Chapter 6

# Inequality Measures


## Summary

Inequality is a broader concept than poverty in that it is defined over the entire population, and does not only focus on the poor.

The simplest measurement of inequality sorts the population from poorest to richest and shows the percentage of expenditure (or income) attributable to each fifth (quintile) or tenth (decile) of the population. The poorest quintile typically accounts for 6–10 percent of all expenditure, the top quintile for 35–50 percent.

A popular measure of inequality is the Gini coefficient, which ranges from 0 (perfect equality) to 1 (perfect inequality), but is typically in the range of 0.3 to 0.5 for per capita expenditures. The Gini coefficient is derived from the Lorenz curve, which sorts the population from poorest to richest, and shows the cumulative proportion of the population on the horizontal axis and the cumulative proportion of expenditure (or income) on the vertical axis. While the Gini coefficient has many desirable properties—mean independence, population size independence, symmetry, and Pigou-Dalton Transfer sensitivity—it cannot easily be decomposed to show the sources of inequality.

The best known entropy measures are Theil's T and Theil's L, both of which allow one to decompose inequality into the part that is due to inequality within areas (for example, urban and rural) and the part that is due to differences between areas (for example, the rural-urban income gap), as well as the sources of changes in inequality over time. Typically, at least three-quarters of inequality in a country is due to within-group inequality, and the remaining quarter to between-group differences.

Atkinson's class of inequality measures is quite general, and is sometimes used. The decile dispersion ratio, defined as the expenditure (or income) of the richest

decile divided by that of the poorest decile, is popular but a very crude measure of inequality.

A Pen's Parade graph can be useful in showing how incomes, and income distribution, change over time. Microsimulation exercises are increasingly used to identify the sources of changes in income distribution, and to identify changes resulting from changes in prices, in endowments, in occupational choice, and in demographics.

## Learning Objectives

After completing the chapter on *Inequality Measures*, you should be able to

1. Explain what inequality is and how it differs from poverty.
2. Compute and display information on expenditure (or income) quintiles.
3. Draw and interpret a Lorenz curve.
4. Compute and explain the Gini coefficient of inequality.
5. Argue that the Gini coefficient satisfies mean independence, population size independence, symmetry, and Pigou-Dalton Transfer sensitivity, but is not easily decomposable.
6. Draw a Pen's Parade for expenditure per capita, and explain why it is useful.
7. Compute and interpret generalized entropy measures, including Theil's T and Theil's L.
8. Compute and interpret Atkinson's inequality measure for different values of the weighting parameter $\varepsilon$.
9. Compute and criticize the decile dispersion ratio.
10. Decompose inequality using Theil's T to distinguish between-group from within-group components of inequality, for separate geographic areas and occupations.
11. Identify the main sources of changes in inequality using Theil's L.
12. Explain how microsimulation techniques can be used to quantify the effect on income distribution of changes in prices, endowments, occupational choice, and demographics.

## Introduction: Definition of Inequality

Much of this handbook focuses on poverty—the situation of individuals or households who find themselves at the bottom of the income distribution. Typically, analyzing poverty requires information both about the mean level of, say, expenditure

**Table 6.1 Breakdown of Expenditure per Capita by Quintile, Vietnam, 1993**

| | Expenditure quintiles | | | | | |
|---|---|---|---|---|---|---|
| Indicator | Lowest | Low-mid | Middle | Mid-upper | Upper | Overall |
| Per capita expenditure (thousand dong/year) | 518 | 756 | 984 | 1,338 | 2,540 | 1,227 |
| Percentage of expenditure | 8.4 | 12.3 | 16.0 | 21.8 | 41.4 | 100.0 |
| Memo: Cumulative percentage of expenditure | 8.4 | 20.7 | 36.7 | 58.5 | 100.0 | |
| Memo: Cumulative percentage of population | 20.0 | 40.0 | 60.0 | 80.0 | 100.0 | |

*Source:* Authors' computations, based on the Vietnam Living Standards Survey 1993.
*Note:* Totals may not add up due to slight rounding errors.

per capita, as well as its distribution at the lower end. But sometimes we are more interested in measuring inequality than poverty, which is why we have included this chapter.

Inequality is a broader concept than poverty in that it is defined over the entire population, not just for the portion of the population below a certain poverty line. Most inequality measures do not depend on the mean of the distribution; this property of mean independence is considered to be a desirable feature of an inequality measure. Of course, inequality measures are often calculated for distributions other than expenditure—for instance, for income, land, assets, tax payments, and many other continuous and cardinal variables.

The simplest way to measure inequality is by dividing the population into fifths (quintiles) from poorest to richest, and reporting the levels or proportions of income (or expenditure) that accrue to each level. Table 6.1 shows the level of expenditure per capita, in thousand dong per year, for Vietnam in 1993, based on data from the Vietnam Living Standards Survey. A fifth of the individuals (not households) included in the survey were allocated to each expenditure quintile. The figures show that 8.4 percent of all expenditures were made by the poorest fifth of individuals, and 41.4 percent by the top fifth. Quintile information is easy to understand, although sometimes a summary measure is needed rather than a whole table of figures.

## Commonly Used Summary Measures of Inequality

Several summary measures of inequality have been developed, and in this section we present the most important of these. For further details, see the classic book by Atkinson (1983); Duclos and Araar (2006) provide a more technical treatment, and Araar and Duclos (2006) summarize the details of DAD, a very useful software package they developed specifically for the accurate measurement of inequality and poverty.

## Decile Dispersion Ratio

A simple and popular measure of inequality is the decile dispersion ratio, which presents the ratio of the average consumption (or income) of the richest 10 percent of the population to the average consumption (or income) of the poorest 10 percent. This ratio can also be calculated for other percentiles (for instance, dividing the average consumption of the richest 5 percent, the 95th percentile, by that of the poorest 5 percent, the 5th percentile).

