The econometrics of inequality and poverty

Ch3 : Welfare functions, inequality and poverty

Michel Lubrano

October 2017

Contents

In this chapter, we develop the pure welfarist approach which means that welfare depends on a single indicator which is taken to be either income or consumption. We thus suppose that our basic observations are individual incomes. These data are usually provided by governmental agencies. Either they cover the entire population and are available every five years or more, or they are just survey data, drawn at random to get a representative sample of the total population.

- I shall use the Chinease Social Survey and more precisely the 2006 wave. We made that choice simply for convenience. I had worked on these data for a paper.

Data concern households, which directly introduce the question of equivalence scales. We have usually access to household composition and to some kind of income decomposition in earnings, financial revenues, rents and transfers. In a subsequent lecture, we shall detail how households of different composition can be made comparable. For the while, we suppose that households have the same size and the same composition.

A good deal of the econometrics of income distribution will be devoted to the estimation of the income distribution, either parametrically or non-parametrically. Indices are a good way of summarizing the dispersion characteristics of a distribution in order to provide comparisons between countries or through time. Why should we take interest in the left tail of the income distribution and thus have a particular attention for the poor? We have to explain the aversion of a society for inequality and poverty. Atkinson (1970) formalized this problem by mean of welfare functions. This is also the approach adopted by Deaton (1997) in his chapter 3, chapter on which we shall draw a lot.

1 Welfare functions

Following Atkinson (1970) or chapter 3 of Deaton (1997), let us consider that society is formed by a collection of n individuals and that we want to measure welfare of this entity considered as a whole. We measure welfare with respect to a univariate variable noted x_i that represents either income or consumption. We have thus a first collection of observations on income

$$X = (x_1, x_2, \cdots, x_n) \tag{1}$$

that represents the income distribution.

1.1 Graphical representation

We indicate here how we can represent graphically this collection of individuals and their income. After the Dutch economist Jan Pen, we propose the Pen's parade. Every individual is given a size proportional to his income, normalized by the mean income of the population. Then each individual is ranked according to his size. The abscise are normalized by the sample size.

We use first the results of an income survey made in the Philippines and available as an R data set. We then use income data for China which come from the CGSS, 2006. We also provide the Gini coefficient: $G = 0.427$ for the Philippines and $G = 0.527$ for China (zero incomes

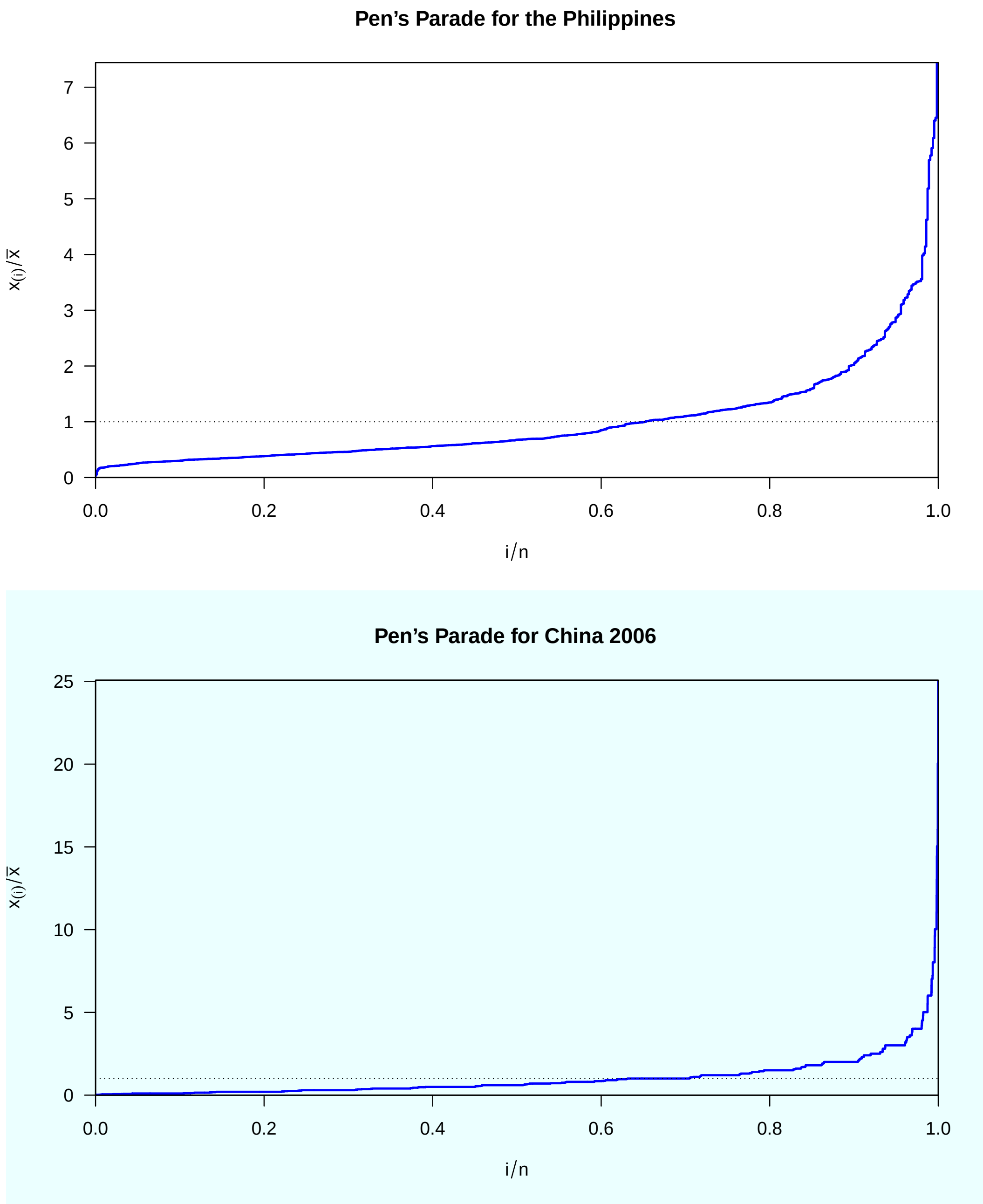

Figure 1: An example of Pen’s parade

were removed). A lot of information are already contained in Figure 1 displaying Pen's Parade for the two countries. For the Philippines, the mean income is reached only at the 6th decile of the population. The richest person earns 7 times the mean income. For China, the mean income is reached further away after the 7th decile, while the richest person earn 25 the mean income. However, the Philippine data set concerns Illicos, which is a small region in the north of the Philippines. The CGSS is representative of the whole China.

