Note VIII.1. Measurement of Income Inequality

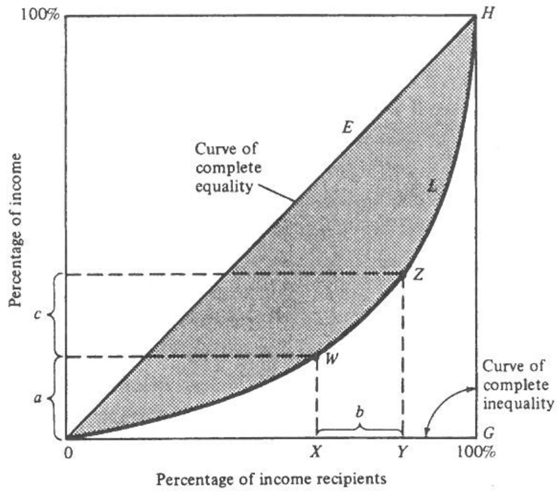
For a given distribution of income there are many ways to measure inequality. In this Note we will discuss only the measures that are used in the selections included in this chapter.

Estimation of the distribution of income presents many difficulties. Ideally we would like to use estimates of the incomes of all households in a country. By contrast, to estimate income or GNP per capita we only need estimates of two quantities: total income and total population. Because the data collection requirements are so great, countries rarely report estimates of their income distributions more frequently than every 10 years, and for many countries there exist no reliable estimates of income distribution at all. The estimates we do have are invariably based on the incomes of households in a given year rather than averaged over many years. (They are also almost always based on before-tax rather than after-tax incomes and are sometimes based on incomes received by individuals rather than households.) To see the problem this presents for measurement of inequality, consider a hypothetical country inhabited entirely by farmers, some of whom grow crops that benefit from above average rainfall and the rest of whom grow crops that benefit from below average rainfall. Suppose that all farmers within each group are identical, and that on average (that is, for average rainfall) these two groups of farmers earn the same household incomes. We would then want to say that there is no inequality in the distribution of income for this country. However, in any given year rainfall will differ from the average, and unequal incomes will be observed. Simon Kuznets lists the specifications for an ideal estimate of income distribution at the beginning of Selection VIII.A.1.

Ignoring these difficulties, we can measure inequality using the estimated income distributions that we have. Two popular measures are the ratio of the share of income received by the tenth *decile* (richest 10 percent) of households to the share of income received by the first decile (poorest 10 percent) of households, and the ratio of the share of income received by the fifth *quintile* (richest 20 percent) of households to the share of income received by the first quintile (poorest 20 percent) of households. Both of these measures are reported in Exhibit VIII.1. The quintile ratio is used in the empirical analysis of Matthew Higgins and Jeffrey G. Williamson in Selection VIII.A.4. The selections by William Easterly (VIII.B.1) and by Alberto Alesina and Roberto Perotti (VIII.B.2) label the shares of income received by the second, third, and fourth quintiles and by the third and fourth quintiles, respectively, as the "middle-class" income shares.

A more rigorous way to measure income inequality is to use *Lorenz curves.* A Lorenz curve plots the percentage of a country's income received by the poorest x percent of households against x. Thus in Figure 1, OX percent of households (the poorest group) receives a percent of income, and so on, giving the Lorenz curve L. Complete equality would occur only if a percent of households received a percent of income, yielding the curve of complete equality E. The curve of perfect inequality is OGH, with a right angle at G. This curve represents the case where one household has 100 percent of the country's income. In practice the Lorenz curve is typically fitted to data giving income shares by decile, that is, the points on the Lorenz curve that are actually observed are the percentage of income received by the poorest 10 percent of households, the percentage of income received by the poorest 20 percent of households, and so on. If the Lorenz curve for one income distribution α lies above that of another income distribution β for at least one point and never lies below it, that is, if distribution α *Lorenz dominates* distribution β, we say that distribution α is more equal than distribution β. It can be shown (see, e.g., Fields and Fei 1978) that if distribution α Lorenz dominates distribution β for the same level of income, α can be obtained from β by transferring positive amounts of income from the relatively rich to the relatively poor. If an additional dollar of income is worth less to the relatively rich than to the relatively poor (i.e., if there is diminishing marginal utility of income), we can judge that this transfer increases social welfare. In short, for a given

Figure 1. The Lorenz Curve



level of income, Lorenz dominance gives us a measure of income inequality that we can link to social welfare, with lower inequality implying higher welfare.

