holds: $O_{ji} = 1$.

The lowest value of O_{ji} is zero (as the cumulative distribution is always a non-decreasing function of X) and it is equal to zero if no part of distribution F_j lies within the range of distribution F_i .

The other limiting case is reached for the distribution of F_i being divided into two parts: all values greater than the expected value should lie beyond the range of distribution F_j , while all values smaller than the expected value should lie beneath the range of distribution F_j . That may be seen as the expression:

$$\text{COV}_i(F_j(X), X) = \int F_j \cdot (X - \mu_X) dF_i \quad (4)$$

assumes maximum value if for positive differences $X - \mu_X > 0$ the value of F_j will be as high as possible (that is, equal to 1), while for negative differences $X - \mu_X < 0$ the value of F_j will be as small as possible (that is, equal to 0).

In each case the maximum value of O_{ji} is at most 2, as (cf. Yitzhaki 1994):

$$\text{COV}_i(F_j(X), X) \leq \frac{1}{2} E(|X - \mu_X|),$$

$$\text{COV}_i(F_i(X), X) = \frac{1}{4} E(|X_i - Z_i|),$$

thus

$$O_{ji} \leq \frac{\frac{1}{2} E(|X - \mu_X|)}{\frac{1}{4} E(|X_i - Z_i|)},$$

and what follows from Jensen's inequality:

$$O_{ji} \leq \frac{2E(|X - \mu_X|)}{E(|X_i - Z_i|)} \leq 2. \quad (5)$$

However, for the case of particular distributions the maximum value might be smaller than 2. As the overlapping index O_{ji} is maximum for the distribution F_j concentrated at the mean value of distribution F_i , this maximum will depend rather on the shape of F_i than F_j (the density of the latter tending to the shape of the Dirac delta located at the mean value of distribution F_j). These more precise restrictions for the overlapping index will be presented for three particular distributions in what follows.

The overall overlap for group i is a sum of all indexes O_{ji} , that is:

$$O_i = \sum_{j=1}^g p_j O_{ji} = p_i + \sum_{\substack{j=1 \\ j \neq i}}^g p_j O_{ji}, \quad (6)$$

as $O_{ii} = 1$.

The concept related to the overlapping index is the index of stratification of group i , denoted by Q_i . It is defined as:

$$Q_i = \frac{\text{COV}_i(F_i(X) - F_{ni}(X), X)}{\text{COV}_i(F_i(X), X)} = 1 - \frac{\text{COV}_i(F_{ni}(X), X)}{\text{COV}_i(F_i(X), X)}, \quad (7)$$

where F_{ni} denotes the cumulative distribution function of the whole population apart from group i , that is:

$$(1 - p_i) F_{ni} = F - p_i F_i, \quad (8)$$

(which follows from $F = p_i F_i + F - p_i F_i = (1 - p_i) F_{ni}$).

Comparing (2), (6) and (7) it is easy to express the overlapping index in terms of the stratification index and vice versa. As:

$$\begin{aligned} O_i &= \frac{\sum_{j=1}^g \text{COV}_i(F_j(X), X)}{\text{COV}_i(F_i(X), X)} = \frac{\text{COV}_i(F(X), X)}{\text{COV}_i(F_i(X), X)} = \\ &= p_i + \frac{(1 - p_i) \text{COV}_i(F_{ni}(X), X)}{\text{COV}_i(F_i(X), X)}, \end{aligned} \quad (9)$$

using (8) one obtains:

$$O_i = p_i + (1 - p_i)(1 - Q_i) = 1 - Q_i(1 - p_i). \quad (10)$$

Now we may proceed to decompose the overall Gini index for the whole population with respect to the Gini indexes for particular groups and two other terms.

The Gini index for the whole population reads:

$$G = \frac{2}{\bar{\mu}} \text{COV}(F, X) \quad (11)$$

while for the i -th group:

$$G_i = \frac{2}{\bar{\mu}_i} \text{COV}_i(F_i, X). \quad (12)$$

Using (1), covariance in (11) may be re-expressed as follows:

$$\text{COV}(F, X) = \int (x - \mu)F(x)dF = \sum_{i=1}^g p_i \int (x - \mu)F(x)dF_i. \quad (13)$$

Subtracting and adding μ_i one gets:

$$\text{COV}(F, X) = \sum_{i=1}^g \left[p_i \int (x - \mu_i)F(x)dF_i + \int (\mu_i - \mu)F(x)dF_i \right]. \quad (14)$$

The first term is a sum of within-group covariances:

$$\sum_{i=1}^g p_i \int (x - \mu_i)F(x)dF_i = \sum_{i=1}^g p_i \text{COV}_i(F, X). \quad (15)$$

The second term is the covariance between the g -tuple of groups' means and the g -tuple of average values of cumulative distribution for particular groups:

$$\begin{aligned} \sum_{i=1}^g \int (\mu_i - \mu)F(x)dF_i &= \sum_{i=1}^g (\mu_i - \mu) \int F(x)dF_i \equiv \\ &\equiv \sum_{i=1}^g (\mu_i - \mu) \bar{F}_i = \text{COV}(\mu_i, \bar{F}_i). \end{aligned} \quad (16)$$

Altogether, the Gini index (11) may be rewritten as:

$$G = \frac{2}{\bar{\mu}} \sum_{i=1}^g p_i \text{COV}_i(F, X) + \frac{2}{\bar{\mu}} \text{COV}(\mu_i, \bar{F}_i). \quad (17)$$

Notice that $\text{COV}_i(F, X)$ is not the one that would appear in the formula for G_i , as cumulative distribution here is measured in the general population, not within the group. To reach the within-groups cumulative distribution, one can use (8) and re-express:

$$\text{COV}_i(F, X) = p_i \text{COV}_i(F_i, X) + (1 - p_i) \text{COV}_i(F_{ni}, X). \quad (18)$$

Expression $\text{COV}_i(F_i, X)$ is related to G_i by:

$$G_i = \frac{2}{\bar{\mu}_i} \text{COV}_i(F_i, X). \quad (19)$$

On the other hand, $\text{COV}_i(F_{ni}, X)$ may be expressed in terms of either the stratification index or the overlapping index, as:

$$\text{COV}_i(F_{ni}, X) = (1 - Q_i) \text{COV}_i(F_i, X) = \frac{\mu_i}{2} (1 - Q_i) G_i, \quad (20)$$

or:

$$\text{COV}_i(F_{ni}, X) = \frac{\mu_i}{2} \left(\frac{O_i - p_i}{1 - p_i} \right) G_i. \quad (21)$$

Thus:

$$G = \frac{1}{\mu} \sum_{i=1}^g p_i^2 \mu_i G_i + \frac{1}{\mu} \sum_{i=1}^g p_i \mu_i (1 - Q_i) (1 - p_i) G_i + \frac{2}{\mu} \text{COV}(\mu_i, \bar{F}_i), \quad (22)$$

or, after rearranging,

$$G = \sum_{i=1}^g s_i G_i + \sum_{i=1}^g s_i G_i Q_i (p_i - 1) + \frac{2}{\mu} \text{COV}(\mu_i, \bar{F}_i), \quad (23)$$

and

$$G = \sum_{i=1}^g p_i^2 \frac{\mu_i}{\mu} G_i + \sum_{i=1}^g p_i \frac{\mu_i}{\mu} (O_i - p_i) G_i + \frac{2}{\mu} \text{COV}(\mu_i, \bar{F}_i), \quad (24)$$

or:

$$G = \sum_{i=1}^g p_i s_i G_i + \sum_{i=1}^g s_i (O_i - p_i) G_i + \frac{2}{\mu} \text{COV}(\mu_i, \bar{F}_i). \quad (25)$$

The three terms are correspondingly within-group Gini, overlap-due Gini, and between-group Gini indexes:

$$G_W \equiv \sum_{i=1}^g p_i s_i G_i, \quad (26)$$

$$G_O \equiv \sum_{i=1}^g s_i (O_i - p_i) G_i, \quad (27)$$

$$G_B \equiv \frac{2}{\mu} \text{COV}(\mu_i, \bar{F}_i). \quad (28)$$

For the discrete case, the counterpart of the cumulative distribution function are the normalised ranks.

3. Three Simple Illustrative Examples

3.1. Preliminary Remarks

The relationship between O_{ij} and O_{ji} is far from obvious. O_{ji} measures the degree to which the distribution j is included within the range of the

distribution i . O_{ij} measures the degree to which the distribution i is included within the range of the distribution j . Thus, it might be suspected that the greater one of them is, the smaller is the other one. This would be indeed the case in some circumstances, e.g. the fixed central tendencies of both distributions and changing dispersions. However, under other circumstances, e.g. fixed dispersions and changing central tendencies, both O_{ij} and O_{ji} will change accordingly, simultaneously increasing or decreasing. Let us visualise these possible effects with simple examples of two distributions of the same kind.

