Review of Income and Wealth
Series 45, Number 1, March 1999

BETWEEN GROUP INEQUALITY AND TARGETED TRANSFERS

BY QUENTIN T. WODON

World Bank (LCSPR)

This paper provides two extensions to the group decomposition of the Gini index by Yitzhaki and Lerman. First, within group, stratification, and between group inequality are analyzed along several dimensions at once. This provides for a better understanding of the determinants of inequality. Second, the impact on the Gini of marginal changes in income or consumption by group is derived. This can be used to evaluate targeted redistributive policies or assess the impact of exogenous shocks by group. The analysis is applied to data from Bangladesh with a focus on the impact of land ownership, education, and occupation on inequality.

Inequality in income and consumption is a concern for policy-makers. Source and group decompositions have been developed to better understand the determinants of inequality and the policies which could be implemented to reduce it. Unfortunately, the methodologies available for group decompositions are less advanced than those available for source decompositions. This paper helps to bridge the gap. Two extensions to Yitzhaki and Lerman's (1991, hereafter YL) group decomposition of the Gini index are derived. They are applied to data from Bangladesh to analyze the inequality between land-owning, education, and occupation groups, and to estimate the impact of targeted transfers by group on the Gini for the overall population.

It is well known that the Gini index is not additively decomposable into within and between group components. The attractiveness of YL's decomposition along a unique dimension (say by education groups) lies in that the remainder of the decomposition has an intuitive interpretation as a measure of stratification or overlap between groups. Unfortunately, without an extension to take into account several dimensions at once, the YL decomposition is able to account for only a small part of total inequality. This is because when the groups are defined along a unique dimension, the within group component—which remains unexplained—typically contributes the most to the overall Gini (from 92 to 97 percent in the decompositions reported by YL). To account for a larger part of total inequality through the stratification and between group components of the decomposition, and to measure inequality by subgroups along one dimension (say land-owning class) within groups defined along another dimension (say education level), a multidimensional extension is needed. This extension is provided.

Note: I am grateful to Robert Lerman for many discussions on this topic. The data was provided by the Bangladesh Bureau of Statistics as part of a long term collaborative effort with the South Asia Country Department 1 at the World Bank. The findings, interpretations, and conclusions of this paper should not be attributed to the World Bank, its executive directors or the countries they represent. The help of Dileni Gunewardena and the comments of two anonymous referees are gratefully acknowledged.

The second contribution of the paper consists in deriving the impact of marginal changes in income or consumption by group on the Gini for total income or consumption. This extension is important for policy purposes. It is not uncommon in developing countries to implement redistribution schemes from large land-owners to the landless. The formulae derived here yield a way to analyze the impact of (marginal) targeted transfers by group on the overall Gini.

The third contribution of the paper is substantive. Little research has been done on income and consumption inequality in Bangladesh.[1] Osmani (1982) presented empirical work based on surveys conducted between 1963–64 and 1973–74. The contributions of Khan (1986), Rahman and Haque (1988), and Rahman (1988) were based on surveys up to the early 1980s. Rahman and Huda (1992) considered inequality between occupational groups using the 1983–84 Household Expenditure Survey (HES) of the Bangladesh Bureau of Statistics (hereafter BBS). The BBS (1995) itself does report more recent estimates of inequality, but only for the urban and rural sectors. Moreover, these measures are over-estimated because the BBS does not adequately take into account price differentials between the two sectors and between areas within each sector (Wodon, 1997). Using group data from the 1991–92 HES and urban/rural price deflators, Ravallion and Sen (1996) found that the urban and rural Ginis for per capita consumption in 1991–92 were equal to 0.319 and 0.255. However, they could not investigate in any detail the impact of household characteristics such as education, land ownership, and occupation on inequality in the absence of household level data. The unit level data of the 1991–92 HES survey have been made available for this study, which makes it possible in this paper to investigate the determinants of inequality in much more detail using a nationally representative sample.

The structure of the paper is as follows. The first section presents YL's unidimensional group decomposition of the Gini. The second section presents a multidimensional extension to the decomposition. The third section derives formulae for the impact on the overall Gini of marginal changes in income or consumption by group. The fourth section applies these extensions to inequality along educational, land, and occupational categories in Bangladesh.

1. Unidimensional Decomposition

Consider a population in which households can be grouped along a dimension i, such as their educational level, their geographical area, or their occupation. Following YL's (1991) notation, define the following:

y_{ih} is the income of household h belonging to group i;

$y_{i.}$ is the mean income of the households in group i;

m_i is the number of households in group i;

$k = \Sigma_i m_i$ is the number of households in the overall population;

$y_{..}$ is the mean income of households in the overall population;

$P_i = m_i/k$ is the population share of group i;

$S_i = P_i y_{i.}/y_{..}$ is the income share of group i;

R_{ih} is the rank of household h in the overall population ranked by income level;

[1]Much more work has been done on poverty (see Wodon, 1997 for references).

$R_{i.} = 1/m_i \, \Sigma_h^{mi} R_{ih}$ is the average rank of group i in the overall population;

$F_i(y_{ih})$ is the normalized rank (taking a value between zero and one) of household h from group i in the group's cumulative distribution of income F_i;

$F_{i.}$ is the normalized mean rank of households in group i in the population;

$F_{ni}(y_{ih})$ is the normalized rank of household h from group i in the cumulative distribution of income F_{ni} of all households except those of group i.

Denoting by $\text{cov}_i\,(x, y)$ the covariance between x and y over the members of group i only, YL expressed the Gini index G_i and the stratification index Q_i of group i as:

(1) $$G_i = 2\,\text{cov}_i\,(y_i, F_i)/y_i, \qquad i = 1, \ldots, n$$

(2) $$Q_i = \text{cov}_i\,[y, (F_i - F_{ni})]/\text{cov}_i(y_i, F_i) \qquad i = 1, \ldots, n.$$

The interpretation of the Gini index and its covariance expression are well-known and need not be recalled here (LY, 1984, 1989). The stratification index, a measure of overlap between the members of a group and the rest of the population, may be less familiar. It is a ratio of two terms. On the numerator, we have the covariance between the income of the households in group i and these households' difference in ranking in their own group and in the rest of the overall population. The denominator, which can be treated here as a normalizing factor, is the covariance for households in group i between the incomes and the rankings in their own group.

