école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

Bayesian Inference for Parametric Growth Incidence Curves

Edwin Fourrier-Nicolai
Michel Lubrano

WP 2021 - Nr 31

Bayesian Inference for Parametric Growth Incidence Curves*

Edwin Fourrier-Nicolai† Michel Lubrano‡

March 2021

Abstract

The growth incidence curve of Ravallion and Chen (2003) is based on the quantile function. Its distribution-free estimator behaves erratically with usual sample sizes leading to problems in the tails. We propose a series of parametric models in a Bayesian framework. A first solution consists in modelling the underlying income distribution using simple densities for which the quantile function has a closed analytical form. This solution is extended by considering a mixture model for the underlying income distribution. However in this case, the quantile function is semi-explicit and has to be evaluated numerically. The alternative solution consists in adjusting directly a functional form for the Lorenz curve and deriving its first order derivative to find the corresponding quantile function. We compare these models first by Monte Carlo simulations and second by using UK data from the Family Expenditure Survey where we devote a particular attention to the analysis of subgroups.

Keywords: Bayesian inference, growth incidence curve, inequality.
JEL codes: C11, D31, I31

---

*We would like to thanks the editor and two anonymous referees for useful comments and remarks. We should also mention nice conversations with Jeff Racine and Karim Abadir. Of course remaining errors are solely ours. This work was supported by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A*MIDEX.

†Aix Marseille Univ, CNRS, AMSE, Marseille, France and Toulouse School of Economics, Université Toulouse Capitole, Toulouse, France

‡School of Economics, Jiangxi University of Finance and Economics, China and Aix Marseille Univ, CNRS, AMSE, Marseille, France

1 Introduction

Since the influential work of Kuznets (1955), considerable efforts have been devoted to comparing how economic growth is distributed among the population. In this attempt, Ravallion and Chen (2003) have introduced the growth incidence curve (GIC) to illustrate how growth is distributed over the quantiles of an income distribution. More precisely, they propose to compute the mean growth rate of every quantile of a base-year distribution. In doing so, the growth incidence curve allows to gauge easily whether growth has been favourable to the poor or whether growth has been inequality-enhancing. Particularly, we can say that growth is pro-poor if the changes in the income of the poor quantiles is greater than the growth rate of the mean (or median) income. A growing applied literature has been making use of the GIC. Examples are the world income distribution with Lakner and Milanovic (2016) or the income distribution for emerging countries with Chancel and Piketty (2019) for India or Novokmet et al. (2018) for Russia. However, what can we conclude from a growth incidence curve and how to compare them? Since the growth incidence curve is based on the quantile function, it is closely related to stochastic dominance at the first-order (Duclos 2009) paving the way for robust welfare comparisons. From a statistical perspective, Araar et al. (2009) develop inference and tests for the growth incidence curve in a distribution-free framework. In doing so, they have no specific assumption for the shape of the income distribution.

In this paper, we show that this distribution-free estimator causes the growth incidence curve to behave erratically with usual sample sizes resulting in a spurious statistical inference in the tails of the distribution. This is of a particular concern when studying poverty and inequality. We make use of the relationship between the growth incidence curve, the quantile function and the Lorenz curve to propose a series of parametric forms for growth incidence curves within a Bayesian framework. A *first solution* consists in modelling the underlying income distribution using simple densities for which the quantile function has a closed analytical form. This solution is extended by considering a mixture model for the underlying income distribution. However in this case, there is no analytical form for the quantile function and numerical optimization has to be introduced. An *alternative solution* consists in adjusting directly a functional form for the Lorenz curve, choosing along the long list proposed in the literature (see e.g. the survey of Sarabia 2008). We provide a variant of this option which consists in adjusting a parametric Lorenz curve to the data and then to compute its first order derivative for obtaining the quantile function. Confidence intervals and statistical tests for the growth incidence curve can be obtained easily within a Bayesian framework as the direct product of a MCMC output. We compare these proposed models by Monte Carlo simulations and on UK data from the Family Expenditure Survey (FES). We devote a particular attention to the analysis of subgroups. Advantages of the proposed approach are twofold. First, by imposing a structure on the data, parametric models compensate for the small number of observations when analysing subgroups. Second, Bayesian inference leads to simple tests of dominance.

