The Bayesian approach to poverty measurement

Lecture 4: Poverty indices and poverty curves

Michel Lubrano

November 2022

Contents

1 Introduction

Poverty indices are a way to summarise the left tail of an income distribution $f(x)$, obeying different types of axioms (see e.g. the survey of Zheng 1997). Poverty indices are thus particular transformations of the income distribution. In a Bayesian framework, the usual route is to consider a parametric model $f(x|\theta)$ for the income distribution. Once we have obtained draws from the posterior distribution of θ, we can transform these draws into draws of various poverty indices. Because there is no universal rule for selecting a particular poverty index, Jenkins and Lambert (1997) introduced TIP curves which document at the same time the three dimensions of poverty for each quantile of $f(x|\theta)$. Later Ravallion and Chen (2003) considered that growth is favourable to the poor if the lower quantiles of $f(x|\theta)$ increase more than its higher quantiles.

2 Preliminary definitions

A person is declared poor if she has an income lower than the poverty line z. For mathematical reasons, it is better to suppose that a person who has an income exactly equal to z is not poor. If x is a vector of incomes in a population of size n, the number of poor is given by:

$$q = \sum_{i=1}^{n} \mathbb{1}(x_i < z),$$

and the poverty rate:

$$H = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(x_i < z).$$

The average poverty gap is defined by:

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - z)_+.$$

A poverty index is a function $P(x, z)$ of the individual income x and the poverty line z into $\mathcal{R}^+$ whose value indicates a degree of poverty. The functional form of a poverty measure depends largely upon what we want to know about poverty: incidence (the number of poor), intensity (the distance to the poverty line) or inequality (the distribution of the poor). Sen (1976) has developed an axiomatic measurement of poverty and many indices have followed since that date.

2.1 Basic axioms

The main axioms of Sen are:

1. Focus axiom: the poverty index does not depend on the income distribution of those who are not poor.
2. Monotonicity axiom: Any increase in the income of a poor should decrease the poverty index.
3. Transfer axiom: a poverty measure should be sensitive to the redistribution of the income within the poor.

Several other axioms were introduced later in the literature which are surveyed in Zheng (1997). He organised the many proposed indices of the literature around the axioms that they fulfill or not. Among the axioms that were later developed, we should pay attention to the decomposability axiom which is mainly due to Foster et al. (1984). This axiom insures that a decomposable measure can decompose overall poverty into that of subgroups according to certain characteristics such as rural and urban for instance. The Increasing Poverty Line Axiom means that an index should increase if the poverty line is increased.

2.2 Main poverty indices

We shall consider essentially three families of poverty indices: FGT, Watts and Sen poverty indices because they are going to be the object of further developments.

Some of the most widely used poverty indices violate the transfer axiom. They are insensitive to the distribution among the poor:

1. The headcount measure

   $$H = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(x_i < z).$$

   which serves to measure a poverty rate. But it is however essential to know the number of poor if we want to implement a policy targeted toward the poor. For Atkinson (1987) a minimum income may be seen as a basic right. So the headcount measures the number of people that are deprived of that right. This measure is decomposable.

2. And the income gap ratio $I(x,z)$:

$$I(x,z) = 1 - \frac{\mu_p(x,z)}{z},$$

   where $\mu_p(x,z)$ is the mean income among the poor. Note that this index has two supplementary drawbacks. It is not decomposable, and it does not increase when z is increased.

Of course the main interest of Sen (1976) was to develop a family of distribution sensitive indices. The idea is to use a Gini index to measure inequality among the poor:

$$S(x,z) = \frac{2}{z\, n\, (q+1)} \sum_{i=1}^{q} (z - x_i)(q + 1 - i).$$

In this formulation, the Gini appears because of the rank weighting function $(q + 1 - i)$. Atkinson (1987) pointed out that the arguments about relative position and ranking are more persuasive for inequality measurement than for poverty measurement. Also this index does not satisfy the increasing poverty line axiom. The Thon index overcome some of the drawbacks of Sen index by changing the weighting function so as to obtain:

$$ST(x,z) = \frac{2}{z\, n\, (n+1)} \sum_{i=1}^{q} (z - x_i)(n + 1 - i).$$

As none of these indices are decomposable, Foster et al. (1984) proposed a new class of indices that are decomposable. Essentially and once again, the weighting function is changed:

$$F(x,z,\alpha) = \frac{1}{n\, z^{\alpha}} \sum_{i=1}^{q} (z - x_i)^{\alpha} = \frac{1}{n} \sum_{i=1}^{q} (1 - x_i/z)^{\alpha},$$

where α is a positive integer. For $\alpha = 0$, we have the headcount measure $F(x,z) = H(x,z)$, for $\alpha = 1$, $F(x,z) = H(x,z)I(x,z)$, for $\alpha = 2$ $F(x,z)$ satisfies the three axioms of Sen (1976) plus decomposability. A major advantage of this index is that $F(x,z,\alpha)$ when it is viewed as a function of z it directly related to stochastic dominance of order $\alpha + 1$.

A final index is the Watts (1968) index which proceed slightly differently:

$$W(x,z) = \frac{1}{n} \sum_{i=1}^{q} (\log z - \log x_i).$$

This is a distribution-sensitive measure, as it can be shown that:

$$W(x,z) = H(x,z)[Th(x,z) - \log(1 - I(x,z))],$$

where $Th(x,z)$ is the Theil inequality index among the poor:

$$Th(x,z) = \frac{1}{q}\sum_{i=1}^{q}(\log \mu_p - \log x_i).$$

And this index is also decomposable. However, it did not benefit of an axiomatic building by its author. However, **?** has provided such an axiomatic approach.

3 Posterior draws for poverty indices

All the indices that we have reviewed can be directly computed once we have observed a sample of incomes x_i. Note however, that these distribution free estimators discard all the observations above the poverty line z. In a Bayesian approach, we are going to take into account all the observations of x in order to make inference on the parameters θ of the complete income distribution. We assume that the latter can be modelled by a mixture of lognormal densities. So we are going to express all the afore mentioned poverty indices assuming that the income is a continuous variable with distribution $f(x)$. We shall model $f(x)$ according to a mixture of a lognormal distributions and express each poverty index in term of the parameters of this mixture. For this we need first to recall some of the properties of the lognormal distribution

3.1 Modelling the income distribution

A mixture of distributions accounts for the fact that the population is made of different groups with specific characteristics while the belonging to a particular group is not observed. When each member is assumed to follow a lognormal, the formulation of a finite mixture with K members is:

$$f(x|\theta) = \sum_{k=1}^{K}\eta_k f_\Lambda(x|\mu_k, \sigma_k^2),$$

where η_k are the weights summing to 1 and μ_k, σ_k^2 the parameters of each member and θ regroups all the parameters. Mixtures have very nice properties due to their linearity. In particular, the mean and the cumulative distribution (CDF) have direct expressions with:

$$\mathrm{E}(x|\theta) = \sum_{k=1}^{K} \eta_k \int_0^\infty x\, f_\Lambda(x|\theta_k)\, dx,$$

and

$$F(x|\theta) = \sum_{k=1}^{K} \eta_k F_\Lambda(x|\theta_k).$$

So the mean is weighted average of the mean of each component and the CDF is the weighted average of each component CDF.

