5

MEASURES OF INEQUALITY

In this short chapter, we take up another fairly technical topic: How can inequality be measured? We will not attempt to give a full technical treatment of this question, but will introduce the reader to some of the most commonly used concepts and techniques, and we hope in accessible language.

What Is a Distribution?

In very simple terms, a distribution is a plot that puts income, expenditure, pay, or wealth on the horizontal, and the number of people at each level of the *x*-variable on the vertical axis. In the cases of income, expenditure, or pay, the distribution starts at zero since there is no such thing as a negative value of these variables.

Measures of inequality concern the *shape* of the distribution of income, or wealth, or pay. Typically the unit of observation in the distribution is the individual, or the household. If you think of a diagram with income levels on the *x*-axis and on the *y*-axis the numbers of people or households at each level, the usual shape of the distribution is to rise sharply from a very low value, to peak at the modal income (that level enjoyed, or not enjoyed, by the largest group of people), and then to tail off to the right. Usually the distribution will have a very long tail, reflecting small numbers of people or households with very high income. If society becomes more equal, the part in the middle becomes taller, and the tail flattens out. If society becomes less equal, then the hump in the middle becomes flatter and the tail tends to become a bit fat. Changes in this shape are subject to statistical measurement.

It is possible to fit mathematical functions to the distribution of income, and the usual result resembles a *log-normal* distribution, meaning that the logarithms of the income values follow a bell curve or Gaussian or normal distribution (see Figure 5.1). Closer examination of actual distributions suggest that in reality they are (a bit) better modeled by thinking of them as a blend of two distributions. For the bottom 95 percent or so, the best fit is a log-linear distribution. For those at the

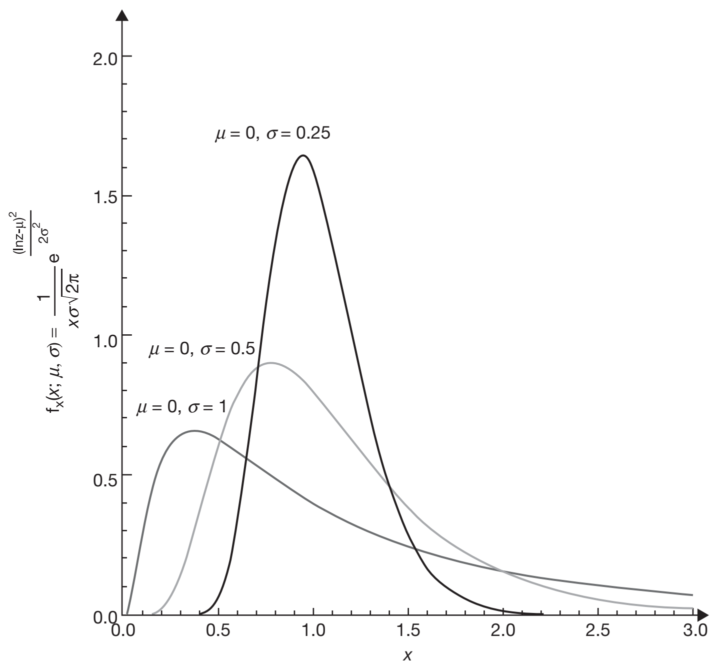

Figure 5.1 Three Log-Normal Distributions

The tall distribution in red is the most equal; that which peaks on the left is the least equal. Note the tails.

very top, the better fit is a Pareto distribution or a power law, which states that the proportion of people with an income above a certain value is a function (for example, a square root) of that value.

A consequence of having a power law distribution at the top is that income may be very unequally shared; there will be more ultra-rich than there would be under a log-normal distribution. (This is known as having a "fat tail.") A power law is the mathematical form underlying the common 80–20 rule, which in its original formulation was the discovery (by Vilfredo Pareto) that 20 percent of Italians held 80 percent of Italian land. Many other variations of this principle have since been proposed, suggesting that power laws are quite common in the real world.

What Is the Range and What Is the Inter-Quartile Range?

The *range* is the simplest possible indicator of a distribution. For income it is the distance, measured in currency units, between the highest and the lowest value in the sample under observation. Simple as that—but not very informative. The low and the high values may be anomalies, far removed from most of the other observations. So to know the range is not to know very much about the population.