In the three following examples it will be assumed that we are dealing with two subpopulations with equal numbers of members, that is, the share of each of these subpopulations is equal to 0.5.

For the case of two subpopulations with equal numbers of members, formula (25) becomes:

$$G = \frac{1}{(n_1 + n_2)(n_1\mu_1 + n_2\mu_2)} [n_1^2\mu_1 G_1 + n_2^2\mu_2 G_2] + \frac{n_1 n_2}{(n_1 + n_2)(n_1\mu_1 + n_2\mu_2)} [\mu_1 G_1 O_{21} + \mu_2 G_2 O_{12}] + \frac{2}{(n_1\mu_1 + n_2\mu_2)} \frac{n_2 n_1 (\mu_1 - \mu_2)}{(n_1 + n_2)} (\bar{F}_1 - \bar{F}_2), \quad (29)$$

which for an equal number of individuals in both subpopulations, $n_1 = n_2 = n$, $N = 2n$, reduces further:

$$G = \frac{1}{2(\mu_1 + \mu_2)} [\mu_1 G_1 + \mu_2 G_2] + \frac{1}{2(\mu_1 + \mu_2)} [\mu_1 G_1 O_{21} + \mu_2 G_2 O_{12}] + \frac{\mu_1 - \mu_2}{\mu_1 + \mu_2} (\bar{F}_1 - \bar{F}_2). \quad (30)$$

In what follows, three simple examples will be examined: uniform distribution, normal distribution, and Weibull distribution. Although the main subject of interest in the context of inequality and stratification is income and related quantities which are known to be far from symmetrically distributed, investigating uniform and normal distributions enables us to gain some intuition of overlapping and stratifications, which are concepts not as familiar as inequality and the Gini index itself.

3.2. Uniform Distribution

Let us start with the simplest example: two uniform distributions with the same median value, m . One of the distributions ranges over a constant

interval, $\langle a_1, b_1 \rangle \left(\frac{a_1 + b_1}{2} = m \right)$, while the other one starts from the wider interval $\langle a_2, b_2 \rangle$ with $a_2 < a_1$, $b_2 > b_1$ (and $\frac{a_2 + b_2}{2} = m$, $a_1 - a_2 = b_2 - b_1$), gradually increasing a_2 through $a_2 = a_1$ and $a_2 > a_1$ while simultaneously decreasing b_2 in a symmetrical way, through $b_2 = b_1$ and $b_2 < b_1$. Symmetry in the change of the second uniform distribution means that $a_2 = a_1 + x$ and $b_2 = b_1 - x$. For illustration, three graphs are presented in Figure 1, one for a basic distribution defined on the interval $\langle 1000, 3000 \rangle$ (corresponding to the case when $x = 0$), and two additional graphs for $x = 500$ and $x = -500$.

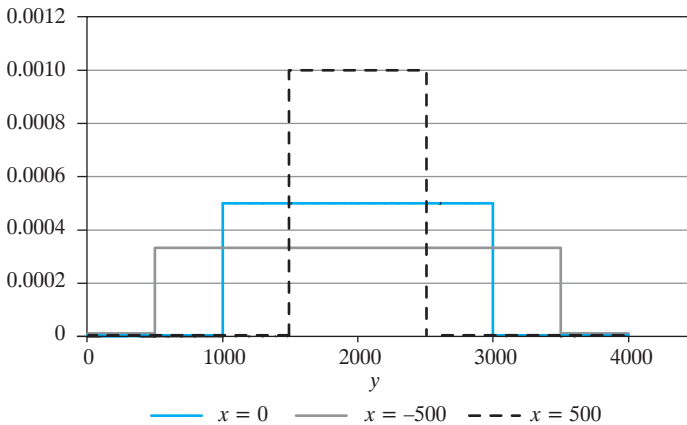


Fig. 1. Uniform Distributions on the Interval $\langle 1000 + x; 3000 - x \rangle$

Source: elaborated by the author.

The relationship between O_{12} and O_{21} is in this case intuitively predictable: the wider the interval $\langle a_2, b_2 \rangle$ as compared to the interval $\langle a_1, b_1 \rangle$ (and thus the smaller fraction of its range is included within the range of the latter) the smaller is O_{21} . Thus, O_{21} is an increasing function of x , increasing linearly with respect to $y = \frac{(b_1 - a_1)}{(b_2 - a_2)}$ up to $y = 1$, then still increasing but in a nonlinear way. On the other hand, O_{12} decreases with y , and from $y = 1$ it is equal to $1/y$ (as it is equal to the fraction of the range of $\langle a_2, b_2 \rangle$ included within $\langle a_1, b_1 \rangle$ – that is, equal to $\frac{(b_2 - a_2)}{(b_1 - a_1)}$). The explicit analytical form of O_{21} is as follows:

$$O_{21}^U = \begin{cases} \frac{b_1 - a_1}{b_1 - a_1 - 2x} & \text{for } x < 0 \\ 1 + \frac{2x(b_1 - a_1 - x)}{(b_1 - a_1)^2} & \text{for } x \geq 0 \end{cases} \quad (31)$$

These are presented in Figure 2 (as function of x).

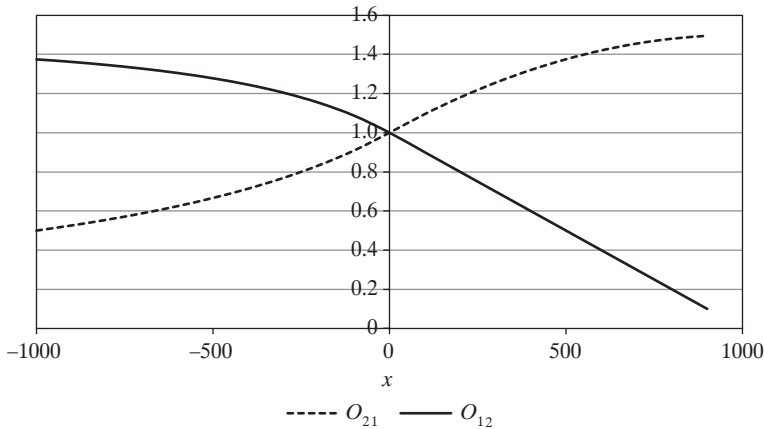


Fig. 2. Graphs of O_{12} and O_{21}

Source: author's own calculations.

As mentioned before, O_{ji} is bounded from above by 2; however, in particular cases yet more precise limitations may be fixed. For the uniform distribution case, even if the other distribution assumed the form of the Dirac delta centred in the expected value of the uniform distribution, it may be easily shown that the overlapping index will not exceed:

$$O_{U,21}^{\max} = 1.5, \quad (32)$$

as for x approaching $\frac{b_1 - a_1}{2}$ (second distribution concentrated at the average value of the first one), (31) assumes the value 1.5. This may be also noticed in Figure 2.

As the dispersion of one of the distributions lowers, the Gini index for the sum of the distributions also decreases with increasing y . Simultaneously, the share of within Gini in the overall inequality increases (although the absolute value of the within Gini index decreases), while – what follows – the

share of the overlap part of the Gini index increases (as the average ranks of both distributions are equal to 0.5, the between Gini is equal to zero here) – cf. Figure 3.

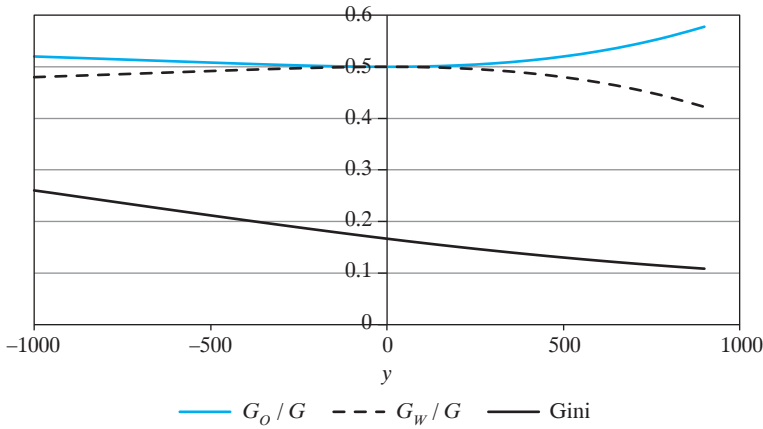


Fig. 3. Graphs of G , G_W/G , and G_O/G

Source: author’s own calculations.

One may also consider two uniform distributions over the intervals of identical width, shifted one with respect to the other: $\langle a_1, b_1 \rangle$ and $\langle a_2, b_2 \rangle$, $a_2 = a_1 + x, b_2 = b_1 + x$ (see Figure 4, with $a_1 = 1000, b_1 = 3000$).

