The econometrics of inequality and poverty
Lecture 3 : Welfare functions, inequality and poverty

Michel Lubrano

September 2016

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. 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 summarising 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) formalised this problem by mean of welfare functions. This is also the approach adopted by Deaton (1997) in his chapiter 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, normalised by the mean income of the population. Then each individual is ranked according to his size. The abscise are normalised by the sample size. We use the results of a income survey made in the Philippines and available as an R data set. A lot of information are already contained in Figure 1 displaying Pen's Parade 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.

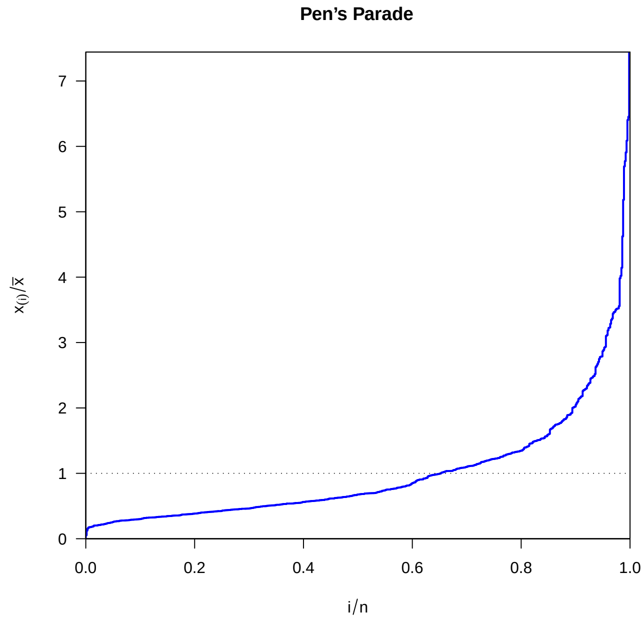

Figure 1: An example of Pen's parade

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, see in particular Carlsson et al. (2005) and the references given in this paper.

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 factorise the mean income μ:

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

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

$$W(x) = \mu(1 - I) \quad (4)$$

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:

$$W = \frac{1}{n} \sum_i \frac{x_i^{1-\epsilon}}{1-\epsilon}, \tag{5}$$

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

$$W = \frac{1}{n} \sum_i \log x_i. \tag{6}$$

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 maximise 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:

$$I_A = 1 - \left(\frac{1}{n} \sum_i (x_i/\mu)^{1-\epsilon}\right)^{1/(1-\epsilon)}. \tag{7}$$

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:

$$I = 1 - \frac{\xi}{\mu}. \tag{8}$$

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

$$\xi(x) = \left[\frac{1}{n}\sum_i x_i^{1-\epsilon}\right]^{1/(1-\epsilon)}. \tag{9}$$

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 $0.30 \times \mu$.

Kolm (1976) consider an alternative axiom of scale independence which requires that

$$W(x_1, \cdots, x_n) = W(\delta + x_1, \cdots, \delta + x_n)$$

where δ is a positive number. The welfare function should not change if the same positive amount is given to everybody. This alternative scale independence definition leads to Kolm inequality index

$$I_K = \frac{1}{\alpha}\log\left(\frac{1}{n}\sum_i \exp(\alpha[x_i - \mu(x)])\right) \tag{10}$$

with $\alpha > 0$.

4 Inequality indices

The responsiveness to transfers is the most fundamental property that an inequality index must verify. Scale invariance is another property. 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 Common inequality indexes

Simple, but ineffective indices were proposed in the literature that do not verify the principle of transfers. The *interquartile range*

$$I_Q = \frac{q_{0.75} - q_{0.25}}{q_{0.50}}$$

where q_α is the α quintile of the income distribution. 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 normalised distance between extreme deciles.

Figure 2: Distribution of annual net wages
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.

The variance of logarithms has also some unwanted properties.

Indexes of the *generalised 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. The Theil coefficient is at value between 0 and $\log n$.

4.2 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 normalise around the mean, which gives:

$$I_G = \frac{1}{\mu n(n-1)} \sum_{j=1}^{n-1} \sum_{i=j+1}^{n} |x_i - x_j|. \tag{11}$$

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 generalise this function as

$$W = \mu(1 - I_G)^\sigma \tag{12}$$

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 ones depicted in Lecture 2, we would have seen that global welfare, as measured by this alternative index had fallen during that period.

