

ECONOMIC GROWTH CENTER

YALE UNIVERSITY

Box 1987, Yale Station
New Haven, Connecticut

CENTER DISCUSSION PAPER NO. 263

DECOMPOSING LDC INEQUALITY

Gary S. Fields

June 1977

Note: Center Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Discussion Papers should be cleared with the author to protect the tentative character of these papers.

DECOMPOSING LDC INEQUALITY

I. Introduction

At the present time, there is great interest among development economists in the problem of economic inequality in less developed countries (LDCs). Studies of the determinants of inequality follow either of two general approaches. The more traditional approach is associated with names like Kuznets (1963), Chenery and associates (1960, 1968, 1975), Adelman and Morris (1973), Ahluwalia (1976) and Chiswick (1971). These studies share a common methodology, consisting basically of looking at a cross-section of countries, and (1) measuring the degree of inequality in each, (2) measuring other characteristics of each country (e.g., level of GNP, its rate of growth, importance of agriculture in total product, etc.), and (3) relating the level of inequality to that economy's characteristics using correlation or regression analysis.

In the last few years, another type of approach has been followed, which looks instead at inequality within a country, and measures the contribution of the various components to total inequality.¹ In this type of approach, using a variety of methodologies, inequality has been decomposed by economic sector (e.g., urban vs. rural), income source (e.g., income from labor vs. capital vs. land vs. transfers), or family characteristics (including attributes of the workers, their jobs, and regional and other locational considerations). This mode of inquiry is potentially of great value for understanding the structure of inequality and identifying which are the most important explanatory factors.

¹The decomposition studies in LDCs include works by Ayub (1977), Chiswick (1976), Fei and Ranis (1974), Fei, Ranis, and Kuo (1977), Fields (forthcoming), Fields and Schultz (1977), Fishlow (1972, 1973), Langoni (1972, 1975), Mangahas (1975), Mehran (1974), Pyatt (1976), Theil (1967, 1972), Uribe (1976), and Van Ginneken (1975).

This study explores the decomposition type of inequality analysis. I summarize the alternative decomposition methodologies which have been set forth in the literature and review the principal findings of empirical studies.

II. Types of Decomposition Problems

Decomposition problems are of three general types: functional decomposition by income source, functional decomposition by economic sector, and microeconomic decomposition by income-determining characteristics. Let us now review each.

A. Decomposition by Income Source

The starting point for source decompositions is the assumption that income determination can best be studied by disaggregation into a small number of functional income sources. Take as an example the familiar functional division of income into income from labor, income from capital, and (at the micro level) income from transfers. The question asked by source decompositions is: of total inequality, how much is attributable to income from labor, how much to income from capital, and how much to income from transfers? Source decomposition procedures quantify these effects and further show how each source's contribution to overall inequality depends positively on the degree of inequality of each income source, the importance of that income source in total income, and the extent of correlation between income from that source and total income.

B. Decomposition by Economic Sector

Sectoral decompositions divide the economy into economic sectors (e.g., agriculture vs. non-agriculture). Generally, these sectors are thought to be mutually exclusive, so that all of the household's income is treated as agricultural or non-agricultural. The question asked by sector decompositions is: of total inequality, how much is attributable to variability in agricultural incomes, how much to variability in non-agricultural incomes, and how much to between-sector inequality?

Sector and source decompositions have been presented independently, here, as is the practice in the literature. This distinction, though convenient, is not necessary. The economy could very easily be divided into segments defined by source-sector combinations, e.g., rural labor income, urban labor income, rural capital income, and so on.

Source and sector decompositions have in common the property that total inequality is completely accounted for by the several components, in much the same way that total national income is completely accounted for by summing income from consumption, investment, government expenditures, and net exports. The characterization of source and sector decompositions as accounting procedures is deliberate. For just as decompositions of national income into consumption, investment, government, and export components cannot explain why national income was what it was, neither can source and sector decompositions explain why national income inequality was what it was. The value of these decompositions is that they gauge the relative importance of various sources and sectors in respect to overall inequality, and thereby direct our attention to potentially fruitful areas of research.

Suppose, for instance, we find, as indeed the data show, that the primary contribution to overall income inequality is made by variation in labor income. This suggests that a valuable next step in understanding overall income inequality would be to study those economic forces which might determine the amount and distribution of labor income. In this connection, many characteristics of family members and their jobs become important. Note that microeconomic data on the individual households and their family members are needed to explore the determinants of income from labor or any other source or sector. Let us now consider what types of decompositions can be performed when such microeconomic data are available.

C. Decomposition by Income Determinants

A now large number of studies of less developed countries have shown that households' overall incomes and labor market earnings are systematically related to a number of family characteristics: the number of labor force participants, their incidence of unemployment, their personal characteristics (such as education and age), the family's location (by region, size of place, or rural vs. urban), the nature of their jobs (including occupation, industry, and employer's characteristics).¹

In a few of these studies (see Section V.C below), attempts have been made to decompose income inequality according to income determinants.