1.2 The welfare function

We define the welfare function as a function with n arguments representing the empirical income distribution:

$$W(x) = V(x_1, \cdots, x_n). \tag{2}$$

The welfare function is a very normative function. It must obey a certain number of axioms that define the comparisons we want to operate between individuals. It represents social preferences over the income distribution.

1. **Pareto axiom:** The welfare function is increasing for all its inputs. This axioms can be weakened so that it is not decreasing for some of its terms while being increasing for the remaining terms. With a weakened axiom, we can construct a welfare function which is increasing for the poor while being constant for the rich.

2. **Symmetry axiom or anonymity:** We can permute the individuals without changing the value of the function. But there are problems when the households have not the same composition. Survey data concern households, while welfare theory deals with individuals. The question of household composition is nontrivial and is usually addressed by equivalence scales. Problems can also arise if agents have different utility functions. Then the aggregation of utilities is not invariant to changes in the order of the arguments.

3. **Principle of transfers:** the quasi concavity of the welfare function implies that if we operate a monetary transfer from a rich to a poor, welfare is increased, provided that the transfer does not modify the ordering of individuals. This is known as the Pigou-Dalton principle. This is a very important principle, which is not always verified. But most of the time we shall try to enforce it.

4. **Other axioms:** there is a large economic literature devoted to building welfare functions and inequality measures or indices. Some axioms are not mutually exclusive. Many papers are devoted to finding the minimal number of necessary axioms when building a welfare function. See in particular the book by Sen (1997).

The main consequence of these axioms is that a welfare function expresses the aversion that a society has for inequality and that the welfare function will be maximal when all individuals have the same income. A whole strand of the empirical literature is devoted the practical measurement of aversion for inequality, the desire for redistribution, the causes of poverty. This kind of opinions can be studied using the CGS for instance for China.

2 Inequality and social welfare

If a social welfare function expresses the aversion of a society for inequality, then it is the natural starting point for inferring inequality measures. Let us suppose that the function is homogenous of degree 1. Using this property, we can factorize the mean income μ:

$$W(x) = \mu V(x_1/\mu, \cdots, x_n/\mu). \tag{3}$$

We then normalize $V(.)$ so that $V(1, \cdots, 1) = 1$. As there is an aversion for inequality, the normalized function reaches its maximum at 1 and thus total welfare cannot be greater than μ. We can thus rewrite the welfare function as:

$$W(x) = \mu(1 - I) \tag{4}$$

where I cannot be greater than 1. I is then interpreted as an inequality measure and μI represents the cost of inequality. Welfare increases with μ, so that we can have at the same time a welfare increase and an increase in inequality. It is essential to note that total welfare is measured by a mix between μ and I, and not only by one minus the degree of inequality I. If the poor get a bit more, and the rich much more, this is a Pareto improvement. And welfare is greater provided μ has risen more than I. The principle of transfers, on the contrary leaves μ unchanged, but decreases I. There is thus a balance to maintain between these two important criteria: Pareto principle and principle of transfers. Note however that μ has a scale while I has none. This might influence the trade-off. We shall discuss the shape of W and the concern for the poor further down in the text. Let us note in passing the famous debate between equity and efficiency, debate initiated by Okun (1975), which is often seen as a trade-off.

3 Welfare function and inequality indices

As $W(x) = \mu(x)(1 - I(x))$, we can start from a welfare function and then solve for the corresponding index of inequality. Or we can do just the reverse. Start from a given inequality measure, verify that it complies with the principle of transfers and then derive the corresponding social welfare function.

3.1 Starting from a welfare function

We illustrate the passage from W to I to derive the inequality index of Atkinson. Let us start from the following welfare function:

$$W = \frac{1}{n} \sum_i \frac{x_i^{1-\epsilon}}{1-\epsilon},$$

where ϵ is the parameter monitoring aversion to inequality. In general, we use values between 0 and 2 for pour ϵ. For $\epsilon = 1$, the above expression is not defined. The indeterminacy is removed (using for instance the de l'Hospital rule which means taking the limit of the ratio of the derivatives) by considering:

$$W = \frac{1}{n}\sum_i \log x_i.$$

This welfare function has important and nice properties. The ratio of marginal social utilities of two individuals has a simple expression:

$$\frac{\partial W/\partial x_i}{\partial W/\partial x_j} = \left(\frac{x_i}{x_j}\right)^{-\epsilon}.$$

As $\epsilon \to \infty$, the marginal utility of the poorest dominates. We are in the Rawlsian situation, Rawls (1971), where the objective of the society is to maximize the situation of the poorest. When $\epsilon \to 0$, more and more concern is put on the situation of rich individuals.

We can derive a measure of inequality from this particular welfare function which is the Atkinson index:

$$I_A = 1 - \left(\frac{1}{n}\sum_i (x_i/\mu)^{1-\epsilon}\right)^{1/(1-\epsilon)}.$$

When $\epsilon = 1$ it has the multiplicative form:

$$I_A = 1 - \prod (x_i/\mu)^{1/n}.$$

3.2 Equally distributed equivalent x

We start again from from the welfare function W and we consider the income distribution $X = (x_1, \cdots, x_n)$. $W(X)$ takes a certain value for this given distribution. Let us now consider another income distribution where everybody has got the same amount, to be determined. We are looking for the equivalent income ξ such that $W(\xi) = W(X)$, which means an income uniformly distributed that provides the same welfare for society. If the principe of transfers applies, then the inequality $\xi \leq \mu$ is always verified. We can then define as an inequality index one minus the ratio ξ/μ:

$$I = 1 - \frac{\xi}{\mu}.$$

We want this index to be independent of the scale of measurement. The usual way of defining scale independence is to require that

$$W(x_1, \cdots, x_n) = W(\lambda x_1, \cdots, \lambda x_n)$$

where λ is a positive number. Using this axiom, the value of ξ is uniquely defined by

$$\xi(x) = \left[\frac{1}{n}\sum_i x_i^{1-\epsilon}\right]^{1/(1-\epsilon)}.$$

which leads naturally to the second index of inequality of Atkinson, $1 - \xi(x)/\mu$. This inequality index is at value in [0,1]. If the computed value of this index is for instance 0.3, this would mean that 70% of the actual total income would be necessary in order to reach the same value of welfare, provided that income is equally distributed. The cost of inequality is $0.30 \times \mu$.

4 Inequality indices

A simple way of comparing income distributions is to summarize those distribution by an index. Of course, in order to produce an adequate summary, those indices have to verify a certain number of axioms.

