Stratification and between-group inequality: a new approach to measurement

Monti, maria and Santoro, alessandro

DSGE, State University of Milan Bicocca, DEAS, State University of Milan, Econpubblica, Bocconi University

January 2010

STRATIFICATION AND BETWEEN-GROUP INEQUALITY: A NEW APPROACH TO MEASUREMENT

Abstract

Traditionally, the literature has seen stratification as linked closely to within-group inequality. More recently, some papers have focused on measuring the impact of stratification on between-group inequality. In this paper, we show that when two groups are involved, such an impact can be measured by a simple comparison of the two cumulative distribution functions. This approach allows an interpretation of stratification in terms of probabilities and paves the way for a neat and simple graphical illustration. We apply it to the analysis of between-continent inequality.

JEL Classification Code: D31

Keywords: between-group inequality, Gini index decomposition, stratification, transvariations.

1. INTRODUCTION

Stratification means a group's isolation from members of other groups (Yitzhaki and Lerman, 1991, p. 319). A group is said to be *stratified* when it tends to form a perfect stratum in the overall distribution. Stratification has been used traditionally in sociological studies, but its rigorous definition and measurement are due to Yitzhaki and Lerman (1991) and to Yitzhaki (1994). To this end, they propose a decomposition of the Gini index into two parts: a component that is a weighted sum of groups' Ginis and a between-group inequality measure. In the first component, the weight depends positively on the value of the overlapping index for each group. In turn, the overlapping index for every group measures the extent to which the members of that group overlap with members of other groups. The less a group is stratified, the more it overlaps with other groups, thus, Yitzhaki and Lerman (1991, p. 323) conclude that "inequality and stratification are inversely related".

Yitzhaki and Lerman argue that, though counterintuitive, this result is consistent with the relative deprivation theory. The idea is that groupings are leagues, so that each person "confines his aspirations to his assigned league" (Yitzhaki and Lerman, 1991, p. 323). To put it differently, since groups are leagues, an individual (i.e. a group member) feels less deprived if members of his group have an income closer to his own, and does not care too much about the income of members of other groups. Thus, "stratified societies can tolerate higher inequality than unstratified societies since, as people become more (less) engaged with each other, they have less (more) tolerance for a given level of inequality" (Yitzhaki and Lerman, 1991, p. 323). Yitzhaki and Lerman (1991, p. 315) reinforce this interpretation observing that "generally a rise in a subgroup's inequality will reduce the subgroup's stratification", so that, in general, if within-group inequality is low, overlapping is also low or, equivalently, stratification is high. Note that in this line of reasoning, the between-group inequality, which is the second component of the Gini decomposition in Yitzhaki (1994), is left completely aside: both Yitzhaki and Lerman (1991) and Yitzhaki (1994) treat stratification and between-group inequality as two completely separate objects. Nevertheless, Milanovic and Yitzhaki

(2002), aware of the impact of stratification on between-group inequality, suggest evaluation by the ratio of the Yitzhaki and Lerman (1991) between-group Gini and the conventional between-group Gini coefficient, as obtained by Pyatt (1976).

Monti and Santoro (2009) propose a Gini decomposition that is based on extensive use of transvariations. Given two groups with different average incomes, a transvariation occurs whenever a member of the poorer (on average) group has an income higher than a member of the richer (on average) group. The number and the extent of transvariations provide natural indicators of the extent of overlapping and of stratification. Transvariations are used to rewrite the Gini index as a sum of three parts: the conventional Gini-within, a new measure of between-inequality and a new residual. The measure of between-inequality thus obtained by Monti and Santoro (2009) is numerically equivalent to the one proposed by Yitzhaki and Lerman (1991) and used by Yitzhaki (1994) and Milanovic and Yitzhaki (2002), among others. By further decomposing this measure, Monti and Santoro (2009) show that it can be expressed as a function of: i) the share of income that should be redistributed to achieve complete between-group equality; ii) a coefficient increasing in the homogeneity of group sizes and iii) an index that measures stratification as a decreasing function of the total number of transvariations. This index is equivalent to the ratio proposed by Milanovic and Yitzhaki (2002), but it has the advantage that it can be expressed as *a part of* a unifying and consistent measure of between-group inequality, rather than as a comparison between a pair of heterogeneous measures.