The decile dispersion ratio is readily interpretable, by expressing the income of the top 10 percent (the "rich") as a multiple of that of those in the poorest decile (the "poor"). However, it ignores information about incomes in the middle of the income distribution, and does not even use information about the distribution of income within the top and bottom deciles.

## Gini Coefficient of Inequality

The most widely used single measure of inequality is the Gini coefficient. It is based on the Lorenz curve, a cumulative frequency curve that compares the distribution of a specific variable (for example, income) with the uniform distribution that represents equality. To construct the Gini coefficient, graph the *cumulative* percentage of households (from poor to rich) on the horizontal axis and the *cumulative* percentage of expenditure (or income) on the vertical axis. The Lorenz curve shown in figure 6.1 is based on the Vietnamese data in table 6.1. The diagonal line represents perfect equality. The Gini coefficient is defined as A/(A + B), where A and B are the areas shown in the figure. If A = 0, the Gini coefficient becomes 0, which means perfect equality, whereas if B = 0, the Gini coefficient becomes 1, which means complete inequality. In this example, the Gini coefficient is about 0.35. Some users, including the World Bank, multiply this number by 100, in which case it would be reported as 35.

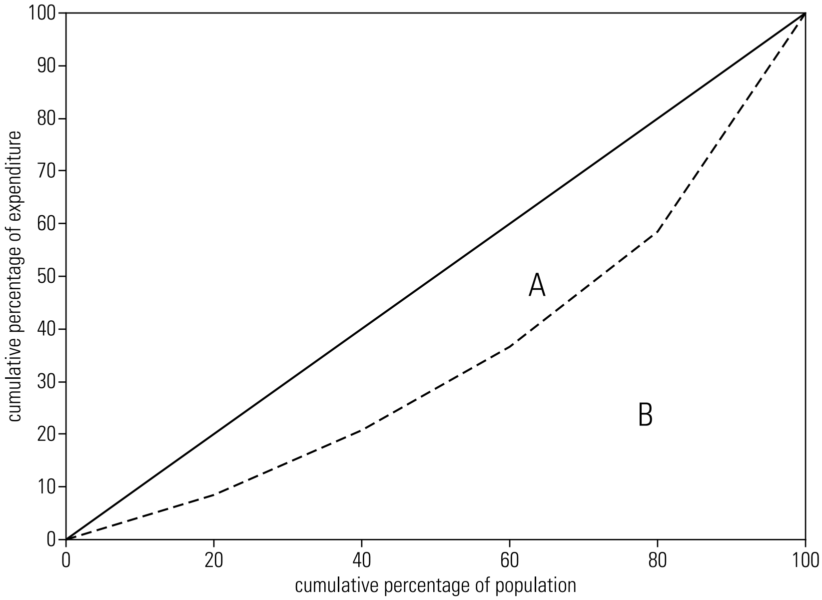
Formally, let $x_i$ be a point on the x-axis, and $y_i$ a point on the y-axis. Then

$$Gini = 1 - \sum_{i=1}^{N} (x_i - x_{i-1})(y_i + y_{i-1}). \tag{6.1}$$

When there are N equal intervals on the x-axis, equation (6.1) simplifies to

$$Gini = 1 - \frac{1}{N}\sum_{i=1}^{N} (y_i + y_{i-1}). \tag{6.2}$$

**Figure 6.1 Lorenz Curve**




*Source:* Authors' illustration.

For users of Stata, there is a "`fastgini`" command that can be downloaded and used directly (see appendix 3). This command also allows weights to be used, a capability not incorporated into equations (6.1) and (6.2). This Stata routine also allows the standard error of the Gini coefficient to be computed using a jackknife procedure.[1] The free, stand-alone DAD software (Araar and Duclos 2006) allows one to measure a wide array of measures of poverty and inequality, including the Gini coefficient.

Table 6.2 shows that the value of the Gini coefficient for expenditure per capita in Vietnam rose from 0.313 in 1993 to 0.350 in 1998. The jackknife standard errors for these estimates are small, and the 95 percent confidence intervals do not overlap; therefore, we can say with some confidence that inequality—as measured by the Gini coefficient, at least—rose during this period. Similarly, it is clear that inequality within the urban areas of Vietnam in 1998 was substantially greater than within rural areas, and this difference is highly statistically significant.

The Gini coefficient is not entirely satisfactory. To see this, consider the criteria that make a good measure of income inequality:

- *Mean independence.* If all incomes were doubled, the measure would not change. The Gini satisfies this.

- *Population size independence.* If the population were to change, the measure of inequality should not change, all else equal. The Gini satisfies this, too.

**Table 6.2 Inequality in Vietnam, as Measured by the Gini Coefficient for Expenditure per Capita, 1993 and 1998**

| | | | 95% confidence interval | |
|---|---|---|---|---|
| Year and area | Gini | Standard error | Lower bound | Upper bound |
| 1993 | 0.3126 | 0.0045 | 0.3039 | 0.3213 |
| 1998 | 0.3501 | 0.0042 | 0.3419 | 0.3584 |
| 1998, urban | 0.3372 | 0.0068 | 0.3238 | 0.3505 |
| 1998, rural | 0.2650 | 0.0037 | 0.2578 | 0.2721 |

*Source:* Authors' calculations based on Vietnam Living Standards Surveys of 1992–93 and 1998.

- *Symmetry.* If any two people swap incomes, there should be no change in the measure of inequality. The Gini satisfies this.
- *Pigou-Dalton Transfer sensitivity.* Under this criterion, the transfer of income from rich to poor reduces measured inequality. The Gini satisfies this, too.