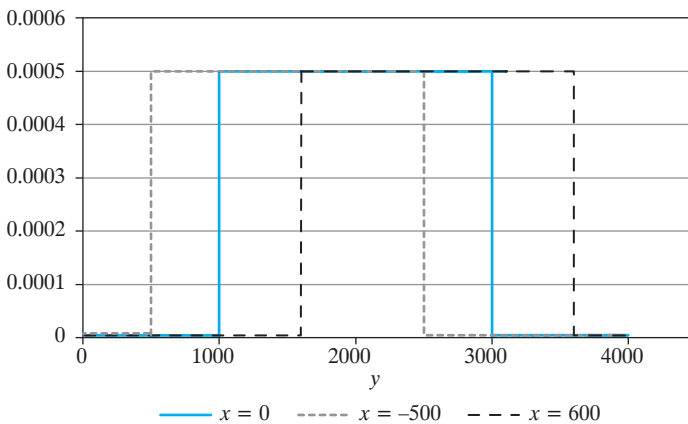


Fig. 4. Uniform Distributions on the Interval $\langle 1000 + x; 3000 + x \rangle$

Source: elaborated by the author.

In this case the overlap indexes will be obviously symmetrical, that is, $O_{12} = O_{21}$, and have the following analytical form:

$$O_{U,21} = \begin{cases} \frac{a_1^3 - 3a_1(b_1^2 + x^2) + 2(b_1^3 - x^3)}{(b_1 - a_1)^2(a_1 + 2b_1)} & \text{for } |x| < b_1 - a_1 \\ 0 & \text{otherwise} \end{cases} \quad (33)$$

If one fixes one of the distributions, shifting the other one towards higher averages, $O_{12} = O_{21}$ will reach the maximum value of 1 for $x = 0$ (see Figure 5).

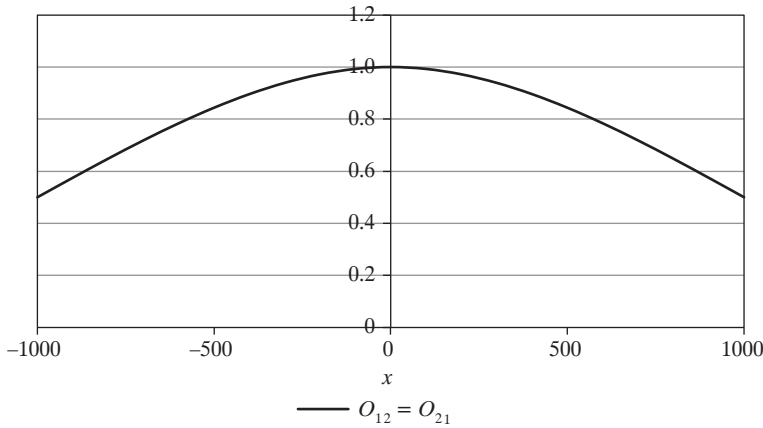


Fig. 5. Graphs of $O_{12} = O_{21}$ for Two Uniform Distributions for $x \in \langle -1000; 1000 \rangle$.

Source: author's own calculations.

The Gini index for the first distribution will be fixed and for the other one will decrease (as the mean value increases). Thus, the within Gini will decrease. The overlap Gini will reach its maximum for $x = 0$, and the between Gini its minimum for the same value. In Figure 6, the functions of x : G_W, G_O, G_B are presented for $x \in \langle -1000; 1000 \rangle$.

To conclude this simplest example: Q_{12} may change inversely or accordingly to O_{21} depending on whether the change in the relative positions of the probability density functions of two distributions results from the change of dispersion of distributions (or one of them) or rather from the relation between their central tendencies. On the other hand,

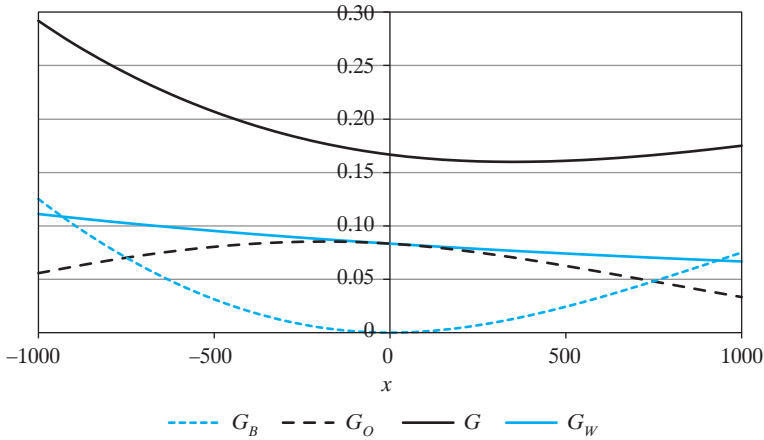


Fig. 6. Graphs of G , G_w , G_O and G_B for $x \in \langle -1000; 1000 \rangle$.

Source: author's own calculations.

the relationship between the overlapping and the share of Gini due to overlapping is not straightforward, as the overlapping indexes are weighted by the group becoming a part of the overall Gini index.

3.3. Normal Distribution

Let us say we have two normally distributed populations with an equal number of members. The distribution of the first population is $N(\mu_1, \sigma_1)$ with the within-group Gini index G_1 , while the second one is distributed according to $N(\mu_2, \sigma_2)$ with the within-group Gini index G_2 . Starting from the case of $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$ we have two identical populations, and, of course, $G_1 = G_2$. Then we change the standard deviation of the second group, keeping its mean value constant.

For $\mu_1 = \mu_2$ formula (23) reduces to the simple form:

$$G = \frac{1}{4}[G_1 + G_2] + \frac{1}{4}[G_1 O_{21} + G_2 O_{12}]. \quad (34)$$

It may be shown, moreover (Schechtman 2005), that for two normal distributions $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$ the overlapping index is equal to:

$$O_{N,21} = \frac{\sqrt{2} \sigma_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \exp \left[\frac{-(\mu_2 - \mu_1)^2}{2(\sigma_1^2 + \sigma_2^2)} \right], \quad (35)$$

with the maximum possible value again lower than 2, namely:

$$O_{N,21}^{\max} = \sqrt{2} \approx 1.414. \quad (36)$$

Figure 7 illustrates Gini indexes and overlapping indexes for the exemplary distributions, for which calculations have been performed. $\mu_1 = \mu_2 = 4000$, $\sigma_1 = 1000$ while σ_2 changes from $\sigma_2 = 100$ to $\sigma_2 = 5000$. Figure 7 presents O_{21} (the degree to which distribution 2 is included in distribution 1) and O_{12} (the degree to which distribution 1 is included in distribution 2).

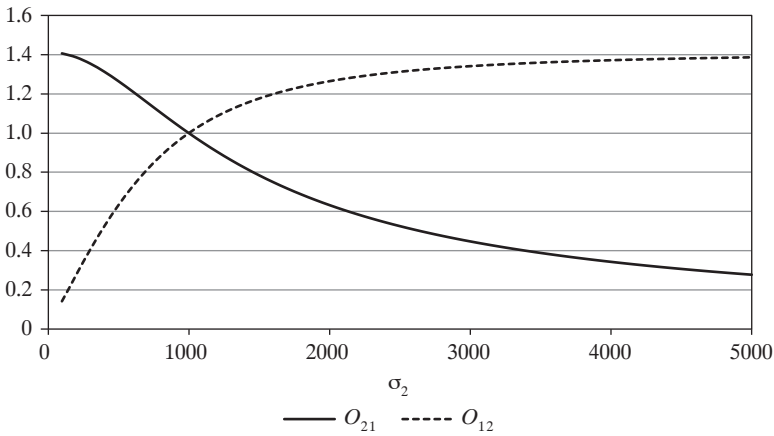


Fig. 7. Graphs of O_{12} and O_{21} for Two Normal Distributions, $\mu_1 = \mu_2 = 4000$, $\sigma_1 = 1000$ while σ_2 Changes from $\sigma_2 = 100$ to $\sigma_2 = 5000$

Source: author's own calculations.

Figure 8 presents the shares of within-group inequality and the overlapping-due term together with the overall Gini index, G . As $\mu_1 = \mu_2$, the between-group term is of course zero.

For this example O_{21} and O_{12} are inversely related, as the greater part of distribution 2 is included inside distribution 1, and the smaller part of distribution 1 is included inside distribution 2.

On the other hand, if we change the mean value of distribution 2 (from $\mu_2 = 5000$ to 20,000, while $\mu_1 = 10,000$) keeping the standard deviations constant (and equal, $\sigma_1 = \sigma_2 = 1000$), O_{21} and O_{12} will be equal (as the overlapping is symmetrical). G_1 is constant, while G_2 monotonously

decreases with increasing mean value (the Gini index decreases if a positive constant is added to each value of the distribution).