As noted by YL, Q_i can take on values between -1 and 1, and its properties make it an insightful index of stratification. When no members of other groups have incomes in the range of the incomes of the households belonging to group i, group i forms a perfect stratum, in which case $Q_i = 1$. At the other extreme, $Q_i = -1$ if the households in group i can be classified into two groups, one at the top of the overall income distribution, and one at the bottom of the distribution, with all the households from the other groups falling between the two subgroups of households in group i. In this case, rather than being homogenous, group i is composed of two heterogenous groups which themselves are perfect strata at the two tails of the distribution. A third special case occurs when $Q_i = 0$. Then, the rank of each household within group i is equal to the household's rank in the overall population and the group i forms no stratum at all.

Using the definitions of G_i and Q_i from equations (1) and (2), LY proved that the Gini of the whole population could be decomposed as a sum of three components:

(3) $$G = \Sigma_i S_i G_i + \Sigma_i S_i G_i Q_i (P_i - 1) + 2\,\text{cov}\,(y_{i.}, F_{i.})/y_{..}.$$

The first term, the within group inequality, is a weighted sum of the within group Ginis with the weights defined as the income shares. The second term accounts for stratification. In general, the Q_i terms are positive, and the more the groups are stratified, the higher the negative value of the stratification component (note that the terms $P_i - 1$ are negative since population shares are less than one). The third term, the between group inequality, is the weighted covariance between the various groups' mean income and their mean rank. It is a direct extension of the covariance-based formulation of the Gini for household level data.

Since the Q_i terms are typically positive, the stratification component is typically negative in the decomposition. Moreover, a higher level of stratification is associated with a larger negative value of the stratification component, and thus with a decrease in inequality. To understand the intuition behind this result, note that stratification implies a relatively low variability in ranks within the groups as these groups tend to form strata. However in this case, the between group component of the decomposition may be higher.[2] Another way to interpret the negative impact of the stratification term is to appeal to relative deprivation theory. According to this theory, the members of a group tend to compare their welfare with the other members of their group, rather than with the members of other groups. The more stratified a society, the less divergences within the groups, and the lower the feeling of inequality (Yitzhaki, 1982).

2. Multidimensional Extension

In this section, we develop a multidimensional extension for the YL decomposition. By multidimensional, we do not mean that we take into account at once different variables whose level of inequality we try to explain. Rather, we use several variables to explain the level of inequality in one unique dimension at a time, such as income or consumption. (The analogy to our multidimensional decomposition in a regression setting is multiple regression, not a system of regressions.) Here, we shall deal with bivariate decompositions. The generalization to more than two dimensions will be straightforward. Consider a population in which the households can be grouped according to two dimensions i and j. For each dimension i and j, we can apply the YL decomposition:

(4a) $$G = \Sigma_i S_i G_i + \Sigma_i S_i G_i Q_i(P_i - 1) + 2\,\text{cov}(y_{i.}, F_{i.})/y_{..}$$

(4b) $$G = \Sigma_j S_j G_j + \Sigma_j S_j G_j Q_j(P_j - 1) + 2\,\text{cov}(y_{j.}, F_{j.})/y_{..}.$$

A first strategy to take both dimensions into account is to define mutually exclusive groups k obtained by the combination of the dimensions i and j. That is, if $i = 1, \ldots, n$, and $j = 1, \ldots, m$, the households in group k ($k = 1, \ldots, n*m$) combine characteristics along both dimensions i and j. We can then apply the YL decomposition along the categories k to obtain:

(5) $$G = \Sigma_k S_k G_k + \Sigma_k S_k G_k Q_k(P_k - 1) + 2\,\text{cov}(y_{k.}, F_{k.})/y_{..}.$$

A more interesting way to approach the bivariate problem is to proceed sequentially. We can analyze the stratification and income inequality within each

[2]A simple example can illustrate this. Consider 2 groups with mean income 1 and 3. Each group has a 50 percent population share. The overall mean income is 2. A case of low stratification could correspond to mean ranks being 0.4 in group 1 and 0.6 in group 2, while higher stratification could correspond to mean ranks of 0.3 and 0.7. The between group term in the first case is equal to 0.2, and in the second case it is equal to 0.4. High stratification can be associated with high between group inequality. Of course, this is only an example, and it is an empirical matter to check what happens in a given setting.

group i by subgroups j. In equation (4a), we can decompose each G_i, $i = 1, \ldots, n$ as follows:

(6) $$G_i = \Sigma_j S_{ij} G_{ij} + \Sigma_j S_{ij} G_{ij} Q_{ij} (P_{ij} - 1) + 2 \operatorname{cov}(y_{ij.}, F_{ij.})/y_{i.} \qquad i = 1, \ldots, n.$$

In this new decomposition, S_{ij} represents the income of all households with both characteristics i and j as a share of the total income received by the households in group i. G_{ij} is the Gini index for group ij which includes only the households with both characteristics i and j. Q_{ij} is the stratification of group ij within group i, and P_{ij} is the population share of group ij within group i. The terms $y_{ij.}$ and $F_{ij.}$ represent the mean income and the average rank (within group i) of all households belonging to the group ij. As before, $y_{i.}$ is the mean income in group i. Using equation (6) in (4a), we obtain the second order decomposition along dimensions i, and then j:

(7)

$G = \Sigma_i S_i \Sigma_j S_{ij} G_{ij}$	Within Groups Component
$+ \Sigma_i S_i G_i Q_i (P_i - 1)$	First Order Stratification Component
$+ \Sigma_i S_i \Sigma_j S_{ij} G_{ij} Q_{ij} (P_{ij} - 1)$	Second Order Stratification Component
$+ 2 \operatorname{cov}(y_{i.}, F_{i.})/y_{..}$	First Order Between Groups Component
$+ \Sigma_i S_i \, 2 \operatorname{cov}(y_{ij.}, F_{ij.})/y_{i.}$	Second Order Between Groups Component

The first term in this second order decomposition is the within group component of total inequality. It is the result of two within group expansions, starting with dimension i, and following with dimension j. The two next terms are stratification components. The first order stratification term measures the stratification within the overall population according to the dimension i. The second order stratification term measures the stratification within the groups i according to the dimension j. Finally, the two last terms are between group components. The first order between group term measures the inequality between groups according to dimension i. The second order between group term measures the extent of the inequality, within the groups i, between the households with different characteristics j.