The lognormal density is noted:

$$f_\Lambda(x|\theta) = \frac{1}{x\sigma\sqrt{2\pi}} \exp -\frac{(\log x - \mu)^2}{2\sigma^2}, \tag{1}$$

with CDF:

$$F_\Lambda(x|\theta) = \int_0^x f_\Lambda(t|\theta)\mathrm{d}t = \Phi\left(\frac{\log x - \mu}{\sigma}\right), \tag{2}$$

where Φ is the Gaussian CDF. The mean and variance are:

$$\mathrm{E}(x|\theta) = \int_0^\infty x f_\Lambda(x|\theta) d\theta = e^{\mu+\sigma^2/2}, \tag{3}$$

$$\mathrm{Var}(x|\theta) = \int_0^\infty x^2 f_\Lambda(x|\theta) d\theta - \mathrm{E}(x|\theta)^2 = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}. \tag{4}$$

However, we need some more indicators, such as partial moments in order to compute the distribution of poverty indices because our interest is limited to integral between zero and z and not between zero and ∞. The first partial moments are (see e.g. Jawitz 2004):

$$\int_0^z x f_\Lambda(x|\theta) = e^{\mu+\sigma^2/2}\Phi\left(\frac{\log(z) - \mu - \sigma^2}{\sigma}\right) \tag{5}$$

$$\int_0^z x^2 f_\Lambda(x|\theta) = e^{2\mu+2\sigma^2}\Phi\left(\frac{\log(z) - \mu - 2\sigma^2}{\sigma}\right) \tag{6}$$

Let us suppose that we have obtained m draws from the MCMC output, obtained using a Gibbs sampler. We call each of these draws $\eta_k^{(j)}$ for the weights and $\mu_k^{(j)}$, $\sigma_k^{(j)}$ for the parameters of the lognormal members.

3.2 FGT indices

We have written poverty indices, assuming that the income distribution was discrete. So we have used sums of discrete elements. Here, we have to give the general expression of these indices, assuming a continuous distribution for x. So, the general class of poverty indices of Foster et al. (1984) writes:

$$FGT(z,\alpha) = \int_0^z (1 - x/z)^\alpha f(x)\, dx, \tag{7}$$

We are going to look for analytical solutions to this integral when $f(x)$ is a mixture of lognormals for each of the usual values of α.

The poverty head-count ratio or poverty rate H corresponds to $\alpha = 0$. In this case

$$FGT(z,0) = \int_0^z f(x)\, dx,$$

and we have just to find the expression of the cumulative distribution of a mixture of lognormals. We have a simple solution to this integral calculus with:

$$H(z|\theta^{(j)}) = \sum_{k=1}^K \eta_k^{(j)} \Phi\left(\frac{\log z - \mu_k^{(j)}}{\sigma_k^{(j)}}\right). \tag{8}$$

For $\alpha = 1$, the poverty gap index can be decomposed into:

$$\int_0^z (1 - x/z) f(x) dx = F(z) - \frac{1}{z}\int_0^z x f(x)\, dx.$$

Using (2) and (5), we have:

$$\begin{aligned} FGT(z|\theta^{(j)},1) \;=\; & \sum_{k=1}^K \eta_k^{(j)} \left[\Phi\left(\frac{\log z - \mu_k^{(j)}}{\sigma_k^{(j)}}\right)\right. \\ & \left. - \frac{1}{z} e^{\mu_k^{(j)} + \sigma_k^{2(j)}/2} \Phi\left(\frac{\log z - \mu_k^{(j)} - \sigma_k^{2(j)}}{\sigma_k^{(j)}}\right)\right]. \end{aligned} \tag{9}$$

For $\alpha = 2$, we have to evaluate

$$\int_0^z f(x)\, dx - \frac{2}{z}\int_0^z x f(x)\, dx + \frac{1}{z^2}\int_0^z x^2 f(x)\, dx.$$

Using (2), (5), (6), we get:

$$\Phi\left(\frac{\log z - \mu}{\sigma}\right) - \frac{2}{z} e^{\mu + \sigma^2/2} \Phi\left(\frac{\log(z) - \mu - \sigma^2}{\sigma}\right) + \frac{1}{z^2} e^{2\mu + 2\sigma^2} \Phi\left(\frac{\log(z) - \mu - 2\sigma^2}{\sigma}\right),$$

so that for a mixture of lognormals we have:

$$\begin{aligned} FGT(z|\theta^{(j)},2) &= \sum_{k=1}^{K} \eta_k^{(j)} \left[\Phi\left(\frac{\log z - \mu_k^{(j)}}{\sigma_k^{(j)}}\right) - \frac{2}{z} e^{\mu + \sigma_k^{2(j)}/2} \Phi\left(\frac{\log z - \mu_k^{(j)} - \sigma_k^{2(j)}}{\sigma_k^{(j)}}\right) \right. \\ &\quad \left. + \frac{1}{z^2} e^{2\mu_k^{(j)} + 2\sigma_k^{2(j)}} \Phi\left(\frac{\log(z) - \mu_k^{(j)} - 2\sigma_k^{2(j)}}{\sigma_k^{(j)}}\right) \right]. \end{aligned} \tag{10}$$

3.3 The Watts index

The Watts (1968) poverty index writes:

$$W(z) = -\int_0^z \log(x/z) f(x)\, dx. \tag{11}$$

Muller (2001) gave its expression when $f(x)$ is a lognormal:

$$W(z) = (\log z - \mu)\Phi\left(\frac{\log z - \mu}{\sigma}\right) + \sigma\phi\left(\frac{\log z - \mu}{\sigma}\right),$$

where ϕ is the Gaussian probability density. The generalisation to mixtures provides:

$$W(z|\theta^{(j)}) = \sum_{k=1}^{K} \eta_k^{(j)} \left[(\log z - \mu_k^{(j)})\Phi\left(\frac{\log z - \mu_k^{(j)}}{\sigma_k^{(j)}}\right) + \sigma\phi\left(\frac{\log z - \mu_k^{(j)}}{\sigma_k^{(j)}}\right) \right].$$

3.4 The Sen index

The revision of Sen index by Shorrocks (1995) leads to:

$$SST(z) = \frac{2}{z} \int_0^z (z - x)(1 - F(x)) f(x)\, dx, \tag{12}$$

as expressed in Davidson (2009). We can decompose it into:

$$SST(z) = 2FGT(z,1) - 2\int_0^z (1 - x/z) F(x) f(x)\, dx.$$

We have already found an expression for $FGT(z, 1)$. The last integral is related to the Gini index and has no analytical solution. In a similar situation, Lubrano and Ndoye (2016) proposed to evaluate numerically the integral for each draw of the parameters, using a Simpson rule.

3.5 Empirical comparison of poverty indices

We have taken the FES data for 1979. We first compute the value of the usual poverty indices, using the distribution free formulae in Table 1. We then make inference for a three member mixture, asking for an ordering of the draws on μ. The number of draws is equal to $m = 1,000$ plus 100 draws for warming the chain, which is certainly not enough, but sufficient for this exercise. The posterior moments are provided in Table 2 and the posterior predictive in Figure 1.

Table 1: Characteristics of the 1979 sample

Pov. line	Head count	FGT1	FGT2	Watts	SST
41.54	0.0937	0.0115	0.00288	0.0139	0.0225

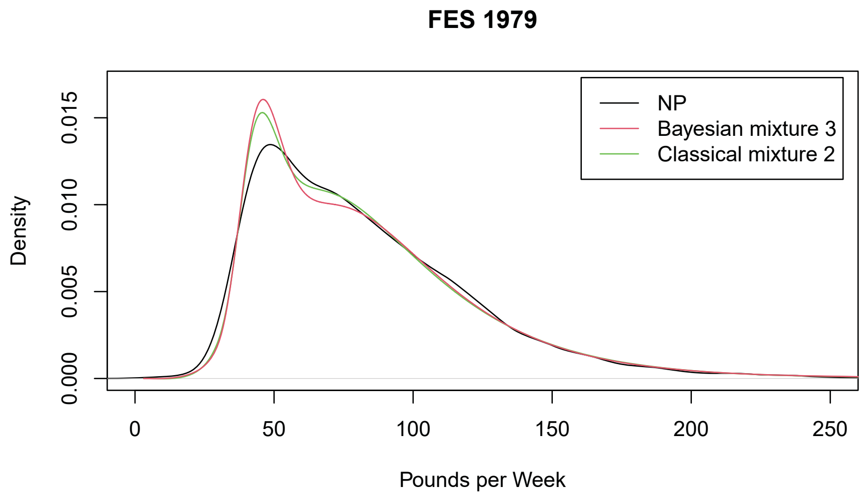


Figure 1: Family Expenditure Survey 1979 data

Table 2: Posterior moments for the mixture parameters

Member	$E(\mu\|y)$	$sd(\mu\|y)$	$E(\sigma^2\|y)$	$sd(\sigma^2\|y)$	$E(\eta\|y)$	$sd(\eta\|y)$
1	3.823	0.012	0.026	0.003	0.198	0.020
2	4.407	0.022	0.251	0.063	0.356	0.140
3	4.457	0.020	0.130	0.054	0.446	0.134

We can then use these posterior draws for deriving the posterior density for several poverty indices. The posterior density is in black and the red vertical line corresponds to the sample value of the corresponding index. For all indices, this vertical line is well within the middle of the density. This is a way to check the calculations. The SST index is slightly longer to compute due to the integral that has to be evaluated for each draw.