It is also desirable to have

- *Decomposability.* Inequality may be broken down by population groups or income sources or in other dimensions. The Gini index is not easily decomposable or additive across groups. That is, the total Gini of society is not equal to the sum of the Gini coefficients of its subgroups.
- *Statistical testability.* One should be able to test for the significance of changes in the index over time. This is less of a problem than it used to be because confidence intervals can typically be generated using bootstrap techniques.

## Generalized Entropy Measures

There are a number of measures of inequality that satisfy all six criteria. Among the most widely used are the Theil indexes and the mean log deviation measure. Both belong to the family of generalized entropy (GE) inequality measures. The general formula is given by

$$GE(\alpha) = \frac{1}{\alpha(\alpha-1)}\left[\frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i}{\bar{y}}\right)^{\alpha} - 1\right], \tag{6.3}$$

where $\bar{y}$ is the mean income per person (or expenditure per capita). The values of GE measures vary between zero and infinity, with zero representing an equal distribution and higher values representing higher levels of inequality. The parameter $\alpha$ in the GE class represents the weight given to distances between incomes at different parts of the income distribution, and can take any real value. For lower values of $\alpha$, GE is more sensitive to changes in the lower tail of the distribution, and for higher values GE is more sensitive to changes that affect the upper tail. The most common values of α used are 0, 1, and 2. GE(1) is Theil's T index, which may be written as

$$GE(1) = \frac{1}{N} \sum_{i=1}^{N} \frac{y_i}{\bar{y}} \ln\left(\frac{y_i}{\bar{y}}\right). \tag{6.4}$$

GE(0), also known as Theil's L, and sometimes referred to as the mean log deviation measure, is given by

$$GE(0) = \frac{1}{N} \sum_{i=1}^{N} \ln\left(\frac{\bar{y}}{y_i}\right). \tag{6.5}$$

Once again, users of Stata do not need to program the computation of such measures from scratch; the "`ineqdeco`" command," explained in appendix 3, allows one to obtain these measures, even when weights need to be used with the data.

## Atkinson's Inequality Measures

Atkinson (1970) has proposed another class of inequality measures that are used from time to time. This class also has a weighting parameter ε (which measures aversion to inequality). The Atkinson class, which may be computed in Stata using the "`ineqdeco`" command, is defined as

$$\begin{aligned} A_\varepsilon &= 1 - \left[\frac{1}{N} \sum_{i=1}^{N} \left(\frac{y_i}{\bar{y}}\right)^{1-\varepsilon}\right]^{1/(1-\varepsilon)}, \quad && \varepsilon \neq 1 \\ &= 1 - \frac{\prod_{i=1}^{N} (y_i^{(1/N)})}{\bar{y}}, && \varepsilon = 1. \end{aligned} \tag{6.6}$$

Table 6.3 sets out in some detail the steps involved in the computation of the GE and Atkinson measures of inequality. The numbers in the first row give the incomes of the 10 individuals who live in a country, in regions 1 and 2. The mean income is 33. To compute Theil's T, first compute $y_i/\bar{y}$ , where $\bar{y}$ is the mean income level; then compute $\ln(y_i/\bar{y})$, take the product, add up the row, and divide by the number of people. Similar procedures yield other GE measures, and the Atkinson measures, too.

Table 6.4 provides some examples of different measures of inequality (Dollar and Glewwe 1998, 40); Gottschalk and Smeeding (2000) summarize evidence on inequality for the world's "industrial" countries. All three measures agree that inequality is lowest in Vietnam, followed closely by Ghana, and is highest in Côte d'Ivoire. This illustrates another point: in practice, the different measures of inequality typically tell the same story, so the choice of one measure over another is not of crucial importance in the discussion of income (or expenditure) distribution.

**Table 6.3 Computing Measures of Inequality**

| Measure | | Region 1 | | | | | Region 2 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Incomes ($y_i$) | | 10 | 15 | 20 | 25 | 40 | 20 | 30 | 35 | 45 | 90 |
| Mean income ($\bar{y}$) | 33.00 | | | | | | | | | | |
| $y_i / \bar{y}$ | | 0.30 | 0.45 | 0.61 | 0.76 | 1.21 | 0.61 | 0.91 | 1.06 | 1.36 | 2.73 |
| $\ln(y_i / \bar{y})$ | | –0.52 | –0.34 | –0.22 | –0.12 | 0.08 | –0.22 | –0.04 | 0.03 | 0.13 | 0.44 |
| Product | | –0.16 | –0.16 | –0.13 | –0.09 | 0.10 | –0.13 | –0.04 | 0.03 | 0.18 | 1.19 |
| **GE(1), Theil's T** | **0.080** | | | | | | | | | | |
| $\ln(\bar{y}/y_i)$ | | 0.52 | 0.34 | 0.22 | 0.12 | –0.08 | 0.22 | 0.04 | –0.03 | –0.13 | –0.44 |
| **GE(0), Theil's L** | **0.078** | | | | | | | | | | |
| $(y_i / \bar{y})^2$ | | 0.09 | 0.21 | 0.37 | 0.57 | 1.47 | 0.37 | 0.83 | 1.12 | 1.86 | 7.44 |
| **GE(2)** | **0.666** | | | | | | | | | | |
| $(y_i / \bar{y})^{0.5}$ | | 0.55 | 0.67 | 0.78 | 0.87 | 1.10 | 0.78 | 0.95 | 1.03 | 1.17 | 1.65 |
| **Atkinson, $\varepsilon$ = 0.5** | **0.087** | | | | | | | | | | |
| $(y_i)^{1/n}$ | | 1.26 | 1.31 | 1.35 | 1.38 | 1.45 | 1.35 | 1.41 | 1.43 | 1.46 | 1.57 |
| **Atkinson, $\varepsilon$ = 1** | **0.164** | | | | | | | | | | |
| $(y_i / \bar{y})^{-1}$ | | 3.30 | 2.20 | 1.65 | 1.32 | 0.83 | 1.65 | 1.10 | 0.94 | 0.73 | 0.37 |
| **Atkinson, $\varepsilon$ = 2** | **0.290** | | | | | | | | | | |

*Source:* Authors' compilation.