Determinant decompositions ask the question: of total inequality, how much inequality is associated with variation in income determinant 1, how much with income determinant 2, etc. and how much is not associated with any of the explanatory variables? The presence of an unexplained component is one important difference between the determinant decompositions and the other types of decompositions. Another important difference is that determinant decompositions provide much more insight into causal factors underlying the distribution of income than is the case with decompositions by source and/or sector.

We now turn to the different types of decomposition methodologies.

¹Among these studies are Fields (1976), McCabe(), Langoni (1975), Johnson (1971), and Chiswick (1976).

III. Decomposition Methodologies

Three different decomposition methodologies are in current use: Gini decompositions, Theil decompositions, and the analysis of variance. We consider these in turn.

A. Gini Decomposition

The Gini coefficient is the most popular measure of relative income inequality, owing to the ease of interpreting it vis-a-vis the Lorenz curve. Gini decomposition procedures have been devised independently by Fei and Ranis (1974, 1977), Pyatt (1974), and Mangahas (1974); in addition to the empirical applications by these authors, Gini decompositions have been applied in research by Mehran (1974) Ayub (1977), and Fields (forthcoming).

For purposes of discussion, let us suppose there are three income sources -- wage income, property income, and transfer income -- and that the sum of these is the total income for each family and for the economy as a whole.

Using the Gini coefficient as our measure of inequality, it might be thought that the overall Gini for the economy as whole would be a weighted average of the Ginis for the individual components, the weights being given by the factor share of that income in the total. This is, however, incorrect, because the Gini coefficient requires the households to be ranked in increasing order of income and the different component incomes (wage, property, transfer) may not be monotonically related to one another or to the total.

To indicate the correct relationship between the overall Gini coefficient and the factor Ginis, let us order the families according to total income. For each factor income source, we may then compute a so-called pseudo-Gini coefficient, i.e., the Gini coefficient that would

be obtained if households in that sector were not ordered with their incomes monotonically increasing. The overall Gini for the economy (G) turns out to be a weighted average of the pseudo-Ginis for the i 'th income source ($\bar{G}_i$) with the weights given by the factor share of that income sources (ϕ_i):

$$(1) \quad G = \bar{G}_1 \phi_1 + \bar{G}_2 \phi_2 + \bar{G}_3 \phi_3.$$

Fei, Ranis, and Kuo (1977, Chapter 7) have shown that the pseudo-Gini for the i 'th source ($\bar{G}_i$) is equal to the product of the true Gini for that source (G_i) and a relative correlation coefficient (R_i), defined below:

$$(2) \quad \bar{G}_i = G_i R_i.$$

For each factor, the relative correlation coefficient is the ratio of two other correlations:

$$(3) \quad R_i = \frac{\text{cor}(Y_i, \rho)}{\text{cor}(Y_i, \rho_i)} = \frac{\text{coefficient of correlation between factor income amount and total income rank}}{\text{coefficient of correlation between factor income amount and factor income rank}}.$$

To further explain (3), consider the R_i for labor income. The numerator of (3) is the correlation between labor income in dollars (Y_i) and the family's total income position (ρ), ordered from lowest to highest. The denominator of (3) relates the dollar labor income figure (Y_i) to that family's labor income rank (ρ_i).

Substituting (2) and (3) into (1) and dividing through by G , we obtain:

$$(4) \quad 100\% = \phi_1 \frac{G_1}{G} \frac{\text{cor}(Y_1, \rho)}{\text{cor}(Y_1, \rho_1)} + \phi_2 \frac{G_2}{G} \frac{\text{cor}(Y_2, \rho)}{\text{cor}(Y_2, \rho_2)} + \phi_3 \frac{G_3}{G} \frac{\text{cor}(Y_3, \rho)}{\text{cor}(Y_3, \rho_3)},$$

$$= \text{FIW}_1 + \text{FIW}_2 + \text{FIW}_3,$$

the FIW's denoting the so-called Factor Inequality Weights of labor, property, and transfer income respectively. Overall inequality in an economy is seen to depend on the degree of inequality of each income source, the extent of correlation between income from ~~that~~ source and total income, and the importance of that income source in the total.

Other decomposition procedures partition total inequality differently. These are reviewed below.

B. Theil Decomposition

A decade ago, Theil (1967) set forth a readily-decomposable inequality measure, which he subsequently (1972) illustrated with a number of empirical applications. Because an exact decomposition is possible, the Theil index has received widespread use. Among the studies of LDCs performing Theil decompositions are those by Fishlow (1972), Van Ginneken (1974), Chiswick (1976) and Uribe (1976).

The Theil index of inequality is derived rigorously from the notion of entropy in information theory. The fundamental idea of information entropy is that occurrences which differ greatly from what was expected should receive more weight than events which conform with prior expectations. The entropy index gauges the expected information content from the various outcomes, with the weights depending on the likelihood of each.