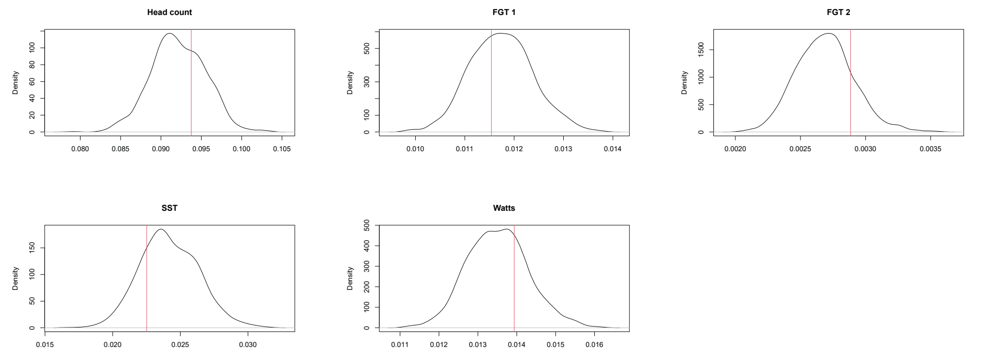


Figure 2: Posterior densities for three FGT indices

4 TIP curves

Let us go back to the FGT poverty indices. How to choose the value of α, which means which aspect of poverty do we want to measure? The TIP curve of Jenkins and Lambert (1997) documents the three dimensions of poverty for each quantile of the income distribution up to the quantile corresponding to the poverty line z.

4.1 A formal definition

Let us consider an income vector of n individuals denoted by $y \in \mathbb{R}^+$, and a poverty line z. We define the relative poverty gap as:

$$\max\{1 - y/z, 0\} = (1 - y/z)\mathbb{1}(y \leq z), \tag{13}$$

where $\mathbb{1}(y \leq z)$ is the indicator function which takes the value one if $y \leq z$ and zero otherwise. A poverty gap measures the asymmetric distance between the poverty line and the income vector. The class of decomposable poverty indices introduced by Foster et al. (1984) corresponds to various partial sums over the relative poverty gap depending on the integer parameter $\alpha \in \{0, 1, 2\}$ as defined in (7) that we recall here for convenience:

$$P^\alpha(z) = \int_0^z (1 - y/z)^\alpha f(y)dy, \tag{14}$$

For $\alpha = 1$ and when $P^\alpha(z)$ is considered as a function of z, we have the normalised deficit curve. The relative TIP curve of Jenkins and Lambert (1997) is defined with respect to the cumulative relative poverty gaps and is thus closely related this curve:

$$TIP(p, z) = \int_0^{F^{-1}(p)} (1 - y/z)\mathbb{1}(y \leq z) f(y) dy, \tag{15}$$

$F^{-1}(p)$ being the quantile function, and p the proportion of individuals. This is the quantile approach to poverty measurement because the integration bound is expressed in term of a quantile.

A distribution free estimator for the TIP curve is easily obtained once we order the observations:

$$\widehat{TIP}(\hat{p}_i = i/n, z) = \frac{1}{n} \sum_{j=1}^{i} (1 - y_{(j)}/z)\mathbb{1}(y_{(j)} < z), \tag{16}$$

where $y_{(j)}$ are the order statistics of the distribution.

For values of p greater than the poverty incidence or headcount ratio $P^0 = F(z)$, the TIP curve saturates and becomes horizontal as illustrated in Figure 3. At this abscissae point P_0, the ordinate value is the poverty intensity P^1 or average poverty gap. Finally the curvature of the curve represents the inequality among the poor P^2. A useful feature of the relative TIP curve is that poverty incidence P^0, poverty intensity P^1 and poverty inequality P^2 are equivalent to the FGT(α) indices when $\alpha = 0, 1, 2$ respectively. A TIP curve is thus a convenient device for displaying at the same time the three essential aspects of poverty, justifying its name the Three Is of Poverty.

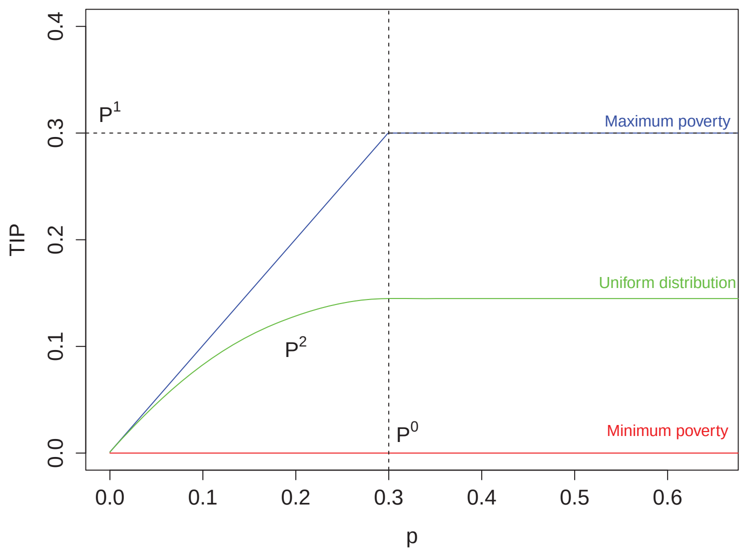


Figure 3: TIP curves from different income distributions

In Figure 3, we have represented three TIP curves for comparing three income distributions that have the same poverty rate $P_0 = 0.3$, using the same poverty line z.

1. The top blue TIP curve represents the case of maximum poverty where all poor people have a zero income. In this case, the average poverty gap P^1 is the highest and $P_1 = P_0$. This is a straight line reflecting the fact that every poor has the same income. Inequality among the poor P^2 is also maximum (maximum slope of 1) and the Gini among the poor is 1.

2. The bottom red TIP curve reflects the perfect opposite situation where all poor people have the the same level of income and are hitting the poverty line z while still being poor. The proportion of poor is 0, the average poverty is zero and the Gini among the poor is also 0. This is a straight horizontal line, and inequality among the poor is minimum (minimum slope of 0). In this case the TIP curve corresponds to the x-axis.

3. Between these two extremes, we have the intermediate green curve, which is drawn here as an example from an uniform distribution with a proportion of poor P^0 being 0.3. The average poverty gap is 0.15. The curvature of the TIP curve is directly related to the inequality among the poor, exhibiting a Gini of 0.33. The closer the income of the poor individuals are to the poverty line, the smaller are the poverty gaps, and the closer the slope is to 0.

Figure 3 clearly shows that an identical P_0 can hide very different situations, all described adequately by a different TIP curve.

4.2 A parametric representation

The TIP curve was defined as:

$$TIP(p,z) = \int_0^{F^{-1}(p)} (1 - x/z)\mathbb{1}(x \le z) f(x) dx,$$

Letting $q = F^{-1}(p)$, we can decompose this equation into:

$$TIP(p,z) = \int_0^q f(x)\,dx - \frac{1}{z}\int_0^q x f(x)\,dx = p - \frac{1}{z}GL(p), \qquad \text{for } p \le F(z),$$

where $GL(p)$ is the generalised Lorenz curve.