The *inter-quartile range* is a way of lopping off the extreme values, in order to look at only the middle part of a distribution, where the "middle 50 percent" of the observations lie. The inter-quartile range is calculated by lining up all the observations in order of income, and then removing both the top and bottom 25 percent. The range for the remainder is the inter-quartile range.

The inter-quartile range conveys some information about the center of the distribution, but not very much. And like the full range, it is measured in the original unit of observation (dollars), so it's very hard to compare to other range measures in other currencies or at different times, when the value of the dollar unit was different from what it is today.

What Are Quantiles and What Are Quantile Ratios?

These simple observations tell us that to measure inequality in a meaningful way, we need a way to do so that is not dependent on the currency unit, which therefore can be compared across space, time, and countries in a consistent manner.

One simple way to do this is again to line up all the observations from low to high, and to count them off in percentage terms, point by point. The resulting groups are called *percentiles*. If instead you count by 10-percent groups, you have *deciles*. If you group in bunches of 20 percent of the population each, you have a standard category known as *a quintile*. Each of these is an instance of a general practice known as taking *quantiles*. (Quintiles, quantiles, let's call the whole thing off.)

A *quantile ratio* is the income level at some particular percentile, divided by the income level at some other percentile. Quantile ratios are commonly known by the two percentiles they involve: thus the 90–10 ratio is the ratio of income at the 90th percentile to income at the 10th percentile. Taking this ratio can be a useful quick-and-dirty way of gauging the changing inequality of a distribution through time, without the potential distortion of a few extreme cases. Similarly, the 90–50 ratio may be taken as an index of inequality among the upper-middle classes, while the 50–10 ratio suggests the extent of relative deprivation at the lower end of the scale. The 75–25 ratio is the inter-quartile ratio, not to be confused with the inter-quartile *range*.

These measures are each useful, and they are unit-free, which means that different distributions can now be compared. But notice that they involve only two pieces of information: the income levels at each of two separated percentiles. All the other information from the underlying distribution—which must be collected in order to compute percentiles in the first place—is simply thrown away.

What Is the Palma Ratio?

The *Palma ratio* is the recent invention of Cambridge University economist Gabriel Palma. It consists of the ratio of the income share of the top decile to that of the bottom two quintiles. The intuition behind choosing this ratio is that the excluded upper-middle, from the third quintile through the ninth decile, seems (according to Palma's research) to maintain a fairly stable share of total income in many countries. Thus changing inequality is substantially a shift from the poorest 40 percent to the top 10 percent, or vice versa.

The Palma ratio is designed to be a simple and sensitive measure of such shifts, and it is enjoying a certain popularity as this is written. Whether it will go on to become a standard summary measure of inequalities remains to be seen. A limitation lies in the fact that in order to calculate it, you must first have income measures from a survey or micro-sample, from which one can measure the deciles of the income distribution. This is a limitation of all percentile-based inequality measures, and also of the Gini coefficient, to which we will turn shortly.

What Are the Top Shares and What Do We Know about Them?

A difficulty of survey-based measures of income inequality is the practice of "top-coding." Usually in administering a survey one cannot ask for respondents to reveal their exact incomes. So instead the survey-taker asks respondents to enter a range within which their income falls, and the analyst later makes some assumption about how incomes are distributed within that range. (Often, the simplest assumption is that "between $50K and $60K in annual income" means $55K.)

The difficulty arises because at the very top of the income scale, it is necessary to leave an open-ended category. So the survey may have as an option, "income greater than $250K." Or, "income greater than $1m." And that leaves open the question, how much greater? While we may imagine that in a given population most people who report annual incomes over a million dollars are actually earning quite close to that sum, there may be a few earning a hundred or even a thousand times as much.

An approach to this problem is to turn to income tax records, where individuals are required to specify their exact taxable incomes. Tax records are confidential, but anonymized files permit researchers to calculate the share of total taxable income earned by small numbers of people at the very top of the distribution: the top 1 percent, the top 0.1 percent, and even the top .01 percent in some cases. This is the approach taken by Thomas Piketty, Emmanuel Saez, and Anthony Atkinson in recent research.