- *Scale invariance* is the easiest property. The index should not change if we change the unit of measure
- The responsiveness to transfers is one of the most fundamental property. When taking money from the rich to redistribute it t the poor (without changing the order) the index should diminish.
- *Population principle*: the value of the index should not depend on the size of the population or of the sample. If we replicate the sample, the index should not change.
- *Fixed range*. If everybody has got the same income (the mean income), then the index should be zero. The index can be bounded above. The Gini index for instance is equal to one in the case of perfect inequality (one individual has all the income, all the other have zero).
- *Subgroup decomposability*. If we can cut x into two exclusive subgroups such that $x = x_1 \bigcup x_2$, decomposability means that inequality in the whole population x can be written as a weighted sum between inequality indices in the subgroups plus a residual depending only on the mean inside the subgroup. This last term represent between group inequality while the weighted sum represents inequality within groups.

Once these properties are verified, we can start from an inequality index and deduce the corresponding welfare function by means of $W = \mu(1 - I)$.

4.1 Inequality indices based on the quantiles

Some authors like very much to describe the income distribution by means of its quantiles. We shall see in a next chapter how to estimate those quantiles. What is a quantile? There are various ways of defining it. Let us suppose that we know the density from which a random variable X is drawn and call it $f(x)$. It integrates to one. We suppose also that X is at values in $[0, \infty[$. Then, the p-quantile is the value x_p such that:

$$\int_0^{x_p} f(x)\, dx = p.$$

If we know the cumulative distribution $F(x)$, the p-quantile can be defined in an explicit way as

$$x_p = F^{-1}(p)$$

Let us define a grid over p, with nine points: $p = (0.1, 0.2, \cdots, 0.9)$. We thus define deciles. The median is the value that separate the sample in two regions of equal probability:

$$x_{med} = F^{-1}(p = 0.50),$$

while the two quarter quantiles correspond to $p = 0.25$ and $p = 0.75$.

This being said, simple indices were proposed in the literature, such as the *interquartile range*:

$$I_Q = \frac{x_{0.75} - x_{0.25}}{x_{0.50}}.$$

However, this index does not verify the principle of transfers. If a transfer is done within a quintile group, the index is left unchanged. This index is nevertheless quite used, especially by official agencies. For instance Insee presents regularly the income distribution in the form of its deciles. A by-product is to measure the normalized distance between extreme deciles.

Table 1: Distribution of annual net wages in France before taxes in euros 2008-2010

	2008	2009	2010
D1	13 595	13 554	13 722
Q1	15 491	15 789	16 037
D5	19 159	19 756	20 107
Q3	26 136	26 869	27 345
D9	38 555	39 046	39 809
D9/D1	2,84	2,88	2,90

Source : Insee, DADS 2010.

For instance Piketty (2017), Piketty et al. (2017) make a great use of quantiles. They have an interpretation of those quantiles in term of social classes:

1. With an income below the median, the individuals are considered to belong to the poor class. This is at contrast with the usual definition of the poverty line we shall see below.
2. With an income between 50% and 90%, we have the middle class which covers 40% of the population
3. The rich individuals are those with an income greater than the top 10% decile.

We must note that this interpretation is specific to that book of Piketty. In particular, defining the middle class is not an easy task and it usually does not corresponds to what is written above.

4.2 Indices based on moments

The coefficient of variation is the square root of the variance of the incomes divided by the mean income:

$$CV = \frac{\sqrt{Variance}}{Mean}.$$

It is easy to compute, bounded at zero, but not bounded from above. It is subgroup decomposable, scale invariant and obeys the transfer principle. It is in fact a particular case of the Generalized Entropy Index.

The variance of logarithms:

$$VL = \text{Var } \log x$$

is often used in relation with wage studies. It is directly related to the lognormal distribution where it represents the parameter σ^2 (we shall detail that distribution later on). It has nevertheless some unwanted properties as underlined in Foster and Ok (1999). In this paper it is explained how the variance of logarithms can contradict a Lorenz ordering.

4.3 The generalized entropy index

Indexes of the *generalized entropy family* have nice properties and is advocated so in Cowell (1995). For a given value of c, they are:

$$I_E = \frac{1}{n\, c(c-1)} \sum \left[\left(\frac{x_i}{\mu} \right)^c - 1 \right].$$

When $c = 0$, a limit argument gives the mean of logarithms:

$$I_E(0) = \frac{1}{n} \sum \log \frac{\mu}{x_i},$$

while for $c = 1$ the same limit argument yields the *Theil index*:

$$I_E(1) = \frac{1}{n} \sum \frac{x_i}{\mu} \log \frac{x_i}{\mu}.$$

There is a one to one mapping between the I_E and Atkinson index I_A for a limited range of c. The generalized entropy index is a subclass of the Atkinson index with $\epsilon = 1 - c$ for $0 \leq c < 1$.

The Theil coefficient is at value between 0 and $\log n$.

4.4 The Gini index and its social welfare function

The most common inequality index is the Gini index. It is based on the mean of every distinct pair of differences of income, taken in absolute value. There are $n(n-1)/2$ different pairs. We normalize around the mean, which gives:

$$I_G = \frac{1}{\mu n(n-1)} \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} |x_i - x_j|. \tag{5}$$

This index is at value in $[0, 1]$. When everybody has got μ, the index is zero. When one has $n\mu$ and the other zero, the index is 1. This index can be costly to compute when n is large. Provided we order the observations, or at least know their rank ρ_i, the Gini index can be computed using a single loop, in the formulation proposed by Angus Deaton:

$$I_G = \frac{n+1}{n-1} - \frac{2}{n(n-1)\mu} \sum \rho_i x_i,$$

where $\rho_i = n$ if x_i is the minimum of the sample and $\rho_j = 1$ if x_j is the max of the sample. If we explicit a bit the rank, we have an expression that is useful for computations:

$$I_G = \frac{n+1}{n-1} - \frac{2}{n(n-1)\mu} \sum x_{[i]}(n+1-i),$$

where $x_{[i]}$ is the order statistics, which means that the observation are ordered by increasing order. A slightly simplified expression for I_G is also used in the literature with

$$I_G = \frac{n+1}{n} - \frac{2}{n^2\mu} \sum x_{[i]}(n+1-i),$$

which can also be written as

$$I_G = \frac{2}{n^2\mu} \sum x_{[i]} i - \frac{n+1}{n}.$$

Despite its weighting scheme, the Gini index focuses its attention to the centre of the income distribution. There are variations around this index, notably by Donaldson and Weymark (1980) who introduce a parameter $\alpha \in [0,1]$ which allows for different weighting schemes of the observations and paying more attention to the tails of the income distribution.