The whole expression has an analytical form for the lognormal distribution. But this is of little use as it is not possible to find the closed expression of $GL(p)$ when $f(x)$ is a mixture. So it is better to consider directly:

$$TIP(p,z) = \sum_{k=1}^K \eta_k \int_0^q f_\Lambda(x|\mu_k, \sigma_k^2) - \frac{1}{z}\sum_{k=1}^K \eta_k \int_0^q x f_\Lambda(x|\mu_k, \sigma_k^2)\,dx,$$

where the quantile q has to be calculated separately. This presentation relies on the two-equation definition of the Lorenz curve, in use before Gastwirth (1971). Both integrals have an analytical solution leading to:

$$\begin{aligned} TIP(p, z|\theta^{(j)}) \quad = \quad & \sum_{k=1}^K \eta_k^{(j)} \left[\Phi\left(\frac{\log q^{(j)} - \mu_k^{(j)}}{\sigma_k^{(j)}} \right) \right. \\ & \left. - \quad \frac{1}{z} e^{\mu_k^{(j)} + \sigma_k^{2(j)}/2} \Phi\left(\frac{\log q^{(j)} - \mu_k^{(j)} - \sigma_k^{2(j)}}{\sigma_k^{(j)}} \right) \right]. \end{aligned} \quad (17)$$

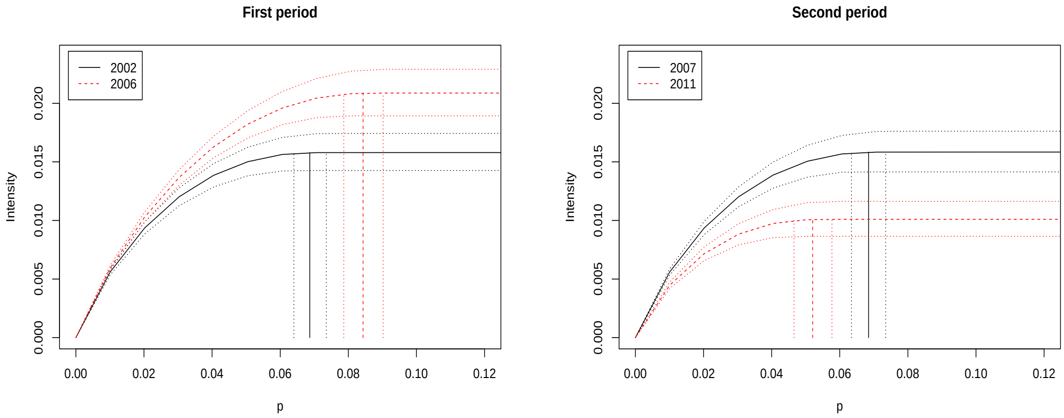
The difficulty is that the left-hand side is a function of p while the right-hand side is a function of q. For each draw of θ, we have to solve numerically the equation:

$$F(q^{(j)}|\theta^{(j)}) = p, \quad (18)$$

for each point of a predefined grid on p. This is a feasible problem because it is of dimension one on a finite interval defined by the range of x. Brent (1971) algorithm is very efficient in this case.

4.3 Child poverty in Germany

The feasibility of the method is illustrated in Figure 4, extracted from Fourrier-Nicolaï and Lubrano (2020). It depicts the evolution of child



90% credible intervals are represented by dotted lines. In black solid line is represented the TIP curve at the beginning of each sub-period. The red dashed line corresponds to the TIP curve of the end of each subperiod.

Figure 4: The three I's of child poverty in Germany

poverty in Germany between 2002 and in 2011. The period has experienced a dramatic change in family social allowances. In each panel, vertical lines represent poverty headcount, horizontal lines poverty intensity and the curvature of the TIP curves poverty inequality. The fact that credible intervals do not overlap indicates that child poverty has significantly changed over the period. It increased a lot between 2002 and 2006 to finally decrease between 2007 and 2011. The change in family social policy has managed to cut the regular increase in child poverty that was documented in Corak et al. (2008).

4.4 The difference between child and adult poverty

Corak et al. (2008) found that there was a strong difference between adult and child poverty between 2000 and 2004 for the whole of Germany. In our sample, the rate of poverty is the same between children and adults in 2000 (6.0% for child and 5.6% for adults). The discrepancy between child and adult poverty rates increases till 2006 with respectively 9.9% and 8.2%. After that date, the adult poverty rate remains constant around 8.2% (between 2006 and 2010) while the child poverty rate decreases to become similar to the rate of adult poverty in 2010 and then becomes even much smaller (6.4% for child poverty versus 7.6% for adult poverty in 2012).[1] In Figure 5, we

[1]We have defined adults as individuals of 18 years and over coming from a household where there was no children. It is just the contrary of the child sample where observations comes from households with children and concern individuals below 18 years. The reported poverty rates were computed using the unbalanced panel, while TIP curves were computed

provide TIP curves for current adult poverty over the two periods. When considering only adults, the remarkable fact is that there is no significant evolution of poverty in all of its dimensions (confidence intervals intersect for the different years), contrary to what happened to child poverty as reported from Figure 4. So the evolution of poverty over the two subperiods concerned

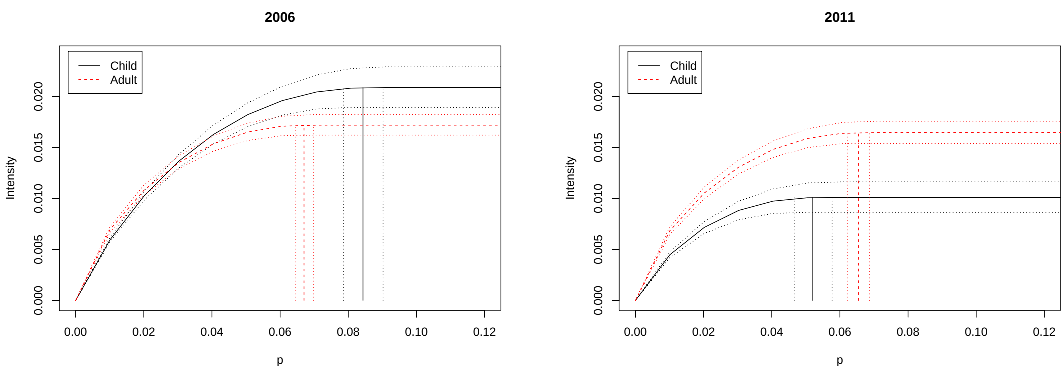


Figure 5: The constancy of current adult poverty over the period

mainly children with an increase and a decrease, while there was no significant effect on the population of adults.

5 Pro-poor growth or trickle-down theory

Ravallion and Chen (2003) try to answer to the question: is growth favourable to the poor? For this they have developed a tool which is the Growth Incidence Curve. For measuring growth, we need two periods when one was sufficient for the TIP curve. What the GIC does is to compare two quantile functions and see which quantile has increased the most.

5.1 A formal definition

Consider two distributions of incomes observed at time $t-1$ and t, and characterised by the respective cumulative distribution functions F_{t-1} and F_t with support contained in the non-negative real line. The growth incidence curve (GIC) measures the growth rate of the p-quantile for every p:

$$g_t(p) = \frac{Q_t(p)}{Q_{t-1}(p)} - 1 \simeq \log Q_t(p) - \log Q_{t-1}(p). \tag{19}$$

using the two balanced panels.

Graphically, the GIC associates the growth rate of income with respect to the proportion p of individuals ordered by increasing income. Thus, we can say that the p-quantile increases if $g_t(p) > 0$ and conversely that the p-quantile decreases when $g_t(p) < 0$. Therefore, growth incidence curves can be used for characterising poverty and inequality changes. For instance, if inequality does not change between $t-1$ and t, then the curve is a flat line i.e. $g_t(p) = \gamma_t$ for all p, γ_t being the average growth rate over the period. It is worth mentioning that, because $Q(p) = GL'(p)$ (by definition of the Lorenz curve, see Gastwirth 1971), the growth incidence curve can be equivalently expressed in terms of the first derivative of the generalised Lorenz curve .