The paper is organized as follows. In section 2, we present the growth incidence curve, its relation to other measures of poverty and inequality. We present its estimation in a distribution-free framework and show the counter-intuitive variability of this estimator. In section 3, we propose a parametric growth incidence curve based on parametric assumptions of the income distribution. We particularly consider the cases of the simple log-normal assumption and that of a mixture of log-normal densities. In section 4, we propose an alternative approach based on a direct parametric modelling of the Lorenz curve, building on existing results of the literature for which we derive the corresponding quantile function. In section 5, we compare these parametric models by examining their implied quantile and GIC functions using a Monte Carlo simulation. In section 6, we illustrate our method to real data from the FES for the United-Kingdom between 1979 and 1996, a period over which the income distribution is known to have experienced many changes. The last section concludes.

2 Measuring pro-poor growth

In this section, we present the growth incidence curve, discuss its relationship with alternative measures of poverty and inequality. Last, we show how its distribution-free estimator causes the growth incidence curve to behave erratically.

2.1 The growth incidence curve

Consider two distributions of incomes observed at time $t-1$ and t, and characterized by the respective cumulative distribution functions F_{t-1} and F_t with support contained in the non-negative real line. We refer to income throughout the paper to signify a measure of individual welfare, which need not to be money income. The growth incidence curve (GIC), introduced in Ravallion and Chen (2003), measures the growth rate of the p-quantile for every p:

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

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 characterizing 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. $g_t(p) = \gamma$ for all p. 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 generalized Lorenz curve .

2.2 Pro-poor judgements using the growth incidence curve

When can we say that distributional changes are favourable to the poor? As a matter of facts, welfare judgements are straightforward when using growth incidence curves as they are based on quantile functions. Because, first-order stochastic dominance of F_{t-1} by F_t implies that $Q_{t-1}(p) \leq Q_t(p)$ for every p (Davidson and Duclos 2000), it follows directly from equation (1) that first-order stochastic dominance of F_{t-1} by F_t is equivalent to:

$$g_t(p) > 0, \quad \forall p \in [0,1]. \tag{2}$$

Graphically, first-order stochastic dominance of the final period over the initial period is verified if and only if the growth incidence curve is positive for every quantile.

Let us now consider a common poverty line z. We can say that growth has decreased poverty if $g_t(p) > 0$ for all $p < F_{t-1}(z)$ where $F_{t-1}(z)$ 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 $p < F_{t-1}(z)$, then we can say that growth has been detrimental for the poor so as the income of the poor have reduced in absolute terms between the two periods. If the growth incidence curve is negative only for some values of $p \leq F_{t-1}(z)$, then we cannot conclude unambiguously about whether the distributional changes have been welfare-improving or not for the poor.

The condition of first-order stochastic dominance essentially says if the income of the poor has increased in absolute terms. However, 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:

$$g_t(p) > \gamma_t, \quad \forall p \in [0, F_{t-1}(z)]. \tag{3}$$

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 greater pace than the average growth.

2.3 Distribution-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:

$$\hat{Q}(p_i = i/n) = y_{(i)}, \tag{4}$$

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

$$\hat{L}(p_i = i/n) = \sum_{j=1}^{i} y_{(j)} / \sum_{j=1}^{n} y_{(j)}. \tag{5}$$

The generalized 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):

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

where K is a kernel density and h the window size to be determined. However this estimator has the same variance as the natural estimator. If $\hat{Q}(p)$ is the estimated quantile and $Q(p)$ its true value, then the asymptotic distribution of the natural estimator is normal with:

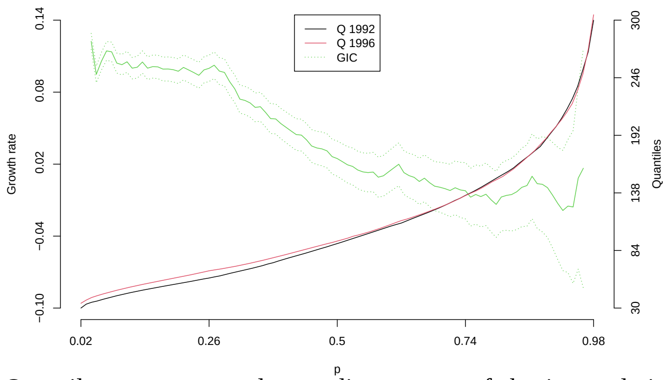
$$\sqrt{n}(\hat{Q}(p) - Q(p)) \sim \mathrm{N}\left(0, \frac{p(1-p)}{f^2(Q(p))}\right), \quad (6)$$

where f is the true density function.[1] The variance is a function of p and of the shape of f. It has a symmetric U-shape in p if f is a symmetric distribution with a minimum at $p = 0.50$. If f is an asymmetric distribution like the Gamma or the log-normal, the position of the minimum is shifted to the left. But whatever f, the variance is always greater in the tails.

