Institute of Economic Development and Research

SCHOOL OF ECONOMICS
University of the Philippines

Discussion Paper No. 72-27 February 27, 1973

A NOTE ON "INCOME INEQUALITY AND ECONOMIC GROWTH: THE POSTWAR EXPERIENCE OF ASIAN COUNTRIES"

(Revised)

by

MAHAR Kelly MANGAHAS, 1944-

[illegible]

by [illegible]

This note has two points to make concerning Harry T. Oshima's recent Malayan Economic Review paper [1]:

1. The proposed Index of Decile Inequality is not clearly superior to traditional measures, among which is the Gini ratio, because the index has its own statistical bias, albeit different from the biases of the other measures. In brief, the index is middle-class biased.

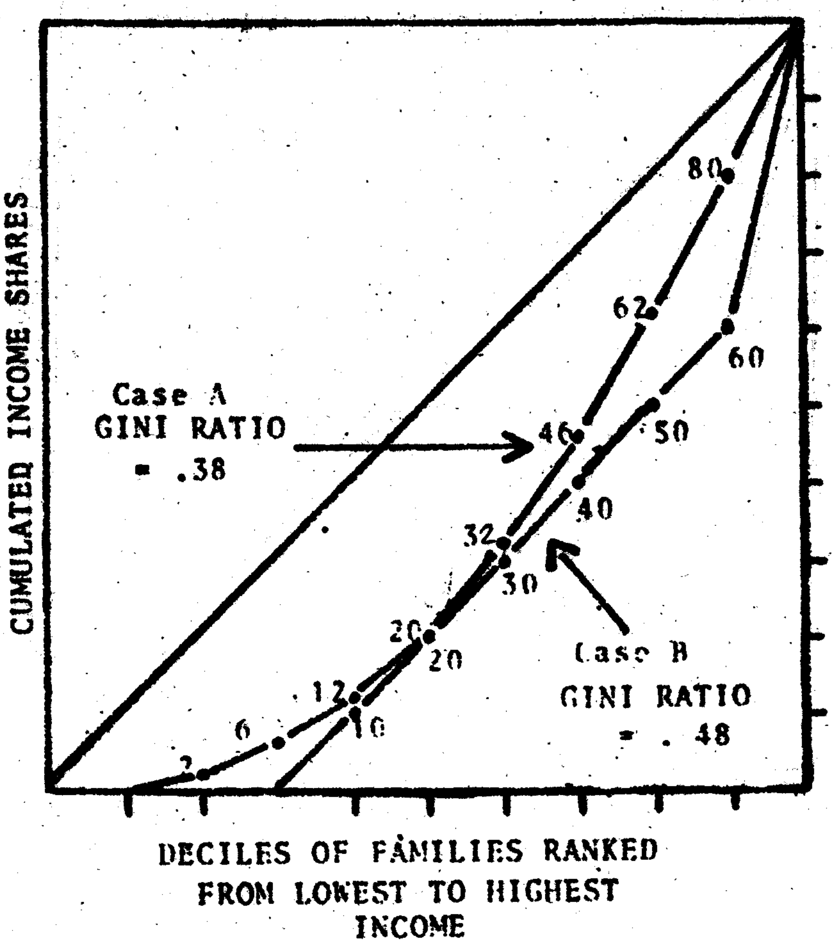
2. In order to accept the decomposition of inequality as to source, therefore, one has to accept the peculiar bias in the absolute-deviations approach. However, a further problem with the decomposition technique used is that it is apt to lead the analyst to conclusions which are inherent in the method rather than in the data.

The Index of Decile Inequality

Deciles are formed by dividing the array of families, arranged according to income in ascending order, into ten equal groups of families. Define d_i as the absolute deviation from 10% of the proportion of total family income held by families in the i^{th} decile. Express d_i in percentage units. The mean absolute deviation is then $\Sigma d_i/10$, and since this expression has range (0, 18), the index of decile inequality is chosen as $\Sigma d_i/180$, which has range (0, 1).

Oshima points out quite rightly that the Gini ratio will tend to be larger than the decile index, because it places relatively high weight on income shares at upper and lower extremes of the distribution. The same observation can be made for measures involving squaring. The decile index, on the other hand, places relatively high weight on the middle-income groups, so that it is not true that it is "without statistical bias [1, p. 11]." This is easiest to show by example:

Decile of Families (Ascending Order)	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	Total
<u>Income share in %</u>											
Case A	0	2	4	6	8	12	14	16	18	20	100
Case B	0	0	0	10	10	10	10	10	10	40	100
Deviation of decile share <u>from 10%, in %</u>											
Case A	10	8	6	4	2	2	4	6	8	10	60
Case B	10	10	10	0	0	0	0	0	0	30	60
Cumulated shares at <u>each decile, in %</u>											
Case A	0	2	6	12	20	32	46	62	80	100	
Case B	0	0	0	10	20	30	40	50	60	100	



TWO LORENZ CURVES HAVING THE SAME
OSHIMA-INDEX OF INEQUALITY (=.30)

These are two distributions having the same index of decile inequality, which is 60/180 = .30, although from the diagram Case B is obviously worse than Case A, as far as ordinary egalitarian standards go. The Gini ratios involved, namely .38 and .48 , are quite in the range of experience, as contrasted to those in Oshima's example (.14 and .18).