In this paper, we take the latter equivalence as our starting point. We show that the ratio proposed by Yitzhaki and Milanovic (2002), i.e. the measure of the impact of stratification on between-group inequality, is indeed a probability function. More precisely, when two groups are considered, we show that this measure depends exclusively on the cumulative distribution functions of the two groups. The value of the ratio proposed by Yitzhaki and Milanovic (2002), or of the equivalent index calculated by Monti and Santoro (2009), can actually be expressed as 1 minus twice the area under the cumulative distribution function for the richer (on average) group, evaluated within a range of values of the cumulative distribution function of the poorer (on average) group. This expression is to some extent very similar to the Gini index expression and it naturally suggests a graphical interpretation. We provide the latter by applying our approach to the analysis of between-continent inequality using the data reported by Yitzhaki and Milanovic (2002).

The paper is organized as follows. Section 2 summarizes the contribution of Yitzhaki and his colleagues to the measurement of stratification, and also summarizes the main results obtained by Monti and Santoro (2009). Section 3 contains the theoretical results of the paper. We derive the measures of the impact of stratification on between-group inequality in terms of the cumulative distribution functions of the compared groups and we provide an analysis of the range of variation of these measures. Section 4 applies these results to the analysis of between-continent inequality using a mainly graphical approach. Section 5 concludes.

2. OVERVIEW OF THE LITERATURE

The concepts of stratification and segmentation are frequently used in social science literature. Their definitions can be traced back at least to Lasswell (1965, p. 10): "A stratum is a horizontal layer. Stratification is the process of forming observable layers [...] where the mass of society is constructed of layer upon layer of congealed population qualities". However, until Yitzhaki and Lerman (1991), a rigorous approach to definition and measurement was missing.

Yitzhaki and Lerman's (1991) contribution is threefold. First, they obtain an index of relative stratification, index Q, which captures the extent of stratification of every group with respect to the entire population, taking into account the size of the group. Second, they derive an index of absolute overlapping, O, which is inversely related to Q. Third, they decompose the Gini index in three parts: a within-inequality component, a component that reflects the impact of stratification and a measure of between-group inequality. When commenting upon the dynamics of these three parts, Yitzhaki and Lerman (1991, p. 323, our emphasis) note that "some changes in Q_i's may leave Ginis unchanged, *and influence only component two*". This implies that, in general, there is no relationship between stratification and between-group inequality and that these can be treated as separate concepts.

Yitzhaki (1994) further develops the index of overlapping *O*, focussing on overlapping between subpopulations. He obtains a decomposition of the Gini index into two components: the between-group inequality measure, defined by Yitzhaki and Lerman (1991), and a term that is the sum of the products of income shares, Ginis and overlaps for all groups. This decomposition is used by Milanovic and Yitzhaki (2002), to measure the world's income inequality. To summarize the main results obtained by Yitzhaki (1994) let us introduce some notation. We consider a population of *n* individuals and, for the sake of simplicity, we confine our attention to the case of two population groups, which we will call "*a*" and "*c*" with *i* = *a, c*. The two groups stand for a given socioeconomic partition of the population based on the individuals' characteristics. The population size is $n_a + n_c = n$, with n, $n_a, n_c \in \mathbb{N}$, where n_a is the number of individuals belonging to group *a*, and n_c is the number of individuals belonging to group *c*. By $y_{ih}, \{y_{ih}\} \in \mathbb{R}^+$ we denote the income of individual *h* belonging to group *i*, and by μ_p, μ_c and μ_a, the overall, the group *c* and the group *a* average income, respectively. Finally, $\{y\}$ is the set of all income units.

Following Yitzhaki (1994), the Gini index decomposes as

(1) $$G(y) = \left(s_c G_c O_c + s_a G_a O_a\right) + G_b,$$

where

(2) $$O_i = p_i + \sum\nolimits_{i \neq j} p_j O_{ji},$$

(3) $$O_{ji} = \mathrm{cov}_i\left(y, F_j\left(y\right)\right) \big/ \mathrm{cov}_i\left(y, F_i\left(y\right)\right),$$

and

(4) $$G_b = 2\,\mathrm{cov}(\mu_i, \overline{F}_{Oi}^{G}) \big/ \mu_p .$$

In (1-3), $F_i(y)$, μ_i, p_i, G_i, represent the cumulative distribution, the average income, the Gini index and the share of group *i* in the overall distribution, respectively. Let $s_i = p_i \mu_i / \mu_p$ denote the share of total income owned by group i, and O_i denote the overlapping index of the same group i with the population's distribution. The index O_i is a function of the overlapping of group j by group i, O_{ji}, which in turn, is equal to the ratio between "the covariance between incomes of group i and their rank, had they been considered as belonging to the group j" (Yitzhaki 1994, p. 149) and the covariance between incomes and own ranking in group i, the latter being a normalizing factor.[1] Finally, in expression (4), G_b is twice the covariance between each groups' average income and groups' average rank in the overall population ($\bar{F}^{G}_{Oi}$), divided by the overall mean income. The overlapping index O_i reflects the overlapping of group i with itself and with the other groups, and can be interpreted as a measure of stratification.