To obtain the decomposition (7), we started with an expansion along dimension i, and followed with dimension j. We can also start with dimension j, and follow with dimension i. Expanding the G_j terms in equation (4b) according to the i dimension yields:

(8) $$G_j = \Sigma_i S_{ji} G_{ji} + \Sigma_i S_{ji} G_{ji} Q_{ji} (P_{ji} - 1) + 2 \operatorname{cov}(y_{ji.}, F_{ji.})/y_{j.} \qquad j = 1, \ldots, m.$$

Proceeding as before and using equation (8) in equation (4b) yields:

(9)

$G = \Sigma_j S_j \Sigma_j S_{ji} G_{ji}$	Within Groups Component
$+ \Sigma_j S_j G_j Q_j (P_j - 1)$	First Order Stratification Component
$+ \Sigma_j S_j \Sigma_i S_{ji} G_{ji} Q_{ji} (P_{ji} - 1)$	Second Order Stratification Component
$+ 2 \operatorname{cov}(y_{j.}, F_{j.})/y_{..}$	First Order Between Groups Component
$+ \Sigma_j S_j \, 2 \operatorname{cov}(y_{ji.}, F_{ji.})/y_{j.}$	Second Order Between Groups Component.

Three decompositions extending the YL methodology have been proposed for the bivariate case, respectively in equations (5), (7), and (9). The within group component in each decomposition remains the same. To prove this, note that $G_{ji} = G_{ij} = G_k$, and that $S_j S_{ji} = S_i S_{ij} = S_k$ (but S_{ji} is not equal to S_{ij}). These identities imply that $\Sigma_j S_j \Sigma_i S_{ji} G_{ji} = \Sigma_i S_i \Sigma_j S_{ij} G_{ij} = \Sigma_k S_k G_k$.

The two sequential approaches provide more information than the mutually exclusive approach. The sequential approaches enable us to analyze the extent of stratification and between group inequality within groups i according to a second dimension j. In equation (7) for example, the second order stratification term tells us about the stratification according to, say, the occupation dimension j, among the households belonging to, say, the various education groups i. By contrast, in equation (9), the second order stratification tells us about the stratification according to the education dimension i, among the households belonging to the various occupation groups j. Similarly, the second order between group term in equation (7) tells us about the between group inequality according to the occupation dimension j within the education groups i, while the second order between group term in equation (9) tells us about the between group inequality according to the education dimension j within the occupation groups i. If the data and sample size permit, the methodology can easily be extended to three (or more) dimensions. To do so, it suffices to replace the within group Ginis G_{ij} (or G_{ji}) by their decomposition according to a third dimension, say l. This would simply yield third (or higher) order stratification and between group terms in the decomposition for the overall population. However, the higher the number of dimensions, the higher the number of terms in the decomposition, and the more difficult its interpretation. To keep things simple, we will use only two dimensions in the decompositions presented in the empirical sections of this paper.

3. Marginal Changes by Group

Consider now a marginal change for the households in group g, such that their income (or consumption) is multiplied by $(1 + e_g)$, where e_g tends to zero. If we consider both the original income of the household and the shock as exogenous, for the unidimensional decomposition (a similar development applies to the multidimensional decomposition), it is proven in the Appendix that the impact on the total Gini of this marginal change for group g is:

$$\frac{\partial G}{\partial e_g} = S_g \left[G_g - \Sigma_i S_i G_i + G_g Q_g (P_g - 1) - \Sigma_i S_i G_i Q_i (P_i - 1) + 2 \frac{y_{g.}}{y_{..} S_g} (F_{g.} - 0.5) - 2 \Sigma_i (F_{i.} - 0.5) \right]. \tag{10}$$

The key assumption to derive equation (10) is that the income of other groups are not affected when the income for the households in group g are modified at the margin. Then, equation (10) accounts for the changes in the overall Gini due to the sum of the changes in the within group, stratification, and between group components. The change in the within group component is $S_g(G_g - \Sigma_i S_i G_i)$. If the group g which sees its income rising has a higher Gini than the within group Gini, the within group inequality will increase. The change in the stratification component is $S_g[G_gQ_g(P_g-1)-\Sigma_i S_i G_i Q_i(P_i-1)]$. Again, if group g has a higher $G_gQ_g(P_g-1)$ than the stratification component, the stratification component will increase. Finally, when multiplied by the share S_g, the last two terms in brackets in equation (10) account for the change in the between group component due to the change in income for group g.

Equation (10) can be interpreted in terms of taxes and transfers. Consider the case in which the decomposition is based on after tax income as the adequate measure of welfare. If group g is taxed at the rate t_g, a household in that group with gross income y will keep $(1-t_g)y$ in after tax income. Alternatively, imagine that households receive transfers whose amount is proportional to their after tax income. If the transfer rate for group g is tr_g, the real standard of living of a household in that group will be $(1+tr_g)y$. Given an initial structure of taxes and/or transfers (t_g or tr_g given), the marginal change e_g can be interpreted as a change in taxes or transfers. For taxes, when e_g is positive (negative), the tax rate is reduced (increased), and the income for households in group g are increased (reduced). For transfers, when e_g is positive (negative), the transfer rate is increased (reduced), and the income for households in group g is increased (reduced). Equation (10) can also be used to estimate the impact on the Gini of exogenous shocks. In Bangladesh, some geographical areas are more subject to floods than others. Assuming that a flood decreases income or consumption in proportion to pre-flood levels (marginally for the sake of the discussion), equation (10) will provide an estimate of post-flooding inequality at the national level in function of the damage caused by floods in the flooded area(s)—the group g in the decomposition, or the various groups affected.

For policy purposes, it is sometimes better to work with absolute changes in income rather than with percentage changes. When income for households in group g are multiplied by $(1+e_g)$, the total change in income for this group is $E_g = e_g y_{..} S_g$. Noting that $(\partial G/\partial e_g) = (\partial G/\partial E_g)(\partial E_g/\partial e_g) = (\partial G/\partial E_g) y_{..} S_g$, we have:

$$\text{(11)} \qquad \frac{\partial G}{\partial E_g} = \frac{1}{y_{..}}\Bigg[G_g - \Sigma_i S_i G_i + G_g Q_g(P_g-1) - \Sigma_i S_i G_i Q_i(P_i-1) + 2\frac{y_{g.}}{y_{..}S_g}(F_{g.}-0.5) - 2\,\Sigma_i(F_{i.}-0.5)\Bigg].$$

Using (11), and following an idea of Lerman and Yitzhaki (1994) applied to the source decomposition of the Gini, we can compute the transfers which would have to be given to a poor group g in order to offset the impact on inequality of, say, an exogenous growth in the incomes of a better off group k. Along an equal inequality curve, we have:

$$\text{(12)} \qquad -\frac{dE_g}{dE_k}\bigg|\, G = \frac{\partial G/\partial E_k}{\partial G/\partial E_g}.$$