The welfare function which is associated to the Gini coefficient is the one which weights every observation using its rank. The poorer will receive the highest weight. We get

$$W = \mu(1 - I_G).$$

This function has been used by Sen (1976b) to rank the India States. We can generalize this function as

$$W = \mu(1 - I_G)^\sigma$$

for σ between 0 and 1. So we can weight the implied trade-off between equity $(1 - I_G)$ and efficiency (μ).

5 From inequality to poverty

When looking at the shape of the welfare function (4), we see that economic growth, e.g. the simultaneous increase of μ and of W can be concomitant with an increase of inequalities: some people can get richer at a greater speed than others.

- That was the case during the Thatcher period in the UK. Atkinson (2003) shows how during the eighties real income of the poorer remained constant while mid-range incomes increased and top incomes increased a lot. Despite this inequality increase, global welfare also increased. However, this is due to the single dimension approach of the social welfare function. If we had used another index such as the new development indices, we would have seen that global welfare, as measured by this alternative index had fallen during that period.

- There is also the example of China with the economic reforms led in a totaly different context. Im (2014, PhD dissertation) comments the famous slogan of Deng Xiaoping, the designer of Chinese economic reforms: "*let some people get rich first*". This slogan, which still has a very important influence within Chinese society, justifies inequality on the ground of efficiency rather than on the ground of deservingness and fairness.

Because of this apparent trade-off between efficiency (μ) and equity (inequality), the interpretation of inequality is not evident. It might be seen as inequity by poor people, those who remain at the bottom of the social ladder or as an opportunity, those who manage to climb the social ladder and are rich. Thus there is the need of another indicator which focusses on the left part of the income distribution. Poverty is felt as a *failure for society* and this feeling justifies that we devote to it a large interest. The welfare function transforms a complete distribution into a single number which allows to analyze the effects of a public economic policy on the whole income distribution. If we want to devote more attention to the poor, we must concentrate our attention to one part of the income distribution, the one which is concerned by the poor, even if we are only interested in counting them. We shall thus move our interest from analyzing inequalities to analyzing poverty by concentrating our attention on the left tail of the income distribution.

Poverty indices are used by official agencies to monitor anti-poverty policies. A lot of different indices were proposed in the literature. Sen (1976a) was the first to propose an axiomatic construction of indices. Zheng (1997) provides an excellent survey. His survey is organized around grouping axioms and examining which index complies to which axiom. It is common to note z the poverty level or line of poverty. With an income below z, a person is said to be poor. Above z, he is no longer poor.

5.1 Poverty lines

For this purpose, we have to defined what is called a poverty line, that is to say a line below which an individual or a household is said to be poor and above which he will no longer be considered as a poor. We feel all the arbitrary character of such a line. We can define it in two different ways.

1. an **absolute line of poverty** is defined with respect to a minimum level of subsistence. For instance, the Indian government has defined a minimum number of calories necessary for subsistence which is different in town and in the countryside. Using a price index, it has defined a monetary level of poverty in town and in the countryside. Using the same food

subsistence, the US government defined an absolute level of poverty, but dividing it by the share of food in the budget of an average household. The French RMI (revenu minimum d'insertion) can also be situated in this framework.

2. In developed countries and more precisely within the EU, one prefer to define a **relative poverty line**. The European Union launched a research programme for measuring poverty where the poverty line is defined with respect to a fraction of the mean or the median of the income distribution. Will be considered as a poor every individual which income is below 50% or 60% of the mean income of his country. This is a notion of relative poverty, which is near from the notion of subjective poverty (pauvreté ressentie). (see also the difference between objective and subjective health status).

3. **At the international level**, there is the desire to define a world poverty line, mainly around the works of the World Bank. There is the famous one-dollar-a-day which has been reevaluated several time, mainly due to changes in PPP. The last value proposed by the World Bank is $1.90. Atkinson and Bourguignon (2001) have promoted the view that an international poverty line should combined both types (relative and absolute). This idea has been illustrated in Xun and Lubrano (2017) for evaluating world poverty.

5.2 Measures of poverty used by official agencies

Two indices are used by most government and by the United Nations: the head count ratio and the income gap ratio. Note the use of the indicator function $\mathbb{1}(\cdot)$ when writing down those indices.

The *headcount ratio* evaluates the number of poor, the number of persons below z:

$$H(x,z) = \frac{1}{n} \sum \mathbb{1}(x_i \leq z) = \frac{q}{n},$$

where q is the number of poor. It is simply the fraction of people in a state of poverty. Despite its appeal (it is always nice to know the number of poor just by multiplying the index by n), this index *does not satisfy the principle of transfers*. If we tax the poorest to redistribute to those just below the poverty line z, the index decreases. This is due to the discontinuity of the index in x_i. However, we can note that Atkinson (1987) argues that a minimum income z is basic right and that it is important to know how many persons are deprived of this right. Its range is between 0 and 1.

The *income gap ratio* $I(x,z)$ measures in percentage the gap between the poverty line z and the mean income among the poor:

$$I(x,z) = \frac{1}{z} \left(z - \frac{1}{q} \sum x_i \mathbb{1}(x_i \leq z) \right) = 1 - \frac{\mu_p}{z},$$

where μ_p the average income of the poor. This second index is also distribution insensitive. This insensitiveness motivates another class of indices, first proposed by Sen (1976a) and which are detailed in the next subsection.

The *poverty gap ratio* is a third index found by multiplying these two indexes:

$$HI(x,z) = \frac{q}{n}\left(1 - \frac{1}{q\,z}\sum x_i \mathbb{1}(x_i \leq z)\right).$$

Despite the fact that it is not distributive sensitive, this index has some good empirical properties.

Watts (1968) was the first to propose a distribution-sensitive index:

$$W = \frac{1}{n}\sum_{i=1}^{n}(\log z - \log x_i)\mathbb{1}(x_i \leq z).$$

This index is related to the Theil inequality index as:

$$W = H[T - \log(1-I)],$$

where:

$$T = \frac{1}{q}\sum_{i=1}^{n}(\log \mu_p - \log x_i)\mathbb{1}(x_i \leq z),$$

H and I being defined above. Its range is between 0 and infinity.