Building on this concept of entropy, the Theil index (T) of income inequality is formally the expected information of the message which transforms population shares into income shares. Mathematically, its algebraic formula is given by

$$(5) \quad T = \sum_{i=1}^n q_i \log \frac{q_i}{1/n},$$

where n = number of individuals or households,

q_i = income share of i 'th individual.

Theil (1972, p. 100) notes that T equals the mean product of income and its own logarithm. Why this should be used as measure of economic inequality is far from transparent.

In any case, the main attraction of the Theil index lies not in its intuitive justification but rather, as remarked above, in its decomposability. Theil decompositions are well-suited for estimating the contribution of different groups to total inequality; examples of such groups are economically distinct regions of a country or population subgroups divided into educational and/or age categories.

Various decomposition formulas are given in Theil (1972, p. 100), Chiswick (1976, p. 9), and Fishlow (1972, p. 395) among other places. Fishlow, for instance, gives two alternative decomposition procedures:

$$(6) \quad I_{ijk} = \sum_i y_{i..} \log \frac{y_{i..}}{x_{i..}} + \sum_i y_{i..} \left\{ \sum_j \frac{y_{ij.}}{y_{i..}} \log \frac{y_{ij.}/y_{i..}}{x_{ij.}/x_{i..}} \right\} + \sum_i \sum_j \frac{y_{ij.}}{y_{i..}} \left\{ \sum_k \frac{y_{ijk}}{y_{ij.}} \log \frac{y_{ijk}/y_{ij.}}{x_{ijk}/x_{ij.}} \right\}$$

and

$$(7) \quad I_{jk} = \sum_j y_{j.} \log \frac{y_{j.}}{x_{j.}} + \sum_k y_{.k} \log \frac{y_{.k}}{x_{.k}} + \left\{ \sum_j \sum_k y_{jk} \log \frac{y_{jk}}{x_{jk}} - \sum_j y_{j.} \log \frac{y_{j.}}{x_{j.}} - \sum_k y_{.k} \log \frac{y_{.k}}{x_{.k}} \right\},$$

where y are the income shares, x the population shares, and the subscripts i , j , and k refer to income class, sector, and education. Equation (6) decomposes total inequality into between-group and within-group components, while (7)

decomposes the between-group component according to the variation among the means of the various groups.

Another decomposition procedure, substantially similar in nature, is the analysis of variance, which we now examine.

C. Analysis of Variance (ANOVA)

ANOVA procedures have a long history in social scientific analysis, but their applications to economic problems are quite limited. In particular, on the problem of economic inequality, work is just beginning; see Schultz (1965), Langoni (1972, 1975), Chiswick (1976), Fields (forthcoming), and Fields and Schultz (1977).

The basic idea of analysis of variance is to decompose the variance of a dependent variable, which is the sum of squared deviations from the over-all mean, into two types of effects: those due to variation between different groups and those due to variation within each of the groups. For example, if the dependent variable is income or its logarithm in each of a number of households and the independent variable is the region of the country in which they live, the total sum of squares (SS) of income is expressed as:

$$(8) \quad SS_y = SS_{\text{between}} + SS_{\text{within}}$$

$$\text{where } SS_y = \sum_j \sum_i (Y_{ji} - \bar{Y})^2$$

in which $\bar{Y}$ is the overall mean of income Y in the entire sample, the i 's are households, and the j 's are various regions

$$SS_{\text{between}} = \sum_j N_j (\bar{Y}_{j.} - \bar{Y})^2$$

regions

in which $\bar{Y}_{j.}$ is the mean income in region j , and N_j is the number of sample households in region j

$$\text{and } SS_{\text{within}} = \sum_j \sum_i (Y_{ji} - \bar{Y}_{j.})^2$$

regions

In other words, equation (8) tells us the relative importance of income inequality within regions as compared with diversity in mean incomes across regions.

In the example of the preceding paragraph, the only explanatory factor was region. ANOVA may also handle multiple explanatory variables, say region and education. We then obtain a breakdown such as:

$$\begin{aligned} (9) \quad SS_y &= \text{SS due to region} \\ &+ \text{SS due to education} \\ &+ \text{SS due to interaction between region and education} \\ &+ \text{SS within region-education groupings.} \end{aligned}$$

A decomposition like (9) tells us whether income inequality is greater across regions or across educational groups, whether the effects of region and education on income are independent of one another, and the relative importance of variations across these groupings as compared with the variations within them. Both gross and marginal effects may be estimated. Additionally, and quite importantly, tests of statistical significance are available for each factor.

Another characteristic of analysis of variance techniques is that because they are very much like multiple regressions they indicate the quantitative importance of each category of the explanatory variables. Thus, we can learn from ANOVA decompositions how much difference it makes to one's income if the family is located in one region rather than another or some family member has more education rather than less.