Finally, one may wish to tax the well-off group k (say, large land-owners) in order to provide transfers to a poorer group g (the landless) without creating a budget deficit. For deficit neutrality, we need $S_g d e_g = -S_k d e_k$. It is immediate to verify that the change in Gini will be:

$$dG = S_g \left[G_g - G_k + G_g Q_g(P_g - 1) - G_k Q_k(P_k - 1) + 2\frac{y_{g.}}{y_{..}S_g} F_{g.} - 2\frac{y_{k.}}{y_{..}S_k} F_k \right] d e_g. \tag{13}$$

4. Land, Education, and Occupation

Three of the most important dimensions affecting inequality in developing countries are the household head's education level, the household head's main occupation or field of employment, and the household's ownership of land. Below, our multidimensional extension of the LY decomposition is applied to bivariate decompositions of inequality by alternative ordered pairs of these three dimensions. The results of the six bivariate decompositions are compared with the three unidimensional decompositions obtained separately by education, land, and occupation only, as well as with the three decompositions obtained through the definition of mutually exclusive groups for each pair of dimensions.[3] Also, the impact of marginal transfers or taxes by land owning, education, and occupation groups are compared. The analysis is applied to per capita income (y) and consumption (x) adjusted for regional price differences. That is, x is the welfare ratio or per capita consumption divided by the poverty line, and y is the income ratio or per capita income divided by the poverty line (for details on the construction of the poverty lines, see Wodon, 1997). If a household has a welfare (income) ratio of one, his per capita consumption (income) is exactly at the level of his regional poverty line. If the ratio is less than one, the household is poor, and if it is more than one, the household is not poor.[4]

4.1. *Unidimensional Decompositions*

Gini decompositions are sensitive to the number of mutually exclusive groups or categories defined along given dimensions (at the limit, if we were to define each household as being a group, the Gini will be equal to the between group component) . To avoid such sensitivity as much as can be, to ensure that groups defined according to a combination of characteristics have a reasonable sample size, and to facilitate the comparisons between the decompositions, five categories were defined for each of the three dimensions of interest. The definition of these

[3]In computing the Ginis and covariances, we used weights to take into account the size of each household or group as given in the sample, and we computed the normalized ranks of each area at mid-point, as suggested by LY (1989). This matters especially for the estimation of the between group component.

[4]Because the poverty lines are used as price deflators in the computation of the welfare ratios, the Gini coefficients are sensitive to the choice of the poverty lines. We used two sets of regional poverty lines to conduct our analysis, a lower and an upper one (the upper poverty lines include a larger allowance for non-food consumption basic needs). The results with the lower poverty lines are reported here. The results with the upper poverty lines are very similar.

TABLE 1

STATISTICS BY EDUCATION, LAND OWNED, AND OCCUPATION FOR INCOME RATIOS

Education Categories (i)	Illiterate	Primary Level, Cannot Write	Primary Level, Can Write	High School Level	Above High School Level	Total
Population share (P_i)	0.441	0.203	0.105	0.118	0.113	1.000
Income share (S_i)	0.350	0.192	0.107	0.133	0.218	1.000
Mean income ratio ($y_{i.}$)	1.099	1.308	1.413	1.568	2.272	—
Gini coefficient (G_i)	0.255	0.264	0.256	0.261	0.311	—
Stratification index (Q_i)	0.056	0.049	0.098	0.160	0.399	—
Mean rank ($F_{i.}$)	0.392	0.492	0.542	0.599	0.749	—
Land ownership categories (j)	**Less than 0.05 acres**	**0.05 to 0.50 acres**	**0.50 to 1.50 acres**	**1.50 to 2.50 acres**	**2.50 acres or more**	**Total**
Population share (P_j)	0.233	0.280	0.192	0.106	0.189	1.000
Income share (S_j)	0.207	0.243	0.190	0.113	0.247	1.000
Mean income ratio ($y_{j.}$)	1.230	1.202	1.373	1.472	1.812	—
Gini coefficient (G_j)	0.322	0.287	0.275	0.250	0.275	—
Stratification index (Q_j)	−0.105	−0.030	0.071	0.154	0.277	—
Mean rank ($F_{j.}$)	0.417	0.424	0.508	0.568	0.668	—
Occupation categories (l)	**Tenants and Agricultural Workers**	**Factory, Industry, Blue Collar Workers**	**Retired Person, Not Working, Student**	**Land-owners, Small and Large**	**Official, Manager, White Collar Workers**	**Total**
Population share (P_l)	0.151	0.300	0.091	0.367	0.091	1.000
Income share (S_l)	0.100	0.304	0.099	0.361	0.135	1.000
Mean income ratio ($y_{l.}$)	0.918	1.404	1.513	1.364	2.054	—
Gini coefficient (G_l)	0.246	0.306	0.310	0.256	0.291	—
Stratification index (Q_l)	0.093	−0.005	0.054	0.127	0.298	—
Mean rank ($F_{l.}$)	0.292	0.502	0.540	0.520	0.718	—

Source: Own computations from HES unit level data. Numbers may not add up due to rounding.

categories and summary statistics by education, land ownership, and occupation are provided in Tables 1 and 2 for, respectively, income and consumption.

Tables 1 and 2 remind us that half of the population of Bangladesh as represented in the 1991–92 HES is illiterate. One-fourth of the population is landless, and another fourth is near landless. In terms of occupations, agricultural labourers and tenant farmers represent together one-sixth of the population. Landowners deriving their income from their own land parcels, be they small or large, make up one-third of the population. Factory, industrial, and other blue collar workers account for another third of the population. White collar workers such as officials, managers, teachers, and public servants represent one-tenth of the population, and so do non-working heads. The mean income and welfare ratios are increasing with the educational level of the household head, the land owned by the household, and the occupation status of the head.

