Income stratification and between-group inequality

Allanson, Paul

Published in:
Economics Letters

DOI:
10.1016/j.econlet.2014.05.025

Publication date:
2014

Economics Letters 124 (2014) 227–230

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Income stratification and between-group inequality

Paul Allanson *

Economic Studies, University of Dundee, 3 Perth Road, Dundee DD1 4HN, United Kingdom

HIGHLIGHTS

- I generalize a result on the decomposition of the Gini index to more than two groups.
- It is shown explicitly how overlapping of groups impacts between-group inequality.
- An overall index of income stratification is identified for the population.
- I tabulate the pairwise contributions of regions to global income stratification.

ARTICLE INFO

Article history:
Received 4 December 2013
Received in revised form
16 May 2014
Accepted 24 May 2014
Available online 2 June 2014

JEL classification:
D31
D63

Keywords:
Gini coefficient
Subgroup decomposability
Income stratification

ABSTRACT

The paper shows explicitly how the overlapping of groups impacts between-group inequality by generalizing a result on the group-wise decomposition of the Gini index to more than two groups. It is demonstrated that the ratio of Yitzhaki's measure of between-group inequality to the conventional measure is in general equal to one minus twice the weighted average probability that a random member of a richer (on average) group is poorer than a random member of a poorer (on average) group, and may therefore be interpreted as an overall index of income stratification in the population. The results are used to tabulate the contribution of each pair of regions in the world to the overall level of global income stratification.

1. Introduction

It is well known that the standard decomposition of the Gini index G by population groups does not yield an exact partition into between-group and within-group components, G_B and G_W respectively, unless the income ranges of the groups are non-overlapping (e.g., Mookherjee and Shorrocks, 1982). This has led both to an extensive literature exploring the nature of the "residual" from the standard decomposition (e.g. Lambert and Aronson, 1993; Lambert and Decoster, 2005) and to a parallel search for alternative decompositions that might prove more amenable to analysis and interpretation. In the latter vein, Yitzhaki and Lerman (1991) provides a partition of the Gini into between-group, within-group and overlapping components – G_b, G_w and G_o respectively – where overlapping is considered as the inverse of the sociological concept of 'stratification'. Yitzhaki (1994) subsequently combines the latter two elements into a single within-group measure G_{wo} that is explicitly written as a function of the degree of inequality within groups and the degree of overlapping between each pair of groups, but G_b is also affected by overlapping and it remains to be shown how this measure relates to the conventional between-group index G_B if there are more than two groups.[1]

2. Group-wise decomposition of the Gini index

Consider a population divided into $K \geq 2$ mutually exclusive and exhaustive groups that are ordered by mean income from the poorest to the richest group. Let Y_k, $F_k(Y_k)$, μ_k, p_k and q_k represent respectively the income (or some other relevant aspect of wellbeing) variable, cumulative distribution function, expected value, population share and income share of group k. The overall population $Y_u = Y_1 \cup Y_2 \cdots \cup Y_K$ is the union of all groups with distribution function $F_u(Y_u) = \sum_k p_k F_k(Y_k)$ and expected value $\mu_u = \sum_k p_k \mu_k$. The (fractional) ranking of group k incomes in

* Tel.: +44 01382 384377.
E-mail address: p.f.allanson@dundee.ac.uk.

[1] See Yitzhaki and Schechtman (2013) for a recent monograph on the Gini methodology.

http://dx.doi.org/10.1016/j.econlet.2014.05.025

the group l and overall income distributions are given as $F_l(Y_k)$ and $F_u(Y_k)$ respectively, with corresponding mean ranks $\bar{F}_{kl}$ and $\bar{F}_{ku}$.

Following Mookherjee and Shorrocks (1982), the conventional group-wise decomposition of the Gini index may be written as $G = 2\mathrm{cov}(Y_u, F_u(Y_u))/\mu_u = G_B + G_W + R$ where $G_B = \sum_k \sum_l p_k p_l |\mu_l - \mu_k| / 2\mu_u$; $G_W = \sum_k p_k q_k G_k$ with $G_k = 2\mathrm{cov}(Y_k, F_k(Y_k))/\mu_k$ denoting the Gini index of group k; and the residual R is interpreted as an 'interaction effect'. The alternative approach of Yitzhaki (1994) yields the exact decomposition $G = G_b + G_{wo}$ where $G_b = 2\sum_k p_k(\mu_k - \mu_u)(\bar{F}_{ku} - 0.5)/\mu_u$; and $G_{wo} = \sum_k q_k G_k O_k$ with O_k denoting the overlapping index of group k with the entire population. In turn $O_k = \sum_l p_l O_{lk}$ where the pairwise overlapping index $O_{lk} = \mathrm{cov}(Y_k, F_l(Y_k))/\mathrm{cov}(Y_k, F_k(Y_k))$ lies in the open interval [0, 2] and is an increasing function of the fraction of group l that is located in the income range of group k, taking a value of zero when there is no overlap between the two groups and of one if the income distributions of the two groups are identical, i.e. $F_l(Y_k) = F_k(Y_k)$.