Yet for all their virtues, income tax records pose challenges of their own. Compared to surveys, they are not good at capturing unofficial and unreported incomes, such as cash earned in the informal sector. So the top shares may be exaggerated for that reason. Or they may be underestimated, if there is a lot of tax evasion by the very rich. More prosaically, comparing top shares across countries is problematic, because tax laws differ and therefore so do definitions of taxable income. The share of total income required to be reported for tax purposes may vary greatly between countries, even if the underlying distribution of actual cash incomes is exactly the same. Even within countries, the definition of taxable income will change when tax laws are rewritten, and this will upset the comparability of taxable-income measures over time.

Then there is the most prosaic problem, which is that one can acquire income-tax records only in countries that actually have income tax. In a recent compilation, Professor Piketty presents data for just 29 countries, not including any of the oil kingdoms of the Middle East, for instance, where it has been rumored that some of the locals are among the world's most prosperous persons.

What Are the Lorenz Curve and the Gini Coefficient?

After exploring the above simple measures, we can see that it would be nice to have a measure of inequality that is drawn from surveys, unit-free, comparable across different populations, and that makes use of all the information that a survey or census may make available about the incomes of the population under study. It should also be the case that transferring a small amount of income from a richer to a poorer person causes the index to decline. The Gini coefficient is a nice example of an inequality measure that meets these tests. It is, by far, the most popular and widely used measure of inequality in use.

The easiest way to understand the Gini coefficient is to envision the Lorenz curve, a simple plot that can be drawn for any distribution. To draw the Lorenz curve, first rank all members of the population or survey in order of income, and count them off into quintiles, deciles, or percentiles (the finer the better, for accuracy's sake). Plot the quantiles on the x-axis. On the y-axis, record the share of total income cumulatively earned up to each quantile. Thus, if the bottom 10 percent of the population earns 2 percent of the income, record a point at (10,2) and so forth. Connect the dots.

The resulting curve will be bowed below a 45-degree line, except for the case—unheard-of in real applications—of actual income equality. It will start at (0,0) and end at (100,100), since none of the people necessarily earns none of the income, and all of the people necessarily earn all of it. At any point of the curve, one can read exactly how much of total income the people below that income level have the right to call their own.

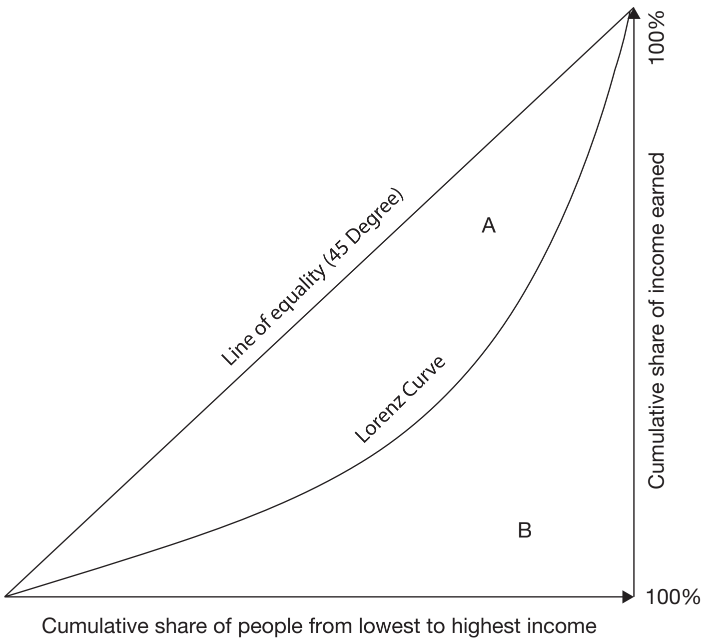

Figure 5.2 The Lorenz Curve

The Gini coefficient—invented by another Italian, Corrado Gini—is a simple geometric representation of the Lorenz curve. First, take the area between the Lorenz curve and the 45-degree line. Then divide that area by the area of the triangle below the 45-degree line—which in the case of a graph that is 100 on a side is the number 5,000. The result is the Gini coefficient. In Figure 5.2 the Gini coefficient is the ratio of the areas A and (A+B), or

$$\text{Gini} = [A / (A + B)] * 100$$

The Gini coefficient will vary from a value of zero—the case of absolute equality—to the value of 100, in the case where the last percentile holds all of the income. In income surveys of advanced countries over the past fifty years, measured Gini coefficients run from the low twenties, for such communist regimes as the German Democratic Republic, to the high sixties for some countries in sub-Saharan Africa. Measures for the major advanced countries have tended to run from the low thirties to the mid-forties, with a rising trend visible in most data sets.