**Table 6.4 Expenditure Inequality in Selected Developing Countries**

| Country | Gini coefficient | Theil's T | Theil's L |
|---|---|---|---|
| Côte d'Ivoire, 1985–86 | 0.435 | 0.353 | 0.325 |
| Ghana, 1987–88 | 0.347 | 0.214 | 0.205 |
| Jamaica, 1989 | — | 0.349 | 0.320 |
| Peru, 1985–86 | 0.430 | 0.353 | 0.319 |
| Vietnam, 1992–93 | 0.344 | 0.200 | 0.169 |

*Source:* Reported in Dollar and Glewwe (1998, 40).

*Note:* — = Not available. The numbers for Vietnam differ from those shown in table 6.2 because a slightly different measure of expenditure per capita was used in the two cases.

One caveat is in order: income is more unequally distributed than expenditure. This is a consequence of household efforts to smooth consumption over time. It follows that when comparing inequality across countries it is important to compare either Gini coefficients based on expenditure, or Gini coefficients based on income, but not mix the two.

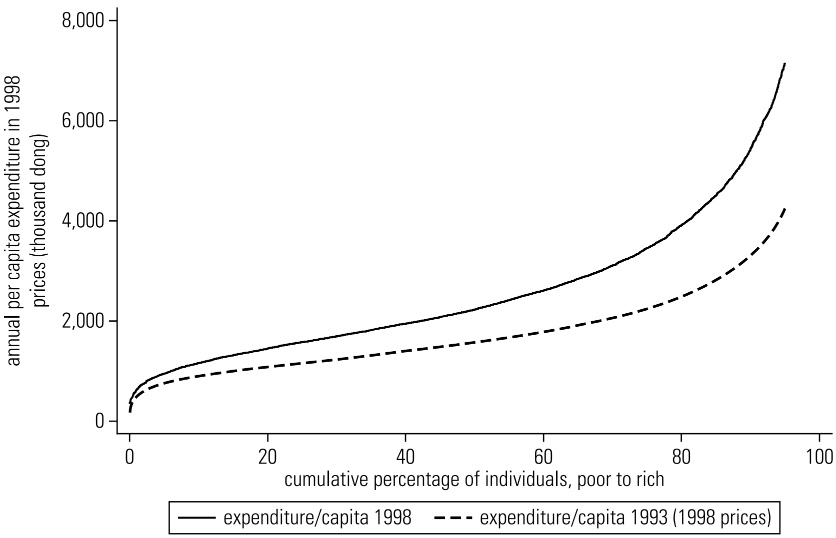
## Inequality Comparisons

Many of the tools used in the analysis of poverty can be similarly used for the analysis of inequality. Analogous to a poverty profile (see chapter 7), one could draw a profile of inequality, which, among other things, would look at the extent of inequality among certain groups of households. This profile provides information on the homogeneity of the various groups, an important element to take into account when designing policy interventions.

The nature of changes in inequality over time can also be analyzed. One could focus on changes for different groups of the population to show whether inequality changes have been similar for all or have taken place, say, in a particular sector of the economy. In rural Tanzania, average incomes increased substantially between 1983 and 1991—apparently tripling over this period—but inequality increased, especially among the poor. Although the nationwide Gini coefficient for income per adult equivalent increased from 0.52 to 0.72 during this period, the poverty rate fell from 65 percent to 51 percent. Ferreira (1996) argues that a major cause of the rise in both rural incomes and rural inequality was a set of reforms in agricultural price policy; despite higher prices, poorer and less efficient farmers found themselves unable to participate in the growth experienced by wealthier, more efficient farmers.

In comparing distributions over time, one of the more useful graphs is a *Pen's Parade,* which is a form of *quantile graph*. On the horizontal axis, every person is lined up from poorest to richest, while the vertical axis shows the level of expenditure (or income) per capita. Often the graph is truncated toward the upper end of the distribution, to focus on changes at the lower end, including the zone in which people are in poverty. If the axes were flipped, this graph would simply be a cumulative density

**Figure 6.2 Pen's Parade (Quantile Function) for Expenditure per Capita, Vietnam, 1993 and 1998**




*Source:* Created by the authors, based on data from the Vietnam Living Standards Surveys of 1992–93 and 1998.

*Note:* This function is truncated at the 95th percentile.

function. The Pen's Parade is most helpful when comparing two different areas or periods. This is clear from figure 6.2, which shows the graphs for expenditure per capita for Vietnam for 1993 and 1998; over this five-year period, incomes (and spending) rose across the board, and although inequality increased, wages were still higher at the bottom of the distribution in 1998 than in 1993. There is nothing inevitable about this; Ferreira and Paes de Barros (2005) show an interesting quantile graph of income per person in urban Brazil; between 1976 and 1996 incomes on average rose slightly and inequality on average was reduced, yet the position of those at the very bottom actually worsened—a feature that appears very clearly on their Pen's Parade.

## Measuring Pro-Poor Growth

As national income (or expenditure) rises, the expenditure of the poor may rise more or less quickly than that of the country overall. A visually compelling way to show this effect is with a *growth incidence curve*, which can be computed as long as data are available from surveys undertaken at two times. The procedure is as follows:

a. Divide the data from the first survey into centiles—for instance, using the `xtile` command in Stata—and compute expenditure per capita for each of the 100 centiles.