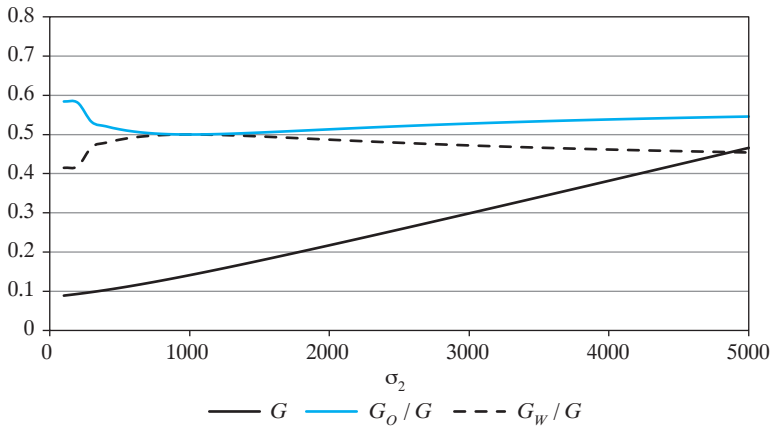


Fig. 8. Graphs of G , G_W/G , and G_O/G

Source: author's own calculations.

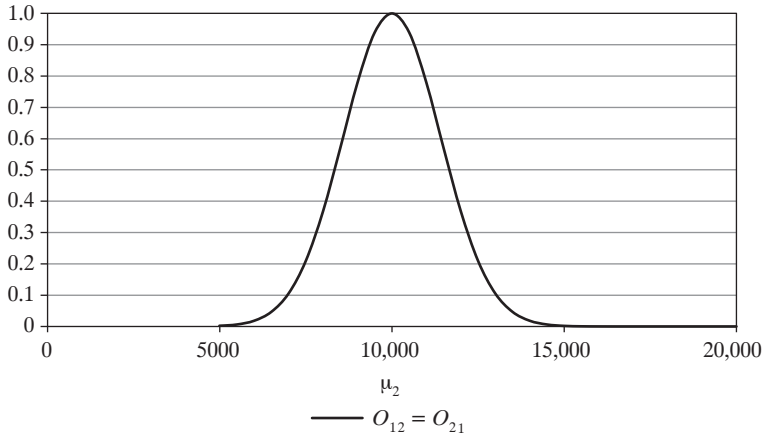


Fig. 9. Graphs of $O_{12} = O_{21}$ for Two Normal Distributions, $\sigma_1 = \sigma_2 = 1000$, $\mu_1 = 10,000$, while μ_2 changes from $\mu_2 = 5000$ to $\mu_2 = 20,000$

Source: author's own calculations.

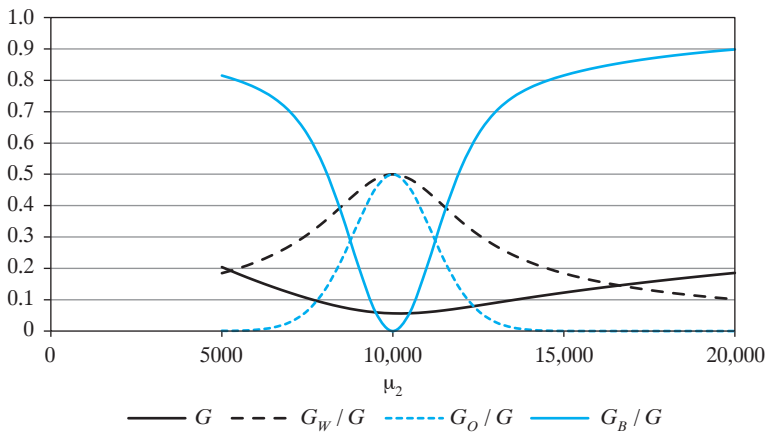


Fig. 10. Graphs of G , $G_W(x)/G$, $G_O(x)/G$ and $G_B(x)/G$

Source: author's own calculations.

Again, it may be concluded that the relationship between O_{12} and O_{21} may be either inverse or strict. While in the former case the share of the overlapping term within the overall Gini is not in a straightforward dependence on the overlapping indexes, in the latter, while they are the same, the behaviour of the overlapping share in overall inequality is qualitatively in accordance with the behaviour of the overlapping indexes.

3.4. Weibull Distribution

Let us now consider two Weibull distributions describing two populations consisting of the same number of individuals. Let us further fix the parameters of one of the distributions as $k_1 = 5$, $\lambda_1 = 5000$ (and thus $\mu_1 = 4590.84$). The behaviour of overlapping indexes is investigated in two cases: 1) if one requires the overall mean to be fixed while changing the inequality of the second distribution 2) keeping the inequalities of both distributions fixed while changing their relative positioning.

Although it is not possible to obtain a simple analytical expression for the overlapping index for the case of the Weibull distribution, one may obtain the maximum value of it. Let us recall that for any case this index cannot exceed the value of 2. However, for particular distributions there might be further restrictions, lowering this maximum. For the case of the Weibull distribution the maximum value of the overlapping index is equal to:

$$O_W^{\max} = \frac{4 \exp[-\Gamma^k((k+1)/k)] [-\Gamma((k+1)/k) + \exp[\Gamma^k((k+1)/k)] \Gamma((k+1)/k, \Gamma^k((k+1)/k))]}{(2 - 2^{(k-1)/k}) \Gamma((k+1)/k)}, \quad (36)$$

which depends on the shape parameter of the Weibull distribution. It starts from 2 for $k \rightarrow 0$, assumes a minimum value of 1.41823 near $k = 49$, tending to 1.41827 with $k \rightarrow \infty$. For $1.5 < k < 2.5$ (ranges of values for fitting the considered data) the maximum of overlapping assumes values between 1.43 and 1.44 (which is close to the maximum overlapping for the normal distribution, $\sqrt{2}$).

Figure 11 presents overlapping terms for the changing shape of one of the Weibull distributions, $k_1 = 5$, $\lambda_1 = 5000$ while k_2 from 2 to 8 (adjusting λ_2 as to keep $\mu_2 = 4590,84$).

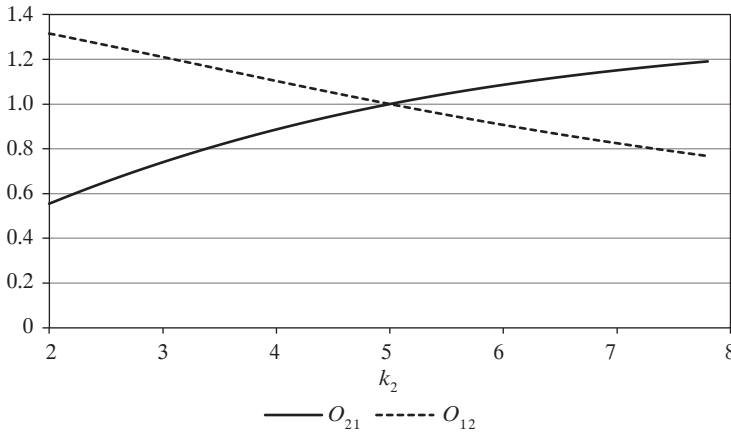


Fig. 11. Graphs of O_{12} and O_{21} for Weibull Distributions: $k_1 = 5$, $\lambda_1 = 5000$ while k_2 changes from $k_2 = 2$ to $k_2 = 8$ (Adjusting λ_2 as to Keep $\mu_2 = 4590,84$)

Source: author's own calculations.

Figure 12 presents the share of within group inequality and the share of the overlapping-due term together with the overall Gini index, G .

On the other hand, if we keep the shapes (i.e. Gini indexes) of both of distributions unchanged, changing only their relative positions, the overlapping indexes change, roughly speaking, in the same direction – as the Weibull distribution is not symmetrical, they are not the same (see Figure 13).

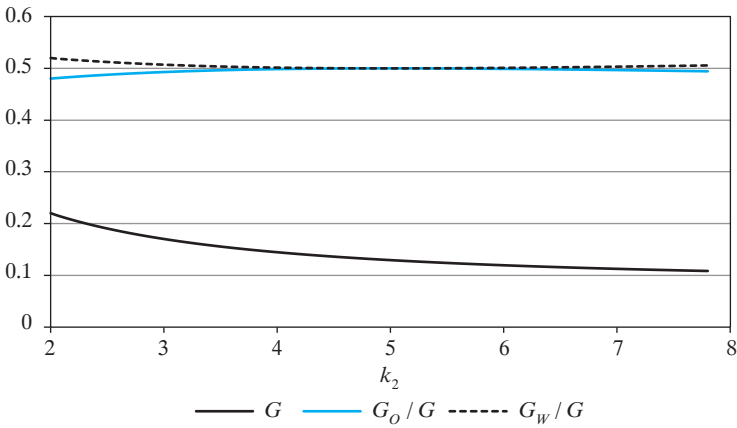


Fig. 12. Graphs of G , G_W/G , and G_O/G

Source: author's own calculations.