In sum, this is what analysis of variance procedures can do:

1. Decompose overall inequality into within-factor and between-factor components;

2. Measure the gross contribution of each explanatory factor to total inequality;
3. Test the statistical significance of these main effects;
4. Measure the marginal contribution of each explanatory factor;
5. Test the statistical significance of the marginal effects;
6. Measure the effects of interactions between pairs of explanatory factors (and higher order combinations if needed);
7. Test the statistical significance of the interaction effects;
8. Estimate the magnitude of each category of each explanatory variable to income.

Theil decompositions do only 1, 2, 4, and 6 and Gini decompositions only 1 and

2. Thus, in comparison with other available decomposition procedures, ANOVA provides richer information on the sources of inequality.

Let us now take up a number of other considerations which are relevant to the choice of decomposition procedure.

IV. Choice of Decomposition Procedure

In weighing the advantages of the various decomposition procedures for empirical research, two central issues arise: the properties of the inequality measure itself, and the suitability of the measure for the different decomposition problems.

A. Properties of the Different Measures

One way of choosing an inequality measure is to consider the measure's basic nature. In this respect, the Gini decomposition and the analysis of variance come out ahead. The Gini coefficient is easily conceptualized in terms of the Lorenz curve, while the variance has a familiar basis in standard statistical analysis. In contrast, the Theil index, as a measure of inequality, has no clear interpretation.

Another selection criterion is the usefulness of the inequality measure in making inequality comparisons. Among the desirable axioms for this purpose are:¹

A1. Axiom of Scale Irrelevance. If one distribution is a scalar multiple of another (i.e., everyone's income in the first case is $x\%$ of their income in the second), then the two distributions have the same degree of inequality. Put somewhat differently, the degree of inequality in the distribution of income is measured independently of the level of income.

A2. Axiom of Symmetry. If two income distributions are identical except that different families receive the income in the two cases, then the two distributions have the same degree of inequality. This follows from the principle of treating all individuals and families alike with regard to income distribution.

¹See Fields and Fei (1978) for an axiomatic development and, for an even more individualistic set of social welfare judgments, Atkinson (1970).

A3. Axiom of Rank-Preserving Equalization. If one distribution is obtained from another by the transfer of a positive amount of income from a relatively rich family to a relatively poor one while preserving their relative rank in the distribution, then the new distribution is more equal than the old. (While few persons are likely to quarrel with this axiom, it should be noted that some additional, non-trivial assumptions about the nature of judgments of social well-being are necessary to guarantee that a "more equal" distribution is always regarded as "better.")

The Gini coefficient and Theil index satisfy these axioms. The variance does not fulfill the Axiom of Scale Irrelevance. However, Scale Irrelevance is satisfied by the variance of the logarithm of income (commonly known as the log-variance). Since the logarithm of income is used in other branches of income distribution research, particularly in earnings functions, ANOVA seems as suitable by the axiomatic criterion for decomposition analysis as are the Gini coefficient and Theil index.

Another consideration of some importance is the sensitivity of the different measures to income changes at various points in the distribution. Persons whose value judgments lead them to give greatest weight to the economic position of the poor may wish to choose that inequality measure which is most sensitive to inequality associated with low income groups. Observations on the several inequality measures may be found in Sen (1973), Weisskoff (1970), Szal and Robinson (1975), and Chiswick (1976) among others, but perhaps the most thorough analysis of this question is in the work of Champernowne (1974). He found, among other things, that the variance of the logarithms of income is most sensitive to inequality associated with poverty, the Theil index is

most sensitive to inequality associated with the very rich, and the Gini coefficient is most sensitive to inequality in the middle of the income distribution. For observers whose main concern is with the low income population, analysis of variance procedures would appear more appropriate on this basis.

All in all, analysis of variance procedures based on the logarithms of income have a number of inherently desirable properties including an axiomatic justification, sensitivity to inequality associated with poverty, quantitative estimation of the magnitude of income determinants, and decomposability. The other decomposition measures are less powerful. We now consider the suitability of the various decomposition procedures for the different problems of interest.

B. Different Decomposition Measures for Different Problems

Consider first the problem of decomposing inequality by functional income source. As described above, procedures for using the Gini coefficient for this problem have been worked out in considerable detail. Particularly helpful is the technique for constructing Factor Inequality Weights and the breakdown of those weights into factor share, factor Gini, and correlational components (see equation (4)). In principle, ANOVA and Theil procedures could be decomposed similarly, but they have not yet been used in this way.

For the sectoral decomposition problem, which analyzes between- and within-sector inequality, each of the three procedures appears satisfactory. The choice among them is therefore partially dependent on the properties discussed in Section A above, and in part a matter of convenience (depending, for example, on the availability of computer programs for the different procedures).