In (1), within the round brackets, the subgroup Gini indices, G_i, and the overlap indices, O_i, have symmetrical impacts on overall inequality, since inequality rises in both. Nevertheless, high stratification implies low overlapping so that if G_b is ignored, one concludes that "inequality and stratification are inversely related" (Yitzhaki and Lerman, 1991, p. 323). According to Milanovic and Yitzhaki (2002, p. 161), however, more overlapping (i.e. less stratification) leads to lower correlation between average income and average rank and this decreases the between-group component. To measure the impact of stratification on between-group inequality, Milanovic and Yitzhaki (2002) refer to the conventional decomposition of the Gini index as proposed by Pyatt (1976)

(5) $$G(y) = G_W + G_B + R,$$

where,

(6) $$G_W = \frac{1}{n^2 \mu_p}\left(G_a n_a^2 \mu_a + G_c n_c^2 \mu_c\right) \text{ and } G_B = \frac{1}{n^2 \mu_p}\left(n_a n_c\left(\mu_a - \mu_c\right)\right); \ \mu_a > \mu_c.$$

In (6), the terms G_a and G_c denote the group a and group c Gini index, respectively, so that G_W measures within-group inequality, G_B captures between-group inequality and R is the residual,

[1]To interpret these definitions recall that Yitzhaki and Lerman (1991, p. 321) estimate the cumulative distribution, $F(y)$, by the rank of y.

which depends on the overlapping between the two group income distributions. In the conventional Gini decomposition, G_B is different from G_b. In both G_B and G_b, each group is jointly represented by its mean income and rank. However, in G_B "the rank is the rankings of the group's mean incomes" while in G_b one "takes account of each observation's raking in the overall distribution by averaging these rankings within each group" (Yitzhaki and Lerman, 1991, p. 322). The two between-group inequality components are equal if there is no overlap between groups and it can be shown that $G_b \leq G_B$ when groups overlap (we return to this in Section 3). The ratio G_b/G_B is suggested by Milanovic and Yitzhaki (2002, p.161) as an index representing the loss in between-group inequality due to an increase (decrease) in overlapping (stratification).

Using the Gini index decomposition proposed by Monti and Santoro (2009), the ratio G_b/G_B can be expressed differently. This decomposition is based on the concept of transvariation. In general, a transvariation occurs whenever a member of the poorer (on average) group is richer than a member of the richer (on average) group (Gini, 1959). If $\mu_a > \mu_c$, a transvariation occurs whenever a member of group *c* is richer than a member of group *a*. It can be immediately noted that, when no transvariation occurs, the two groups do not overlap at all, i.e. in our context, the first n_a richest individuals in the overall distribution are the n_a members of group *a*. Equally, one can say that when there are no transvariations between the two groups, the groups are perfect strata. It can be shown that G_b rewrites as:

$$G_b = \frac{(\mu_a - \mu_c)}{\mu_p} \cdot \frac{n_a n_c}{n^2} \cdot I, \tag{7}$$

with

$$I = 1 - \frac{2N^{TR}}{n_a n_c}. \tag{8}$$

where N^{TR} is the total number of transvariations. From (6) and (7), it is immediate to note that

$$I = \frac{G_b}{G_B}. \tag{9}$$

3 INDEX I

Monti and Santoro (2009) have commented on expression (7). Here, we focus on I and on expression (9).

Since $n_a n_c$ is the total number of comparisons between members of the two groups, the ratio $N^{TR}/n_a n_c$ can be interpreted as the probability that the sign of a difference between two incomes belonging to different groups is opposite with respect to the difference between the means of the two groups. In other words, this ratio corresponds to the probability of a transvariation.[2] Then, assuming $\mu_a > \mu_c$, we represent the income sets of the two groups as two discrete random variables denoted by Y_c and Y_a, respectively, so that [3]

(10) $$\text{Prob}\left[Y_c > Y_a\right] \equiv \frac{N^{TR}}{n_a n_c}$$

is the probability that a member of the poorer (on average) group is richer than a member of the richer (on average) group. Thus, we can write

(11) $$I = 1 - 2\text{Prob}\left[Y_c > Y_a\right].$$

We can now derive a number of results concerning the range of I. First, we note that index I assumes its maximum, I = 1 (G_b = G_B), when stratification is perfect, i.e. when there are no transvariations ($N^{TR} = 0$).