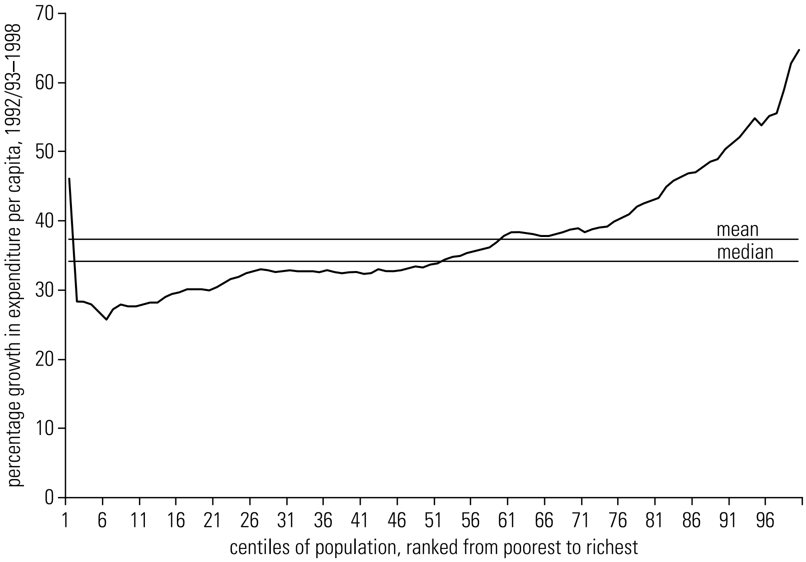
b. Divide the data from the second survey into centiles, and again compute expenditure per capita for each centile.

c. After adjusting for inflation, compute the percentage change in (real) expenditure per capita for each centile and graph the results.

Figure 6.3 shows just such a graph for Vietnam and compares the outcomes of the Vietnam Living Standards Surveys of 1992–93 and 1998; it uses the same data as figure 6.2. During this period, the mean increase in real expenditure per capita was 37 percent, and the median increase was 34 percent. It is clear from the graph that, with the exception of the very poorest centile, expenditure rose less quickly for those in the lower part of the expenditure distribution than for those who were better off.

Even so, expenditure rose substantially even for the poor. One way to measure the "rate of pro-poor growth," suggested by Ravallion and Chen (2003), is to compute the mean growth rate of expenditure per capita experienced by the poor. In 1992–93 the headcount poverty rate was 57 percent in Vietnam; averaging the growth in expenditure for this group using the data that underlie the growth incidence curve yields 32 percent. This means that, although expenditure per capita rose by 37 percent nationwide over the five-year interval between the two surveys, the increase was 32 percent for the poor.

**Figure 6.3 Poverty Incidence Curve for Expenditure per Capita, Vietnam, 1993 and 1998**




*Source:* Created by the authors, based on data from the Vietnam Living Standards Surveys of 1992–93 and 1998.

The growth incidence curve reflects averages. The incomes of the poor might rise on average, but some poor households might still find themselves worse off. The examination of poverty dynamics of this nature requires panel data; this topic is discussed in some detail in chapter 11.

## Decomposition of Income Inequality

The common inequality indicators mentioned above can be used to assess the major contributors to inequality, by different subgroups of the population and by region. For example, average income may vary from region to region, and this alone implies some inequality "between groups." Moreover, incomes vary inside each region, adding a "within-group" component to total inequality. For policy purposes, it is useful to be able to decompose these sources of inequality: if most inequality is due to disparities across regions, for instance, then the focus of policy may need to be on regional economic development, with special attention to helping the poorer regions.

More generally, household income is determined by household and personal characteristics, such as education, gender, and occupation, as well as geographic factors including urban and regional location. Some overall inequality is due to differences in such characteristics—this is the "between-group" component—and some occurs because there is inequality within each group, for instance, among people with a given level of education or in a given occupation. The generalized entropy (GE) class of indicators, including the Theil indexes, can be decomposed across these partitions in an additive way, but the Gini index cannot.

To decompose Theil's T index (that is, GE(1)), let $Y$ be the total income of all $N$ individuals in the sample, and $\bar{y} = Y/N$ be mean income. Likewise, $Y_j$ is the total income of a subgroup (for example, the urban population) with $N_j$ members, and $\bar{y}_j = Y_j/N_j$ is the mean income of this subgroup. Using $T$ to represent GE(1),

$$\begin{aligned} T &= \sum_{i=1}^{N} \frac{y_i}{N\bar{y}} \ln\left(\frac{y_i N}{\bar{y} N}\right) \\ &= \sum_{i=1}^{N} \frac{y_i}{Y} \ln\left(\frac{y_i N}{Y}\right) \\ &= \sum_{j} \left(\frac{Y_j}{Y}\right) T_j + \sum_{j} \left(\frac{Y_j}{Y}\right) \ln\left(\frac{Y_j/Y}{N_j/N}\right), \end{aligned} \tag{6.7}$$

where $T_j$ is the value of GE(1) for subgroup $j$. Equation (6.7) separates the inequality measure into two components, the first of which represents within-group inequality while the second term measures the between-group inequality.

> ***Exercise:*** Decompose Theil's T measure of inequality into "within" and "between" components, using the income data provided in table 6.2. (Hint: "Within" inequality should account for 69.1 percent of all inequality.)