5.2 Pro-poor judgements using the growth incidence curve

When can we say that distributional changes are favourable to the poor? Let us consider a common poverty line z. We can say that growth has decreased poverty if $g_t(p) > 0$ for all $p < F_t(z)$ where $F_t(z)$ is the proportion of individuals below the poverty line (the headcount ratio). This means that the proportion of individuals below the poverty line is always greater in F_{t-1} than in F_t, for any poverty line lower than z. Conversely, if $g_t(p) < 0$ for all $p < F_t(z)$, then we can say that growth has been detrimental for the poor so as the income of the poor has reduced in absolute terms between the two periods. If the growth incidence curve is negative only for some values of $p \leq F_t(z)$, then we cannot conclude unambiguously about whether the distributional changes have been welfare-improving or not for the poor.

One can be concerned about the relative situation of the poor regardless of the initial distribution F_{t-1}. Indeed, there are normative conditions for promoting a relative pro-poor growth so that the poor benefit proportionately more from growth than the rich (Kakwani and Pernia 2000). In this attempt, Duclos (2009) and Araar et al. (2009) go a step further and state that growth is relatively pro-poor if:

$$g_t(p) > \gamma_t, \quad \forall p \in [0, F_t(z)].$$

Thus, in order to determine whether growth has been relatively pro-poor, we only have to compare the growth incidence curve with the average growth rate γ_t. This is equivalent to the statement that the quantiles of the poor increase at a pace greater than the average growth.

5.3 Parameter-free inference

Suppose that we have n observations of income and that we have ordered these observations so as to obtain n order statistics denoted $y_{(i)}$. Then, the empirical quantile function is obtained as:

$$\hat{Q}(p_i = i/n) = y_{(i)}, \tag{20}$$

If we normalize this graph by the mean, we get the well-known Pen's parade while the Lorenz curve corresponds to :

$$\hat{L}(p_i = i/n) = \sum_{j=1}^{i} y_{(j)} / \sum_{j=1}^{n} y_{(j)}. \tag{21}$$

The generalised Lorenz curve is obtained by multiplying the Lorenz curve by the empirical mean $\bar{y} = n^{-1} \sum_{i=1}^{n} y_i$. As the empirical quantile function is not continuous, a Kernel estimator for $Q(.)$ has been proposed in the literature (see e.g. Yang 1985):

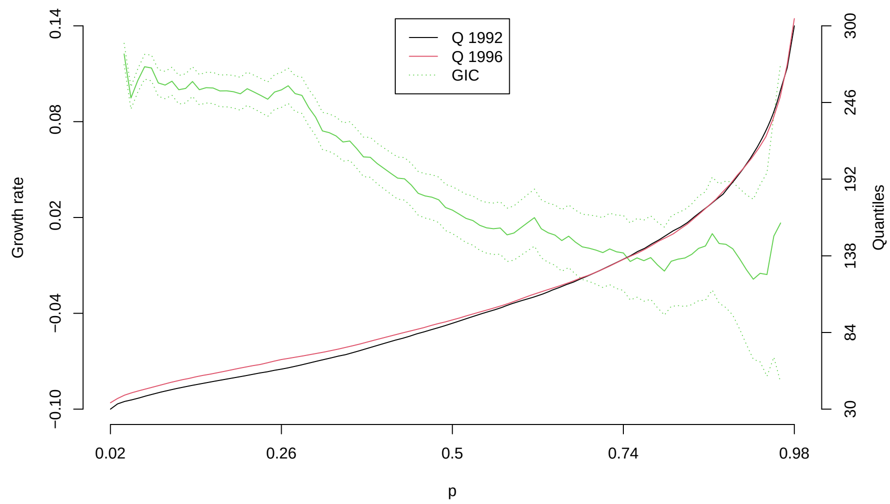
$$S_n(\tau) = \frac{1}{n\,h} \sum_{i=1}^{n} y_{(j)} K\left(\frac{i/n - \tau}{h}\right),$$

where K is a kernel density and h the window size which has to be determined. However this estimator has the same variance as the natural estimator.

Considering two income series of the same length n, the distribution-free empirical growth incidence curve is obtained as:

$$\hat{g}_t(p_i) = \log \hat{Q}_t(p_i) - \log \hat{Q}_{t-1}(p_i).$$

Thus, the empirical GIC is defined as the difference between two empirical step functions for which the variance is likely to be incidentally increased in the tails of the distribution. This is a particular concern when studying poverty or inequality. This causes the GIC to behave erratically as illustrated in Figure 6. Particularly, we estimate a distribution-free GIC based on empirical data for the UK in 1992 and 1996. Although, the two empirical quantile functions appear to be quite smooth with small differences, the difference of their logs (at value between -0.019 and 0.12) is too much erratic. It is worth mentioning that this variability appears even with 6,595 and 9,043 observations for 1992 and 1996, respectively. In addition, this variability is incidental in that it originates from a pure sample variation that has no economic interpretation.



Quantiles are computed using an interpolation method. Dots correspond to a 90% confidence interval.

Figure 6: Quantile functions and GIC, FES data 1992-1996

5.4 The growth incidence curve for the log-normal model

In this section, we propose parametric growth incidence curves based on parametric assumptions for the income distribution. First, we model the income distribution using the log-normal assumption. This is an introduction for a more complex model where the income distribution is modelled as a mixture of K log-normal densities. This approach allows to increase substantially the flexibility of the model as the number of mixture components increases. Then, we propose Bayesian inference for these parametric model and thus for the growth incidence curves.

Let us recall that the lognormal distribution provides a direct formula for the quantile function:

$$Q(p) = \exp(\mu + \sigma\Phi^{-1}(p)), \tag{22}$$

where Φ is the cumulative distribution function of the standard normal distribution.

Consequently, the growth incidence curve for the *log-normal distribution* can be obtained easily. Suppose that we have estimated (μ_1, σ_1^2) and (μ_2, σ_2^2) for the log-normal distributions in the first and the second periods, respectively. The growth incidence curve based on the log-normal distribution is:

$$g_t(p) = (\mu_2 - \mu_1) + (\sigma_2 - \sigma_1)\Phi^{-1}(p). \tag{23}$$

This parametric form is very much constrained as its shape depends entirely on the shape of cumulative distribution of the Gaussian density. For constant values of μ_2 and μ_1, because $\Phi^{-1}(0.50) = 0.0$, the GIC curve turns around a fixed point, and its slope depends only on the difference $\sigma_2 - \sigma_1$.

A mixture model allows to increase substantially the flexibility of the distributional assumption as the number of components K increases. They are a natural way to improve the flexibility of the very constrained GIC curve given in (23). If it is easy to derive the CDF of a mixture, the quantile function of a mixture has no closed analytical form. $F(y|\theta)$ has to be inverted numerically. Precisely, for each point p and for each value of θ, we have to solve numerically in q the equation:[2]

$$p = F(q|\theta). \tag{24}$$

Substituting the computed quantiles into equation (25), we get the growth incidence curve for a mixture of K log-normal densities, without further analytical result.

Bayesian inference for mixtures of log-normal densities relies on a Gibbs sampler, a procedure that we have detailed in the previous chapter. Once we have obtained posterior draws of the mixture parameters, we can get an evaluation of the quantile function which is the necessary ingredient for evaluating the posterior distribution of the growth incidence curve. More

[2] The operation is very quick and secure because the dimension of q is one. The `R` function `uniroot` searches a predefined interval using the Brent algorithm without derivatives.

precisely, for each point of the grid p and for each draw, we solve numerically in q the equation (24) so as to obtain the quantile function $Q(p|\theta)$, and then obtain a draw of the growth incidence using equation (25). The posterior mean of $g_t(p)$ is obtained as the mean of all these m posterior draws for each value of p. The 0.05 and 0.95 quantiles of these m draws provide an evaluation of a 90% confidence interval of the growth incidence curve. However, the use of mixture models raises concern about over-fitting and about the choice of the number of log-normal components knowing that the number of parameters increases in this case by a multiple of 3 when the number of components K increases.

6 An empirical illustration using UK data

As an empirical illustration, we use data from the FES for four years: 1979, 1988, 1992 and 1996. This period has been extensively studied in the literature (e.g. Charpentier and Flachaire 2015, Jenkins 1995) as it covers a period of considerable changes in the UK income distribution. Particularly, prime minister Margaret Thatcher (1979-1990) introduced a series of neoliberal major economic policies, so important and adverse for the poor, that they finally led to a political crisis and to her demission in 1990. She was replaced by John Major (1990-1997) who introduced some measures to reduce economic inequalities, till the victory of Tony Blair in 1997.