The national Ginis for per capita income and consumption are respectively 0.300 and 0.274. The within group Gini coefficients G_i, G_j, and G_l range from 0.246 to 0.322 for income, and from 0.215 to 0.299 for consumption. The coefficients are increasing with the education level and the occupation status of

TABLE 2

STATISTICS BY EDUCATION, LAND OWNED, AND OCCUPATION FOR WELFARE RATIOS

Education Categories (i)	Illiterate	Primary Level, Cannot Write	Primary Level, Can Write	High School Level	Above High School Level	Total
Population share (P_i)	0.441	0.203	0.105	0.118	0.133	1.000
Consumption share (S_i)	0.354	0.192	0.108	0.132	0.215	1.000
Mean welfare ratio ($x_{i.}$)	1.005	1.181	1.285	1.340	2.021	—
Gini coefficient (G_i)	0.226	0.241	0.235	0.231	0.290	—
Stratification index (Q_i)	0.084	0.044	0.100	0.187	0.416	—
Mean rank ($F_{i.}$)	0.390	0.488	0.545	0.601	0.760	—
Land ownership categories (j)	**Less than 0.05 acres**	**0.05 to 0.50 acres**	**0.50 to 1.50 acres**	**1.50 to 2.50 acres**	**2.50 acres or more**	**Total**
Population share (P_j)	0.233	0.280	0.192	0.106	0.189	1.000
Consumption share (S_j)	0.212	0.248	0.191	0.114	0.235	1.000
Mean welfare ratio ($x_{j.}$)	1.138	1.109	1.244	1.345	1.558	—
Gini coefficient (G_j)	0.299	0.264	0.256	0.232	0.247	—
Stratification index (Q_j)	−0.114	−0.028	0.057	0.154	0.268	—
Mean rank ($F_{i.}$)	0.423	0.426	0.505	0.575	0.657	—
Occupation categories (l)	**Tenants and Agricul-tural Workers**	**Factory, Industry, Blue Collar Workers**	**Retired Person, Not Working, Student**	**Land-owners, Small and Large**	**Official, Manager, White Collar Workers**	**Total**
Population share (P_l)	0.151	0.300	0.091	0.367	0.091	1.000
Consumption share (S_l)	0.102	0.308	0.100	0.358	0.133	1.000
Mean welfare ratio ($x_{l.}$)	0.842	1.285	1.371	1.222	1.818	—
Gini coefficient (G_l)	0.215	0.287	0.280	0.225	0.269	—
Stratification index (Q_l)	0.158	−0.014	0.056	0.145	0.311	—
Mean rank ($F_{l.}$)	0.282	0.504	0.553	0.518	0.725	—

Source: Own computations from HES unit level data. Numbers may not add up due to rounding.

the head of the household. This was to be expected as the spectrum of earnings (and consumption opportunities) among better educated households is wider than that among less-educated households. The incomes and consumption patterns of a well-educated manager and a well-educated teacher are likely to diverge more than the incomes and consumption patterns of two illiterate agricultural workers. Interestingly, the Ginis are decreasing with the amount of land owned. This could be due to the fact that large land-owners derive similar levels of income and consumption from the cultivation of their land. (If returns to scale are decreasing, as it has been argued in the case of Bangladesh, differences in land holdings among large proprietors will not lead to large differences in standards of living.) By contrast, the landless and near landless groups are much more heterogenous categories of households, this heterogeneity resulting in higher inequality.

The stratification indices Q_i, Q_j, and Q_l are almost always positive, indicating that most groups are stratified. Stratification increases with the level of education, land ownership, and occupation. The stratification indices are negative for the landless and near landless, as well as for factory, industry, and blue collar workers, suggesting that within these categories, diverging levels of education,

TABLE 3

GROUP DECOMPOSITIONS FOR INCOME AND WELFARE RATIOS

	Education		Land		Occupation	
Dimension (i, j, or l)	y	x	y	x	y	x
Within group component	0.270	0.244	0.285	0.262	0.280	0.255
Stratification component	−0.036	−0.035	−0.015	−0.012	−0.021	−0.021
Between groups component	0.065	0.065	0.030	0.021	0.041	0.040
Overall Gini	0.300	0.274	0.300	0.274	0.300	0.274

Source: Own computations from HES unit level data. Numbers may not add up due to rounding.

occupation, and other characteristics render the groups less homogenous.[5] This is congruent with the higher within group Gini coefficients for these groups.

Table 3 provides the unidimensional Gini decompositions along the three dimensions. The within group components are highest for land, medium for occupation, and lowest for education. The rankings in terms of the stratification and between group components are reversed: stratification and between group inequality are highest along the education dimension, medium along the occupation dimension, and lowest along the land dimension. Table 3 indicates that at the aggregate level, households with similar levels of education tend to enjoy similar levels of consumption and to form strata, with relatively large differences in consumption patterns across the strata. To the contrary, households with similar levels of land ownership tend to form less strata and to experience more inequality within their group and less inequality between groups. Note that the decrease in inequality observed when shifting from the income to the consumption space almost fully occurs through a decrease in the within group components. Income is not associated with a significant rise in stratification and between group inequality.

Figure 1 may provide more insight in the working of the unidimensional decompositions. On the figure, using the decomposition based on the welfare ratios, the land, education, and occupation categories are identified by numbers from 0 to 4, with 0 standing for the least and 4 for the most favourable categories—Land0 represents the households with less than 0.5 acres of property, and Land4 the households with more than 2.5 acres. The contributions S_iG_i of the fifteen groups (five for each of the three decompositions) to the within group components are given on the horizontal axis. The larger the Gini and the income share, the larger the contribution to the within group Gini. For example, despite their lower consumption share, and due to their higher within group inequality, factory, industry and blue collar workers (Occup1) contribute more to the within group component of the occupation decomposition than land-owners (Occup3). The same applies to the comparison of the contributions to the within group component of the landless and near-landless households in the land decomposition.

[5]The heterogeneity of categories such as the landless and near landless apparent in the negative stratification for these groups indicates also that using land as a targeting category would result in large targeting errors. In other words, at the national level, poverty cannot be well characterized by one policy instrument. This need not be true within rural areas where land tends to be a better indicator of standards of living.