A similar decomposition is possible for GE(0); this breakdown of Theil's L is given by

$$L = \sum_{i-1}^{N} \frac{1}{N} \ln\left(\frac{\bar{y}}{Y_i}\right) = \sum_{j} \left(\frac{N_j}{N}\right) L_j + \sum_{j} \frac{N_j}{N} \ln\left(\frac{\bar{y}}{\bar{y}_j}\right). \tag{6.8}$$

When we have information on a welfare measure for two time points, we are often interested in identifying the components of the change in inequality. Defining $n_j$ = N$_j$/N, which is the proportion of those in the sample who are in the $j$th subgroup, and adding the time subscripts 1 (for initial period) and 2 (for the second period), where appropriate, we have, for Theil's L

$$\Delta L \approx \sum_{j} n_j \left[\ln\left(\frac{\bar{y}_{,2}}{\bar{y}_{,1}}\right) - \ln\left(\frac{\bar{y}_{j,2}}{\bar{y}_{j,1}}\right)\right] + \sum_{j} \left[L_j + \ln\left(\frac{\bar{y}}{\bar{y}_j}\right)\right] \Delta n_j + \sum_{j} n_j \Delta L_j \,. \tag{6.9}$$

This decomposition is accurate if the changes are relatively small, and if average values across the two periods (for example, of $n_j$ or $L_j$) are used. The first term on the right-hand side measures the effect on inequality of changes in relative mean incomes; if the income of a small, rich group grows particularly rapidly, for instance, greater inequality is likely to result. The second term measures the effects of shifts in population from one group to another. Finally, the third term in equation (6.9) measures the size of changes in within-group inequality.

These decompositions may be illustrated with data on expenditure per capita for Vietnam, as set out in table 6.5. Using Theil's L, measured inequality rose appreciably between 1993 and 1998. In 1993, about a fifth of inequality was attributable to the urban-rural gap in expenditure levels (after correcting for price differences); by 1998, almost a third of inequality arose from the urban-rural gap, which widened considerably during this period. This shows up in the breakdown of the change in inequality; following equation 6.9 we have

$$\begin{array}{llll} 0.039 & \approx 0.050 & -\ 0.016 & +\ 0.005. \\ \text{[change in L]} & \approx \text{[effect of change in incomes]} & +\ \text{[population shift effect]} & +\ \text{[change in within-group inequality]} \end{array}$$

From this breakdown it appears that the rise in inequality in Vietnam between 1993 and 1998 was mainly due to a disproportionately rapid rise in urban, relative to rural, incomes. This increase was attenuated by a rise in the relative size of the urban population, but exacerbated by a modest increase in inequality within both urban and rural populations.

Similar results were found for Zimbabwe in 1995–96. A decomposition of Theil's T there showed that the within-area (within rural areas and within urban areas) contribution to inequality was 72 percent, while the between-area (between urban and

**Table 6.5 Decomposition of Inequality in Expenditure per Capita by Area, Vietnam, 1993 and 1998**

| | Theil's L (GE(0)) | | |
|---|---|---|---|
| Area | 1993 | 1998 | Change |
| All Vietnam | 0.160 | 0.199 | **0.039** |
| Urban only | 0.173 | 0.189 | 0.016 |
| Rural only | 0.118 | 0.120 | 0.002 |
| Decomposition | | | |
| "Within" inequality | 0.129 | 0.135 | |
| "Between" inequality | 0.031 | 0.064 | |
| Memo: "Between" inequality as percentage of total inequality | 20 | 32 | |

*Source:* Authors' calculations based on Vietnam Living Standards Surveys of 1992–93 and 1998.

rural areas) component was 28 percent. In many Latin American countries, the between-area component of inequality explains an even higher share of total inequality, reflecting wide differences in living standards between one region and another in countries such as Brazil and Peru.

We are often interested in which of the different income sources, or components of a measure of well-being, are primarily responsible for the observed level of inequality. For example, if total income can be divided into self-employment income, wages, transfers, and property income, the distribution of each income source can be examined. If one of the income sources were raised by 1 percent, what would happen to overall inequality? The simplest and most commonly used procedure is to compute the measure of inequality using the initial data, and then to simulate a new distribution (for instance, by raising wages by 1 percent) and recompute the measure of inequality.

Table 6.6 shows the results for the Gini coefficient for income sources in Peru (1997). As the table shows, self-employment income is the most equalizing income source. Thus, a 1 percent increase in self-employment income (for everyone that receives such income) would lower the Gini coefficient by 4.9 percent, which represents a reduction in overall inequality. However, a rise in property income would be associated with an increase in inequality.

Generally, results such as these depend on two factors:

- The importance of the income source in total income (for larger income sources, a given percentage increase will have a larger effect on overall inequality)
- The distribution of that income source (if it is more unequal than overall income, an increase in that source will lead to an increase in overall inequality).

Table 6.6 also shows the effect on the inequality of the distribution of *wealth* of changes in the value of different sources of wealth.

A final example, in the same spirit, comes from the Arab Republic of Egypt. In 1997, agricultural income represented the most important inequality-increasing source of income, while nonfarm income had the greatest inequality-reducing

**Table 6.6 Expected Change in Income Inequality Resulting from a 1 Percent Change in Income (or Wealth) Source, 1997 (as Percentage of Change in Gini Coefficient), Peru**

| Income source | Expected change | Wealth sources | Expected change |
|---|---|---|---|
| Self-employment income | –4.9 | Housing | 1.9 |
| | | Durable goods | –1.5 |
| Wages | 0.6 | Urban property | 1.3 |
| Transfers | 2.2 | Agricultural property | –1.6 |
| Property income | 2.1 | Enterprises | 0 |

*Source:* Rodriguez 1998.

**Table 6.7 Decomposition of Income Inequality in Rural Egypt, 1997**

| Income source | Percentage of households receiving income from this source | Share in total income (%) | Concentration index for the income source | Percentage contribution to overall income inequality |
|---|---|---|---|---|
| Nonfarm | 61 | 42 | 0.63 | 30 |
| Agricultural | 67 | 25 | 1.16 | 40 |
| Transfer | 51 | 15 | 0.85 | 12 |
| Livestock | 70 | 9 | 0.94 | 6 |
| Rent | 32 | 8 | 0.92 | 12 |
| All sources | 100 | 100 | | 100 |

*Source:* Adams 1999, 32.

potential. Table 6.7 sets out this decomposition and shows that while agricultural income represents only 25 percent of total income in rural areas, it accounts for 40 percent of the inequality.