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 analyse 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 analysing inequalities to analysing poverty by concentrating our attention on the left tail of the income distribution.

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).

5.2 Use and misuse

The distinction between these two types of poverty lines is not anodyne. The recent report to the Parliament Bachelot (2011) illustrates the confusions that can arise from a misuse of these two definitions of a poverty line, especially when international comparisons are involved. An absolute poverty line will move only with the price index. Consequently the number of poor will have a tendency to mechanically decrease as soon as economic growth is positive. A relative poverty line will indicate the same number of poor whatever the growth rate of the economy, provided the shape of the income distribution remains constant.

In the report Bachelot (2011) devoted to examining the former government objective to reduce poverty of 33% in five years, the definition which is taken is pauvreté ancrée dans le temps. This means that a poverty line is chosen, a relative one can be chosen. Then it evolves over time only by changing the price index, irrespective to the change in the income distribution. But when doing international comparisons with other European countries, they take, for the other countries the official definition recommended by the EU, the relative poverty line.

6 Poverty measures

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 organised 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.

6.1 Official measures

Two indices are used by most government and by the United Nations: the head count ratio and the income gap ratio.

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.

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:

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

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

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

This index is related to the Theil inequality index as

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

where

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

H and I being defined above.

We have defined these indices by summation over the whole sample, using the indicator function $\mathbb{1}(x_i \le z)$. 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.

6.2 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.

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 normalised 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 normalised 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

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

which will give rise to interesting developments in a next chapter. 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:

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

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.

6.3 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

$$P_\alpha = \frac{1}{n} \sum_i (1 - x_i/z)^\alpha \mathbb{I}(x_i \leq z), \tag{13}$$

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:

$$P_0 = \frac{1}{n} \sum_i \mathbb{I}(x_i \leq z) = \frac{q}{n}. \tag{14}$$

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

$$P_1 = \frac{1}{n} \sum_i (1 - x_i/z) \mathbb{I}(x_i \leq z). \tag{15}$$

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

$$P_2 = \frac{1}{n}\sum_i (1 - x_i/z)^2 \mathbb{1}(x_i \leq z). \tag{16}$$

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

$$\begin{aligned} P_\alpha &= p\frac{1}{n}\sum_{i=1}^{n_U}\left(\frac{z - x_i^U}{z}\right)^\alpha \mathbb{1}(x_i \leq z) + (1-p)\frac{1}{n}\sum_{i=1}^{n_R}\left(\frac{z - x_i^R}{z}\right)^\alpha \mathbb{1}(x_i \leq z) \\ &= p\, P_\alpha^U + (1-p)\, P_\alpha^R. \end{aligned} \tag{17}$$

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

7 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 maximise 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 maximises

$$W(x) = \mu(1 - I), \tag{18}$$

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

$$W(x) = \mu - \mu P - \mu I = \mu(1 - P - I). \tag{19}$$

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

$$W(x) = \mu - \mu P = \mu(1 - P). \tag{20}$$

Finally the **last option** consists in using a trade-off between inequality and poverty. The welfare function is identical to that given in (19)

$$W(x) = \mu - \mu I - \mu P. \tag{21}$$

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.

8 Empirical illustrations

We are going to illustrate some of the above notions using data of the net income drawn from British Family Expenditure Survey covering the dates 1979, 1988, 1992, 1996. The survey covers the period when Margaret Thatcher was at power (1979-1990). We are going to use these data to illustrate the debate between poverty and inequality. The data provide a sample of household current disposable income, together with household characteristics such has the number of children. The consumer price index was 223.5 in 1979, 421.7 in 1988, 546.4 in 1992 and 602.4 in 1996 for a base of 100 in January 1974 as given in O'Donoghue et al. (2004).

8.1 The software `R`

`R` is free software which can be used easily for analysing 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. When you run the software, it opens a first window, called `R Console` where your results will be displayed. Within the `File` menu, you open in a second window either a new script or an existing script where the command you need are entered or displayed in the case of an existing script. To run your code, you must highlight it first (`Ctrl A` is a convenient short cut to highlight your whole code), then you run the selected code with `Ctrl R`.