Finally, with respect to decompositions by income-determining factors, ANOVA and Theil techniques come out ahead. The strength of these procedures is that they give a clear picture of the importance of each explanatory factor in determining overall inequality, while at the same time gauging the unexplained residual. Gini decompositions, on the other hand, deal with deviations from predicted values in a quite cumbersome way, the difficulty being inherent in the Gini coefficient itself.¹

In the income determinant problem, how do we choose between analysis of variance and Theil decompositions? I would say that two considerations work strongly in favor of ANOVA. One is the use of log-variance as the measure of inequality. The parallel between ANOVA and multiple regressions explaining the logarithm of income permits a richer characterization of the income determination process than does Theil.² A second overriding consideration is the availability of statistical significance tests for ANOVA but not for Theil. Thus, using ANOVA, we can measure the likelihood that the estimated contribution of an explanatory variable like region or education is a "true" effect compared with the alternative possibility that the apparent relationship is due to chance sampling. This permits us to bring the full logic of conventional statistical analysis to bear on the problem of

¹From equation (1), for an exact Gini decomposition, we must calculate the Gini coefficient of the residual errors ϵ_i in the linear model $Y_i = \alpha_i + \sum_j \beta_j X_{ij} + \epsilon_i$. But roughly half the ϵ_i are negative. The Gini coefficient of a variate with negative values is undefined.

²See Fields and Schultz (1977) for a direct combination of ANOVA and regression results.

ascertaining the determinants of inequality. From a causal (as versus an accounting) perspective, this is valuable indeed.

In sum, on the choice of decomposition procedures for the types of problem under consideration, we may conclude: (1) The Gini decomposition technique is a proven method for the source problem; (2) For the sector problem, the choice of technique is a matter of some indifference, possibly, the available computer software proving decisive, and (3) Analysis of variance dominates for decomposing inequality into the contributions of various determinantal factors.

V. Survey of Empirical Findings in LDCs

The various techniques for decomposing inequality have been applied to analyses of the structure of inequality by income source, economic sector, and income-determining characteristics in a number of LDCs.¹ Some patterns seem to be emerging from these studies. This section reviews the major results.

A. Source Decompositions

The pioneering work on the source decomposition problem is that of Fei and Ranis (1974) and Fei-Ranis-Kuo (1977) in their study of Taiwan. Their methodology was followed in subsequent research on Pakistan by Ayub (1977) and on Colombia by Fields (forthcoming).

The source decompositions are based on the Gini coefficient. Taiwan's overall Gini is 0.28, which is among the lowest of all countries in the world.² The source decomposition tells us which of five income sources (wage, mixed,³ property, gifts, and other) accounts for how much of the overall inequality. The natural place to start is by looking at the Gini coefficients of the individual income sources. In the absence of micro-economic data, these were computed across income groups. The results are given in Table 1. Fei and Ranis report that property and gift income have the highest factor Ginis and therefore are least equally distributed, mixed and other are in an intermediate position, while wage income is most equally distributed. From this, we might be inclined to conclude that property and gift income account for the largest part of overall

¹There is also some literature examining changes over time in one more of these problems (e.g., Fei-Ranis-Kuo (1977), Ayub (1977)) but that work lies outside the purview of this paper.

²Cf. Paukert (1973).

³Mixed income includes agricultural income, business income, and similar mixtures of returns to capital and labor.

inequality and wage income the least. In actuality, these inferences would be mistaken, the reason being that we have omitted two important factors from consideration, namely, (1) the factor shares, which tell us the importance of that factor in total income, and (2) the correlations between factor income and total income, which tell us whether that factor augments total inequality or offsets the inequality attributable to other sources.

When one looks at the factor shares in Row 2 of the Table, wage income is seen to be the most important source of income by far, mixed income is in an intermediate position, and property and gift income are relatively unimportant. As the decomposition procedure (equation (4)) showed, total inequality is a weighted average of inequality in the individual factor incomes. In the case of Taiwan, wage income is relatively equally distributed but has the largest factor share, property and gift income are relatively unequally distributed but have small factor shares, and mixed and other sources are in the middle in both respects.

The Factor Inequality Weights presented in Row 3 measure each factor's contribution to total inequality. The data show that wage income is the source of more than half of total inequality, while property and gifts combined account for less than 20%; the rest is accounted for by mixed income, some substantial but unknown part of which reflects returns to labor.

The same basic decomposition methodology has been applied to the cases of Pakistan and urban Colombia with quite similar results. Both Ayub (1977) and Fields (forthcoming) report: (1) The highest factor Gini coefficients for non-labor income sources than for labor incomes;¹

¹In Pakistan, non-labor income refers to income from property. In Colombia, income from capital and income from transfers are distinguished, capital income including an imputation for the value of owner-occupied housing.

Table 1

Decomposition of Inequality in Taiwan, 1972

	<u>Wage</u>	<u>Mixed</u>	<u>Property</u>	<u>Gifts</u>	<u>Other</u>	<u>Total</u>
1 Factor Gini	.2518	.2968	.4020	.3965	.2925	
2 Factor Share	.582	.275	.093	.046	.004	1.000
3 Factor Inequality Weight	.5187	.2882	.1322	.0584	.0024	1.000

Source: Fei and Ranis (1974).