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

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

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



Quantiles are computed according to one of the interpolation methods described in Hyndman and Fan (1996). Standard errors are computed using (6) and the Delta method for logs. Dots correspond to a 90% confidence interval.

Figure 1: Quantile functions and GIC, FES data 1992-1996

is a particular concern when studying poverty or inequality. This causes the GIC to behave erratically as illustrated in Figure 1. Particularly, we estimate a distribution-free GIC based on empirical data (presented in section 6) 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.

Son (2004) has proposed an alternative measure for assessing distributional changes based on the difference between two generalized Lorenz curves (instead of their derivatives), namely

[1]We are grateful to Jeffrey S. Racine for providing us this reference.

the poverty growth curve (PGC). The distribution-free empirical PGC is defined as:

$$\hat{G}_t(p_i) = \log \hat{G}L_t(p_i) - \log \hat{G}L_{t-1}(p_i).$$

It turns out that the empirical PGC, as displayed in Figure 2, provides a much smoother shape. Intuitively, this stems from the fact that the PGC is based on partial moments and is therefore

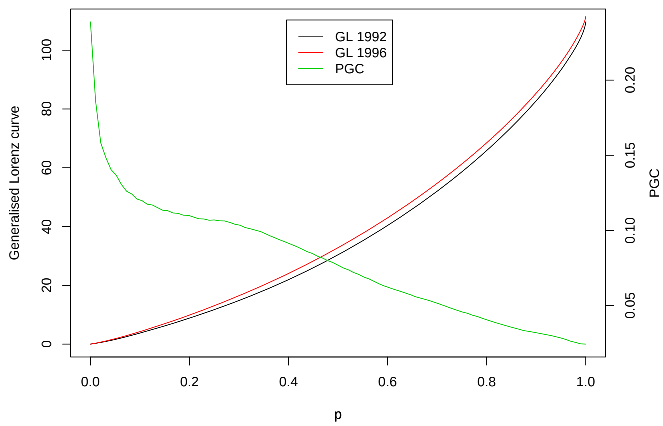


Figure 2: Generalised Lorenz curves and PGC, FES data 1992-1996

more constrained than the quantile function. However, the PGC considers whether the poor enjoy the benefits of growth more than the others and not the growth rate of the p-quantile per se. Because the growth incidence curve is of particular interest in itself, we propose in this paper a series of parametric growth incidence curves in order to get a more precise inference for the GIC.

3 An approach based on parametric assumptions of the income distribution

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 models and thus for the growth incidence curves.

3.1 The growth incidence curve for the log-normal model

The literature on poverty and inequality has for long considered mainly the log-normal assumption to model the income distribution.[2] The log-normal probability density function is defined such as:

$$f_\Lambda(y|\mu,\sigma^2) = \frac{1}{y\sigma\sqrt{2\pi}} \exp -\frac{(\log(y)-\mu)^2}{2\sigma^2}, \tag{7}$$

with mean:

$$\mathrm{E}(y) = \exp(\mu + \sigma^2/2). \tag{8}$$

The main advantage of this distribution is that we have both an analytical expression for its cumulative distribution function:

$$F_\Lambda = \Phi\left(\frac{\log x - \mu}{\sigma}\right), \tag{9}$$

and for its quantile function:

$$Q(p) = \exp(\mu + \sigma\Phi^{-1}(p)), \tag{10}$$

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:

$$g_t(p) \simeq (\mu_2 - \mu_1) + (\sigma_2 - \sigma_1)\Phi^{-1}(p). \tag{11}$$

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 $\sigma_2 - \sigma_1$.

3.2 The growth incidence curve for mixtures of log-normal densities

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 constrained GIC curve given in (11). A finite mixture of log-normal densities $f(y|\theta)$ is a linear combination of K log-normal densities $f_{\Lambda k}(y|\mu,\sigma^2)$ such that:

$$f(y|\theta) = \sum_{k=1}^{K} \eta_k f_{\Lambda k}(y|\mu_k, \sigma_k^2), \qquad 0 \le \eta_k < 1, \quad \sum_{k=1}^{K} \eta_k = 1, \tag{12}$$

[2] See, for instance, Aitchison and Brown (1957) or Anderson et al. (2014) in the context of a mixture of log-normal densities.