6.1 Data and context

We consider the disposable household income (i.e. post-tax and transfer income), normalised by the McClements adult-equivalence scale and deflated by the corresponding relative consumer price index. We consider three mutually exclusive sub-groups: households where the head is retired, households where both heads are working, households where both heads are unemployed. There are of course intermediate cases which are regrouped in the category *others* and which represents between 15% and 20% of the sample. We can question if growth over the period has benefited in a similar way to these three different groups with different characteristics and what were the major breaks.

We consider a poverty line defined as 60% of the median income (the official UK definition). In Table 3, we report the proportion of each group, the headcount ratio (i.e. the proportion of households with income below the poverty line based on the mean) and the Theil index as a measure of inequality.

The general pattern, inferred from the total sample, is that inequality and poverty have increased from 1979 to 1992 and then decreased between 1992 and 1996 but not till the low levels of 1979. This general pattern is reproduced for the unemployed group. In retired group, the evolution of poverty is slowly decreasing. For the working group, poverty fluctuates with no specific trend.[3] Inequality has an evolution which is quite similar between the groups. It increases till 1992 and then decreases to a level which depends on the groups.

Table 3: Poverty and inequality by subgroups
Poverty line 60% of median income

	Total sample		Retired			Both working			Both unemployed		
	H.C.	Theil	Prop	H.C.	Theil	Prop	H.C.	Theil	Prop	H.C.	Theil
1979	0.135	0.107	0.293	0.309	0.085	0.449	0.017	0.069	0.059	0.459	0.164
1988	0.180	0.162	0.307	0.309	0.135	0.401	0.016	0.096	0.122	0.511	0.168
1992	0.196	0.179	0.301	0.293	0.153	0.380	0.029	0.112	0.151	0.515	0.240
1996	0.151	0.151	0.299	0.200	0.124	0.390	0.021	0.102	0.157	0.412	0.142

H.C. is the head count ratio measuring the proportion of people under the poverty line. Theil is the Theil index. Prop. indicates the proportion of the concerned group in the total population.

Table 3 provides a partial portray of the evolution of poverty and inequality although these results might be sensitive to the choice of the measure of inequality but also to the choice of the poverty line. The growth incidence curve provides a more complete portray of the evolution of the income distribution over time which is not sensitive to the choice of the measures of poverty and inequality.

6.2 Growth incidence curves for the whole population

Figure 8 confronts the performance of three parametric growth incidence curves to the distribution free estimator for real data. The first parametric model is based on the simple log-normal assumption for the income distribution represented by the green line and using 5 000 posterior draws of the GIC under a non-informative prior. The second model adjusts a mixture of three log-normals - represented by the red line - for the income distribution using 5 000 and 500 draws for warming the chain and an informative prior.

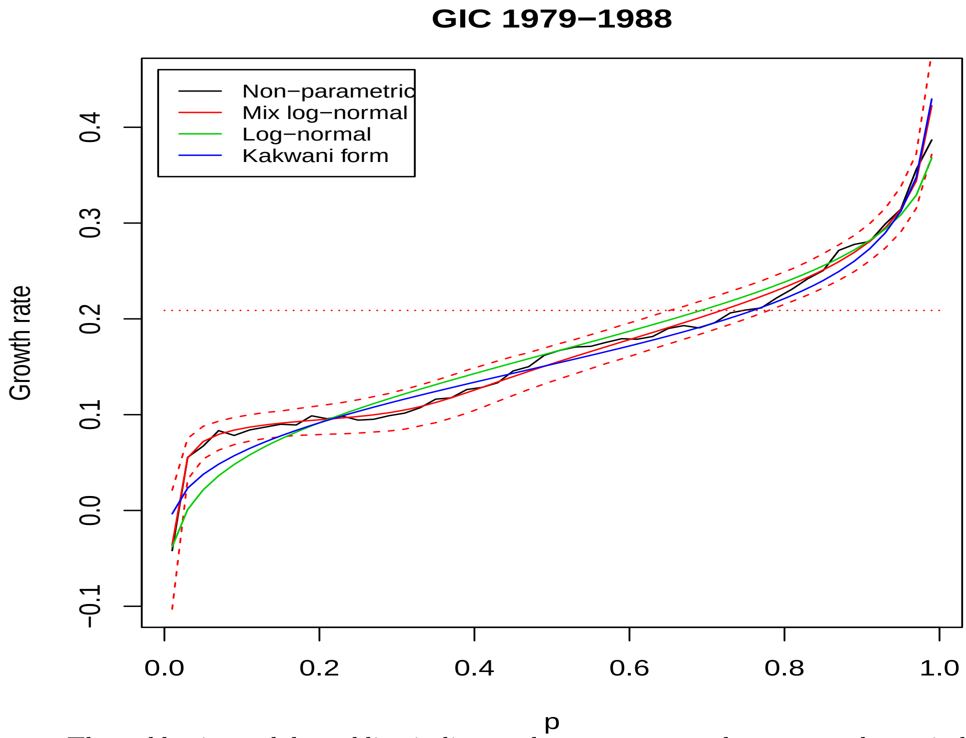
[3]If we had used a poverty line based on 50% of the mean, the evolution of poverty within each group would have followed a different pattern, because the mean income has increased a lot over the period.

Remark:

> A prior was devised as follows. Equal weights were assumed for the Dirichlet prior on η with $\nu_0 = 5$. For the normal prior on μ, the prior expectation was set equal to the mean of the logs of the observations with $n_0 = 1$. A great care was devoted to the elicitation of the prior on σ^2, a parameter which is important for the resulting shape of the mixture. The prior expectation was chosen equal to 0.01 for the first member, increased by a factor of $\text{Var}(\log(y))/K$ for the next members and the number of degrees of freedom was chosen equal to 50. This type of prior is a way to avoid label switching as explained in Lubrano and Ndoye (2016). Note that the number of components have been chosen so as to minimise the Bayesian information criterion (BIC).

The third model - in blue - makes use of the Kakwani (1980)'s functional form using 5 000 posterior draws of the GIC and a non-informative prior. These parametric forms of the GIC are compared to the distribution-free estimate provided in black for 1979-1988 and 1992-1996. For the sake of clarity, we only display the 90% probability interval for the mixture model as a baseline case.

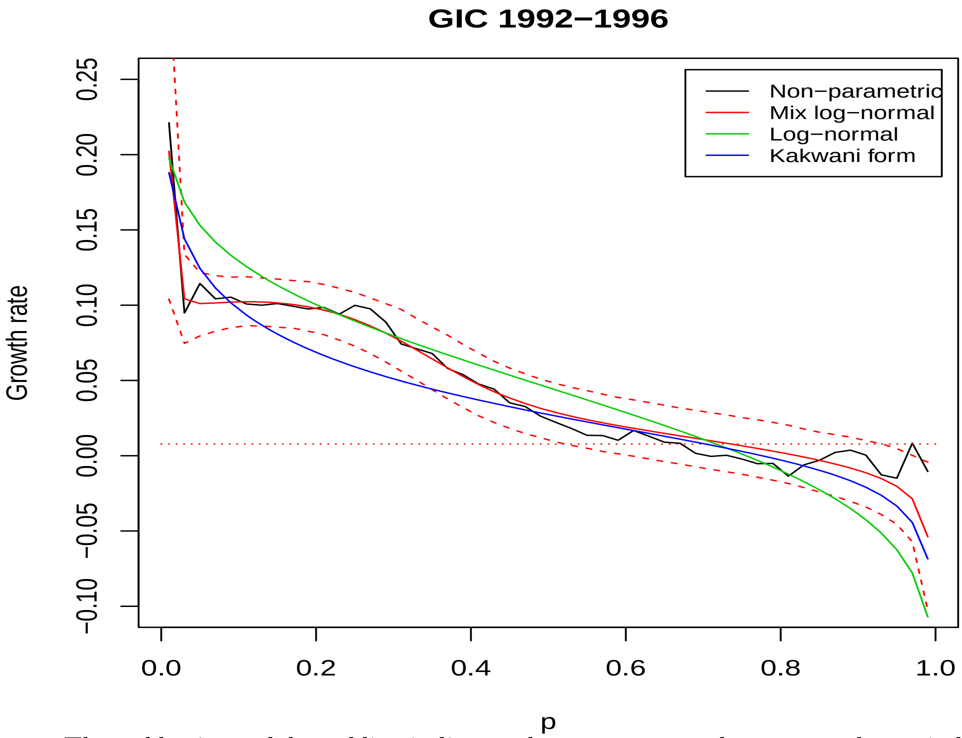
Overall, all methods provide essentially the same message: an increase of inequality during the first period and the inverse movement during the second period.