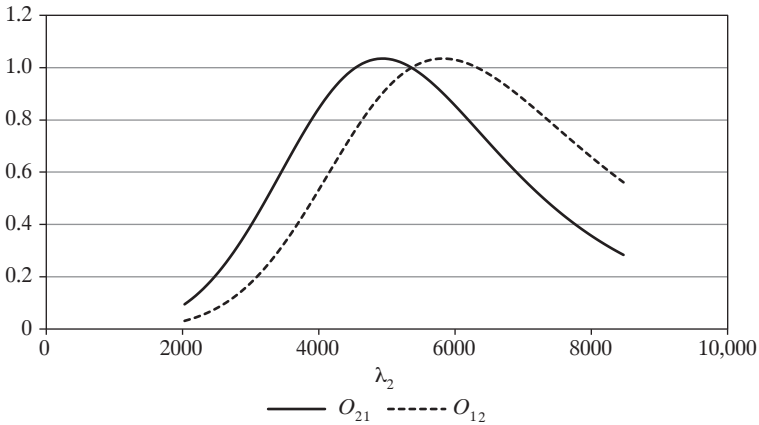


Fig. 13. Graphs of O_{12} and O_{21} for Weibull distributions: $k_1 = 5$, $\lambda_1 = 5000$ while λ_2 Changing from 2000 to 8000 (Adjusting k_2 as to Keep $G_2 = G_1 = 0.1295$)

Source: author's own calculations.

Still, the qualitative changes in the share of overlap-due input to inequality behaves in accordance with the behaviour of overlapping indexes (see Figure 14).

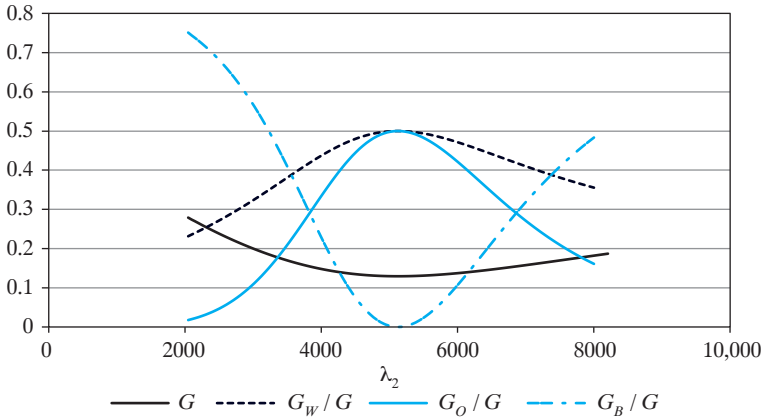


Fig. 14. Graphs of G , $G_W(x)/G$, $G_O(x)/G$ and $G_B(x)/G$

Source: author's own calculations.

4. Results for Poland

The data investigated in this section comes from the Household Budget Survey for 2011 (HBS 2011) and consists of 37,167 households of different (and specified below) types from all Polish regions.

Although it is known that some areas of Poland are richer than others, one would not expect provinces to be overly stratified. Moreover, it would be more interesting to look at the potential stratification of different types of households as this would throw light on the situation of different parts of Polish society and their input into overall inequality in Poland.

To this end, all households were divided into 22 groups. Of these, 21 consisted of households of the same type, while the final 22nd group was composed of all the remaining households, which in total comprises 1.1% of households. The OECD equivalence scale has been adopted, i.e. the first adult person is the reference point and is equal to 1, all other adults (that is, over 14 years of age) are counted as 0.7, while each child below 24 is equivalent to 0.5 of a single adult. Table 1 presents the number of adults and number of children within each type of household, its equivalence scale, and the number of households of that type, together with the average equivalent value for the equivalent adult person in the household. The quantities that will be considered here are: incomes, expenditures, and the difference between incomes and expenditures, i.e. the “money balance” (which may also assume negative values).

Table 1. Summary Statistics

Adults	Children	Households	Equivalence Scale	\bar{x}_{in}	\bar{x}_{sp}	\bar{x}_b	G_{in}	G_{ex}	G_b
1	0	6619	1.0	1804.70	1633.33	171.37	0.32	0.32	2.51
1	1	314	1.5	1741.60	1610.15	131.45	0.34	0.33	2.89
2	0	10,841	1.7	1924.86	1592.74	332.12	0.30	0.31	1.35
1	2	126	2.0	1329.76	1371.48	-41.71	0.32	0.35	-8.01
2	1	3007	2.2	1941.33	1558.29	383.04	0.32	0.30	1.22
3	0	4779	2.4	1635.16	1309.12	326.04	0.29	0.29	1.12
2	2	2601	2.7	1624.2	1334.48	289.72	0.33	0.31	1.39
3	1	1649	2.9	1427.37	1169.52	257.85	0.30	0.29	1.25
4	0	2458	3.1	1442.24	1138.32	303.92	0.28	0.28	1.09
2	3	553	3.2	1200.85	1034.8	166.06	0.33	0.30	1.94
3	2	641	3.4	1267.72	973.81	293.90	0.32	0.26	1.19
4	1	928	3.6	1287.04	1006.33	280.71	0.29	0.27	1.15
2	4	107	3.7	935.479	844.75	90.73	0.27	0.28	1.93
5	0	713	3.8	1283.85	1008.99	274.86	0.28	0.28	1.17
3	3	132	3.9	1047.26	864.00	183.26	0.27	0.29	1.28
4	2	410	4.1	1229.31	897.25	332.05	0.34	0.26	1.15
5	1	327	4.3	1140.65	897.60	243.06	0.27	0.24	1.09
6	0	197	4.5	1203.67	873.79	329.88	0.30	0.27	1.09
4	3	99	4.6	1165.07	755.12	409.95	0.38	0.26	0.99
5	2	137	4.8	1058.21	813.39	244.82	0.27	0.25	1.01
6	1	112	5.0	1041.5	862.32	179.18	0.26	0.29	1.48
Remaining		417		941.69	745.13	196.56	0.30	0.26	1.18

Source: author's own calculations based on (Central Statistical Office 2011).

The overall Gini index for incomes equals 0.321, for expenditures – 0.320, and for the money balance – 1.321.

It is known that while negative values are present, the Gini index is no longer restricted to the range $\langle 0, 1 \rangle$ but may exceed 1 as well as being negative. There are some methods to deal with this problem, e.g. to standardise this value in a proper way (see, for instance, Raffinetti, Siletti & Vernizzi 2015; Ostasiewicz & Vernizzi 2017). Here, however, as the decomposition is what is being investigated, there will simply be presented the shares of each of the terms of the decomposition within the overall inequality, and with respect to the shares the non-standardised character of inequality does not matter.

Although it is slightly outside the scope of this paper, let us briefly comment on a discrepancy between the inequality of these particular components (incomes and expenditure) and the inequality of the money balance. According to the standard decomposition of inequality with respect to different sources, overall inequality can be composed of inequalities from these two sources according to:

$$G = \frac{1}{\bar{x}} (\bar{x}_{in} R_{in} G_{in} + \bar{x}_{ex} R_{ex} G_{ex}),$$

where \bar{x} , \bar{x}_{in} and \bar{x}_{ex} are the average values of the money balance, incomes, and expenditures respectively, and R_{in} and R_{ex} are Gini correlations of incomes and expenditures, defined as:

$$R_q = \frac{\text{COV}(x_q, F_b(x_q))}{\text{COV}(x_q, F_q(x_q))},$$

where q stands for a quantity being either incomes or expenditures, x_q denotes its value, $F_b(x_q)$ – the value of the cumulative distribution of the “money balance” for the value of the balance corresponding to a particular value of quantity q and $F_q(x_q)$ – the cumulative distribution of quantity q .

Gini correlations, like conventional correlation coefficients, range between -1 and 1 and measure the degree to which the order of incomes/expenditures is in accordance with the order of the money balance. To estimate very roughly the order of values which overall inequality may acquire, let us assume that the inequalities of both incomes and expenditures are the same (which is lightly deviated in the real data), $G_{in} \approx |G_{ex}|$. Denoting by η the average fraction of incomes that are spent, $\bar{x}_{ex} = -\eta \bar{x}_{in}$ (as expenditures are taken as negative values), one gets:

$$G = \frac{\bar{x}_{in} G_{in}}{\bar{x}_{in} (1 - \eta)} (R_{in} - \eta R_{ex}),$$

where the sign between the terms in the bracket is “minus” as G_{ex} has a negative sign.