where the η_k are the weights of the mixture, $\theta = (\eta, \mu, \sigma^2)$ represents the collection of all the parameters of the mixture. The expectation of y is found by linear property of the mixture:

$$\mathrm{E}(y|\theta) = \sum_{k=1}^{K} \eta_k \exp(\mu_k + \sigma_k^2/2). \tag{13}$$

The cumulative distribution is also found easily by the linear property of the mixture. It is the weighted sum of the components' CDF:

$$F(y|\theta) = \sum_{k=1}^{K} \eta_k F_{\Lambda k}(y|\mu_k, \sigma_k^2). \tag{14}$$

If it is easy to derive the CDF of a mixture, it is however more difficult to derive the the corresponding quantile function as it 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:

$$p = F(q|\theta). \tag{15}$$

The operation is very quick and secure because the dimension of q is one. Its range is fully determined by the range of y so that Brent (1973) algorithm is the natural and efficient solution to this problem.[3] Substituting the computed quantiles into equation (1), we get the growth incidence curve for a mixture of K log-normal densities, without further analytical result.

3.3 Bayesian inference and tests for the GIC

The log-normal model is a simple transformation of the ordinary linear regression. Bayesian inference for this model is recalled in Appendix A which displays useful results for the next sections. Bayesian inference for mixtures of log-normal densities relies on a Gibbs sampler. We can find in Lubrano and Ndoye (2016) or Fourrier-Nicolai and Lubrano (2020) its detailed implementation which is recalled in Appendix B together with the prior densities used. Note however that 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.

For both choices, the important point is to be able to get m random draws for the posterior densities of the parameters of the income distribution for $t = 1, 2$. For the log-normal model, these draws can be directly converted into draws of the quantile function (10) for a predefined grid of np values of p between 0 and 1. When the income distribution is modelled using a mixture of K log-normal densities, the case is slightly more complex as there is no analytical formula corresponding to (10). For each point of the grid p and for each draw, we have to solve numerically in q equation (15) so as to obtain a random draw for the quantile function $Q(p|\theta)$ for the two periods. Then, applying formula (1) for each draw, we summarize the posterior density of the GIC in a matrix of m rows and np columns. A generic line j of this matrix is noted $g_t^{(j)}(p)$.

[3]The R function `uniroot` searches a predefined interval using the Brent algorithm without derivatives.

The posterior GIC can be estimated by taking the mean over j for each point p of the grid; that is, taking the mean over each column of the matrix. Confidence intervals at 90% for the GIC can be obtained by taking the 0.05 and 0.95 quantiles for each column.

Several hypotheses can be easily tested as proposed in Fourrier-Nicolai and Lubrano (2020). First, we can test if growth has been welfare-improving in terms of first-order stochastic dominance i.e. $g_t(p) > 0$ for all p. For each point p of the grid, we evaluate the following probability:

$$\Pr(g_t(p) > 0) \simeq \frac{1}{m} \sum_{j=1}^{m} \mathbb{I}(g_t^{(j)}(p) > 0), \tag{16}$$

where $\mathbb{I}(.)$ is the indicator function which equals one if the condition $g_t(p) > 0$ is true and zero otherwise. So the evaluation relies on a sampling mean. Second, we can test whether growth has been favourable to the poor i.e. $g_t(p) > \gamma_t$ for all $p \leq F(z)$. We have m draws of $g_t^{(j)}(p)$ to be compared to m draws of $g^{(j)}$ computed using either (8) or (13). We evaluate the following probability:

$$\Pr(g_t(p) > \gamma_t) \simeq \frac{1}{m} \sum_{j=1}^{m} \mathbb{I}(g_t^{(j)}(p) - \gamma_t^{(j)} > 0). \tag{17}$$

Note that because γ_t appears also in the definition of $g_t(p)$, we have to reorder randomly the draws when evaluating $g_t^{(j)}(p)$ and $\gamma_t^{(j)}$.

4 A direct approach based on parametric Lorenz curves

Instead of modelling the income distribution and then deriving the associated growth incidence curve, we can exploit the relationship between the Lorenz curve and the quantile function in order to propose a new class of parametric growth incidence curves. Particularly, a vast set of functional parametric Lorenz curves have been proposed in the literature, see Sarabia (2008) for a survey. In this section, we present this class of parametric Lorenz curves, then we present a survey of the existing functional forms, and finally we derive the growth incidence curve for a sample of these functional forms.