(12) $$I = 1 \Leftrightarrow \text{Prob}\left[Y_c > Y_a\right] = 0 \Leftrightarrow N^{TR} = 0.$$

Second, we note that

(13) $$I = 0 \Leftrightarrow \text{Prob}\left[Y_c > Y_a\right] = 1/2 \Leftrightarrow N^{TR} = n_a n_c/2.$$

[2]See Gini (1959) page 8. on this point

[3]We observe that in the definition of the Gini index and in its decomposition, there is an implicit assumption of independence between Y_a and Y_c. For all j and l, the probability of the difference (y_{aj}-y_{cl}) is the product of the probability to observe y_{aj} in the distribution Y_a and the probability to observe y_{cl} in distribution Y_c. That is, given $p\left(y_{aj}\right) = 1/n_a$ and $p\left(y_{cl}\right) = 1/n_c$,one has $p(y_{aj}, y_{cl}) = p\left(y_{aj}\right) \cdot p\left(y_{cl}\right) = \frac{1}{n_a} \cdot \frac{1}{n_c}$.

The index I is equal to zero if, and only if, the probability of transvariation is equal to 0.5 or, in other words, when the number of transvariations is equal to the number of non-transvariations. [4] Third, we note that the index $I = G_b/G_B$ is minimum when the probability of transvariation reaches its maximum, i.e.

$$\min\left\{I = \frac{G_b}{G_B}\right\} = 1 - 2\max\left\{\text{Prob}\left[Y_c > Y_a\right]\right\}, \tag{14}$$

where

$$\max\left\{\text{Prob}\left[Y_c > Y_a\right]\right\} = \frac{1}{n_a n_c}\max N^{TR}, \tag{15}$$

and max N^{TR} is the maximum number of transvariations between the two groups under the assumption $\mu_a > \mu_c$. We can obtain the expression of max N^{TR} by considering the difference between the total number of comparisons between members of the two groups, $n_a\, n_c$, and the minimum number of non-transvariation. By denoting q_a as the number of members of group a whose income is higher than μ_a, and p_c as the number of members of group c whose income is (weakly) lower than μ_c, one obtains [5]

$$\max N^{TR} = n_a n_c - q_a\, p_c, \tag{16}$$

and the minimum value of $I = G_b/G_B$ is

$$\min\left\{I = \frac{G_b}{G_B}\right\} = -1 + \frac{2q_a p_c}{n_a n_c}. \tag{17}$$

Expression (17) suggests immediately to write both the index I and its minimum in a continuous form. To do this, one has to express the number of transvariations and its maximum as functions of continuous cumulative distributions. Assuming a continuous approximation of the random variables Y_a and Y_c, Monti and Santoro (2007) show that the *expected value of transvariations* (p_{ac}) is

[4] A non-transvariation is the difference between two incomes, belonging to different groups, that has the same sign with respect to the difference between the means of the two groups.

[5] The proof of this result can be obtained from the authors upon request.

(18) $$p_{ac} = \int_{-\infty}^{\infty} F(y)\left(1 - G(y)\right) dy$$

where $F(y)$ and $G(y)$ are the cumulative distributions of Y_a and Y_c, respectively. The result in (18) is obtained starting from a continuous definition of the Gini mean difference extended to two different variables.[6] Here, we show that the *number of transvariations* and its maximum also rewrite in terms of cumulative distribution functions. In so doing, we follow a different approach with respect to Monti and Santoro (2007). Our results are derived in a very simple way starting from the definition of the probability of transvariation.

In what follows, we assume that (see footnote 3)

(19) $$\mathrm{Prob}\left(Y_a < y_a, Y_c < y_c\right) = \mathrm{Prob}\left(Y_a < y_a\right)\mathrm{Prob}\left(Y_c < y_c\right) = F\left(y_a\right)G\left(y_c\right),$$

where $F(y_a)$ and $G(y_c)$ are the cumulative distribution of Y_a and Y_c, respectively. Then, because

(20) $$\mathrm{Prob}\left(Y_a > \mu_a, Y_c < \mu_c\right) = \left[1 - F\left(y_a\right)\right]G\left(y_c\right),$$

expression (15) rewrites as

(21) $$\max\left\{\mathrm{Prob}\left[Y_c > Y_a\right]\right\} = 1 - \left[1 - F(\mu_a)\right]G(\mu_c),$$

and expression (17) becomes

(22) $$\min\left\{I = \frac{G_b}{G_B}\right\} = 2\left[1 - F(\mu_a)\right]G(\mu_c) - 1.$$