(2) The reverse ordering for factor income shares;¹ and (3) The overwhelming importance of labor incomes (including wage employment and self-employment) in accounting for overall inequality.²

Individually and together, the results for Taiwan, Pakistan, and Colombia give a common impression about the contribution of the various income sources to overall inequality: the bulk of income inequality is attributable to labor income. The high Factor Inequality Weights for labor incomes suggest that the principal inequality-producing factor is that some people receive a great deal more income for their work than do others. The intuitive prior notion that the most unequally-distributed factors contribute the most to total inequality is found to be false in each case.

B. Sector Decompositions

Sector decomposition studies do three things: they measure the inequality within each sector or region of an economy, indicate the importance of within-sector inequality for all sectors taken together, and determine the amount of inequality accounted for by between-sector variation. The available studies decompose inequality within a country and within regions of the world.

Within-country sector decompositions have been carried out using the Gini coefficient by Mehran (1974) for Iranian cities, by Mangahas (1975) for areas and regions of the Philippines, and by Pyatt (1976) for urban and rural locations in Sri Lanka. In other studies---by Fishlow (1972, 1973) and Langoni (1972, 1975) in Brazil, van Ginneken (1975)

¹Self-employment income in Pakistan accounted for 65% of total income in 1971/72, and an additional 19% was provided by wages and salaries. In urban Colombia, they were 35% each in 1967/68.

²Factor Inequality Weights were not computed for Pakistan. For urban Colombia, they were: labor income, 69%; capital income (including imputed rent), 27%; and transfer income, 4%.

in Mexico, Chiswick (1976) in Thailand, and Fields (forthcoming) and Fields and Schultz (1977) in Colombia---regional or urban-rural decompositions were undertaken as part of a larger exercise; these studies used Theil decompositions or analysis of variance. Without exception, the result emerges that variations within sectors or regions are far more important in accounting for inequality than variations between sectors.

Another result of the within-country sector decompositions is that inequality is found to be greater within urban than within rural areas. See, for example, Mangahas (1975, p. 295) for the Philippines, Pyatt (1976, Table 3) for Sri Lanka, Fei-Kuo-Ranis (1977, Diagram 2) for Taiwan, Ayub (1977, Table XII) for Pakistan, and Fields and Schultz (1977, Table 4) for Colombia. These results accord with the findings of Kuznets (1955) and many other income inequality studies.

Sector decompositions have also been applied to studies of inequality in the world. First Theil (1972) and after him Uribe (1976) using the same methodology examined the structure of inequality within a number of countries and across countries. Theil's analysis covered all parts of the world, while Uribe's was limited to Latin America only. Both studies found more inequality within countries than across them.

In summary, the sector decomposition studies report more inequality within sectors or countries than across them. As with the source decomposition literature, these studies clearly demonstrate the importance of going down to the household level in order to understand the determinants of incomes and income inequality.

C. Determinant Decompositions

Seven studies decomposing inequality in less developed countries by income determinants are in existence. The countries covered are Brazil

(two studies), Mexico, Thailand, Taiwan, and Colombia (two studies). Each of the three statistical decomposition methodologies have been used. The results of these studies are summarized in Table 2.

The available studies exhibit several similarities: (1) Greater effects are found for personal attributes than for employment or locational aspects. (2) Of the personal attributes considered, education and age contribute roughly equal explanatory power. (3) Regional effects are found to be of some importance, but these effects are not major ones. (4) Intra-regional inequality dominates inter-regional inequality.

The considerable importance of personal attributes in the decomposition studies and the lesser importance of employment and locational information accords with the findings of income- and earnings-generating functions; see, for example, Fields (1976) and the references cited therein, McCabe (), Langoni (1975), Johnson (1971), and Chiswick (1976). In those studies, personal characteristics were found to explain as much as 60% of the variance in the logarithms of income, while little was gained by adding information on the employer or the place of residence.

Other sources also suggest the limitations of analyses of income distribution at the sectoral level. Webb (1976), for instance, reports that the poor in Lima are found scattered in many different sectors--commerce, manufacturing, transport, construction, public service, modern sector firms or occupations, and miscellaneous services--each sector containing at least 10% of the poor. More generally, it would appear that to predict an individual's income, we can do much better knowing his education and age than which economic sector he is located in and