This feature makes measurement of income inequality using Lorenz curves very attractive. Unfortunately, it does not allow us to compare inequality of income distributions whose Lorenz curves cross. One way to deal with this problem is to measure inequality using the *Gini coefficient,* computation of which is illustrated in Figure 1. The shaded area in the figure, enclosed by the theoretical line of equality *E* and the observed Lorenz curve *L,* is known as the concentration area or area of inequality. The Gini coefficient is the ratio of this area to the total area under the line of equality. One way to compute the Gini coefficient is to take the sum of the areas of all "trapezoids" such as *WXYZ* and subtract it from the area under *E* to obtain the concentration area. The required ratio then follows. As a measure of income inequality, the Gini coefficient ranges from 0 to 1—the larger the coefficient, the greater the inequality. Thus 0 represents perfect equality, and 1 represents perfect inequality. It is clear that if one income distribution Lorenz dominates another, its Gini coefficient will be smaller, but the Gini coefficient also allows us to rank inequality for income distributions whose Lorenz curves cross. The Gini coefficient is reported in Exhibit VIII.1 and is a measure of income inequality used in the selections here by Higgins and Williamson (VIII.A.4), by the World Bank (VIII.C.1), and by Hal Hill (VIII.C.2).

A disadvantage of the Gini coefficient is that it cannot be decomposed into inequality within and inequality between groups, that is, the sum of inequality within and inequality between groups will not add up exactly to the overall Gini coefficient. This disadvantage is especially important when one wants to analyze a change in inequality that occurs over time, since one can often obtain a great deal of insight by analyzing how much of the change can be attributed to a change in within-group inequality and how much can be attributed to a change in between-group inequality. Many groupings are interesting, for example, by levels of education of the heads of households, but in the selections in this chapter attention is focused primarily on the division between rural and urban or agricultural and nonagricultural households.

The decomposable measure of income inequality that is used to analyze changes in inequality in the selections by Sherman Robinson (VIII.A.2) and by James Rauch (VIII.A.3) is the variance in the logarithm of income. Here we give the formula for the log variance using Robinson's notation. Let y_i equal the income of household or individual i and $Y_i = \ln(y_i)$. If there are n households or individuals in the country, the average of the logarithm of income or log mean is given by

$$Y = \sum_{i=1}^{n} Y_i/n$$

and the log variance is given by

$$\sigma^2 = \sum_{i=1}^{n} (Y_i - Y)^2/n$$

One reason for using the log variance rather than the variance is that it does not change when the units of measurement for income change, e.g., it is the same whether incomes are measured in dollars or cents. In Selection VIII.A.2 Robinson gives the decomposition formula for the log variance in the case of two groups. Another decomposable measure of income inequality is the Theil index, used by the World Bank in Selection VIII.C.1. The Theil index is given by

$$T = \sum_{i=1}^{n} s_i[\ln(s_i) - \ln(1/n)],$$

where s_i is the share of household or individual i in total national income. Note that if every household or individual has the same income, both the log variance and the Theil index equal zero. For a more thorough discussion of the advantages and disadvantages of various measures of inequality, see Foster (1985).

References

Fields, Gary S., and John C. H. Fei. 1978. "On Inequality Comparisons." *Econometrica* 46 (March): 303–16.

Foster, James. 1985. "Inequality Measurement." In Peyton Young, ed., *Fair Allocation* (Providence, R.I.: American Mathematical Society).