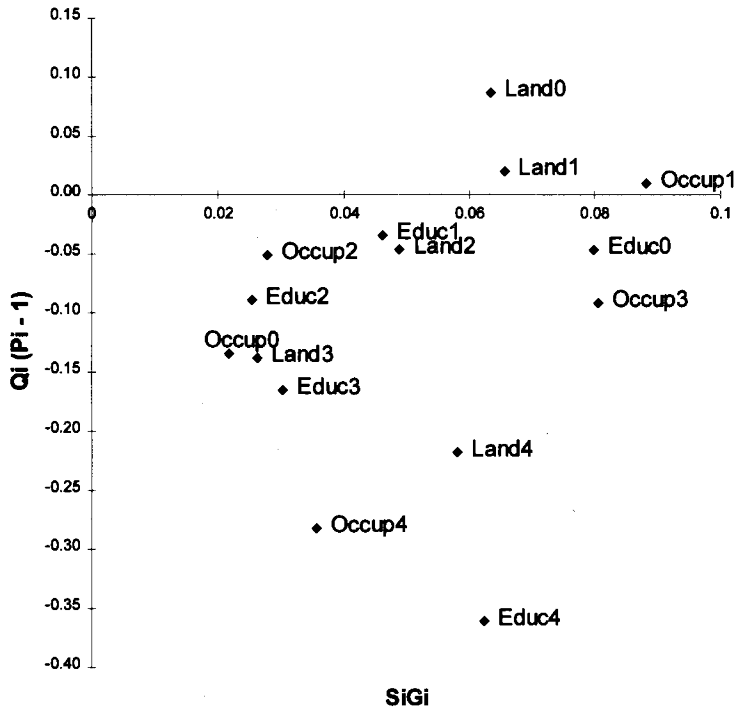

Figure 1. Unidimensional Group Decompositions (for Welfare Ratios)

When we multiply S_iG_i on the horizontal axis by $Q_i(P_i - 1)$ given on the vertical axis, we obtain the contributions to the stratification components. For a given contribution to the within group component S_iG_i, the larger the stratification index and the lower the population share, the larger the absolute contribution to the stratification term (remember that $P_i - 1$ is negative). It is clear from the figure that the highest two categories for each decomposition (Land4, Occup4, and Educ4, and to a lesser extent Land3, Occup3, and Educ3) contribute the most to the stratification component due to both large stratification indices and small population shares.

4.2. *Multidimensional Extension*

Table 4 provides the mutually exclusive bivariate Gini decompositions by pairs of dimensions. These decompositions are based on the 25 rather than 5 mutually exclusive categories. For example, households who are both illiterate and landless form one category, and households who are illiterate and near landless form another category. Since the definition of 25 categories enables us to better track the standard of living of households, the within group component of the mutually exclusive bivariate decompositions are smaller, and the stratification and between group components are larger than their counterparts in the unidimensional cases. In other words, the part of the overall Gini which is not

TABLE 4

MUTUALLY EXCLUSIVE BIDIMENSIONAL DECOMPOSITIONS FOR INCOME AND WELFARE RATIOS

Pairs of Dimension	Education and Land		Land and Occupation		Education and Occupation	
(*i*, *j*, or *l*)	*y*	*x*	*y*	*x*	*y*	*x*
Within group component	0.259	0.236	0.267	0.244	0.262	0.236
Stratification component	–0.044	–0.042	–0.033	–0.030	–0.042	–0.041
Between groups component	0.084	0.080	0.065	0.060	0.080	0.079
Overall Gini	0.300	0.274	0.300	0.274	0.300	0.274

Source: Own computations from HES unit level data. Numbers may not add up due to rounding.

accounted for by the decomposition is reduced. The decompositions for education and land, and education and occupation give fairly similar results in terms of their within group, stratification, and between group components. Yet, as we shall see, differences appear between these two decompositions once the sequential approach is adopted, which shows how the sequential approach provides more information.

Table 5 provides the results of the sequential bivariate decompositions. As noted earlier, the within group components of the sequential decompositions are equal to the within group components of the mutually exclusive decompositions for the same pairs of dimensions. Moreover, the stratification and between group terms for the mutually exclusive decompositions are fairly close to, respectively, the sum of the first and second order stratification and between group terms of the sequential decompositions for the same pairs of dimensions.

Consider for example the stratification and between group components in the consumption case along the education and occupation dimensions. The stratification and between group components from the mutually exclusive decomposition are respectively −0.041 and 0.079. In the sequential bivariate decompositions, if we start with occupation, and then pursue the decomposition by education level, the overall stratification and between group components are split evenly between the first and second order terms (−0.021 for both the first and second order stratification terms, and 0.041 and 0.039 for the first and second order between group terms). At first sight, this would suggest that education and occupation are equally important factors driving inequality. However, if we start the decomposition with education, and pursue the decomposition with occupation, we see that the first order stratification and between group terms along the education dimension, −0.035 and 0.065, are much larger than the second order stratification and between group terms along the occupation dimension, −0.008 and 0.016.

Once groups have been created according to education levels, considering the occupation dimension does not add much explanatory power (through the second order stratification and between group inequality components). However the reverse is not true if we start with occupation, and continue with education. In a nutshell, education appears to be more powerful in driving consumption inequality than occupation. The same applies to the decomposition of the income rather than welfare ratio. The weight of education in the sequential land and education decompositions for consumption and income is even stronger. Finally,

TABLE 5

SEQUENTIAL BIDIMENSIONAL DECOMPOSITIONS FOR INCOME AND WELFARE RATIOS

Ordered Pairs of Dimension	Education and Land		Land and Education		Education Occupation		Occupation Education		Occupation and Land		Land and Occupation	
	y	*x*	*y*	*x*	*y*	*x*	*y*	*x*	*y*	*x*	*y*	*x*
Within group component	0.259	0.236	0.259	0.236	0.262	0.236	0.262	0.236	0.267	0.244	0.267	0.244
First order stratification component	–0.036	–0.035	–0.015	–0.012	–0.036	–0.035	–0.021	–0.021	–0.021	–0.019	–0.015	–0.010
Second order stratification component	–0.010	–0.008	–0.029	–0.029	–0.007	–0.008	–0.021	–0.021	–0.013	–0.012	–0.019	–0.021
First order between group component	0.066	0.065	0.030	0.023	0.066	0.065	0.041	0.040	0.041	0.037	0.030	0.021
Second order between group component	0.020	0.017	0.054	0.056	0.015	0.016	0.039	0.039	–0.026	0.023	0.036	0.040
Overall Gini	0.300	0.274	0.300	0.274	0.300	0.274	0.300	0.274	0.300	0.274	0.300	0.274

Source: Own computations from HES unit level data. Numbers may not add up due to rounding.

in comparing land ownership and occupation categories, it appears that occupation has more of an impact on stratification and between group inequality than land ownership.

To sum up, the sequential bivariate decompositions indicate that education appears to be a more powerful force driving consumption and income inequality than occupation or land ownership in Bangladesh. It could be that this ranking of characteristics in terms of their impact on inequality (measured by the between group and stratification components) would be modified if we considered a third dimension, such as the urban or the rural sector. In rural areas, land ownership could be as important, or perhaps more important than education, while in urban areas, education could be even more prominent than it appears to be at the national level. To investigate the existence of such reversals in ranking, it would suffice to introduce a third dimension, namely the sectoral characteristic, into the decomposition. In any case, a key advantage of multidimensional decompositions is that they allow policy-makers to analyze the impact of tax and transfer policies targeting groups defined along several dimensions at once.