Table 2

Decomposition of Inequality in Five Less Developed Countries

By Income Determinants, Major Findings

Study, Country and Decomposition Methodology	Factors Considered, in Order of Importance			
Fishlow (1972)-Brazil, 1960 Theil Decomposition	1. Education 2. Age 3. Sector 4. Region			
Langoni (1975)-Brazil 1960 & 1970 Multiple Regression Approach to Analysis of Variance (ANOVA)		1960 Rank	1970 Rank	
	Education	1	1	
	Region	2	3+	
	Age	3	2	
	Sex	4	3+	
	Activity	5	3+	
van Ginneken (1975)-Mexico, 1968 Theil Decomposition	1. Education 2. Urban-Rural 3. Age 4. Sector of Activity 5. Occupation			
Chiswick (1976) - Thailand, 1971 Theil Decomposition and Analysis of Variance (ANOVA)	Theil, Gross Explanation	Theil, Marginal Explanation	ANOVA, Gross Explanation	ANOVA, Marginal Explanation
	Age	1	2**	3**
	Education	2	1**	1**
	Sex	3	3*	2**
	Farm Occup.	4	4x	4**
	Type of Earnings	5	5x	5**
Fei, Ranis, and Kuo (1977) Taiwan, 1966 Gini Decomposition	1. Education 2. Age 3. Sex			
Fields (forthcoming)--Colombia 1967/68 Analysis of Variance (ANOVA)	1. Education 2. Age 3. City			
Fields and Schultz (1977)-- Colombia, 1973 Analysis of Variance (ANOVA)	1. Education** 2. Age** 3. Region** 4. Urban/Rural** 5. Type of Employment**			

Notes to Table 2

** = Statistically significant effect at .01 level

* = Statistically significant effect at .05 level

x = Not statistically significant effect at .05 level

If no **, *, or x appears, no test of statistical significance is possible.

† Marginal contributions of these variables were virtually identical.

whether his income comes only from his labor or whether he has property and/or transfer income also.

Decompositions of inequality by income-determining characteristics, such as those summarized in Table 2, are potentially of great usefulness in analyzing LDC wage structures. Economic theory does not yet offer a comprehensive explanation for income inequality. However, we do have partial explanations based on considerations of labor demand, labor supply, technological variability, and institutional influences. Attempts to integrate these various strands of analysis into a unified theory of the determinants of wages and size distribution of income and to implement such a theory empirically have met with only partial success.¹ The empirical results of decomposition studies may aid in the inductive development of a more comprehensive view of this vitally important process.

¹Particularly interesting in this regard in an LDC context is the study by Heady (1976).

VI. Conclusions

This paper has considered three types of decompositions of inequality and three methodologies for decomposition analysis and reviewed the findings from empirical studies in less developed countries. Several methodological and empirical conclusions emerge:

(1) The three different decompositions (by functional income source, by economic sectors, and by income-determining characteristics) are basically quite different. The first two types of decompositions give a total accounting for inequality, whereas determinant decompositions allow for an unexplained residual component. Also, source and sector decompositions are of an accounting nature, while determinant decompositions are causal. Finally, an important difference between source decompositions and sector decompositions is that many households receive income from more than one source, but not ordinarily from more than one sector.

(2) The various decomposition methodologies (by Gini coefficient, Theil index, and analysis of variance) are suited for different types of problems. For the source problem, the Gini decomposition technique is a proven method. In analysis of inequality within and among mutually exclusive sectors, any of the available techniques will serve satisfactorily, although if tests of statistical significance are of interest, analysis of variance may be preferable. For gauging the causal importance of various explanatory factors, analysis of variance can do more than either of the other approaches. ANOVA may also be preferred for its greater sensitivity to income inequality associated with the poverty population.

(3) Source decomposition studies point to variation in labor incomes as the predominant factor accounting for income inequality. To understand the structure of income inequality in LDCs, knowledge of the determinants of incomes from wages and self-employment becomes paramount, as does an understanding of the functioning of LDC labor markets.

(4) Sector decomposition studies indicate substantially more inequality within regions than across them. This implies the need to look within regions for other sources of income variability, at the level of either the worker or his job. Empirically, simple dualistic models will not do.

(5) From studies which decompose inequality by income-determining characteristics, we find that more inequality is attributable to variation in personal characteristics than to the sector of employment or locational aspects. The most powerful personal characteristics explaining inequality are education and age. Economic sector and location make some contribution to explaining inequality, but these variables have lesser effects.

(6) Singly and together, decomposition studies in less developed countries lead to an inescapable conclusion: the overwhelming importance of income variation according to attributes of individuals and the secondary role of variation between economic segments grouped according to sector of the economy or functional income source. Given this overall conclusion, the need for further microeconomic income determination studies at the level of the household stands out. Sectoral considerations may have a role to play in determining LDC inequality too, explaining why some individuals with a given set of personal attributes (education, age, sex, etc.) receive higher incomes than others. These studies, when combined with more macro-economic analyses, may shed some light on the systemic forces generating inequality in LDCs.