Using (22), we can see that min I can be negative and that its value depends on the skewness of the two distributions. If the two distributions are both symmetric with respect to their mean, the minimum value of index I is *-1/2*. On the other hand, one has min I>*-1/2* if either the distribution of Y_a is symmetric with respect to μ_a and the distribution of Y_c has positive asymmetry (μ_c >median, right obliquity), or if the distribution of Y_a has negative asymmetry (μ_a<median, left obliquity) and

[6] We observe that (18) provides an alternative path to derive the expression of the residual R as presented in Lambert and Decoster (2005).

the distribution of Y_c is either symmetric or asymmetric with positive asymmetry.

Moreover, one has min $I<-1/2$ if either the distribution of Y_a is symmetric with respect to μ_a and the distribution of Y_c has negative asymmetry, or if the distribution of Y_a has positive asymmetry (μ_a>median, right obliquity) and the distribution of Y_c is either symmetric or asymmetric with negative asymmetry. Nothing can be said about min I if the two distributions are asymmetric with the same asymmetry.

Let us now consider the continuous expression of index I (expressions (10) and (11)). Observe that

(23) $$\text{Prob}(Y_a < Y_c) = \text{Prob}(Y_a - Y_c < 0).$$

Then, if the difference variable (Y_a - Y_c) is denoted by Z, and its cumulative distribution is denoted by $H_{Y_a - Y_c}(z)$, one has

(24) $$\text{Prob}(Y_a - Y_c < z) = H_{Y_a - Y_c}(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z+y_c} dF(y_a)dG(y_c) = \int_{-\infty}^{\infty} F(z+y_c)\,dG(y_c),$$

and expression (10) becomes

(25) $$\text{Prob}(Y_a - Y_c < 0) = H_{Y_a - Y_c}(0) = \int_{-\infty}^{\infty} F(y_c)\,dG(y_c) \ .$$

Using (25) back in (11), we can write

(26) $$I = 1 - 2\int_{-\infty}^{\infty} F(y_c)\,dG(y_c).$$

Expression (26) says that the measure of impact of overlapping on between-inequality can be expressed as a function of the cumulative distribution functions of the two groups. More precisely, expression (26) says that $I = G_b/G_B$ is equal to 1 minus twice the area under the cumulative distribution function of the richer group, evaluated as a function of the cumulative distribution function of the poorer group.

To sum up, Yitzhaki and Milanovic (2002) propose measuring the impact of stratification on between-group inequality considering a ratio of two between-inequality coefficients derived from two different decompositions of the Gini index. Monti and Santoro (2009) propose an alternative Gini decomposition. In this decomposition, the between component is function of an index, denoted by I, which is a measure of the impact of overlapping. The two indices, I and G_b/G_B , are equivalent. Expression (26) represents index I, and then represents the ratio between the two between-inequality coefficients as an area. This area is evaluated in term of cumulative distributions. This allows a further and deeper insight into the impact of the two groups' overlap on between-inequality.

We discuss the implications of expression (26) using a graphical approach in Section 4, where we analyse between-continent inequality in a way that immediately relates to the research of Yitzhaki and Milanovic (2002).

4. A GRAPHICAL INTERPRETATION

Using the national income/expenditure distribution data from 111 countries, Yitzhaki and Milanovic (2002) decomposed total inequality between individuals in the world, by continents and regions. In particular, they partitioned the world into five continents: Africa; Asia; Western Europe, North America and Oceania (WENAO); Eastern Europe and the Former Soviet Union (EUFSU) and Latin America and the Caribbean (LAC). Commenting on the results concerning between-continent inequality, they note that "between-continent inequality Gini is 0.309…had we used Pyatt's between-group component, we would have gotten a between-continent Gini of 0.398 which means that overlapping has decreased the between-continent component by about 9 Gini points" (Yitzhaki and Milanovic, 2002, p. 163). It is interesting to verify how continent-by-continent comparisons have contributed to this result. We treat countries as units of observation, and continents as groups. Ordering the continents by their per capita average income in international dollars, in Table 1, we report the values of I, i.e. the values of the ratio G_b/G_B, for each pair of continents.