4.1 Alternative parametric models for Lorenz curves

We discussed how parametric Lorenz curves can be derived from an underlying parametric modelling of the income distribution. Even if simple models like the log-normal distribution have simple corresponding quantile function, considering richer models for the income distribution necessitates an heavier technology for finding the quantile function and then to express the GIC. Instead of modelling the underlying income distribution and then deriving the associated GIC, we can exploit the relationship between the Lorenz curve and the quantile function in order to propose a new class of parametric growth incidence curves. Therefore, instead of expressing a Lorenz curve in terms of the parameters of an income distribution, we can consider directly a parametric model for the Lorenz curve itself provided the later fulfils certain properties. In order to be a Lorenz curve, a parametric form has to verify some properties detailed in Sarabia et al. (1999) or in Krause (2014) for instance. In addition, Sarabia et al. (1999) provide conditions under which some transformations of a Lorenz curve $L(p)$ are still Lorenz curves, namely $p^\alpha L(p)$ with $\alpha \geq 0$ and $L(p)^\gamma$ with $\gamma \geq 1$. This opens the way to the generation of new class of functional forms. This approach was used by Sarabia et al. (1999) to build a hierarchical family of Lorenz curves around the Lorenz curve of the Pareto distribution:

$$L(p|\alpha) = 1 - (1-p)^\alpha.$$

In fact, any other initial one-parameter Lorenz curve could have been chosen, for instance that corresponding to the log-normal distribution or the Weibull distribution, see Sarabia (2008) for a survey. Basically, these functional forms are estimated on transformations of the data which are provided by the empirical distribution-free estimate of the Lorenz curve given in (5). This class of functional Lorenz curves paves the way for correcting the lack of flexibility of Lorenz curves associated to usual distributions. In some particular cases, the link between the empirical distribution-free estimate of the Lorenz curve and the parametric Lorenz curve can be established by mean of a simple linear regression. In the general case, a non-linear regression has to be used.

4.2 Three usual functional forms

Functional forms can be classified according to their number of parameters, to which corresponds an increased flexibility. Beyond the one-parameter family of Chotikapanich (1993), Kakwani and Podder (1973) were the first to propose a two-parameters functional form with:

$$L(p|\alpha, \beta) = p^\alpha \exp(-\beta(1-p)), \qquad \alpha > 1,\ \beta > 0.$$

Parameters of this functional form can be estimated by ordinary least squares using:

$$\log(\hat{L}_i) = \alpha \log(p_i) - \beta(1-p_i) + \epsilon_i, \tag{18}$$

where ϵ_i is a Gaussian noise with zero mean and variance σ^2 and $\hat{L}$ the natural estimate of the Lorenz curve. Although, this functional form provides a better adjustment than that of Chotikapanich (1993) for instance, there is a systematic bias for higher quantiles which is of a particular concern when studying inequality. Thus, we hope that a functional form with more parameters might correct for this bias.

In his paper on poverty indices, Kakwani (1980) incidentally makes use of a three-parameters functional form for the Lorenz curve which is based on a variation of the Beta density:

$$L(p|\alpha_0, \alpha_1, \alpha_2) = p - \alpha_0 p^{\alpha_1}(1-p)^{\alpha_2}, \qquad \alpha_0 > 0, \quad 0 < \alpha_1, \quad \alpha_2 \leq 1.$$

It takes advantage of the fact that the support of the Beta density is [0,1]. The two conditions $0 < \alpha_1$ and $\alpha_2 \leq 1$ are sufficient for $L(p|\alpha)$ to be convex with respect to the p-axis. Kakwani (1980) proposes to estimate the parameters of this functional form by ordinary least squares in a linear regression obtained after taking the logs:

$$\log(p_i - \hat{L}_i) = \log(\alpha_0) + \alpha_1 \log(p_i) + \alpha_2 \log(1-p_i) + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma^2). \tag{19}$$

This functional form is not often reported in surveys. For instance, it is neither mentioned in Sarabia (2008) nor in Krause (2014), but it can be found in Darvas (2016) where it is the recommended choice.