4.3. *Targeted Transfers*

To illustrate the impact of marginal changes in income and consumption by group, we shall use the unidimensional decomposition for simplicity, even though similar formulae can be derived for multidimensional decompositions. Table 6 provides the estimates of the marginal changes in the Gini due to marginal percentage changes in income or consumption by group.

The signs of the marginal changes in the within group Ginis tend to be negative for the worst off groups, and positive for the better off groups. Remember that these signs depend on whether the Gini for the group whose income or consumption is affected is larger or smaller than the within group component of the decomposition. It is for the highest level groups in the decompositions that the group-specific Ginis tend to be the largest (with the exception of large landowners as noted earlier), and therefore it is for marginal changes in these groups' income or consumption that the within group component increases. The same holds for the stratification component in absolute terms (note that the stratification component in the decomposition is negative). When the incomes of better-off groups increases, stratification is also increased in absolute terms. Now, for education as well as for land ownership and for occupation, the bulk of the change in the Gini comes through the between group component, which was expected since the marginal changes in income or consumption are group specific. In total, when summing the changes in the three components in the decomposition, a marginal increase in the income or consumption of a relatively poor group decreases the overall Gini, while a raise in the standard of living of a relatively better-off groups increases inequality in total income or consumption.

From a policy point of view, two redistributive strategies appear to be the most efficient in reducing inequality. The first strategy is to provide a transfer to tenants and agricultural workers while taxing officials, managers, and white collar workers. A second strategy is to help the illiterate while requiring a redistributive effort from the well-educated. These simulations are indicative only, and they do

TABLE 6
MARGINAL CONTRIBUTIONS (UNIDIMENSIONAL DECOMPOSITIONS)

Education Categories (i)		Illiterate	Primary Level, Cannot Write	Primary Level, Can Write	High School Level	Above High School Level
Marginal change	y	−0.0053	−0.0012	−0.0015	−0.0012	0.0089
within group component	x	−0.0064	−0.0006	−0.0010	−0.0017	0.0099
Marginal change	y	0.0098	0.0049	0.0014	−0.0001	−0.0156
stratification component	x	0.0086	0.0051	0.0015	−0.0004	−0.0150
Marginal change	y	−0.3631	−0.1203	0.0270	0.1511	0.6969
between group component	x	−0.3778	−0.1317	0.0311	0.1506	0.7179
Total marginal change	y	−0.3585	−0.1165	0.0270	0.1498	0.6902
	x	−0.3756	−0.1272	0.0316	0.1485	0.7187
Land ownership categories (j)		**Less than 0.05 acres**	**0.05 to 0.50 acres**	**0.50 to 1.50 acres**	**1.50 to 2.50 acres**	**2.50 acres or more**
Marginal change	y	0.0077	0.0005	−0.0019	−0.0040	−0.0025
within group component	x	0.0078	0.0005	−0.0012	−0.0032	−0.0028
Marginal change	y	0.0085	0.0052	−0.0002	−0.0022	−0.0112
stratification component	x	0.0081	0.0043	0.0000	−0.0021	−0.0079
Marginal change	y	−0.1825	−0.1731	−0.0165	0.1252	0.3973
between group component	x	−0.1766	−0.1739	−0.0231	0.1430	0.3585
Total marginal change	y	−0.1664	−0.1675	−0.0185	0.1191	0.3833
	x	−0.1606	−0.1691	−0.0242	0.1377	0.1478
Occupational categories (l)		**Tenants and Agricultural Workers**	**Factory, Industry, Blue Collar Workers**	**Retired Person, Not Working, Student**	**Land-owners, Small, and Large**	**Official, Manager, White Collar Workers**
Marginal change	y	−0.0034	0.0079	−0.0030	−0.0088	−0.0010
within group component	x	−0.0041	0.0096	0.0023	−0.0110	−0.0013
Marginal change	y	0.0002	0.0067	0.0006	0.0002	−0.0053
stratification component	x	−0.0008	0.0071	0.0006	0.0001	−0.0050
Marginal change	y	−0.2899	−0.0397	0.0731	−0.0135	0.6330
between group component	x	−0.3102	−0.0410	0.1012	−0.0250	0.6390
Total marginal change	y	−0.2932	−0.0251	0.0766	−0.0221	0.6288
	x	−0.3151	−0.0242	0.1041	−0.0359	0.6353

Source: Own computations from HES unit level data. Numbers may not add up due to rounding.

not take into account incentive effects. However they provide a first estimate of the gains that can be achieved in terms of equality through targeted transfers and taxes. Note that the reason why taxing large land-owners to provide for the landless gives lesser results is that the decompositions have been applied at the national level. A significant number of households in urban areas have a good standard of living even if they do not own any land. If separate decompositions had been conducted for the urban and rural sectors, we might have obtained different results. One way to do this would be to include a third dimension in the decomposition along the lines discussed earlier. In general, using the multidimensional decomposition to assess the impact of marginal changes in income or consumption to more and more precisely defined groups (say, illiterate landless rural households working in the farm sector) will provide for better targeting in the implementation of redistributive policies.

5. Summary

This paper has proposed two extensions to the methodology developed by Yitzhaki and Lerman. The first extension indicated how to analyze within group, stratification, and between group inequality in a multidimensional context. By taking into account several dimensions at once, the within group components of the decompositions, which represent the proportion of the overall inequality which remains unaccounted for, are reduced. Moreover, the sequential decompositions along two (or more) dimensions were shown to provide more information than the decompositions based on mutually exclusive categories along these dimensions. Specifically, the sequential decompositions enable policy-makers to analyze the extent of stratification and between group inequality along one dimension within groups defined according to another dimension. The second extension consisted in deriving formulae for estimating the impact on overall inequality of marginal changes in income or consumption by group. These formulae may help policy-makers to target more finely transfers and taxes in order to reduce inequality.

The two methodological extensions were applied to the analysis of inequality in Bangladesh by land-holding, education, and occupation groups. Education appeared to be a stronger determinant of inequality than occupation, with land ownership ranking third. Marginal targeted transfers and taxes would have a larger redistributive impact when applied to education (from the well-educated to the illiterate) or occupation groups (from officials and managers to tenants and agricultural workers). These results, which were obtained at the national level, may be affected when considering the rural and urban sector separately. To check for this, a next step in the research could consist in applying the multidimensional extension to these two sectors.