The above expression will assume the maximum value for $R_{in} = 1$ and $R_{ex} = -1$, and this maximum value will be:

$$G_{\max} = \frac{G_{in} (1 + \eta)}{(1 - \eta)},$$

which can take arbitrarily large values, depending on η .

As for the Polish data considered here $\eta \approx 0.8$, G could almost assume the value of 3 – thus the value of 1.32 does not seem high.

On the other hand, if both components were positive and both component Gini indexes were positive, the maximum possible value would be

$$G = \frac{\bar{x}_{in} G_{in}}{\bar{x}_{in} (1 + \eta)} (R_{in} + \eta R_{ex}) = G_{in} \text{ (which would be a result for } R_{in} = R_{ex} = 1),$$

thus, for positive components the overall Gini index cannot exceed the larger of the two Gini indexes from which the overall value is composed.

For the data examined here one has: $R_{in} = 0.585324$, $R_{ex} = -0.201955$, and what follows:

$$\begin{aligned} G &= \frac{1}{292.343} (1603.97 \cdot 0.585324 \cdot 0.321125 - (1311.63) \cdot (-0.201955) \cdot 0.319525) = \\ &= 1.03127 + 0.289518 = 1.32079, \end{aligned}$$

thus, it is far from being maximum.

Each of the Gini indexes, G_{in} , G_{ex} and G , may be decomposed into a within-groups part, between-groups part, and the term resulting from overlapping. The results are presented in Table 2.

While the share of within-groups inequality in overall inequality is more or less the same for all three quantities (incomes, expenditures, and balance) and is about 11%, there is a huge difference in the share of between-group inequalities between incomes and expenditures on the one hand and the money balance on the other – of magnitude 10.

The component due to overlapping may be represented as the weighted sum of overlapping indexes between each pair of groups. These overlapping indexes between particular pairs, O_{ij} and O_{ji} , are presented in Tables A1–A3 in the Appendix.

Table 2. Results of the Gini Index and Its Component for Incomes, Expenditures and Money Balance

Gini	Incomes		Expenditures		Money balance	
	0.321158		0.319525		1.320791	
	Value	Share	Value	Share	Value	Share
Within-group	0.035715	0.111218	0.036328	0.113692	0.151776	0.114913
Between-group	0.035741	0.111298	0.042892	0.134237	0.013563	0.010269
Overlapping	0.249703	0.777506	0.240305	0.752071	1.155452	0.874818

Source: author's own calculations.

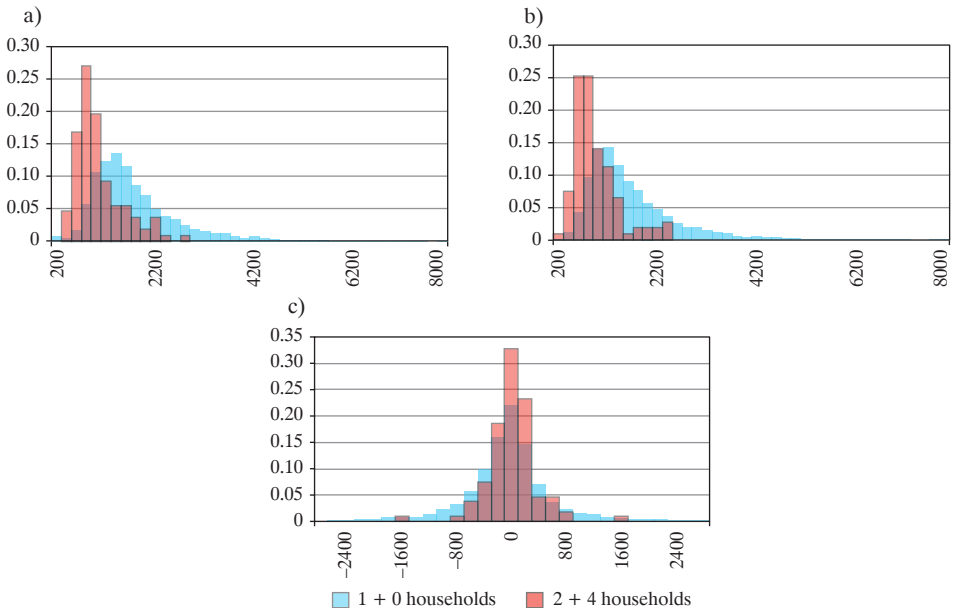


Fig. 15. Histograms for 1 + 0 and 2 + 4 Households for a) Incomes, b) Expenditures, c) Money Balance.

Source: elaborated by the author.

The values of O_{ij} and O_{ji} summarise numerically what might be visualised in the form of histograms. Let us picture the distributions of incomes, expenditures and balance for 1 + 0 and 2 + 4 households. As $O_{21} = 0.56$, one may expect that two histograms will be to a visible degree separated one from another, which can indeed be noticed in Figure 15a. Likewise for expenditures: $O_{21} = 0.57$ and that is again justified by the visualisation (see Figure 15b). Proceeding to the result for the overlapping of the

money balance, it changes significantly to $O_{21} = 1.12$, which may again be corroborated by the visual estimation, cf. Figure 15c.

A similar effect might be observed for many pairs of distributions for the investigated data, hence the noticeable change in the share of the overlapping-due term in overall inequality when passing from incomes/expenditures to the money balance.

5. Discussion and Conclusions

What may be observed, for most types of households, is that the inequalities of incomes are higher than those of expenditures. This seems intuitive as some basic needs have to be satisfied regardless of whether this leads to debt. One of the few exceptions is the 1 + 2 type of household. It is also the one with the worst balance (negative) and huge inequality of this balance (with respect to the absolute value, but with a negative sign). The negative sign of the Gini index is of course the result of the negative average value (as the sum of absolute differences is always positive). Although the sample size is not very large (126 households), the situation of single parents with two children is obviously worth further investigation.

The merit of calculating overlapping indexes for particular pairs of groups and for the overlapping of a single group by all the others, is the simplicity of such a presentation as compared with the (imprecise, in addition) visual estimation of the overlapping of two distributions, which is far more advanced information than merely comparing the average values for two groups. Comparing the situations of different types of households in Poland is beyond the scope of this paper and rather a task for sociologists. However, the tool of the set of overlapping indexes seems to be very helpful.

The between-group term within the Yitzhaki-Lerman decomposition has a share about ten times less for the money balance than for incomes and expenditures. On the other hand, within the Dagum decomposition, the corresponding between-group term for balance has a share of 9%, *versus* 32% and 36% for incomes and expenditures, respectively. Thus, the difference is not so huge, less than four times. It is known that the between-group term for the Dagum decomposition in the presence of non-zero overlapping is always greater than the between-group term within the scope of the Yitzhaki-Lerman method. Thus, one may ask what the “real” between-group inequality is, or rather, how to make the best use of both pieces of information available, in which situations they differ the most, and in which situations it is more illustrative to use one rather than the other, if it is not possible to complement both. This will be the subject of a future study.

Appendix

Table A1. Overlapping Indexes O_{ji} for Incomes for Households of Different Equivalence Scales