The red horizontal dotted line indicates the average growth rate over the period.
Dashed red curves represent the 90% probability interval for the mixture model.

Figure 7: Growth incidence curves for 1979-1988 and 1992-1996

The first period 1979-1988 is clearly characterised by an anti-poor and inequality-increasing growth. We must go up to the 0.70^{th} quantile in order to have an average growth rate of income greater than the average growth.



The red horizontal dotted line indicates the average growth rate over the period.
Dashed red curves represent the 90% probability interval for the mixture model.

Figure 8: Growth incidence curves for 1979-1988 and 1992-1996

The second period 1992-1996 becomes pro-poor and inequality-reducing. Households up to the 0.70^{th} quantile experienced an income increase greater than average. But the average growth rate is much lower than during the first period. Very top incomes had a negative growth rate

6.3 Growth incidence curves for subgroups

Let us now try to gauge whether growth affected the different groups of the population in a similar way with respect to their occupational status:

1. households where the head is retired,

2. households where both heads are working,

3. households where both heads are unemployed.

4. In addition, we also consider lone parenthood, that is single headed families with children.

Figures 9-10 display the subgroups growth incidence curves for both periods, using this time only Kakwani (1980)'s model, without indicating confidence bounds for clarity of the graph. Considering sub-groups implies that sample sizes are reduced, ranging from roughly 400 for the lone parents group to 2 500 for the both parents working group; a case where parametric models should be much more at ease than a distribution-free estimate.

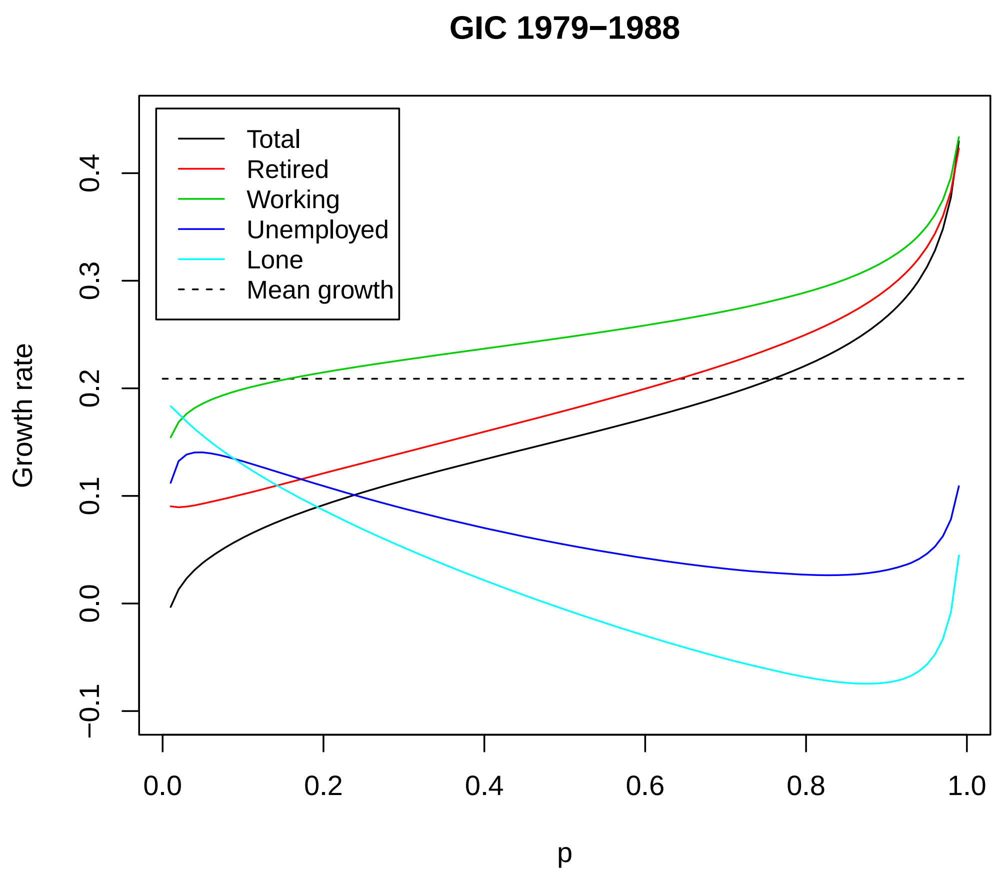


Figure 9: Subgroups GIC for 1979-1988 and 1992-1996

The first period is characteristic of a competitive growing society wherein the working households benefited the most from growth. To a lesser extent, the rich retired households also benefited from growth. On the contrary, the situation of the unemployed and that of single headed families with children worsen dramatically. There were left out of the benefit from growth and lone parents have seen their situation deteriorated in absolute term.

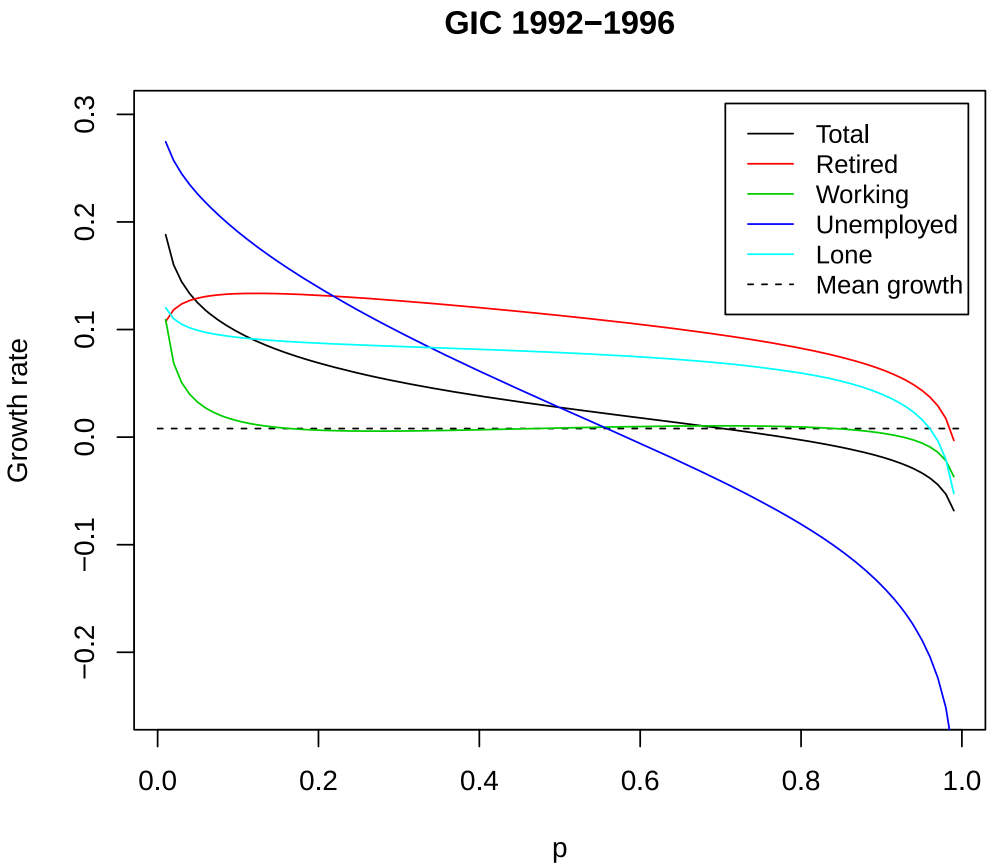


Figure 10: Subgroups GIC for 1979-1988 and 1992-1996

The second period is characterised by a strong equalisation. All curves are slightly downward sloping. Mean growth is much lower than in the first period. Most of the working group just gained the mean. Lone parents and retired people gained more than the mean. The situation of the unemployed is more contrasted. The lower deciles gained more than the mean while the upper deciles experienced a negative growth.