Remark 1 *We have defined these indices by summation over the whole sample, using the indicator function $\mathbb{1}(x_i \leq z)$. The summation can be done only over the sample of the poor, provided the observation are order by increasing value. If q is the number of poor, the sum of the first q observations refers to the population of the poor.*

5.3 Sen family of poverty indices

Sen (1976b) has proposed an axiomatic construction of a poverty index, named after the Sen poverty index. It represents one solution to take into account of inequality among the poor. It combines the **three I's of poverty**, namely

1. Incidence (a head count measure)
2. Intensity (the poverty gap measure)
3. Inequality (a Gini index among the poor stating that the importance given to a poor is its rank)

This index can be defined by reference to the previous indexes H and I, adding G_P as the Gini coefficient of the poor:

$$S(x,z) = H(x,z)(I(x,z) + (1 - I(x,z))G_P).$$

When there is no inequality among the poor, $G_P = 0$ and then $S = HI$. When inequality is extreme $(G_P = 1)$, we are back to the headcount measure. Of course, this index has to be calculated and it can be expressed in term of weighted order statistics. Replacing each element by its analytical expression, we get:

$$S = \frac{2}{(q+1)n} \sum_{i=1}^{q} \frac{z - x_{[i]}}{z} (q + 1 - i),$$

provided we order the observations by increasing order. The ordering is implicit in this writing because we used the *order statistics* $x_{[i]}$. Each observation in this measure is weighted by its relative rank $q + 1 - i$. The poorest have the highest weight. This index precludes the possibility that an anti-poverty policy could decrease a poverty index just by giving transfers to individuals who are just below the poverty line z, leaving the situation unchanged for individuals that are in a state of extreme poverty. Its range is between 0 and 1.

Because it includes a Gini index, S cannot be decomposed into groups, or its decomposition includes a residual which is hard to interpret. It also violates the principle of transfers and is not continuous in x. Shorrocks (1995) proposed a modification of this index which partially solves some of the difficulties raised by the Sen index.

Shorrocks (1995) starts from the fact that the Sen index is simplified in $S = HI$ when $G_P = 0$. If we restrict that property to hold only when $H = 1$, we get a modified index of the form

$$\frac{1}{n^2} \sum_{i=1}^{n} \frac{z - x_{[i]}}{z} (2n - 2i + 1).$$

Introducing now the *focusing axiom* which says that the index is sensitive only to the income of the poor, this new index that we call SST is:

$$SST = \frac{1}{n^2} \sum_{i=1}^{q} \frac{z - x_{[i]}}{z} (2n - 2i + 1).$$

This index shares common features with the Sen index. It is symmetric, replication invariant, monotonic, homogeneous of degree zero in x and z, and normalized to take values in the range $[0, 1]$. But is has the additional properties of being continuous and consistent with the transfer axiom.

Let us now define the variable $\tilde{x}_i$ which is the normalized poverty gap:

$$\tilde{x}_i = \frac{z - x_i}{z} \mathbb{1}(x_i < z).$$

Then it is possible to show that the SST index can take a very simple form:[1]

$$SST = \mu(\tilde{x})(1 + G(\tilde{x})).$$

[1] We find another expression in footnote 9 of Shorrocks (1995), which is more obscure when we want to relate that index to the previous official indices:

$$SST = (2 - H)H\,I + H^2(1 - I)G_P.$$

We give its expression only as a reference because it can appear as thus in some articles or textbooks.

Its range is between 0 and 1.

The modified Sen index was later called in the literature the Sen-Shorrocks-Thon index because this index can be viewed as a variation of the Thon (1979) index. This is the reason why we used the acronym SST. The Thon index is

$$Th = \frac{2}{(n+1)n} \sum_{i=1}^{q} \frac{z - x_{[i]}}{z}(n+1-i).$$

The SST index converges to the Th index when the population x is successively replicated. However, Shorrocks (1995) underlines that the SST index verifies a greater number of axioms than the Thon index.

Finally, it is interesting to note with the end of the paper of Shorrocks (1995) that the SST index is related to the poverty gap profile, later called the TIP curve by Jenkins and Lambert (1997). We shall come back to this notion in Chapter 9.

5.4 FGT indices

Foster et al. (1984) propose a class of poverty indices which have the main property of being decomposable. They are linear, simple to understand and to manipulate. Because of their linearity they are decomposable, a notion that we shall illustrate in a next chapter. These indexes are based on partial moments, built from the income distribution. They have the general form In fact all of these indices can be expressed in a general form

$$P_\alpha = \frac{1}{n} \sum_i (1 - x_i/z)^\alpha \mathbb{1}(x_i \leq z),$$

where α is a parameter that be set to 0,1,2 or more. This class of index is particularly important and we shall come back to it in the next chapter. For the while let us detail the expression of this index for various values of α.

For $\alpha = 0$, we get the usual headcount measure:

$$P_0 = \frac{1}{n} \sum_i \mathbb{1}(x_i \leq z) = \frac{q}{n}.$$

For $\alpha = 1$, the index takes into account the distance of an individual to the poverty line, using the notion of poverty gap $z - x_i$

$$P_1 = \frac{1}{n} \sum_i (1 - x_i/z) \mathbb{1}(x_i \leq z).$$

The contribution of an individual to the value of the index is larger the poorer he is. This index is a continuous function of x which respect the principle of transfers. But this index is not sensitive the distribution of income among the poor. So it is not sensitive to certain types of transfers among the poor. This index is very near from the HI index detailed above.

For $\alpha = 2$, we recover a sensibility to the distribution of income among the poor

$$P_2 = \frac{1}{n}\sum_i (1 - x_i/z)^2 \mathbb{1}(x_i \leq z).$$

The range of these indices is between 0 and 1.

The index of Foster et al. (1984) is decomposable because of its linear structure. Let us consider the decomposition of a population between rural and urban. If X represents all income of the population, the partition of X is defined as $X = X^U + X^R$. Let us call p the proportion of X^U in X. Then the total index can be decomposed into

$$\begin{aligned} P_\alpha &= p\frac{1}{n}\sum_{i=1}^{n_U}\left(\frac{z - x_i^U}{z}\right)^\alpha \mathbb{1}(x_i \leq z) + (1-p)\frac{1}{n}\sum_{i=1}^{n_R}\left(\frac{z - x_i^R}{z}\right)^\alpha \mathbb{1}(x_i \leq z) \\ &= p\,P_\alpha^U + (1-p)\,P_\alpha^R. \end{aligned}$$

where P_α^U is the index computed for the urban population and P_α^R the index computed for the rural population.

6 Poverty and inequality in social welfare functions

The initial formulation of the welfare function (4) implies that a welfare increase can very well occur together with an increase of inequality. How can we propose a formulation of the welfare function so that a better concern for inequality is accounted for? In other words, which form should we give to $W(x)$ if we want to maximize welfare while insisting on poverty. Atkinson (1987) treat this question in section 3 of his paper, while distinguishing four possible options.