## Income Distribution Dynamics

There is a longstanding, if inconclusive, debate about the links between income distribution and economic growth. Simon Kuznets (1966), based on his analysis of the historical experience of the United Kingdom and the United States, believed that in the course of economic development, inequality first rises and then falls. Although there are other cases where this pattern has been observed, it is by no means inevitable. There are many components of inequality, and they may interact very differently depending on the country. Bourguignon, Ferreira, and Lustig (2005, 2) emphasize this diversity of outcomes, and argue that changes in income distribution are largely due to three "fundamental forces":

- Changes in the distribution of assets and the personal characteristics of the population (for example, educational levels, gender, ethnicity, capital accumulation)—the *endowment effects*
- Changes in the returns to these assets and characteristics (for example, the wage rate, or profit rate)—the *price effects*
- Changes in how people deploy their assets, especially in the labor market (for example, whether they work, and if they do, in what kind of job)—the *occupational choice effects*.

To these three one might also add *demographic effects*; for instance, if households have fewer children, the earnings of working members will stretch further, and measured income per capita will rise.

In *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, Bourguignon, Ferreira, and Lustig (2005) set out and apply an approach that is designed to allow quantification of the effects on the whole income distribution of the various changes in "fundamental forces" that occur between two time points. Over time, the way that households choose their jobs, the returns from different types of employment, and the assets (especially education) that households bring to the labor market, all change, and as they do, so does the distribution of income. Thus, the idea is to set up basic parametric models of occupational choice and earnings, to measure these at different times, and then to separate out the effects of changing returns from the effects of changing endowments.

To illustrate, consider the following simplified and stylized example. Suppose that in 1995 we find that wages are related to years of education as follows:

$$\text{In}(w_i) = 4 + 0.03ed_i - 0.00075(ed_i)^2 + \varepsilon_i. \tag{6.10}$$

Here $ed_i$ measures years of schooling for individual $i$, $w_i$ is the wage rate, and $\varepsilon_i$ is an error term that picks up measurement error and the influences of unobserved variables (such as ability, for instance). In this case, an individual with six years of schooling could expect to earn a wage of 63.6; if the person were particularly vigorous and able, such that their actual wage were 85.9, this expression implies that for this person, the observed residual would be $e_i = 0.3$. Now suppose that in 2006 we find, on the basis of new survey data, that

$$\text{In}(w_i) = 3.9 + 0.027ed_i - 0.00085(ed_i)^2 + \varepsilon_i. \tag{6.11}$$

If the individual still has six years of schooling, and is as vigorous and able as before, we could expect his wage to be 76.1. This reflects a reduction in the return on education that appears to have occurred in this society between 1995 and 2006. However, if the individual now has nine years of education, the new wage could be expected to be 79.4. If we perform similar calculations for all individuals in a survey, we can simulate the effects on income distribution of the changes in the "fundamental forces."

The more sophisticated applications of income distribution dynamics are relatively intricate; Bourguignon, Ferreira, and Lustig (2005) provide more details. But the result of this extra effort is that one gains more insight into the factors that drive income distribution. Consider the numbers for urban Brazil shown in table 6.8. Between 1976 and 1996, inequality in income per capita fell very slightly—the Gini coefficient fell from 0.595 to 0.591—and incomes grew slightly. Yet, the incidence of severe poverty rose significantly. Ferreira and Paes de Barros (2005) argue that during this period a number of people were trapped at the bottom of the income distribution, excluded from labor markets (note the rise in the open unemployment rate), and not covered by formal safety nets. At the same time, the rate of return on education fell, but the average level of schooling rose sharply (from 3.2 to 5.3 years per person).

**Table 6.8 Economic Indicators for Brazil, 1976 and 1996**

| Indicator | 1976 | 1996 |
|---|---|---|
| Gini, urban areas, income per capita | 0.595 | 0.591 |
| Poverty headcount rate: poverty line of 30 reais/month in 1976 prices | 0.068 | 0.092 |
| Poverty headcount rate: poverty line of 60 reais/month in 1976 prices | 0.221 | 0.218 |
| Household income per capita per month, in 1976 reais | 265 | 276 |
| Open unemployment rate (%) | 1.8 | 7.0 |
| Percentage employed in formal-sector jobs | 58 | 32 |
| Average years of schooling | 3.2 | 5.3 |

*Source:* Ferreira and Paes de Barros 2005.

The net effect was to create greater equality of incomes in the middle and upper ends of the distribution, while those at the very bottom became worse off.

This type of microsimulation exercise allows one to ask questions such as how much would the Gini coefficient have changed if the only fundamental force to vary between 1976 and 1996 were a change in the return to education for wage earners? The answer in this particular case is that the Gini would have risen from 0.595 to 0.598. Or again, if the only change had been the rise in education, the Gini would have fallen from 0.595 to 0.571. By identifying effects such as these, it is possible to construct a fuller and clearer story about what has driven changes in the distribution of income over time.

## Review Questions

1. You are provided with the following information:

| | | | | | | |
|---|---|---|---|---|---|---|
| a. | Quintile | 20 | 20 | 20 | 20 | 20 |
| b. | % expenditure | 7 | 12 | 15 | 20 | 46 |
| c. | Cumulative % expenditure | 7 | 19 | 34 | 54 | 100 |
| d. | Expenditure/capita | 350 | 600 | 750 | 1,00 | 2,300 |

The Lorenz curve graphs:

- A. a. on the horizontal axis, b. on the vertical axis.
- B. a. on the horizontal axis, c. on the vertical axis.
- C. a. on the horizontal axis, d. on the vertical axis.
- D. None of the above.