TABLE 1 AROUND HERE

Recalling the discussion in Section 3, we note that Table 1 covers the whole range of possible values of *I*. When WENAO is compared with other continents, *I* reaches very high values. In particular, *I* is very close to unity when WENAO is contrasted with Africa (99%), EUFSU (96.4%) and LAC (92.7%). According to expression (12) above, in all these cases, the probability of a transvariation is close to zero, i.e. stratification dominates and there is virtually no overlapping between WENAO countries and countries belonging to other continents. Thus, in these cases, using Pyatt's (1976) decomposition and Yitzhaki's (1994) decomposition, we would yield almost equivalent measures of between-continent inequality. The only case in which WENAO countries are to some extent involved in transvariations, is when they are compared with Asian countries. In this case, *I* equals 72.2%, so that we know from (11) that the probability of an Asian country having a mean income higher than a WENAO country is equal to 13.9%.

At the other extreme, *I* reaches a (small) negative value (i.e. -4.5%) when EUFSU and Asia are compared. According to expression (13), this means that the probability of a transvariation among countries belonging to these two continents is higher than 50%; and using (11), this probability amounts to 52.2%. Thus, although the mean income of EUFSU is 75% higher than the mean income of Asia, it is more likely that an Asian country has a mean income higher than a EUFSU country than the reverse. It is this result that is associated clearly with a high polarization within both these continents, which generates a *negative* value of *I*. This signals low stratification and high overlapping.

Finally, remaining comparisons are somewhat in between these the two extremes. For example, when LAC and EUFSU are compared, the value of *I* is close to 50%, which indicates, again using (11), that the probability of a EUFSU country having a mean income higher than a LAC country is around 25%. According to Table 1 and to expression (11), the probability of an African country having a mean income higher than either an Asian or a EUFSU country, and the probability of an Asian country having a mean income higher than a LAC country, is ranged between 25% and 50%.

These, and more results, can be illustrated considering the probability of transvariation, as in expression (25) and representing, for each pair of continents, the cumulative distribution of the richer continent as a function of the cumulative distribution of the poorer one. In Figure 1, we present four comparisons that represent the graphical counterpart of expression (25).

FIGURE 1 AROUND HERE

In each of the four diagrams, the kinked line represents the cumulative distribution of the richer continent (function $F(y)$ using the notation of expression (25)) plotted against the cumulative distribution of the poorer continent (function $G(y)$). It follows that the coordinates of each point on the kinked line are given by i) the proportion of countries, which, in the poorer continent, has an average income smaller than $\bar{y}$ on the horizontal axis, ii) the percentage of countries, which, in the richer continent, has an average income smaller than $\bar{y}$ on the vertical axis. In each diagram, areas under the $F(y)$ curves are equal to the probability of a transvariation between countries belonging to the corresponding continents, as can be verified by simple numerical computation.

The 45° line (when present) indicates the values that the cumulative distribution of the richer continent should possess, to be exactly equal to the cumulative distribution of the poorer continent [$F(y) = G(y)$] *at the same average income level*. Where the slope of $F(y)$ is higher than 1 (the slope of the 45° line) in a given range of $G(y)$, it means that, for *each point* belonging to that region, the percentage of countries having an average income smaller than $\bar{y}$ in the richer continent is higher than the percentage of countries having an average income smaller than the same value $\bar{y}$ in the poorer continent. In that region transvariations are being originated, because, at the same levels of average income, there are countries belonging to the richer continent whose incomes are lower than at least one of the countries of the poorer continent.

We choose to represent two comparisons involving EUFSU countries and two comparisons involving WENAO countries; these comparisons are both compared with Asian and African countries. The polarization among EUFSU countries is visible in the shape of its cumulative distribution when plotted against Asia and, to some extent, against Africa. In both these cases, at low income levels, $F(y)$ is above the 45° line since many of the absolute poorest countries belong to EUFSU (Georgia, Uzbekistan, Armenia; see Yitzhaki and Milanovic, 2002, p.176). This means that the minimum average income level, i.e. the lower boundary of the region R, where the integral in (25) is evaluated, belongs to the richer (on average) continent, EUFSU in both cases. As higher income levels are considered, the cumulative distribution for EUFSU falls behind the 45° degree line with respect to both Africa and Asia. This signals that the probability of a EUFSU country having an average income below a given level is lower than the probability of finding an African or an Asian country having an average income below the same level. Consequently, in this region, there are no (or few) transvariations. However, in the comparison with Asia, there is another region, at middle-high income levels, where transvariations appear again, so that $F(y)$ lies above the 45°line at the top. The latter reflects polarization of income across Asian countries, namely the presence of high-income countries, such as Singapore, Taiwan, Korea, Japan and Hong Kong.