BIBLIOGRAPHY

- Adelman, I. and Morris, C.T., Economic Growth and Social Equity in Developing Countries, Stanford, 1971.
- Ahluwalia, M.S., "Inequality, Poverty and Development," Journal of Development Economics, December, 1976.
- Atkinson, A.B., "On the Measurement of Inequality," Journal of Economic Theory, Vol. 2, 1970.
- Ayub, M., Income Inequality in a Growth-Theoretic Context: The Case of Pakistan, Unpublished doctoral dissertation, Yale University, 1977.
- Champernowne, D.G., "A Comparison of Measures of Inequality of Income Distribution," The Economic Journal, December, 1974.
- Chenery, H.B., "Patterns of Industrial Growth," American Economic Review, September, 1960.
- Chenery, H.B. and Taylor, L., "Development Patterns: Among Countries and Over Time," Review of Economics and Statistics, November, 1968.
- Chenery, H.B. and Syrquin, M., Patterns of Development, 1950-1970, London, 1975.
- Chiswick, B., "Earnings Inequality and Economic Development," Quarterly Journal of Economics, February, 1971.
- Chiswick, C.U., "Application of the Theil Index to Income Inequality," I.B.R.D. Working Paper Series B-2, July, 1976.
- Fei, J.C.H. and Ranis, G., "Income Inequality by Additive Factor Components," Economic Growth Center, Yale University, Center Discussion Paper No. 207, June, 1974.
- Fei, J.C.H., Ranis, G., and Kuo, S.O., Equity with Growth: The Taiwan Case, unpublished manuscript, 1977.
- Fields, G.S., "Education and Economic Mobility in a Less Developed Country," Economic Growth Center, Yale University, mimeo, June, 1976.
- Fields, G.S., "Income Inequality in Urban Colombia: A Decomposition Analysis," Economic Growth Center, Yale University, Discussion Paper, forthcoming.
- Fields, G.S. and Schultz, T.P., "Sources of Income Variation in Colombia: Personal and Regional Effects," Economic Growth Center, Yale University, Center Discussion Paper No. 262, June, 1977.
- Fields, G.S. and Fei, J.C.H., "On Inequality Comparisons," Econometrica, January, 1978.
- Fishlow, A., "Brazilian Size Distribution of Income," American Economic Review, Vol. 62, No. 2, May, 1972.

- Fishlow, A., "Distribuicao da Renda no Brasil: Um Novo Exame," Dados, No. 11, 1973.
- Heady, C., "Alternative Theories of Wages in Less Developed Countries: An Empirical Test," Economic Growth Center, Yale University, Discussion Paper No. 254, October, 1976.
- Johnson, G., "The Determination of Individual Hourly Earnings in Urban Kenya," Institute for Development Studies, University of Nairobi, Discussion Paper No. 115, September, 1971.
- Kuznets, S., "Economic Growth and Income Inequality," American Economic Review, Vol. 45, No. 2, March, 1955.
- Kuznets, S., "Quantitative Aspects of the Economic Growth of Nations: VIII. Distribution of Income by Size," Economic Development and Cultural Change, Part II, vol. 11, no. 2, January, 1963.
- Langoni, C., "Distribuicao da Renda e Desenvolvimento Economico do Brasil," Estudos Economicos, Outubro, 1972.
- Langoni, C., "Review of Income Distribution Data: Brazil," Research Program in Economic Development, Woodrow Wilson School of Public and International Affairs, Princeton, Discussion Paper no. 60, April, 1975.
- McCabe, J., "The Distribution of Labor Incomes in Urban Zaire," Review of Income and Wealth, vol. 20, March, 1974.
- Mangahas, M., "Income Inequality in the Philippines: A Decomposition Analysis," in Income Distribution, Employment, and Economic Development in Southeast and East Asia, Papers and Proceedings of the seminar sponsored jointly by the Japan Economic Research Center and the Council for Asian Manpower Studies, July, 1975.
- Mehran, F., "Decomposition of the Gini Index: A Statistical Analysis of Income Inequality," International Labour Office, mimeo, 1974.
- Paukert, F., "Income Distribution at Different Levels of Development: A Survey of Evidence," International Labour Review, August-September, 1973.
- Pyatt, G., "On the Interpretation and Disaggregation of Gini Co-efficients," The Economic Journal, vol. 86, June, 1976.
- Schultz, T.P., "The Distribution of Personal Income: Case Study of the Netherlands," unpublished Ph.D. dissertation, MIT, Cambridge, Mass., 1965.
- Sen, A.K., On Economic Inequality, Norton, 1973.
- Szal, R.J. and Robinson, S., "Measuring Income Inequality," The Brookings Institution and Princeton University, mimeo, 1971.
- Theil, H., Economics and Information Theory, Amsterdam: North Holland Press, 1967.

Theil, H., Statistical Decomposition Analysis, Amsterdam: North Holland Press, 1972.

Uribe, P., "Estructura de la Desigualdad del Ingreso en America Latina," Demografia y Economia, vol. X, Num. 1, 1976.

van Ginneken, Wouter, "Analisis de Descomposicion del Indice de Theil Aplicado a la Distribucion del Ingreso Familiar en Mexico," Demografia y Economia, Vol. IX, Num. 1, 1975.

Webb, R., "On the Statistical Mapping of Urban Poverty and Employment," World Bank Staff Working Paper No. 227, January, 1976.

Weisskoff, R., "Income Distribution and Economic Growth in Puerto Rico, Argentina and Mexico," Review of Income and Wealth, December, 1970.