You can also decide to use a very convenient wrapper which is called R Studio and which can be downloaded at

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

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

```
library(ineq)
data1=read.table("fes79.csv",header=F,sep=";")
```

```
data2=read.table("fes88.csv",header=F,sep=";")
data3=read.table("fes92.csv",header=F,sep=";")
data4=read.table("fes96.csv",header=F,sep=";")

y79 = data1[,1]/223.5*223.5
y88 = data2[,1]/421.7*223.5
y92 = data3[,1]/546.4*223.5
y96 = data4[,1]/602.4*223.5
```

So we have read the data, extracted the first column which contains the current disposable income and corrected for inflation.

8.2 Non parametric estimation of densities

A first global indication is given by estimating the income distribution for the different points of observation.

```
plot(density(y79),type="l",xlim=c(0,400),main="Income
     distribution in the UK")
text(20,0.013,"1979")
lines(density(y88),col="red")
text(70,0.011,"1988")
lines(density(y92),col="blue")
text(75,0.0065,"1992")
lines(density(y96),col="green")
text(150,0.0050,"1996")
```

There is a large increase in inequality. The 1979 distribution is concentrated, despite its soft secondary mode. In 1988, the main mode is much lower, the right tail has increased, the secondary mode has increased, showing an heterogeneity in the population.

We have added vertical lines indicating the first and last percentile (1% and 99%) with the following code

```
p1 = 0.01
p2 = 0.99
q179 = quantile(y79,p1);q188 = quantile(y88,p1);
q192 = quantile(y92,p1);q196 = quantile(y96,p1)
q279 = quantile(y79,p2);q288 = quantile(y88,p2);
q292 = quantile(y92,p2);q296 = quantile(y96,p2)

lines(c(q179,q179),c(0,0.005))
lines(c(q279,q279),c(0,0.005))
```