In fairness to Oshima, he does say, "...it may be asked if it is not desirable to give greater weight, as the GR does, to extreme values instead of giving any unit of deviation in any portion of an income distribution the same weight, as in the index of decile inequality [1, p. 10]." The point here is that giving the same weight in itself expresses another sort of bias, which may be a less preferable bias to many people.

Oshima's decomposition

Let X_i be the midpoint income level in the k^{th} income class. This is applied to all families in the class, regardless of whether the families belong to Sector One or Sector Two. Families in the i^{th} income class of Sector One number f_{1i} , and those in the i^{th} income class of Sector Two number f_{2i} , the total in the class being $f_i = f_{1i} + f_{2i}$. The absolute deviation from the national mean income $\overline{X}_N$ of income from a representative family from the i^{th} income class is

$$d_i = \left| X_i - \overline{X}_N \right|$$

and the total of such deviations is $\Sigma f_i d_i$ (summing over income classes), which decomposes into

$$(1) \quad \Sigma f_i d_i = \Sigma f_{1i} d_i + \Sigma f_{2i} d_i$$

Then define, for Sectors One and Two respectively,

$$d_{1i} = \left| X_i - \overline{X}_1 \right|$$

$$d_{2i} = \left| X_i - \overline{X}_2 \right|$$

as absolute deviations from sectoral mean income of income from a representative family of the i^{th} income class, where $\overline{X}_1$ and $\overline{X}_2$ are the sectoral mean incomes.

This gives

$$\left.\begin{array}{ll} (2) & \Sigma f_{1i} d_{1i} \\ \\ (3) & \Sigma f_{2i} d_{2i} \end{array}\right\} = \text{sectoral totals of absolute deviations from sectoral means;}$$

and

$$\left.\begin{array}{l} \Sigma f_{1i} d_{1i} / \Sigma f_{1i} \\ \\ \Sigma f_{2i} d_{2i} / \Sigma f_{2i} \end{array}\right\} = \text{per family absolute deviations from sectoral means.}$$

Taking $\Sigma f_i / 2$ as the "standardized" number of families per sector in a ~~two-sector~~ framework, we have

$$\left.\begin{array}{ll}(4) & \dfrac{\Sigma f_{1i} d_{1i}}{\Sigma f_{1i}} \cdot \dfrac{\Sigma f_i}{2} \\ \\ (5) & \dfrac{\Sigma f_{2i} d_{2i}}{\Sigma f_{2i}} \cdot \dfrac{\Sigma f_i}{2}\end{array}\right\} = \text{standardized sectoral totals of absolute deviations from sectoral means}$$

To the r.h.s. of equation (1) we then subtract expressions (2) and (3), add expressions (4) and (5), and work out the appropriate balancing expression to maintain the equality:

$$(6)\quad \Sigma f_i d_i = \underbrace{(\Sigma f_{1i} d_i - \Sigma f_{1i} d_{1i})}_{\text{I}} + \underbrace{\left[\frac{\Sigma f_{1i} d_{1i}}{\Sigma f_{1i}} \frac{\Sigma f_i}{2}\right]}_{\text{II}} + \underbrace{\left[\frac{\Sigma f_{1i} d_{1i}}{\Sigma f_{1i}} \left(\Sigma f_{1i} - \frac{\Sigma f_i}{2}\right)\right]}_{\text{III}}$$

$$+ \underbrace{(\Sigma f_{2i} d_i - \Sigma f_{2i} d_{2i})}_{\text{IV}} + \underbrace{\left[\frac{\Sigma f_{2i} d_{2i}}{\Sigma f_{2i}} \frac{\Sigma f_i}{2}\right]}_{\text{V}} + \underbrace{\left[\frac{\Sigma f_{2i} d_{2i}}{\Sigma f_{2i}} \left(\Sigma f_{2i} - \frac{\Sigma f_i}{2}\right)\right]}_{\text{VI}}$$