Diagrams involving WENAO are much more conventional. Since the poorest WENAO country (Turkey) is always far richer than the poorest African or Asian country, $F(y)$ (in these cases) lies on the horizontal axes for a large interval of values of $G(y)$. More precisely, there is no WENAO country with a mean income lower than an Asian country up until the 7th decile of the Asian distribution. At higher mean income levels, some transvariations are originated by the rich Asian countries mentioned above, although the kinked line never crosses the 45° line. Overall the probability of a transvariation amounts to 13.8%, again, a value that can be approximated by calculating the area under the kinked curve. The comparison between WENAO and African countries is much more dramatic, since the stratification, as indicated by the value of I at 99% in Table 1, is almost absolute. There are only three transvariations, and the probability of a transvariation is only marginally different from 0; thus the 45° degree line cannot be represented and $F(y)$ is almost everywhere lying on the horizontal axis.

5. CONCLUDING REMARKS

The traditional literature on stratification measurement (Yitzhaki and Lerman, 1991; Yitzhaki, 1994) tends to see stratification as inversely related to inequality. This view arises from the fact that higher stratification, i.e. lower overlapping, is usually associated with lower within-group inequality. However, the most recent literature focuses on the impact of stratification on between-group inequality and proposes a measure to evaluate it (Yitzhaki and Milanovic, 2002; Monti and Santoro, 2009). This measure is such that, *ceteris paribus*, a higher stratification is associated with higher between-group inequality.

In this paper, we interpret this measure as a function of the probability of a transvariation. We show that when two groups are considered, this approach leads to rewrite such a measure as 1 minus twice the area under the cumulative distribution function of the richer group, expressed as a function of the cumulative distribution function of the poorer group. This formula is, to some extent, similar to the expression of the Gini index and naturally suggests a graphical illustration that we provide analysing between-continent inequality. The major advantage of our approach is that a lot of information about the impact of stratification on between-group inequality can be obtained by a simple graphical inspection of the plot of the cumulative distribution function of the group with a higher mean income against the cumulative distribution function of the group with a lower mean income.

What rationale can be provided for this interpretation of the impact of stratification on between-group inequality? We think an answer to this question can be provided by the concept of *group deprivation* and it applies when group members share a strong body of common moral, social and cultural values. By *group deprivation* we mean the feeling that a group has to be deprived whenever any of its members' income is lower than any members' income of the other group. In such a case, every member of a group feels empathy for any other member of his own group the group as a whole is affected by the probability that any of its members is richer than any of the member of the other group. Group deprivation, thus, increases in this probability and this drives the impact of stratification on between-group inequality.

To provide an example, we refer back to the analysis of between-continent inequality and consider the viewpoint of a representative individual of an African country. By representative individual, we mean an individual whose income is exactly equal to the mean income of his their country. Suppose this individual feels to belong to the African continent, not only to his own country. Then, when comparing Africa to any other continent, this individual would care about the possibility that any representative African is richer than the representative individual of a Western or of an Asian country. This possibility corresponds to the probability of a transvariation, i.e. to the probability that any representative African is richer than a representative individual of another (richer) continent. The higher this probability is then the lower the feeling of group deprivation and between-group inequality.

REFERENCES

Gini, C., " *Il Concetto di Transvariazione e le Sue Prime Applicazioni*", 1916, Reproduced by "Corrado Gini Memorie di metodologia statistica". Volume secondo: *Transvariazione*, pp. 1-55, Università degli Studi di Roma (in Italian), 1959.

Lambert, P.J. and A. Decoster, "The Gini coefficient reveals more", *Metron-International Journal of Statistics,* 63, 3, 373-400, 1965.

Lasswell, T.E., *Class and Stratum*, Houghton Mifflin Company, Boston, M.A, 1965.

Milanovic, B. and Yitzhaki S. "Decomposing world income distribution: does the world have a middle class?", *Review of Income and Wealth*, 48, 2, 155-178, 2002.

Monti, M. and A. Santoro, "The Gini decomposition: an alternative formulation with an application to tax reform", DEAS, Working Paper n. 2009-30, available at http://www.economia.unimi.it/uploads/wp/Deas2009_30wp.pdf.

Monti, M. and Santoro A., "A note on the measurement of stratification and between-group inequality", *Journal of Income Distribution*, 18, 3-4, 2009, available at http://www-1.unipv.it/websiep/wp/629.pdf

Pyatt, G., “On the interpretation and disaggregation of Gini coefficient”, *Economic Journal*, 86, 243-255, 1976.