$j \backslash i$	1.00	1.50	1.70	2.00	2.20	2.40	2.70	2.90	3.10	3.20	3.40	3.60	3.70	3.80	3.90	4.10	4.30	4.50	4.60	4.80	5.00	R
1.00	1.00	1.05	0.94	1.10	0.98	1.00	1.07	1.02	1.01	1.07	1.05	1.01	0.83	0.98	0.91	1.09	0.91	1.01	1.14	0.92	0.82	0.91
1.50	0.94	1.00	0.89	1.05	0.93	0.95	1.02	0.97	0.96	1.03	1.01	0.96	0.82	0.94	0.89	1.05	0.89	0.97	1.10	0.89	0.81	0.90
1.70	1.05	1.09	1.00	1.08	1.04	1.02	1.09	1.02	1.00	1.03	1.02	0.98	0.73	0.94	0.83	1.07	0.85	0.97	1.12	0.85	0.74	0.84
2.00	0.77	0.87	0.72	1.00	0.77	0.84	0.91	0.91	0.90	1.03	0.97	0.93	0.97	0.93	0.98	1.01	0.95	0.98	1.10	0.95	0.91	1.01
2.20	1.01	1.05	0.96	1.05	1.00	0.97	1.05	0.98	0.96	1.01	1.00	0.95	0.73	0.91	0.81	1.04	0.83	0.94	1.10	0.84	0.73	0.83
2.40	1.00	1.05	0.94	1.10	0.98	1.00	1.07	1.02	1.01	1.07	1.04	1.00	0.84	0.98	0.92	1.08	0.91	1.01	1.13	0.92	0.83	0.92
2.70	0.92	0.98	0.86	1.04	0.90	0.93	1.00	0.96	0.95	1.03	1.00	0.96	0.84	0.94	0.90	1.04	0.90	0.97	1.11	0.90	0.82	0.91
2.90	0.90	0.97	0.84	1.09	0.88	0.94	1.02	1.00	0.99	1.09	1.04	1.00	0.95	1.00	0.99	1.08	0.98	1.03	1.14	0.98	0.92	0.99
3.10	0.92	0.99	0.86	1.09	0.91	0.96	1.03	1.01	1.00	1.08	1.04	1.00	0.92	1.00	0.97	1.08	0.97	1.02	1.14	0.97	0.90	0.97
3.20	0.70	0.80	0.65	0.95	0.70	0.78	0.85	0.86	0.86	1.00	0.93	0.89	0.98	0.90	0.97	0.97	0.94	0.95	1.06	0.94	0.93	1.01
3.40	0.77	0.87	0.71	1.03	0.76	0.85	0.92	0.93	0.93	1.06	1.00	0.96	1.01	0.97	1.02	1.04	1.00	1.01	1.11	0.99	0.96	1.04
3.60	0.82	0.91	0.76	1.08	0.81	0.90	0.97	0.98	0.98	1.10	1.04	1.00	1.03	1.01	1.05	1.08	1.03	1.05	1.14	1.02	0.98	1.05
3.70	0.56	0.68	0.52	0.85	0.56	0.67	0.74	0.76	0.76	0.94	0.85	0.80	1.00	0.84	0.96	0.89	0.91	0.89	1.00	0.93	0.95	1.03
3.80	0.85	0.93	0.78	1.07	0.83	0.91	0.98	0.98	0.98	1.08	1.03	0.99	0.98	1.00	1.01	1.07	1.00	1.03	1.13	0.99	0.95	1.02
3.90	0.66	0.77	0.61	0.94	0.65	0.75	0.83	0.85	0.85	1.01	0.93	0.89	1.02	0.91	1.00	0.97	0.97	0.96	1.06	0.97	0.97	1.05
4.10	0.70	0.81	0.65	0.98	0.70	0.79	0.87	0.88	0.88	1.03	0.96	0.92	1.01	0.94	1.00	1.00	0.98	0.98	1.08	0.98	0.97	1.04
4.30	0.76	0.85	0.70	1.03	0.75	0.85	0.92	0.93	0.93	1.06	1.00	0.95	1.02	0.97	1.02	1.03	1.00	1.01	1.10	0.99	0.97	1.04
4.50	0.75	0.85	0.69	1.02	0.74	0.83	0.91	0.92	0.92	1.05	0.99	0.95	1.01	0.96	1.01	1.03	0.99	1.00	1.10	0.98	0.96	1.04
4.60	0.61	0.71	0.56	0.87	0.60	0.70	0.76	0.79	0.79	0.95	0.87	0.83	0.97	0.85	0.94	0.91	0.91	0.90	1.00	0.91	0.92	0.99
4.80	0.66	0.77	0.60	0.97	0.65	0.76	0.84	0.86	0.86	1.03	0.96	0.91	1.05	0.94	1.03	1.00	1.00	0.99	1.09	1.00	1.00	1.07
5.00	0.70	0.80	0.63	1.01	0.68	0.80	0.87	0.90	0.90	1.05	0.98	0.94	1.05	0.96	1.04	1.02	1.02	1.01	1.09	1.01	1.00	1.06
R	0.53	0.64	0.49	0.83	0.53	0.63	0.71	0.73	0.74	0.91	0.83	0.78	0.97	0.81	0.93	0.87	0.89	0.87	0.97	0.90	0.93	1.00

Note: "R" stands for "remaining part".

Source: author's own calculations.

Table A2. Overlapping Indexes O_{ji} for Expenditures for Households of Different Equivalence Scales

$j \backslash i$	1.00	1.50	1.70	2.00	2.20	2.40	2.70	2.90	3.10	3.20	3.40	3.60	3.70	3.80	3.90	4.10	4.30	4.50	4.60	4.80	5.00	R
1.00	1.00	1.01	0.99	1.09	0.98	1.00	1.04	1.01	0.98	1.01	0.90	0.92	0.87	0.95	0.93	0.86	0.81	0.89	0.72	0.75	0.92	0.69
1.50	0.99	1.00	0.99	1.07	0.98	0.99	1.03	0.99	0.96	0.98	0.87	0.89	0.85	0.92	0.90	0.83	0.78	0.86	0.70	0.73	0.90	0.67
1.70	1.01	1.01	1.00	1.09	0.99	1.01	1.05	1.01	0.99	1.01	0.90	0.92	0.87	0.95	0.93	0.86	0.81	0.89	0.72	0.75	0.92	0.69
2.00	0.86	0.89	0.87	1.00	0.86	0.93	0.97	0.97	0.97	1.01	0.95	0.94	0.95	0.96	0.99	0.93	0.87	0.95	0.85	0.86	0.99	0.83
2.20	1.01	1.02	1.01	1.11	1.00	1.02	1.06	1.03	1.00	1.02	0.92	0.93	0.89	0.97	0.94	0.87	0.82	0.90	0.73	0.76	0.94	0.70
2.40	0.93	0.95	0.94	1.06	0.93	1.00	1.03	1.03	1.03	1.06	0.98	0.98	0.98	1.01	1.02	0.95	0.90	0.96	0.85	0.87	1.01	0.82
2.70	0.90	0.92	0.91	1.04	0.90	0.97	1.00	1.00	1.00	1.04	0.96	0.96	0.97	0.99	1.01	0.94	0.89	0.96	0.85	0.86	1.00	0.82
2.90	0.84	0.87	0.85	1.00	0.84	0.95	0.97	1.00	1.01	1.06	1.00	1.00	1.03	1.02	1.06	1.00	0.95	1.01	0.93	0.94	1.05	0.91
3.10	0.83	0.87	0.84	0.99	0.84	0.94	0.96	0.99	1.00	1.04	0.99	0.98	1.03	1.00	1.04	0.99	0.94	0.99	0.92	0.93	1.05	0.90
3.20	0.71	0.76	0.73	0.89	0.73	0.85	0.87	0.92	0.94	1.00	0.96	0.95	1.04	0.97	1.05	0.99	0.94	1.00	0.97	0.97	1.05	0.96
3.40	0.71	0.76	0.73	0.90	0.73	0.87	0.88	0.94	0.97	1.03	1.00	0.98	1.08	1.00	1.08	1.03	0.98	1.03	1.01	1.01	1.09	1.00
3.60	0.74	0.79	0.76	0.93	0.76	0.90	0.91	0.96	0.99	1.05	1.02	1.00	1.09	1.02	1.09	1.04	0.99	1.04	1.01	1.01	1.09	0.99
3.70	0.57	0.62	0.58	0.75	0.58	0.71	0.73	0.79	0.82	0.89	0.87	0.86	1.00	0.87	0.98	0.92	0.88	0.94	0.97	0.95	1.00	0.98
3.80	0.72	0.77	0.74	0.91	0.74	0.87	0.89	0.94	0.97	1.03	1.00	0.98	1.07	1.00	1.07	1.02	0.97	1.02	0.99	0.99	1.08	0.98
3.90	0.56	0.62	0.58	0.75	0.58	0.71	0.73	0.80	0.83	0.90	0.89	0.87	1.02	0.88	1.00	0.94	0.90	0.96	1.00	0.98	1.02	1.00
4.10	0.64	0.69	0.65	0.83	0.65	0.80	0.81	0.88	0.91	0.98	0.96	0.94	1.07	0.96	1.06	1.00	0.95	1.01	1.02	1.01	1.07	1.02
4.30	0.65	0.71	0.67	0.87	0.68	0.83	0.85	0.92	0.95	1.02	1.01	0.98	1.11	1.00	1.10	1.05	1.00	1.05	1.05	1.04	1.11	1.05
4.50	0.59	0.65	0.61	0.79	0.61	0.76	0.77	0.85	0.89	0.96	0.95	0.92	1.07	0.94	1.05	0.99	0.95	1.00	1.03	1.02	1.06	1.03
4.60	0.48	0.54	0.50	0.68	0.50	0.65	0.66	0.74	0.78	0.86	0.86	0.83	1.02	0.85	0.98	0.92	0.88	0.93	1.00	0.98	1.00	1.02
4.80	0.56	0.62	0.58	0.77	0.58	0.73	0.75	0.82	0.86	0.94	0.92	0.90	1.05	0.92	1.03	0.97	0.94	0.98	1.01	1.00	1.04	1.02
5.00	0.56	0.62	0.58	0.75	0.58	0.71	0.73	0.79	0.83	0.90	0.88	0.86	1.00	0.88	0.99	0.92	0.88	0.94	0.98	0.96	1.00	0.98
R	0.48	0.54	0.50	0.68	0.50	0.65	0.66	0.74	0.78	0.85	0.85	0.82	1.00	0.84	0.97	0.91	0.87	0.92	0.99	0.96	0.99	1.00

Note: "R" stands for "remaining part".