Thus the Gini has the virtue of a scale that does not depend either on the unit or on the size of the population under study. You can measure it in your classroom, and compare the result to the Gini coefficient for the United States, or the entire world. You can also estimate the Gini from quintile, decile, or percentile shares, using the line segments connecting the information for these intermediate points, rather than a continuous curve. The error involved will be very small.

The Gini coefficient is very useful, fairly easy to compute, and easy to understand; no wonder it is the most popular general measure of inequality out there. But—as always!—there are limitations. The most important one is that the Gini requires a survey or a census with the underlying data; it must in principle be possible to rank the people or households from low to high and to calculate distinct quantiles of the distribution. But surveys are expensive and census records are rare; even in the United States, a census is taken only once a decade. In many countries, the historical record of income surveys is sparse, irregular, and inconsistent, and—as a result—so is the statistical record of Gini coefficients.

A subtler drawback of the Lorenz-Gini approach is that it cannot be used easily to add two populations together, or to break them apart. Suppose, for instance, that one has Gini measures for every country in Europe, alongside the average income and population of each European country. Could one compute a measure of inequality for Europe, taken as a single population? Not with any great confidence in the accuracy or reliability of the result. One cannot ever just average the Gini coefficients of two countries to get the inequality that would pertain if they were merged!

Similarly, there are many problems for which it would be useful to divide a population into constituent groups, so as to work out the inequality between the groups and the inequalities within them. It is often interesting, for instance, to do this by race or gender: to ask (for instance) how much of total inequality can be attributed to inequalities among women, among men, and between the genders? The Gini coefficient is not well suited to this calculation, even if you have underlying data that permit Gini measures for the two genders to be computed separately.

What Are Theil Statistics?

We have just one more approach to discuss. It is based on the work of a University of Chicago econometrician of the mid-twentieth century, Henri Theil (pronounced "Tile"). Theil was interested in the theory of information underlying modern computer science, and in measures of *entropy* that are developed in statistical thermodynamics, and that are closely related to the information problems of signal and noise. Theil's insight and contribution were that a measure of entropy or information content could be converted quite easily into a measure of economic inequality, and the result is a family of *Theil statistics*. One of Theil's statistics, known as *Theil's T*, is in especially common use.

The peculiar advantage of the Theil statistics is that they can be broken apart and added up. If you have two groups (say, men and women) and if you know the population and the average income of each group, then you can compute three inequality measures: inequality among women, among men, and between the two genders. The overall population inequality, computed as if you started with both genders in the same pool, will then be exactly the same as a weighted average of within-group inequalities, plus the between-group inequality. Similarly, if you have T statistics for each country in a region, and the populations and average income of each region, you can compute an inequality measure for the region as if it were one population—even if no unified survey had ever been taken.

And then there is one more advantage. Suppose (as is often the case) you don't have any survey-based evidence on incomes or expenditures at all? Suppose that all you have, for a given country or set of years, is a table, published by the government, that reports (say) payroll and employment in a classification of economic sectors, or (say) income and population by provinces, counties, and precincts? What then?

This type of data is *grouped data*. There are many different ways that groups can be defined, divided, and subdivided: by geographic boundaries, by industrial classification, by personal characteristics. Data of this type are very widely available in published form, often with quite consistent category structures, over many years for many countries. There are even international data sets at the continental and global levels that record data of this kind for industrial pay in many countries. With the Theil statistic, one can calculate a measure of inequality between groups, using whatever administrative data and group structures one may have on hand. This measure is the between-groups component of Theil's T statistic, and it has proved to be a very good instrument for the movement of inequality in many different situations.