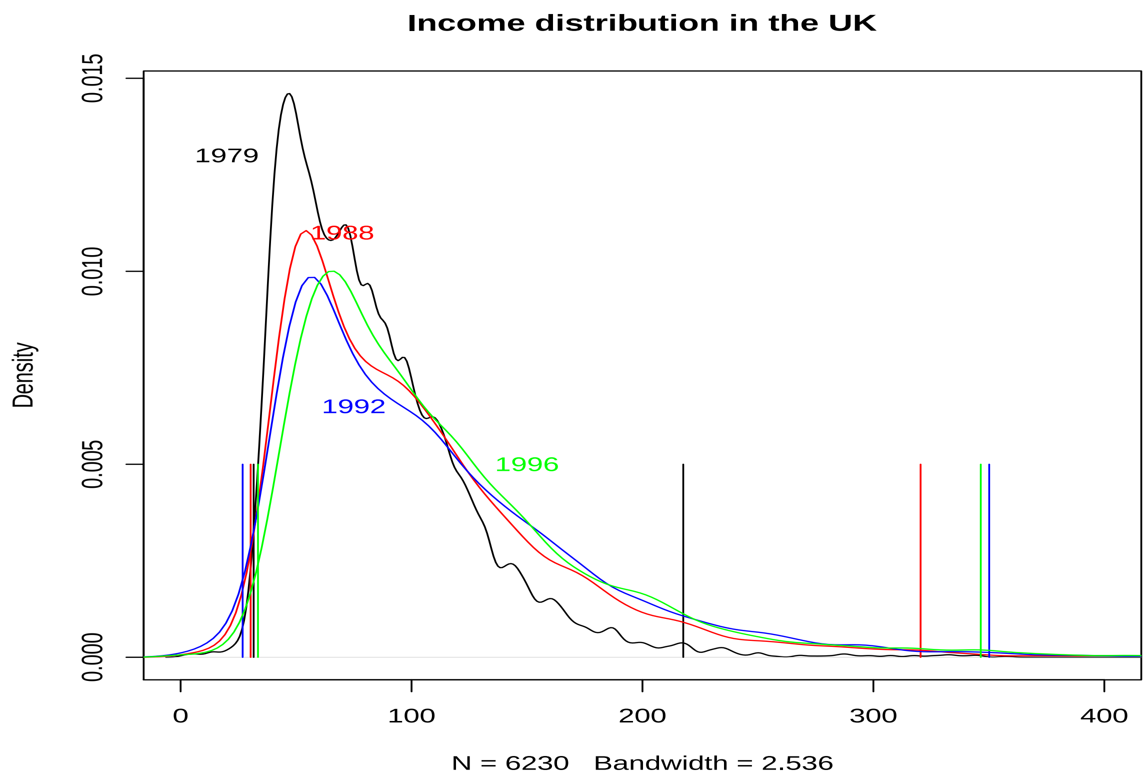

Figure 3: Density estimate of the UK income distribution

```
lines(c(q188,q188),c(0,0.005),col="red")
lines(c(q288,q288),c(0,0.005),col="red")

lines(c(q192,q192),c(0,0.005),col="blue")
lines(c(q292,q292),c(0,0.005),col="blue")

lines(c(q196,q196),c(0,0.005),col="green")
lines(c(q296,q296),c(0,0.005),col="green")
```

These lines depict well what happened during that period. There was a large increase in the income of the 1% richer part of the population in 1988. This increase went on till 1996 where there was a very slight decease. For the 1% poorer part of the population, poverty was made more severe in 1988, went further on in 1992, and was corrected in 1996, where it went back to its position of 1979. These movement are less marked if we take the first and last deciles of the distribution.

Another way of writing legends in a graph is to use for instance

```
legend(10,0.6,
        legend=c("   ","Beta<theta","", "Beta>theta","   "),
        col=c(0,4,0,2,0), lty=1:1, cex=0.9)
```

In this command, 10 and 0.6 refer the coordinates where to start the legend. The legend is given in the next keyword. Note that there are implicitly three curves to be commented. But we wanted to include blanks. The colours are indicated next, 0 corresponding to a dummy colour. `lty` indicates the type of lines which is used, here a plain line. `cex` is a scale factor for the size of the characters. This command was used to produce this graph 4

```
plot(density(S[id]/1000),col=4,xlim=c(2,15),ylim=c(0,0.7),
    main="",xlab="Instructional and total expenditure")
lines(density(S[!id]/1000),col=2)
lines(density(TS[idt]/1000),col=4))
lines(density(TS[!idt]/1000),col=2)
text(3.5,0.6,"Instructional")
text(10,0.3,"Total")
```

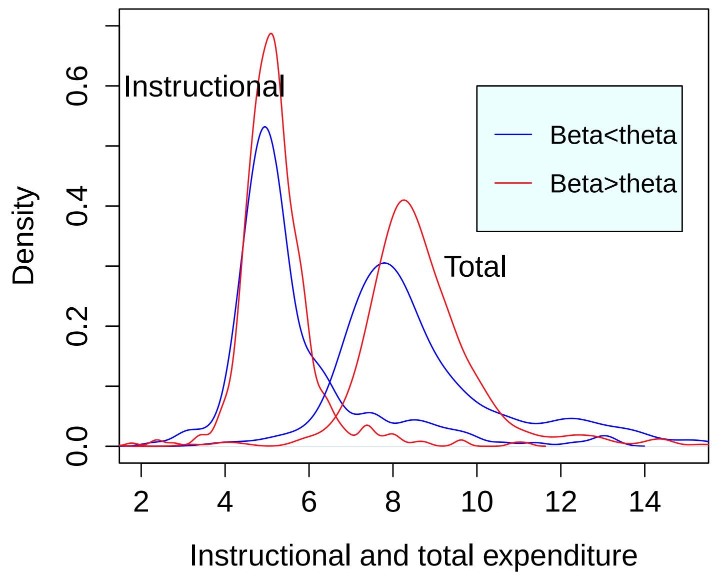

Figure 4: Instructional expenditures

8.3 Estimation of the mean and of the welfare

In the data set, we have

Table 1: Summary statistics and welfare function

	1979	1988	1992	1996
n	6230	6456	6597	6043
μ	83.083	102.885	109.629	111.457
median	74.23	87.363	91.912	94.204
welfare	4.31	4.47	4.52	4.56

We have used the Atkinson welfare function with $\epsilon = 1$, which means a medium concern for inequality. The welfare has risen with this measure, mainly because the mean has increased strongly between 1979 and 1988. The median is lower than the mean, due to the asymmetry of the distribution.

```
NROW(y79);NROW(y88);NROW(y92);NROW(y96)
mean(y79);mean(y88);mean(y92);mean(y96)
```

```
median(y79);median(y88);median(y92);median(y96)
mean(log(y79));mean(log(y88));mean(log(y92));mean(log(y96))
```

8.4 Inequality measures

Let us now compute inequality indices for these data sets.