Obviously the first three terms of the above sum pertain to Sector One while the last three pertain to Sector Two. Oshima refers to terms (I) and (IV) as <u>contributions to inequality on account of the deviation of the sectoral mean income from the national mean income</u>. Terms (II) and (V) <u>are contributions to inequality on account of the variation or dispersion of the sectoral curve of income distribution, using a standardized sectoral size</u>. And terms (III) and (VI) <u>are contributions to inequality on account of the difference of the actual sector size from the standardized size.</u>

The measure of total dispersion, on the l.h.s., is necessarily a positive number. Oshima measures relative contributions of the sources of inequality by taking ratios of each of the terms (I) to (VI) to $\Sigma f_j d_j$. Note that terms (II) and (V) are always positive, whereas (I), (III), (IV) and (VI) can be negative. Negative terms and corresponding negative ratios are termed by Oshima as contributions to <u>equality</u> rather than inequality. In fact, either term (III) or term (VI), and not both, <u>must</u> be negative; so the conclusion is always drawn that the <u>smaller</u> sector has a relative-size effect which contributes to <u>equality</u>, although this is a conclusion inherent in the method rather than in the data. If Sector One is the smaller, then term (III) is negative, and therefore term (I) will automatically be algebraically smaller than term (II), i.e., the method dictates that for the smaller sector the effect of the sectoral/national difference in means is always smaller than the effect of sectoral variation per se. Conversely, the larger sector <u>always</u> has a relative-size effect which contributes to inequality, and <u>always</u> has a sectoral/national means-effect larger than its sectoral variation effect. Thus a close look at the decomposition technique tells us that the following statement merely means that in Southeast Asia the larger sector is agriculture whereas in East Asia and the U.S. the larger sector is non-agriculture:

> "Accordingly Southeast Asian rural (or agricultural) means and weights contribute to national inequality while in the other countries [East Asia and U.S.] their contribution is mainly to national equality. Since a low mean implies a low rural and agricultural productivity and low productivity requires more agriculturists to feed the national population (hence the interrelation of low means and high frequencies), we conclude that the poverty of the rural-agricultural sector is a major contribution to Southeast Asian inequality, while for East Asia, this is not the case [1, p. 23]."

It is not too clear why a decomposition ought to have a size-effect component distinct from the dispersion and mean-deviation components. The above decomposition clearly intends that the size-component should be zero when sectors are of equal size. But when is the size-component largest? Suppose the sectors have equal variances and means, but are of different size. Should one sector be regarded as being a larger source of inequality than the other? It would not seem so. Equalizing the sector sizes, other things equal, certainly would not reduce inequality either. All things considered, it appears that inequality should be equated with some measure of total dispersion, which can then be decomposed (see below) into measures of dispersion within sectors (e.g. sectoral variance) and among sectors (e.g., deviation of sectoral means from the national mean).

Decomposition of squared deviations

Determination of the contributions of various sources of inequality remains an important task. The purpose of the following final remarks is to demonstrate that one who prefers a quadratic measure of inequality can pursue the task as well.

For two sectors A and B, define

$$(7) \quad SST = \Sigma f_i \, (X_i - \overline{X}_N)^2$$

$$(8) \quad SSA = \Sigma f_{Ai} \, (X_i - \overline{X}_A)^2$$

$$(9) \quad SSB = \Sigma f_{Bi} \, (X_i - \overline{X}_B)^2$$

where f_{Ai} and f_{Bi} are numbers of families in the i^{th} income class of Sectors A and B respectively, and $f_i = f_{Ai} + f_{Bi}$, and $\overline{X}_A$ and $\overline{X}_B$ are the sectoral means. Expressions (7) - (9) are related by

$$SST = SSA + (\overline{X}_A - \overline{X}_N)^2 \, \Sigma f_{Ai} + SSB + (\overline{X}_B - \overline{X}_N)^2 \, \Sigma f_{Bi},$$

from which follows

$$(10) \quad \frac{SST}{\Sigma f_i} = \frac{\Sigma f_{Ai}}{\Sigma f_i} \, \frac{SSA}{\Sigma f_{Ai}} + \frac{\Sigma f_{Ai}}{\Sigma f_i} \, (\overline{X}_A - \overline{X}_N)^2 + $$

$$+ \frac{\Sigma f_{Bi}}{\Sigma f_i} \, \frac{SSB}{\Sigma f_{Bi}} + \frac{\Sigma f_{Bi}}{\Sigma f_i} \, (\overline{X}_B - \overline{X}_N)^2 .$$

This decomposes total variance into two sectoral variance terms and two sectoral mean-deviation terms, each properly weighted. There are no negative terms in the decomposition. If one chooses, the X's can stand for logs of income intead of income itself.

REFERENCE

[1] Harry T. Oshima, "Income Inequality and Economic Growth: The Postwar Experience of Asian Countries," Malayan Economic Review, Vol. XV, no. 2 (October 1970), pp. 7-41.