Yitzhaki, S. and R. Lerman, “Income stratification and income inequality”, *Review of Income and Wealth*, 37, 313-329, 1991.

Yitzhaki, S., “Economic distance and overlapping of distributions”, *Journal of Econometrics*, 61, 147-159, 1994.

TABLE 1
Values of *I* for Continent-by-Continent Comparisons

	AFRICA (1,310)	ASIA (1,594.6)	EUFSU (2,780.9)	LAC (3,639.8)	WENAO (10,012.4)
AFRICA					
ASIA	32.6%				
EUFSU	32.0%	-4.5%			
LAC	77.4%	30.0%	45.0%		
WENAO	99.0%	72.2%	96.4%	92.7%	

Note: Mean income in $PPP within brackets (1993)
EUFSU = Eastern Europe and Former Soviet Union
LAC=Latin America and Caribbean
WENAO = Western Europe, North America and Oceania
Source: authors' calculation from Yitzhaki and Milanovic (2002).

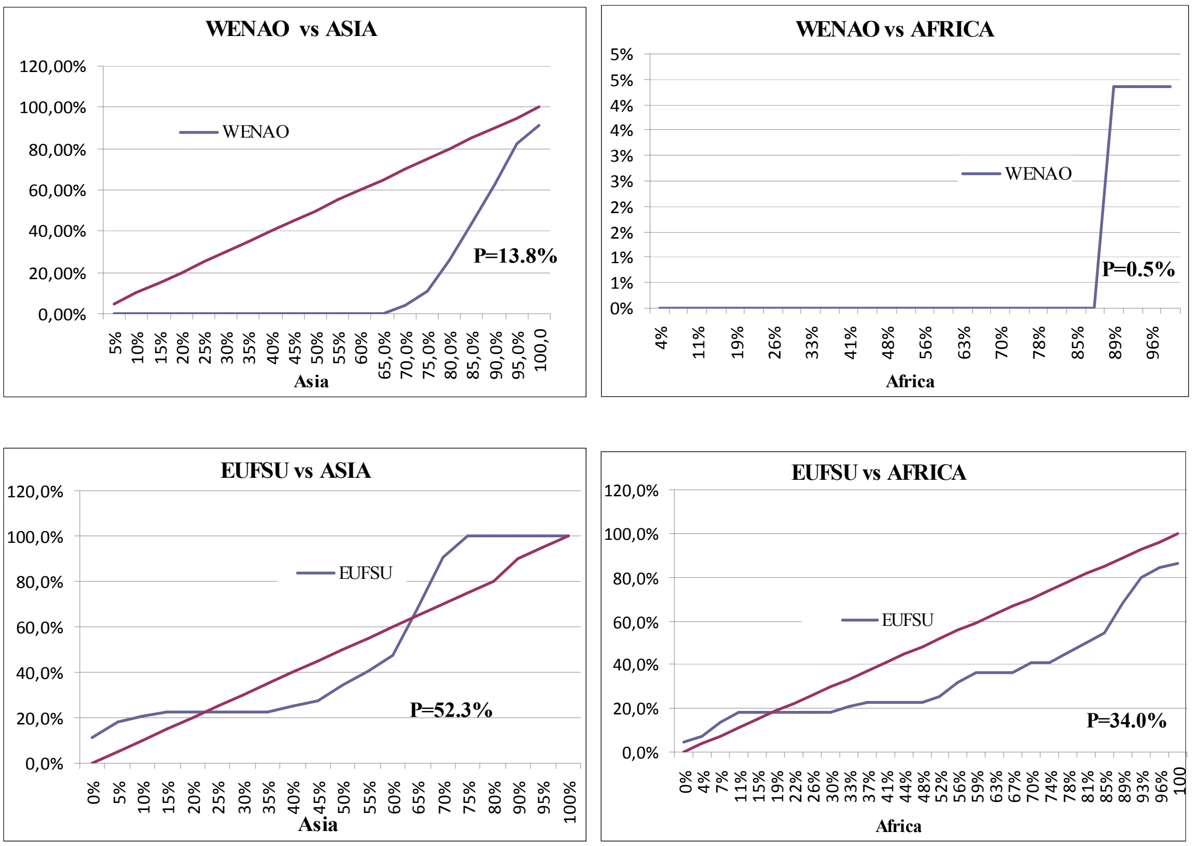

Figure 1. Continent-by-Continent Comparisons