Why does this work? A moment's thought can help make it clear. Suppose we take a large country like China. It is well known that inequality in China rose during the period of economic reforms, as the cities grew rich much more rapidly than the countryside. This will show up clearly in an inter-provincial measure of inequality in China! Similarly, certain economic sectors (finance, transport, utilities) had income gains far exceeding those of farmers or factory workers. This too will show up in an inter-sectoral measure. Taking the two together, and using sectors-within-provinces as the category structure, one can obtain a very detailed picture of the trends that dominate the movement of Chinese inequalities over time (Figure 5.3).

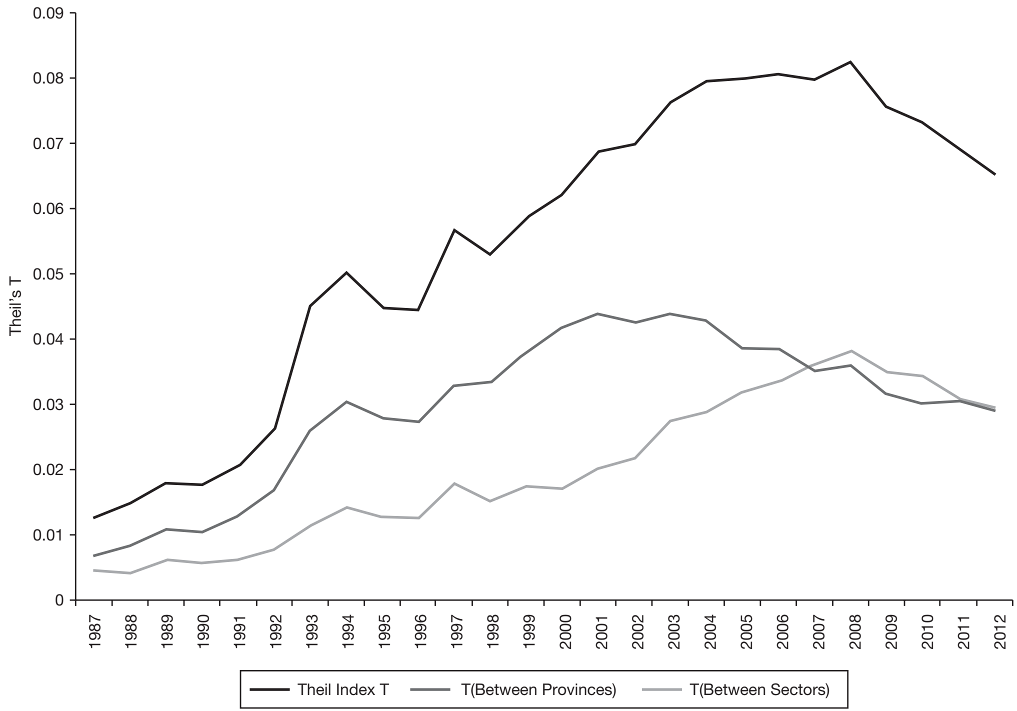

Figure 5.3 Inequality across Sectors and Provinces in China
Calculation by Wenjie Zhang. Used with permission.

Yes, it may be true that inequalities within a sector-within-a-province—among schoolteachers in Hebei, for instance—were also growing. But that is clearly going to be a small matter compared to the differences between bank executives in Shanghai and herdsmen in Tibet.

So long as one has a consistent grouping scheme that stays relatively constant over time, the between-groups Theil T method can be used to generate useful time series of inequality for many countries. In this respect, geographical data work well, since borders tend to change only rarely and since all competent governments are interested in collecting tariffs and taxes and in knowing what is going on in their economies. What is less clear is why the same trick should work for sectoral data measured using the same category scheme in different countries—with data taken (for instance) from the Industrial Statistics of the United Nations Industrial Development Organization or from Eurostat's REGIO data base. But it appears that the standardization of categories does have the effect of making the between-groups component of Theil's T comparable across countries, since the upper bound of the between-groups component is determined by the number of groups. And this is a great advantage, even over computing the underlying Theil from population data, since the underlying Theil (unlike the Gini) is not bounded by 100 or any other value, but will generally rise with an increasing population size.

For about twenty years, this author has been running a research project that involves computing Theil's T statistics and turning them into global inequality data sets. We still haven't run out of fresh ideas.