7 Summary and discussion

With this fourth lecture, we entered the core of the topic: the help of Bayesian statistics to measure poverty. Based on a unique modelling of the income distribution by means of a mixture of several lognormal densities, we derive the posterior distribution of three kinds of mathematical beings.

1. Poverty indices. The economic theory devoted a lot of attention to axiomatic building of poverty indices. If they distribution free estimation is quite easy, deriving their small sample distribution is much more complicated. Once we admit a parametric formulation of the income distribution, poverty indices are essentially transformations of the parameters of this parametric distribution. It is then much easier, using simulation techniques to derive their small sample properties.

2. Poverty indices represent each one aspect of poverty: incidence, intensity, inequality. The TIP curve summarises in one graphic three aspects of poverty. Here again, the TIP curve is a transformation of the parameters of the income distribution. However, because the TIP curve can also be analysed in the framework of the Generalised Lorenz curve, another approach can be introduced, which was not developed here: the direct modelling of the Lorenz curve by means of a simplified model such as the one proposed in Kakwani (1980) which, using three parameters provides an interesting flexibility.

3. Finally, the great question is the nature of growth and its impact on poverty. Is growth pro-poor or is it inequality increasing. We all know the trickle-down theory according to which when some people get richer more quickly than the others, people at the bottom of the income distribution can still get some advantage. The GIC allows us to visualise how the benefit of growth are distributed and who is getting what. The UK example has shown us that there is no automaticity in the distribution of growth and that there is the need of a political will to make growth pro-poor.

We have nevertheless not exhausted the topic for several reasons that motivate further developments.

1. We have provided plots and confidence bands, but we did not give the necessary tools to make formal tests. This will be the object of the next chapter.

2. Our presentation was mainly static, even if the GIC provides a certain vision of dynamics. However, the GIC describe the changes in the quantile of the income distribution, but it does not follow the same individual over time. It does not provide an analysis of income dynamics as could do for instance a transition matrix in a Markov process. The GIC we are presented were very simple because they were anonymous.

The non-anonymous GIC (naGIC) was the object of recent developments, but is more difficult to explain. In the last chapter, we shall detail income dynamics and poverty dynamics, but not the naGIC.

Appendix

A Kakwani (1980)'s functional form

The Growth Incidence Curve (GIC) of Ravallion and Chen (2003) can be approximated by the difference between the logs of two quantile functions:

$$g_t(p) = \log Q_t(p|\theta_t) - \log Q_{t-1}(p|\theta_{t-1}). \tag{25}$$

Because the quantile function corresponds to the first derivative of the generalised Lorenz curve, Fourrier-Nicolaï and Lubrano (2021) proposed two alternative ways for finding a parametric formulation for the GIC curve. The first method relies on finding the quantile function associated to a mixture of lognormal distributions. This requires solving (18) as seen above. The second method uses a direct modelling of the Lorenz curve. Several parametric forms were proposed in the literature, using one parameter (Chotikapanich 1993), two parameters (Kakwani and Podder 1973) or three parameters with Villasenor and Arnold (1989) or Kakwani (1980). The latter is built around the Beta density with:

$$L(p|\alpha) = p - \alpha_0 p^{\alpha_1}(1-p)^{\alpha_2}, \tag{26}$$

leading to the quantile function:

$$Q(p|\alpha) = \bar{y} \times (1 - \alpha_0\alpha_1 p^{\alpha_1 - 1}(1-p)^{\alpha_2} + \alpha_0\alpha_2 p^{\alpha_1}(1-p)^{\alpha_2 - 1}).$$

Bayesian inference on the parameters of (26) is obtained by considering the linear regression:

$$\log(p_i - \hat{L}_i) = \log(\alpha_0) + \alpha_1 \log(p_i) + \alpha_2 \log(1-p_i) + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2), \tag{27}$$

where $\hat{L}_i = L(p_i = i/n) = \sum_{j=1}^{i} y[j]/\bar{y}$, $y[j]$ being the order statistics. Obtaining random draws from this quantile function requires some care as:

$$\begin{aligned} Q(p|\alpha^{(j)}, y) &= \bar{y}\exp(u^{(j)}) \times (1 - \exp(\alpha_0^{(j)})\alpha_1^{(j)} p^{\alpha_1^{(j)}-1}(1-p)^{\alpha_2^{(j)}} \\ &\quad + \exp(\alpha_0^{(j)})\alpha_2^{(j)} p^{\alpha_1^{(j)}}(1-p)^{\alpha_2^{(j)}-1}), \quad u^{(j)} \sim N(0, \sigma^{2(j)}). \end{aligned}$$

References

Araar, A., Duclos, J.-Y., Audet, M., and Makdissi, P. (2009). Testing for pro-poorness of growth, with an application to Mexico. *Review of Income and Wealth*, 55(4):853–881.

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

Brent, R. P. (1971). An algorithm with guaranteed convergence for finding a zero of a function. *The Computer Journal*, 14(4):422–425.

Charpentier, A. and Flachaire, E. (2015). Log-transform kernel density estimation of income distribution. *L'Actualité Economique*, 91(1-2):141–159.

Chotikapanich, D. (1993). A comparison of alternative functional forms for the Lorenz curve. *Economics Letters*, 41(2):129 – 138.

Corak, M., Fertig, M., and Tamm, M. (2008). A portrait of child poverty in Germany. *Review of Income and Wealth*, 54(4):547–571.

Davidson, R. (2009). Reliable inference for the Gini index. *Journal of Econometrics*, 150(1):30 – 40.

Duclos, J.-Y. (2009). What is pro-poor? *Social Choice and Welfare*, 32(1):37–58.

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

Fourrier-Nicolaï, E. and Lubrano, M. (2020). Bayesian inference for TIP curves: an application to child poverty in Germany. *The Journal of Economic Inequality*, 18:91–111.

Fourrier-Nicolaï, E. and Lubrano, M. (2021). Bayesian inference for parametric growth incidence curves. In Bandyopadhyay, S., editor, *Research on Economic Inequality*, volume 29. Emerald Group Publishing Limited.

Gastwirth, J. (1971). A general definition of the Lorenz curve. *Econometrica*, 39(6):1037–1039.

Jawitz, J. W. (2004). Moments of truncated continuous univariate distributions. *Advances in Water Resources*, 27(3):269–281.

Jenkins, S. P. (1995). Accounting for inequality trends: Decomposition analyses for the UK, 1971- 86. *Economica*, 62(245):29–63.

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.

Kakwani, N. (1980). On a class of poverty measures. *Econometrica*, 48(2):437–446.

Kakwani, N. and Pernia, E. (2000). What is pro-poor growth? *Asian Development Review*, 18:1–16.

Kakwani, N. and Podder, N. (1973). On the estimation of Lorenz curves from grouped observations. *International Economic Review*, 14:278–292.

Lubrano, M. and Ndoye, A. A. J. (2016). Income inequality decomposition using a finite mixture of log-normal distributions: A Bayesian approach. *Computational Statistics and Data Analysis*, 100:830 – 846.

Muller, C. (2001). The properties of the Watts poverty index under lognormality. *Economics Bulletin*, 9(1):1–9.

Ravallion, M. and Chen, S. (2003). Measuring pro-poor growth. *Economics Letters*, 78:93–99.

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

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

Villasenor, J. A. and Arnold, B. C. (1989). Elliptical Lorenz curves. *Journal of Econometrics*, 40:327–338.

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

Yang, S.-S. (1985). A smooth nonparametric estimator of a quantile function. *Journal of the American Statistical Association*, 80(392):1004–1011.

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