The first option consists in neglecting poverty. The social welfare function simply maximizes

$$W(x) = \mu(1 - I),$$

where I is an inequality measure and μI measures the cost of inequality. If the welfare function is adequately chosen, we can decompose the inequality measure so that the group of poor people can be separated from the rest of the population. We can thus measure the evolution of poverty without having poverty reduction as a major objective.

In a **second option**, we seek to introduce a priority on the cost of poverty $C_P = \mu P$ where P is a poverty index, while leaving aside the cost of inequality. The corresponding welfare function is:

$$W(x) = \mu - \mu P - \mu I = \mu(1 - P - I). \tag{6}$$

Atkinson (1987) indicates that in this case, it is sensible to use a counting measure for P and a measure satisfying the principle of transfers for I.

The **third option** consists in focusing one's attention only on poverty. The corresponding welfare function is of the form:

$$W(x) = \mu - \mu P = \mu(1 - P).$$

Finally the **last option** consists in using a trade-off between inequality and poverty. The welfare function is identical to that given in (6):

$$W(x) = \mu - \mu I - \mu P.$$

But this time, justice arguments lead to use for I a Gini coefficient computed on the whole population and for P a modified Sen (1976a) poverty measure.

These considerations show that building a social welfare function can be relatively complex when considering its properties and the way individuals are aggregated. The simple form (4) presented above is thus maybe too simple.

7 Empirical illustrations using Chinese survey data

We are going to illustrate some of the above notions using annual income data of the Chinese Social Survey for 2006. All calculations are done using the software `R` and the package `Ineq` when possible.

7.1 The software `R`

`R` is free software which can be used easily for analyzing the income distribution. You can get it for free at:

$$http://www.r-project.org/$$

You can make computations of your own, while a lot of packages are available for estimation purposes. The basic package allows you to estimate density non parametrically, plot the corresponding density, eventually doing multiplots. The package *ineq* is useful for estimating poverty and inequality indices.

The `R` wrapper `Rstudio`, already documented in the introduction is especially convenient. It can be downloaded at:

$$https://www.rstudio.com/$$

When you run `Rstudio`, there a first window in the upper left part of your screen where you can type and edit the file where your code is located. Write you code there and save it in a file (you will be asked for a name in the form *myfile.r*). For running your code, entirely or just a part of it, you have to highlight it. *Ctrl A* is a good way for highlighting it all. Then press *Ctrl R* to run the code. Compilation and numerical results will appear in a lower left window. If there is a graph, it weill appear in the lower right panel.

Here is the `R` code that we used in the remaining paragraphs.

Table 2: Summary statistics for annual income

Min	Q_{25}	Q_{50}	Mean	Q_{75}	Max
20	3 000	6 000	9 972	12 000	250 000

```
rm(list = ls())  # to erase everything from the working space
library(ineq)    # the ineq library
library(weights) # for means, variances and quantiles with weights
library(reldist) # for gini with weights

setwd("E:\\Cours Nanchang\\Calculs")

CGSS = read.table("CGSS2006.csv",header=T,sep=";")
names(CGSS)
attach(CGSS)

income = qd35a
income[income>=455555]=NA
income[income==0] = NA
id = !is.na(income)
y = income[id]
n = length(y)
summary(y)
```

So we have read the data. Names are displayed. Income is the *qd35a* variable. A description of the data is available in the release notes of the survey. Some values are missing. So there are replaces by *NA*. We have declared as missing zero incomes and values greater than 455 555 which have a special meaning. All these values are discarded from the working sample which is *y*. It remains 7 709 observations (the value of n) out of 9 517 (the length of *qd35a*).

It is important to have a first idea of what is in the sample. This is the *summary* command. From Table 2, we see that the distribution is very asymmetric due to the large distance between the Median and the Mean. And the Max is quite far away from the mean. Note that we have not used weights that will be detailed later on. Using sample weights can make a significant difference.

There is a second difficulty with those income data which is revealed by using the `table(y)` command. This command is used to count the number of observations which have specific values. We are not going to display its result, but there are 296 different values, which means that the variable is not continuous. There are some rounded values that have a significant frequency, while intermediate values were used by only very few individuals. There was a rounding mechanism when individuals were answering this question that most individuals used, but not all.

7.2 Non parametric estimation of densities

A first global indication is given by estimating the income distribution.

```
h = 1.06*sd(y)/n^0.2)
plot(density(y,bw=h,xlim=c(0,100000),
   main="China: annual income in 2006",
   xlab="Income in yuans")
q01 = quantile(y,0.01)
q10 = quantile(y,0.10)
q50 = quantile(y,0.50)
q90 = quantile(y,0.90)
q99 = quantile(y,0.99)
lines(c(q01,q01),c(0,0.00005),col=1)
lines(c(q10,q10),c(0,0.00005),col=2)
lines(c(q50,q50),c(0,0.00005),col=3)
lines(c(q90,q90),c(0,0.00005),col=4)
lines(c(q99,q99),c(0,0.00005),col=5)
```

We have used the Silverman rule to determine the bandwidth. The econometrics of density estimation will be explained later in the lectures. We have displayed vertical lines the locate the 1%, 10% quantiles, the median and the 90% and 99% quantiles. The result appears in Figure 2. There are several facts that we can retain from this Figure. First, there is a tiny gap between the

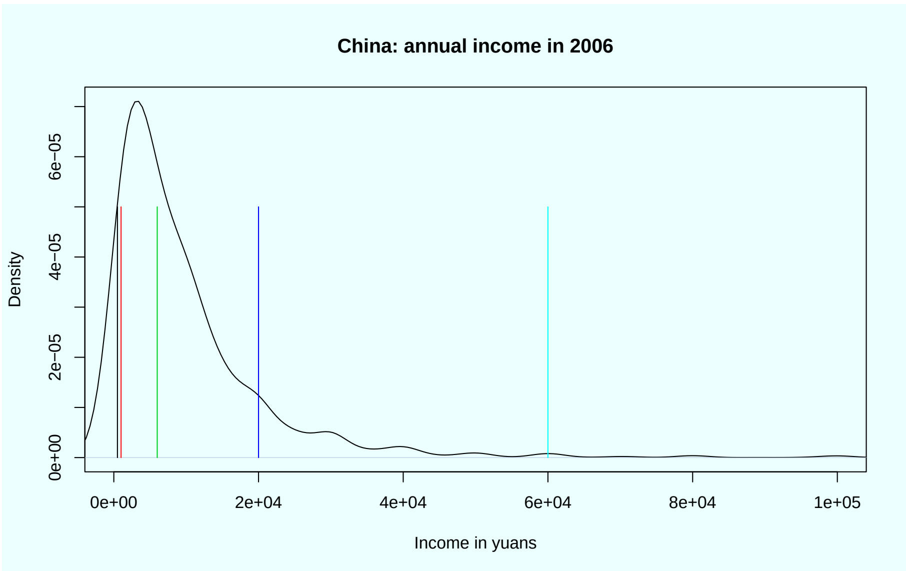

Figure 2: Density estimate of the Chinese income distribution

1% and 10% quantiles while the distance between the 90% and the 99% is huge. The density is very much concentrated around the median (the green vertical line). We can thus expect a large inequality in this very asymmetric distribution. In order to get a realistic gap, we were obliged to discard the part of the graph which was greater than 100 000 yuans, which means the part between 100 000 and 250 000 yuans.