Thus $G_{wo} = G_W$ if there is perfect stratification in the sense of Lasswell (1965), since $O_{kk} = 1$ by definition, whereas $G_{wo} > G_W$ if the income ranges of the various groups overlap to any extent with the difference $R_W = G_{wo} - G_W$ given as:

$$
\begin{aligned}
R_W &= \sum_k q_k G_k \left(\sum_{l \neq k} p_l O_{lk} \right) \\
&= 2 \sum_k p_k \left(\sum_{l \neq k} p_l \mathrm{cov}(Y_k, F_l(Y_k)) \right) / \mu_u \geqslant 0. \qquad (1)
\end{aligned}
$$

Yitzhaki and Lerman (1991, p. 323) conclude that "inequality and stratification are inversely related", arguing that this relationship is consistent with relative deprivation theory in that "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". However, as Monti and Santori (2011) observe, this conclusion ignores the effect of overlapping on the between-group component G_b, which will also affect the overall level of inequality perceived by the society.

Yitzhaki and Lerman (1991, p. 322) note that $G_b = G_B$ if there is no overlapping and $G_b < G_B$ otherwise. Monti and Santori (2011) further demonstrate in the two group case that the ratio of G_b to G_B is equal to:

$$
I = G_b/G_B = 1 - 2\mathrm{Prob}(Y_1 > Y_2) \qquad (2)
$$

where $\mathrm{Prob}(Y_1 > Y_2)$ is the probability of transvariation, i.e. the probability that the income of a random member of the poorer (on average) group is more than that of a random member of the richer (on average) group. To extend this result to the general case of $K \geq 2$ groups, note that G_b may also be expressed as:

$$
\begin{aligned}
G_b &= 2 \sum_k p_k \mu_k \left(\sum_{l \neq k} p_l (\bar{F}_{kl} - 0.5) \right) / \mu_u \\
&= \sum_k \sum_{l>k} (p_k + p_l) \frac{(p_k \mu_k + p_l \mu_l)}{\mu_u} \\
&\quad \times \left(\frac{2\left(p_k \mu_k p_l (\bar{F}_{kl} - 0.5) + p_l \mu_l p_k (\bar{F}_{lk} - 0.5)\right)}{(p_k + p_l)(p_k \mu_k + p_l \mu_l)} \right) \\
&= \sum_k \sum_{l>k} (p_k + p_l)(q_k + q_l) G_b^{kl} \qquad (3)
\end{aligned}
$$

where the first line follows since $\bar{F}_{ku} = \sum_l p_l \bar{F}_{kl}$ and $\bar{F}_{kk} = 0.5$, while G_b^{kl} denotes the Yitzhaki (1994) between-group index in the sub-population consisting only of groups k and l. Similarly, G_B can be written as:

$$
\begin{aligned}
G_B &= \sum_k \sum_{l>k} p_k p_l (\mu_l - \mu_k) / \mu_u \\
&= \sum_k \sum_{l>k} (p_k + p_l)(q_k + q_l) G_B^{kl} \qquad (4)
\end{aligned}
$$

where G_B^{kl} denotes the between-group Gini in the sub-population consisting of groups k and l only. Using (2) and (4), (3) may be rewritten as:

$$
\begin{aligned}
G_b &= \sum_k \sum_{l>k} (p_k + p_l)(q_k + q_l) G_B^{kl} \left(\frac{G_b^{kl}}{G_B^{kl}} \right) \\
&= \sum_k \sum_{l>k} p_k p_l (\mu_l - \mu_k) \{1 - 2\mathrm{Prob}(Y_k > Y_l)\} / \mu_u \qquad (5)
\end{aligned}
$$

from which it follows immediately that I will in general be equal to:

$$
\begin{aligned}
I &= G_b/G_B = \sum_k \sum_{l>k} w_{kl} (1 - 2\mathrm{Prob}(Y_k > Y_l)) \\
&= 1 - 2 \sum_k \sum_{l>k} w_{kl} \mathrm{Prob}(Y_k > Y_l) \\
&= \sum_k \Bigg\{ \sum_{l<k} w_{kl} (0.5 - (1 - \mathrm{Prob}(Y_k < Y_l))) \\
&\quad + \sum_{l>k} w_{kl} (0.5 - (\mathrm{Prob}(Y_k > Y_l))) \Bigg\} \qquad (6)
\end{aligned}
$$