Source: author's own calculations.

Table A3. Overlapping Indexes O_{ji} for Money Balance for Households of Different Equivalence Scales

$j \setminus i$	1.00	1.50	1.70	2.00	2.20	2.40	2.70	2.90	3.10	3.20	3.40	3.60	3.70	3.80	3.90	4.10	4.30	4.50	4.60	4.80	5.00	R	
1.00	1.00	0.99	1.00	0.90	0.99	0.95	0.99	0.94	0.94	0.94	0.93	0.91	0.78	0.93	0.87	0.93	0.89	0.92	0.91	0.87	0.91	0.85	
1.50	1.01	1.00	0.98	0.95	0.96	0.93	0.97	0.93	0.92	0.94	0.91	0.89	0.82	0.91	0.86	0.89	0.87	0.89	0.86	0.85	0.85	0.92	0.83
1.70	0.98	0.97	1.00	0.84	1.00	0.96	0.99	0.93	0.94	0.92	0.95	0.92	0.70	0.93	0.83	0.96	0.89	0.94	0.96	0.86	0.86	0.86	0.85
2.00	1.02	1.01	0.96	1.00	0.93	0.90	0.95	0.90	0.89	0.93	0.87	0.85	0.86	0.88	0.87	0.83	0.84	0.84	0.78	0.82	0.92	0.80	0.80
2.20	0.97	0.96	0.99	0.83	1.00	0.96	0.98	0.93	0.94	0.92	0.95	0.92	0.69	0.92	0.82	0.97	0.89	0.94	0.97	0.86	0.86	0.86	0.85
2.40	1.01	1.01	1.03	0.88	1.03	1.00	1.02	0.97	0.98	0.96	0.98	0.95	0.76	0.97	0.88	0.99	0.93	0.98	0.98	0.91	0.91	0.91	0.89
2.70	1.00	0.99	1.01	0.88	1.01	0.97	1.00	0.95	0.95	0.94	0.96	0.93	0.76	0.94	0.86	0.96	0.90	0.95	0.95	0.88	0.90	0.87	0.87
2.90	1.04	1.04	1.05	0.92	1.04	1.02	1.04	1.00	1.00	0.99	1.00	0.97	0.82	0.99	0.92	0.99	0.96	0.99	0.98	0.94	0.96	0.91	0.91
3.10	1.03	1.02	1.05	0.89	1.05	1.02	1.04	0.99	1.00	0.98	1.00	0.97	0.78	0.98	0.90	1.00	0.95	0.99	1.00	0.93	0.94	0.91	0.91
3.20	1.06	1.06	1.04	0.99	1.02	1.00	1.04	1.00	0.99	1.00	0.98	0.95	0.89	0.98	0.94	0.95	0.95	0.95	0.92	0.93	0.99	0.91	0.91
3.40	1.05	1.04	1.06	0.93	1.05	1.02	1.05	1.01	1.01	0.99	1.00	0.97	0.83	0.99	0.93	1.00	0.96	0.99	0.97	0.94	0.97	0.92	0.92
3.60	1.06	1.06	1.07	0.94	1.06	1.05	1.07	1.03	1.03	1.02	1.03	1.00	0.86	1.02	0.95	1.02	0.99	1.01	1.00	0.98	1.00	0.95	0.95
3.70	1.12	1.12	1.09	1.04	1.06	1.06	1.09	1.07	1.06	1.07	1.04	1.01	1.00	1.04	1.03	1.00	1.02	1.00	0.95	1.01	1.09	0.99	0.99
3.80	1.04	1.04	1.06	0.91	1.05	1.03	1.05	1.01	1.02	1.00	1.01	0.99	0.83	1.00	0.93	1.01	0.97	1.00	1.00	0.96	0.97	0.93	0.93
3.90	1.09	1.10	1.10	0.98	1.08	1.07	1.09	1.07	1.06	1.05	1.05	1.03	0.93	1.05	1.00	1.04	1.03	1.03	1.01	1.02	1.05	0.99	0.99
4.10	1.05	1.05	1.06	0.94	1.04	1.03	1.05	1.01	1.01	1.01	1.01	0.98	0.86	1.00	0.94	1.00	0.98	0.99	0.98	0.96	0.98	0.94	0.94
4.30	1.07	1.07	1.08	0.96	1.07	1.05	1.08	1.04	1.04	1.03	1.03	1.01	0.89	1.03	0.97	1.02	1.00	1.01	1.00	0.99	1.02	0.96	0.96
4.50	1.04	1.04	1.06	0.91	1.05	1.03	1.05	1.01	1.01	1.00	1.02	0.99	0.83	1.00	0.93	1.02	0.97	1.00	1.00	0.96	0.97	0.94	0.94
4.60	1.05	1.05	1.06	0.92	1.06	1.04	1.06	1.02	1.02	1.01	1.03	0.99	0.85	1.01	0.94	1.02	0.98	1.01	1.00	0.97	0.98	0.95	0.95
4.80	1.08	1.08	1.09	0.96	1.08	1.06	1.09	1.05	1.05	1.04	1.05	1.02	0.89	1.04	0.98	1.04	1.01	1.03	1.01	1.00	1.03	0.98	0.98
5.00	1.07	1.07	1.07	0.96	1.06	1.04	1.07	1.03	1.03	1.02	1.02	0.99	0.87	1.01	0.95	1.00	0.98	1.00	0.98	1.00	0.96	1.00	0.94
R	1.11	1.11	1.10	1.02	1.08	1.07	1.10	1.07	1.07	1.07	1.06	1.03	0.98	1.05	1.02	1.03	1.03	1.02	0.99	1.02	1.08	1.00	1.00

Note: "R" stands for "remaining part".

Source: author's own calculations.

Bibliography

- Bhattacharya, N. and Mahalanobis, B. (1967) "Regional Disparities in Household Consumption in India". *Journal of the American Statistical Association* 62(317): 143–61, <https://doi.org/10.1080/01621459.1967.10482896>.
- Dagum, C. (1997) "A New Approach to the Decomposition of the Gini Income Inequality Ratio". *Empirical Economics* 22(4): 551–31, <https://doi.org/10.1007/bf01205777>.
- Central Statistical Office (2011) "Household Budget Survey for Poland". Warsaw.
- Jędrzejczak, A. (2014) "Income Inequality and Income Stratification in Poland". *STATISTICS* 269.
- Kot, S. M. (2008) *Polaryzacja ekonomiczna: teoria i zastosowanie* [Economic Polarisation: Theory and Applications]. Warszawa: PWN.
- Ostasiewicz, K. and Vernizzi, A. (2017) "Decomposition and Normalization of Absolute Differences, when Positive and Negative Values are Considered: Applications to the Gini Coefficient". *Quantitative Methods in Economics XVIII*: 427–91, <https://doi.org/10.22630/mibe.2017.18.3.44>.
- Pyatt, G. (1976) "On the Interpretation and Disaggregation of Gini Coefficients". *The Economic Journal* 86(342): 243–55, <https://doi.org/10.2307/2230745>.
- Raffinetti, E., Siletti, E. and Vernizzi, A. (2015) "On the Gini Coefficient Normalization when Attributes with Negative Values are Considered". *Statistical Methods & Applications* 24(3): 507–21, <https://doi.org/10.1007/s10260-014-0293-4>.
- Shorrocks, A. F. (1984) "Inequality Decomposition by Population Subgroups". *Econometrica: Journal of the Econometric Society* 52(6): 1369–85, <https://doi.org/10.2307/1913511>.
- Yitzhaki, S. (1994) "Economic Distance and Overlapping of Distributions". *Journal of Econometrics* 61(1): 147–59, [https://doi.org/10.1016/0304-4076\(94\)90081-7](https://doi.org/10.1016/0304-4076(94)90081-7).
- Yitzhaki, S. and Lerman, R. I. (1991) "Income Stratification and Income Inequality". *Review of Income and Wealth* 37(3): 313–29, <https://doi.org/10.1111/j.1475-4991.1991.tb00374.x>.

Abstract

Nierówności przychodów i wydatków wśród polskich gospodarstw domowych

Stratyfikacja oraz pojęcie do niej odwrotne, przekrywanie, są ściśle powiązane z miarami nierówności i odgrywają kluczową rolę w jednym ze sposobów dekompozycji współczynnika Giniego. W niniejszej pracy badane są własności współczynników przekrywania. Analizowana jest dekompozycja nierówności w Polsce (przychodów, rozchodów oraz różnic pomiędzy przychodami i rozchodami gospodarstw domowych), w której pojawia się człon zależny od przekrywania się rozkładów dla poszczególnych grup.

Słowa kluczowe: stratyfikacja, przekrywanie, współczynnik Giniego, dekompozycja, nierówności międzygrupowe.