7.3 Inequality measures

Let us now compute inequality indices for this data set. Those indices indicate a high level of

Table 3: Inequality measures

Gini	Theil	Atkin 0.5	Atkin 1.0	Atkin 1.5
0.527	0.511	0.232	0.425	0.582

inequality, starting with the Gini. The Atkinson index increases with ϵ which takes more and more account of the poor individuals. Figure 3 describes the evolution of the two indices for increasing values of ϵ and c.

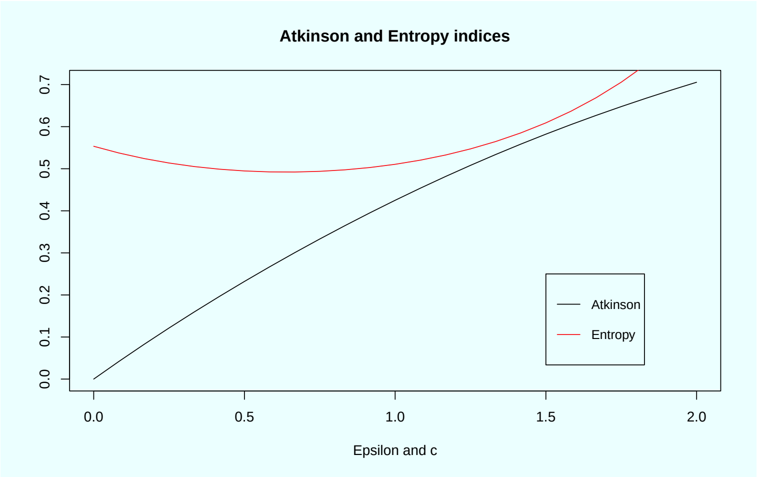

Figure 3: Atkinson and Entropy measures for varying ϵ and c

The code for plotting this Figure is as follows:

```
nk = 25
At = rep(0,nk)
En = rep(0,nk)
e = seq(0,2,length=nk)
for (i in 1:nk){
  At[i] = Atkinson(y,parameter=e[i])
  En[i] = Entropy(y,e[i])
}

plot(e,At,type="l",main="Atkinson and Entropy indices",
xlab="Epsilon and c",ylab="")
lines(e,En,col=2)
legend(1.50,0.25,
       legend=c(" ","Atkinson","", "Entropy"," "),
       col=c(0,1,0,2,0), lty=1:1, cex=0.9)
```

Note the different code which was used for the legend. The command entropy gave errors. So we reprogrammed it.

```
Entropy = function(y,c){
  n = length(y)
  mu = mean(y)
  if (c==0){En = mean(-log(y/mu))}
  else if(c==1){En = mean(y/mu*log(y/mu))}
  else {En = mean((y/mu)^c-1)/c/(c-1)}
  return( En)
}
```

7.4 Poverty measures

In China, like in India or in the Philippines, it is important to distinguish between rural and urban when analysis poverty. For instance, in December 2011 India fixed the urban poverty line at INR 32 (USD 0.60, EUR 0.46) per day per capita and to INR 26 in rural areas.

In 2011 China has raised the official rural poverty line to 2,300 yuan a year which makes around $400 which makes around $1.10 a day to be compared to the World Bank international poverty line of $1.25 at 2005 prices. Chen and Ravallion (2008) analyze the evolution of poverty in China before that revision and indicate a poverty line varying from 600 to 1400 yuans per year per capita, depending on the provinces, to take into account differences in prices.

The fact of being classified as urban or rural is determined in China by the Hukou system. In the CGSS survey the variable *qa03a* indicates that status. When $qa03a = 1$, the individual has a rural status, when $qa03a = 3$, he is classified as urban. An other status corresponds to $qa03a = 2$. For individuals that have a positive income, we have the following statistics, reported in Table

Table 4: Income statistics by Hukou status

Status	Pop %	Min	Q_{25}	Q_{50}	Mean	Q_{75}	Max
Rural	50	60	2 000	3 000	6 111	7 000	250 000
Other	3	720	6 000	10 000	12 630	18 000	100 000
Urban	47	20	6 000	10 000	13 980	17 000	200 000
Total	100	20	3 000	6 000	9 972	12 000	250 000

5. We see that Urbans and Others have very similar characteristics. So we can join the two categories and keep only the distinction *rural* versus *urban*.

```
rural = qa03a[id]==1
urban = qa03a[id]==3
other = qa03a[id]==2
table(qa03a[id])/n
summary(y[rural])
summary(y[other])
summary(y[urban])
urban = (qa03a[id]==3)|(qa03a[id]==2)
```

We are now faced to the choice of a poverty line. There is the new official poverty line of 2 300 yuans which corresponds to 77% of the median rural income or to 38% of the average rural income. The alternative is to consider a relative poverty line, which is more adapted for urban areas. The usual practice (see e.g. Atkinson 1998) is to take either 50% of the mean income, which makes here 4 986 or 60% of the median income, which makes here 3 600.

```
z1 = 2300
z2 = 0.6*median(y)
z3 =0.5*mean(y)
z = c(z1,z2,z3)
for (i in 1:3){
cat(Foster(y[rural],z[i], parameter = 1),"")
  cat(Foster(y[rural],z[i], parameter = 2),"")
      cat(Sen(y[rural],z[i]),"")
          cat(SST(y[rural],z[i]),"")
              cat(Watts(y[rural],z[i]),"\n")
}
```

Table 5 clearly shows the importance of a correct poverty line, and the huge difference between rural and urban areas. The second fact is that the two relative lines do not give the same poverty rates: the poverty line based on the median is lower than that based on the mean. This is due to the very large asymmetry on the income distribution in China and implicitly to the large inequality.