where $w_{kl} = p_k p_l (\mu_l - \mu_k) / \left(\sum_k \sum_{l>k} p_k p_l (\mu_l - \mu_k) \right) \geqslant 0$, with $\sum_k \sum_{l>k} w_{kl} = 1$ by definition, and the final line holds since $\mathrm{Prob}(Y_k > Y_l) = (1 - \mathrm{Prob}(Y_k < Y_l))$.

Hence I is in general equal to one less twice the weighted average probability of transvariation between the various pairs of groups in the population. In his study of earnings differentials Gastwirth (1975) proposes $TPROB = 2\mathrm{Prob}(Y_1 > Y_2)$ as an index of overlapping between two groups, taking an "ideal" value of one when the two distributions are identical since $\mathrm{Prob}(Y_1 > Y_2) = 0.5$ in this case. Thus I in (2) may be interpreted as the complementary index of non-overlapping or stratification, with (6) providing a generalization to two or more groups. I is a unit-free index that will take a maximum value of one when there is no overlap between any of the groups such that $\mathrm{Prob}(Y_k > Y_l) = 0 \, \forall k, \ l > k$; and will equal zero when the income distributions of all the groups are identical.[2] For $K > 2$, the extent to which non-overlapping between any pair of groups contributes to the overall level of stratification is an increasing function of their population shares and the difference in mean incomes between them. I is invariant to both the scaling and translation of incomes. It is also invariant to replication both of the population within existing groups and of groups.

I has previously been identified by Milanovic and Yitzhaki (2002, p. 161) "as an index indicating the loss of between group inequality due to overlapping". The difference $R_B = G_b - G_B$ can be written from (6) as:

$$
\begin{aligned}
R_B &= -2G_B \sum_k \sum_{l>k} w_{kl} \mathrm{Prob}(Y_k > Y_l) \\
&= -2 \sum_k \sum_{l>k} p_k p_l (\mu_l - \mu_k) \mathrm{Prob}(Y_k > Y_l) / \mu_u \leq 0 \qquad (7)
\end{aligned}
$$

2 Negative values of I are also possible when mean incomes by group are negatively correlated with mean ranks.

Table 1
Income stratification between regions of the world.

	Population share (%)	Mean income ($PPP)	Mean rank in income distribution of:					
			Africa	Asia	EFSU	LAC	WENAO	World
Africa	10.0	1310.0	0.500	0.515	0.275	0.261	0.049	0.407
Asia	59.5	1594.6	0.485	0.500	0.265	0.247	0.064	0.397
EFSU	7.8	2780.9	0.725	0.735	0.500	0.483	0.136	0.609
LAC	8.4	3639.8	0.739	0.753	0.517	0.500	0.172	0.629
WENAO	14.3	10012.4	0.951	0.936	0.864	0.828	0.500	0.861
World	100.0	3031.8						0.500
			Pairwise contribution to I					Sum
Africa			–	−0.000	0.002	0.004	0.047	0.052
Asia			−0.000	–	0.011	0.021	0.258	0.290
EFSU			0.002	0.011	–	0.000	0.024	0.037
LAC			0.004	0.021	0.000	–	0.021	0.046
WENAO			0.047	0.258	0.024	0.021	–	0.350
World								0.776

Notes: Top panel. Source: Milanovic and Yitzhaki (2002) Tables 4 and 7—see also Table 1 for the list of countries in each region (EFSU—Eastern Europe and Former Soviet Union; LAC—Latin America and Caribbean; WENAO—Western Europe, North America and Oceania). Bottom panel. Author's own calculations.

on which basis it may be argued, in contrast to Yitzhaki and Lerman (1991), that unstratified societies can tolerate more between-group inequality than stratified societies because individuals' positions within society are less narrowly determined by group membership. Nevertheless, holding G_W and G_B constant, overlapping *per se* must increase overall inequality since $R \geq 0$ by definition (Pyatt, 1976), with (1) and (7) yielding a novel expression for $R = R_W + R_B$ as:

$$R = 2\sum_k p_k \left(\sum_{l>k} (p_l (\mathrm{cov}(Y_k, F_l(Y_k)) + \mathrm{cov}(Y_l, F_k(Y_l)) - (\mu_l - \mu_k)\,\mathrm{Prob}\,(Y_k > Y_l))) \right) / \mu_u$$
$$= 2\sum_k p_k \left(\sum_{l>k} p_l \int (1 - F_k(y_k))\, F_l(y_k) \partial y_k \right) / \mu_u \quad (8)$$

where the final line makes use of the expression for R presented in Lambert and Decoster (2005) for the two group case.[3]