Appendix

This Appendix proves equation (10). The first steps follow the analysis of Stark *et al.* (1986) for source decompositions of the Gini. We consider the impact on the within group component of multiplying the income of the members of group g by $(1 + e_g)$ and assume that the income of households in other groups are not affected when the income for group g is modified.

Consider first the within group component of the decomposition. Note that the Gini G_g of group g is not modified when all households in this group see their income multiplied by $(1 + e_g)$. The Ginis of other groups also remain unchanged. The only change in the within group component results from changes in shares. By definition, for any group $i \neq g$, we have:

$$\text{(A.1)} \qquad \Delta S_i = \frac{y_{i.}}{\Sigma_i y_{i.} + e_g y_{g.}} - \frac{y_{i.}}{\Sigma_i y_{i.}} = \frac{-e_g S_i S_g}{1 + e_g S_g}$$

For $i = g$, we have instead:

(A.2) $$\Delta S_i = \frac{(1+e_g)y_{g.}}{\Sigma_i y_{i.} + e_g y_{g.}} - \frac{y_{g.}}{\Sigma_i y_{i.}} = \frac{e_g S_g - e_g S_g^2}{1 + e_g S_g}$$

Substituting (A.1) and (A.2) in the decomposition yields for the within group component:

(A.3) $$\Delta(\Sigma_i S_i G_i) = \Sigma_i \frac{-e_g S_i S_g}{1+e_g S_g} G_i + \frac{e_g S_g}{1+e_g S_g} G_g.$$

Taking the limit for e_g tending to zero yields:

(A.4) $$\frac{\partial(\Sigma_i S_i G_i)}{\partial e_g} = S_g(G_g - \Sigma_i S_i G_i).$$

Consider now the stratification term and assume that no two households in the sample have exactly the same income. Then, any infinitesimal change in incomes for group i will not affect the ranking of households in their group and in the rest of the population. For any i, Q_i will be unaffected when income source g is multiplied by $(1 + e_g)$. Noting that population shares remain constant, following the same steps for shares as for the within group Gini, we have:

(A.5) $$\frac{\partial\,[\Sigma_i S_i G_i Q_i(P_i - 1)]}{\partial e_g} = S_g[G_g Q_g(P_g - 1) - \Sigma_i S_i G_i(Q_i(P_i - 1)].$$

For the between group component, note first that the new between group term NBG is:

(A.6) $$\text{NBG} = \frac{2}{y_{..}}[\Sigma_{i\neq g}(y_{i.} - y_{..}(1+e_g S_g))(F_{i.} - 0.5) + (y_{g.}(1+e_g) - y_{..}(1+e_g S_g))(F_{g.} - 0.5)].$$

Rearranging terms, we have:

(A.7) $$\text{NBG} = \frac{2}{y_{..}}[\Sigma_i(y_{i.} - y_{..})(F_{i.} - 0.5) + y_{g.}e_g(F_{g.} - 0.5) - \Sigma_i y_{..}e_g S_g(F_{i.} - 0.5)].$$

Denoting the old between group term by OBG,

(A.8) $$\text{NBG} - \text{OBG} = 2\frac{e_g}{y_{..}}[y_{g.}(F_{g.} - 0.5) - \Sigma_i y_{..}S_g(F_{i.} - 0.5)].$$

Taking the limit for e_g tending to zero and denoting the between group by BG yields:

(A.9) $$\frac{\partial \text{BG}}{\partial e_g} = 2S_g\left[\frac{y_{g.}}{y_{..}S_g}(F_{g.} - 0.5) - \Sigma_i(F_{i.} - 0.5)\right].$$

Overall, the change in the Gini due to a change in incomes for group g is:

$$\text{(A.10)} \qquad \frac{\partial G}{\partial e_g} = S_g\Big[G_g - \Sigma_i S_i G_i + G_g Q_g (P_g - 1) - \Sigma_i S_i G_i Q_i (P_i - 1) + \frac{2y_{g.}}{y_{..} S_g}(F_{g.} - 0.5) - 2\Sigma_i (F_{i.} - 0.5)\Big].$$

REFERENCES

Bangladesh Bureau of Statistics (BBS), *Report on the Household Expenditure Survey 1988–89*, Bangladesh Bureau of Statistics, Dhaka, 1991.

———, *Summary Report of Household Expenditure Survey 1991–92*, Bangladesh Bureau of Statistics, Dhaka, 1995.

Hossain, Z. R. and M. Hossain (eds.), *Rethinking Rural Poverty: Bangladesh as a Case Study*, Sage Publications, New York, 1995.

Khan H., Income Inequality, Poverty and Socio-Economic Development in Bangladesh: An Empirical Investigation, *Bangladesh Development Studies*, 14, 75–92, 1986.

Lerman, R. I. and S. Yitzhaki, A Note on the Calculation and Interpretation of the Gini Index, *Economic Letters*, 15, 363–68, 1984.

———, Improving the Accuracy of Gini Estimates, *Journal of Econometrics*, 42, 43–7, 1989.

———, Effect of Marginal Changes in Income Sources on U.S. Income Inequality, *Public Finance Quarterly*, 22, 403–17, 1994.

Osmani, S., *Economic Inequality and Group Welfare: A Theory of Comparison with Application to Bangladesh*, Clarendon Press, Oxford, 1982.

Rahman, M., Some Aspects of Income Distribution in Rural Bangladesh, *Applied Economics*, 20, 1007–15, 1988.

——— and S. Huda, Decomposition of Income Inequality in Rural Bangladesh, *Modern Asian Studies*, 26, 83–93, 1992.

Rahman, A. and T. Haque, Poverty and Inequality in Bangladesh in the Eighties: An Analysis of Some Recent Evidence, Research Report No. 91, BIDS, Dhaka, 1988.

Ravallion, M. and B. Sen, When Method Matters; Toward A Resolution of the Debate about Bangladesh's Poverty Measures, *Economic Development and Cultural Change*, 44, 761-92, 1996.

Stark, O., J. E. Taylor, and S. Yitzhaki, Remittances and Inequality, *The Economic Journal*, 96, 722–40, 1986.

Wodon, Q. Food Energy Intake and Cost of Basic Needs: Measuring Poverty in Bangladesh, *Journal of Development Studies*, 34, 66–101, 1997.

Yitzhaki, S., Relative Deprivation and Economic Welfare, *European Economic Review*, 17, 99–113, 1982.

——— and R. I. Lerman, Income Stratification and Income Inequality, *Review of Income and Wealth*, 37, 313-29, 1991.