Table 5: Poverty measures by Hukou status

Status	P-line	H.C.	FGT	Sen	SST	Watts
Rural						
	Official	0.388	0.160	0.214	0.279	0.253
	60% Med	0.531	0.269	0.341	0.433	0.457
	50% mean	0.617	0.358	0.437	0.543	0.648
Urban						
	Official	0.054	0.021	0.029	0.042	0.033
	60% Med	0.096	0.041	0.054	0.080	0.067
	50% mean	0.162	0.068	0.091	0.128	0.111
Total						
	Official	0.223	0.092	0.122	0.169	0.144
	60% Med	0.316	0.156	0.199	0.277	0.264
	50% mean	0.392	0.214	0.268	0.366	0.382

Finally, what is the poverty rate in China. It has dropped a lot in the recent years. However, we note that the rate (Head Count) is still very important in rural areas. It is comparable to OECD rates for urban areas (between 10% and 16% using a relative poverty line). At the country scale, poverty remains relatively high, due to the weight of rural areas.

8 Exercises

8.1 Computing limits

Consider two functions $f(x)$ and $g(x)$ and x_0 such that $f(x_0) = g(x_0) = 0$. The limit $\lim_{x \to x_0} f(x)/g(x)$ is not defined. However. L'Hospital rule give a solution to remove this indeterminacy. The rule says that it is equivalent to compute the limit of $\lim_{x \to x_0} f'(x)/g'(x)$ where $f'(.)$ is the first order derivative. Using this rule, gives the expression of

- the generalized entropy for $c = 0$ and $c = 1$
- the Atkinson index for $\epsilon = 1$
- the Atkinson welfare function for $\epsilon = 1$

8.2 Properties of indices

- Show that the range of the Atkinson index is [0,1].
- Detail the relation between the income gap ratio and the P_1 index of FGT.
- Show that for $c = 2$ the GE coefficient is equal to the coefficient of variation.

8.3 Poverty indices

The Sen-Shorrocks-Thon index has several expressions, which are more or less manageable. The usual way for computing the index is

$$SST = \frac{1}{n^2} \sum_{i=1}^{q} \frac{z - x_{[i]}}{z} (2n - 2i + 1).$$

Let us define the variable Let us define the variable

$$\tilde{x}_i = \frac{z - x_i}{z} \mathbb{1}(x_i < z).$$

- Show that the SST index can take a very simple form

$$SST = \mu(\tilde{x})(1 + G(\tilde{x})).$$

- Show that the SST index can also be written as $H\, I(1 + G(\tilde{x}))$.
- What is the difference between this index and the Thon index given by

$$Th = \frac{2}{(n+1)n\, z} \sum_{i=1}^{q} (z - x_{[i]})(n + 1 - i).$$

- Show that the SST index is asymptotically equivalent to the Thon index when the same sample is replicated an infinite number of times.

8.4 Decomposable poverty indices

A poverty index P is decomposable if it can be written as a weighted sum of partial indices. More precisely, let $x = (x_1, x_2)$ and let n_1 and n_2 be the respective sizes of the subsamples x_1 and x_2. Then P is decomposable if $P = \frac{n_1}{n} P_1 + \frac{n_2}{n} P_2$.

- Show that the Watts index

$$P_W = \frac{1}{n} \sum_{i=1}^{q} (log(z) - log(x_{(i)}))$$

  is decomposable.

- Show the headcount measure is decomposable.

8.5 Empirics

Explore the software R and load the library *ineq*. In this library there is a data base coming from the Philippines and called *Ilocos*. Describe this data base (*help("Ilocos")*). Draw the Pen's parade corresponding to this data set. Explain your results.

Use the previous data base to decompose poverty between urban and rural regions in the Philippines.

References

Atkinson, A. (1970). The measurement of inequality. *Journal of Economic Theory*, 2:244–263.

Atkinson, A. (1987). On the measurement of poverty. *Econometrica*, 55:749–764.

Atkinson, A. (1998). *Poverty in Europe*. Blackwell, Oxford.

Atkinson, A. (2003). Income inequality in oecd countries: Data and explanations. Working Paper 49: 479-513, CESifo Economic Studies, CESifo.

Atkinson, A. B. and Bourguignon, F. (2001). Poverty and inclusion from a world perspective. In Stiglitz, J. E. and Muet, P.-A., editors, *Governance, Equity and Global Markets*. Oxford University Press.

Chen, S. and Ravallion, M. (2008). China is poorer than we thought, but no less successful in the fight against poverty. Technical Report WPS 4621, The World Bank Development Research Group.

Cowell, F. (1995). *Measuring Inequality*. LSE Handbooks on Economics Series. Prentice Hall, London.

Deaton, A. (1997). *The Analysis of Household Surveys*. The John Hopkins University Press, Baltimore and London.

Donaldson, D. and Weymark, J. (1980). A single-parameter generalization of the gini indices of inequality. *Journal of Economic Theory*, 22(1):67–86.

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 52:761–765.

Foster, J. E. and Ok, E. A. (1999). Lorenz dominance and the variance of logarithms. *Econometrica*, 67(4):901–907.

Im, D.-K. (2014). *Attitudes and Beliefs about Distributive Justice in China*. PhD thesis, Harvard University.

Jenkins, S. P. and Lambert, P. J. (1997). Three Is of poverty curves, with an analysis of UK poverty trends. *Oxford Economic Papers*, 49(3):317–327.

Okun, A. (1975). *Equality and Efficiency: The Big Tradeoff*. The Brookings Institution, Washington (D.C.).

Piketty, T. (2017). *Capital in the Twenty-First Century*. Harvard University Press.

Piketty, T., Yang, L., and Zucman, G. (2017). Capital accumulation, private property and rising inequality in China, 1978-2015. Technical Report Working Paper 23368, NBER.

Rawls, J. (1971). *A Theory of Justice*. Oxford University Press, London.

Sen, A. (1976a). Poverty: an ordinal approach to measurement. *Econometrica*, 44(2):219–231.

Sen, A. (1976b). Real national income. *Review of Economic Studies*, 43:19–39.

Sen, A. K. (1997). *On Economic Inequality*. Clarendon Press, Oxford, expanded edition with a substantial annexe by j. e. foster and a.k. sen. edition.

Shorrocks, A. F. (1995). Revisitng the sen poverty index. *Econometrica*, 63(5):1225–1230.

Thon, D. (1979). On measuring poverty. *Review of Income and Wealth*, 25:429–440.

Watts, H. (1968). An economic definition of poverty. In Moynihan, D., editor, *On Understanding Poverty*. Basic Books, New-York.

Xun, Z. and Lubrano, M. (2017). A bayesian measure of poverty in the developing world. *Review of Income and Wealth*, pages n/a–n/a.

Zheng, B. (1997). Aggregate poverty measures. *Journal of Economic Surveys*, 11(2):123–162.