Table 2: Inequality measures

	1979	1988	1992	1996
Gini	0.256	0.307	0.321	0.298
Theil	0.107	0.162	0.179	0.151
Atkin 0.5	0.052	0.076	0.084	0.071
Atkin 1.0	0.100	0.143	0.160	0.136
Atkin 1.5	0.147	0.205	0.271	0.203

All indices deliver the same message. Inequality has sharply risen from 1979 to 1988. It continued to rise till 1992 and then it has fallen without reaching its low level of 1979. The simple welfare function said that the rise in inequality was offset by the rise of mean income.

```
Gini(y79);Gini(y88);Gini(y92);Gini(y96);
Theil(y79);Theil(y88);Theil(y92);Theil(y96);
Atkinson(y79, parameter = 0.5);Atkinson(y88, parameter = 0.5);
    Atkinson(y92, parameter = 0.5);Atkinson(y96, parameter = 0.5);
Atkinson(y79, parameter = 1.0);Atkinson(y88, parameter = 1.0);
    Atkinson(y92, parameter = 1.0);Atkinson(y96, parameter = 1.0);
Atkinson(y79, parameter = 1.5);Atkinson(y88, parameter = 1.5);
    Atkinson(y92, parameter = 1.5);Atkinson(y96, parameter = 1.5);
```

Remark:

> The Kolm index gives results which are totaly at variance with the other indices. Try to explain why.

8.5 Poverty measures

We have first to define a poverty line. We took half of the mean income.

The number of poor rose dramatically between 1979 and 1988, as given by the headcount measure. It then dropped below its 1979 level, to rise slightly again 1996. This was certainly a

Table 3: Poverty measures

	1979	1988	1992	1996
z	41.54	51.44	54.81	55.73
Watts	0.014	0.035	0.052	0.033
Sen	0.017	0.041	0.055	0.036
FGT 0.0	4.00	7.90	3.36	4.36
FGT 1.0	0.093	0.171	0.192	0.142
FGT 2.0	0.011	0.028	0.038	0.025

consequence of an economic policy targeting at the number of poor after 1988. It we now take into account the income distribution of the poor, the result of that economic policy is less clear. The Watts, Sen and FGT($\alpha = 2$) rise till 1992. There was a drop 1996. But the level reached is still greater than the value of 1979.

```
z79 = mean(y79)/2; z88 = mean(y88)/2;
z92 = mean(y92)/2;z96 = mean(y96)/2;
z79;z88;z92;z96;
Watts(y79,z79);Watts(y88,z88);Watts(y92,z92);Watts(y96,z96)
Sen(y79,z79);Sen(y88,z88);Sen(y92,z92);Sen(y96,z96);
Foster(y79,z79,parameter=0);Foster(y88,z88,parameter=0);
    Foster(y92,z92,parameter=0);Foster(y96,z96,parameter=0);
Foster(y79,z79,parameter=1);Foster(y88,z88,parameter=1);
    Foster(y92,z92,parameter=1);Foster(y96,z96,parameter=1);
Foster(y79,z79,parameter=2);Foster(y88,z88,parameter=2);
    Foster(y92,z92,parameter=2);Foster(y96,z96,parameter=2);
```

9 Exercises

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

9.2 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.

9.3 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.

9.4 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.

9.5 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. (2003). Income inequality in oecd countries: Data and explanations. Working Paper 49: 479–513, CESifo Economic Studies, CESifo.

Bachelot, R. (2011). Suivi de l'objectif de baisse d'un tiers de la pauvreté en cinq ans. Rapport au parlement, Ministère des solidarités et de la cohésion sociale, Paris.

Carlsson, F., Daruvala, D., and Johansson-Stenman, O. (2005). Are people inequality-averse, or just risk-averse? *Economica*, 72:375–396.

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.

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

Kolm, S.-C. (1976). Unequal inequalities i. *Journal of Economic Theory*, 12:416–442.

O'Donoghue, J., Goulding, L., and Allen, G. (2004). Consumer price inflation since 1750. *Economic Trends*, 604:38–46.

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

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.

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