2. If income is transferred from someone who is better off to someone who is less well off, the Gini coefficient always rises (Pigou-Dalton transfer sensitivity).

- True
- False

3. A country has five residents, whose incomes are as follows: 350, 600, 750, 1000, and 2300. Then

- A. Theil's T is 0.114 and Theil's L is 0.344.
- B. Theil's T is 0.444 and Theil's L is 0.344.
- C. Theil's T is 0.114 and Theil's L is 0.113.
- D. Theil's T is 0.444 and Theil's L is 0.344.

4. A Pen's Parade is used for comparing

- A. Income distributions at two points in time.
- B. Expenditure distributions at two points in time.
- C. Income distributions in two different areas.
- D. All of the above.

5. According to table 6.8, between 1976 and 1996 in Brazil, which did *not* occur:

- A. Inequality rose (as measured by the Gini coefficient).
- B. Deep poverty worsened.
- C. The unemployment rate rose.
- D. Income rose.

6. In decomposing inequality, which of the following is *not* true?

- A. A change in Theil's L = the effect of a change in income + population shift effect + change in within-group inequality.
- B. A change in Theil's L = the change in between group inequality + change in within-group inequality.
- C. $L = \sum_{i-1}^{N} \frac{1}{N} \ln\left(\frac{\bar{y}}{Y_i}\right) = \sum_j \left(\frac{N_j}{N}\right) L_j + \sum_j \frac{N_j}{N} \ln\left(\frac{\bar{y}}{\bar{y}_j}\right).$
- D. The change in the Gini coefficient equals $= \sum_j \left(\frac{Y_j}{Y}\right) T_j + \sum_j \left(\frac{Y_j}{Y}\right) \ln\left(\frac{Y_j/Y}{N_j/N}\right),$

7. Suppose that between 2005 and 2007, the urban-rural gap widened but inequality within urban areas stayed the same, and inequality within rural areas did not change either. This implies that between-group inequality rose relative to within-group inequality.

- True
- False

| 8. According to the Kuznets curve, in the course of economic development, |
|---|
| ○ A. Inequality first falls and then rises.<br>○ B. Inequality first rises and then falls.<br>○ C. Inequality begins high and then falls to a lower sustainable level.<br>○ D. Inequality begins low and then rises to a higher sustainable level. |

## Note

1. Suppose that we have a statistic, $\theta$, and would like to calculate its standard error. The statistic could be as simple as a mean, or as complex as a Gini coefficient. Using the full sample, our estimate of the statistic is $\hat{\theta}$. We could also estimate the statistic leaving out the *i*th observation, representing it as $\hat{\theta}_{(i)}$. If there are N observations in the sample, then the jackknife standard error of the statistic is given by $\widehat{se} = [(N-1/N)\sum_{i=1}^{N}(\theta_{(i)} - \theta)^2]^{1/2}$. Provided the statistic of interest is not highly nonlinear, the jackknife estimate typically gives a satisfactory approximation, and it is useful in cases, such as the Gini coefficient, where analytic standard errors may not exist.

## References

Adams, Richard H., Jr. 1999. "Nonfarm Income, Inequality, and Land in Rural Egypt." Policy Research Working Paper No. 2178, World Bank, Washington, DC.

Araar, Abdelkrim, and Jean-Yves Duclos. 2006. "DAD: A Software for Poverty and Distributive Analysis." PMMA Working Paper 2006-10, Université Laval, Quebec.

Atkinson, A. B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory* 2 (3): 244–63.

———. 1983. *The Economics of Inequality*, 2nd edition. Oxford: Clarendon Press.

Bourguignon, François, Francisco Ferreira, and Nora Lustig, eds. 2005. *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. Washington, DC: World Bank and Oxford University Press.

Dollar, David, and Paul Glewwe. 1998. "Poverty and Inequality: The Initial Conditions." In *Household Welfare and Vietnam's Transition*, ed. David Dollar, Paul Glewwe, and Jennie Litvack. World Bank Regional and Sectoral Studies. Washington, DC: World Bank.

Duclos, Jean-Yves, and Abdelkrim Araar. 2006. *Poverty and Equity: Measurement, Policy and Estimation with DAD*. New York: Springer, and Ottawa: International Development Research Centre.

Ferreira, Francisco, and Ricardo Paes de Barros. 2005. "The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil, 1976–1996." In *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, ed. François Bourguignon, Francisco Ferreira, and Nora Lustig. Washington, DC: World Bank and Oxford University Press.

Ferreira, M. Luisa. 1996. "Poverty and Inequality during Structural Adjustment in Rural Tanzania." Policy Research Working Paper No. 1641, World Bank, Washington, DC.

Gottschalk, P., and T. Smeeding. 2000. "Empirical Evidence on Income Inequality in Industrial Countries." In *Handbook of Income Distribution. Volume 1. Handbooks in Economics, vol. 16,* ed. A. B. Atkinson and F. Bourguignon. Amsterdam, New York, and Oxford: Elsevier Science/North-Holland.

Kuznets, Simon. 1966. *Modern Economic Growth: Rate, Structure and Spread*. New Haven and London: Yale University Press.

Ravallion, Martin, and Shaohua Chen. 2003. "Measuring Pro-Poor Growth." *Economics Letters* 78 (1): 93–9.

Rodriguez, Edgard. 1998. "Toward a More Equal Income Distribution? The Case of Peru 1994-1997." Poverty Reduction and Economic Management Network, Poverty Group, World Bank, Washington, DC.

Vietnam General Statistical Office. 2000. *Viet Nam Living Standards Survey 1997–1998*. Hanoi: Statistical Publishing House.

World Bank. 1994. "Viet Nam Living Standards Survey (VNLSS) 1992–93." Poverty and Human Resources Division, World Bank, Washington, DC.