3. Empirical illustration

By way of illustration, this section elaborates the empirical analysis presented in Milanovic and Yitzhaki (2002) of world inequality in 1993 by regions.[4] The top panel in Table 1 presents estimates from their Tables 4 and 7 of population shares, p_k; mean incomes, μ_k; and mean rankings in the income distributions of each region, $\bar{F}_{kl}$, and the world $\bar{F}_{ku}$. The lower panel reports the pairwise components of I identified in the final line of (6), where the calculation of these estimates makes use of the identity $\bar{F}_{kl} = \mathrm{Prob}(Y_k > Y_l)$. The components sum to give the value of the stratification index $I = 0.776$, which is equal to the ratio of their reported estimates of $G_b = 0.309$ and $G_B = 0.398$.[5] Examination of the individual entries shows that the main contribution to stratification, accounting for as much as two thirds of the total, is due to the Asia and Western Europe/North America/Oceania (WENAO) pair as a result of a combination of the low degree of income overlap, the populousness of the two regions and the large difference in mean incomes between them. In contrast, the Africa and Asia pair contributes negatively to stratification, although the magnitude of this effect is negligible, because an African chosen at random is likely to be better off than a randomly chosen Asian despite the fact that average incomes are lower in Africa. Given that the value of I implies a weighted average probability of transvariation of 11.2%, only the Africa and WENAO pair and Asia and WENAO pair contribute more to R_B than to G_B.

4. Conclusion

The paper demonstrates how the residual from the conventional decomposition of the Gini index is fully absorbed into the between-group and within-group components proposed by Yitzhaki (1994). In particular, it is established that $I = G_b/G_B$ is in general equal to one minus twice the weighted average probability of transvariation and may therefore be interpreted as an overall index of income stratification in the population. Using this result it is shown that the main source of stratification between regions of the world in 1993 was the limited overlap between the income distributions of Asia and WENAO given the relative populousness of the two regions and the difference in mean incomes between them. High per capita growth rates in some poorer Asian countries, most notably China and India, may be expected to have reduced levels of both stratification and inequality between regions in more recent years.[6]

References

Gastwirth, J.L., 1975. Statistical measures of earnings differentials. Amer. Statist. 29 (1), 32–35.
Lambert, P.J., Aronson, J.R., 1993. Inequality decomposition analysis and the gini coefficient revisited. Econ. J. 103 (420), 1221–1227.
Lambert, P.J., Decoster, A., 2005. The gini coefficient reveals more. Metron 63 (3), 373–400.
Lasswell, T.E., 1965. Class and Stratum. Houghton Mifflin, Boston.
Milanovic, B., 2012. Global inequality recalculated and updated: the effect of new PPP estimates on global inequality and 2005 estimates. J. Econom. Inequal. 10 (1), 1–18.
Milanovic, B., Yitzhaki, S., 2002. Decomposing World income distribution: does the World have a middle class?. Rev. Income Wealth 48 (2), 155–178.
Monti, M., Santori, A., 2011. Stratification and between-group inequality: a new interpretation. Rev. Income Wealth 57 (3), 412–427.
Mookherjee, D., Shorrocks, A.F., 1982. A decomposition analysis of the trend in UK income inequality. Econ. J. 92 (368), 886–902.

3 Lambert and Decoster (2005) state that their attention is confined to the case of two population subgroups "for ease of presentation, but the results can clearly be extended".

4 These regions are described as 'continents' in Milanovic and Yitzhaki (2002) though the correspondence is not exact.

5 Note that this is not the case with the results presented in Monti and Santori (2011) who base their analysis on country-level mean income data.

6 See Milanovic (2012) for further discussion and evidence on trends in between-country inequality.

Pyatt, G., 1976. On the interpretation and disaggregation of Gini coefficients. Econ. J. 86 (342), 243–255.
Yitzhaki, S., 1994. Economic distance and overlapping of distributions. J. Econometrics 61 (1), 147–159.
Yitzhaki, S., Lerman, R., 1991. Income stratification and income inequality. Rev. Income Wealth 37 (3), 313–329.
Yitzhaki, S., Schechtman, E., 2013. The Gini Methodology: A Primer on a Statistical Methodology. In: Springer Series in Statistics, vol. 272. Springer, New York.