The starting point of Villasenor and Arnold (1989) is to remark that the portion of an ellipse can be a good candidate for fitting a Lorenz curve. When using three parameters, the equation of an ellipse in (p, L) with parameters a, b, d can be:

$$ap^2 + bpL + L^2 + dp - (a + b + d + 1)L = 0,$$

which when solved in L leads to:

$$L(p|\alpha, \beta, \delta) = \frac{1}{2}\left[(a - \beta p) - \sqrt{a^2 + bp + cp^2}\right],$$

where:

$$a = \alpha + \beta + \delta + 1, \quad b = -2a\beta - 4\delta, \quad c = \beta^2 - 4\alpha.$$

The following conditions need to be imposed $\alpha + \delta < 1$ and $\delta \geq 0$ so as to have an ellipse. The parameters can be estimated using the linear regression (without a constant term):

$$\hat{L}_i(1 - \hat{L}_i) = \alpha(p_i^2 - \hat{L}_i) + \beta\hat{L}_i(p_i - 1) + \delta(p_i - \hat{L}_i) + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma^2). \tag{20}$$

This functional form is one of the models favoured in Datt and Ravallion (1992) and it has been used extensively by the World Bank for the quality of its adjustment.

4.3 Functional forms and implied quantile functions

The quantile function corresponds to $Q(p) = \bar{y}L'(p)$, thus it is obtained by multiplying the mean income by the derivative of the Lorenz curve. Then, we can get the implied quantile function (for a given mean income) for any functional form mentioned before by differentiating its Lorenz curve.

The quantile function associated to the Kakwani and Podder (1973)'s Lorenz curve is:

$$Q(p|\alpha, \beta, \bar{y}) = \bar{y} \times \frac{\alpha\, p^{\alpha-1} + p^{\alpha}\, \beta}{\exp(\beta(1-p))}. \tag{21}$$

The popular functional form proposed by Villasenor and Arnold (1989) leads to the following quantile function:

$$Q(p|\alpha, \beta, \delta, \bar{y}) = \bar{y} \times \frac{1}{2}\left[-\beta - \frac{b + 2\,c\,p}{2\sqrt{a^2 + b\,p + c\,p^2}}\right], \tag{22}$$

whereas, the Lorenz curve introduced in Kakwani (1980) leads to:

$$Q(p|\alpha_0, \alpha_1, \alpha_2, \bar{y}) = \bar{y} \times (1 - \alpha_0\alpha_1 p^{\alpha_1 - 1}(1-p)^{\alpha_2} + \alpha_0\alpha_2 p^{\alpha_1}(1-p)^{\alpha_2 - 1}). \tag{23}$$

Substituting these associated quantile functions into equation (1), we obtain a new class of parametric growth incidence curves.

4.4 Bayesian inference for the GIC

For Bayesian inference, we have first to run one of the linear regressions (18), (20) or (19) so as to obtain m draws from the regression coefficient say α and for the variance of the error term σ^2. Then, for each draw and each of the np points of a predefined grid for p, we simulate an error term $u^{(j)}$ from a Gaussian with zero mean and standard deviation $\sigma^{(j)}$ so as to obtain draws from the posterior quantile function. If we take the example of the quantile function (23) derived from Kakwani (1980)'s model, we get:

$$\begin{aligned} Q(p|\alpha^{(j)}, y) &= \bar{y} \exp(u^{(j)}) \times (1 - \exp(\alpha_0^{(j)}) \alpha_1^{(j)} p^{\alpha_1^{(j)}-1} (1-p)^{\alpha_2^{(j)}} \\ &\quad + \exp(\alpha_0^{(j)}) \alpha_2^{(j)} p^{\alpha_1^{(j)}} (1-p)^{\alpha_2^{(j)}-1}) \end{aligned} \tag{24}$$

Let us now consider two periods with two vectors of observations y_1 and y_2. The growth incidence curve is obtained with the following algorithm:

1. Choose a grid of np points for p.
2. Compute the sufficient statistics for the two samples $y_i, i = 1, 2$ as well as the sample means $\bar{y}_1$, $\bar{y}_2$.
3. For $j = 1, ..., m$:
   (a) Draw $\alpha_i^{(j)} \sim f_t(\alpha|y_i), i = 1, 2$ and $\sigma_i^{2(j)} \sim f_{IG}(\sigma^2|y_i), i = 1, 2$
   (b) Compute $Q(p|\alpha_i^{(j)}, \sigma_i^{2(j)}, y_i), i = 1, 2$
   (c) Evaluate $g^{(j)} = \log Q(p|\alpha_2^{(j)}, \sigma_2^{2(j)}, y_2) - \log Q(p|\alpha_1^{(j)}, \sigma_1^{2(j)}, y_1)$
   (d) Store the vector $g(p)^{(j)}$.
4. End loop

This algorithm produces the same $m \times np$ matrix of posterior draws for the GIC as in the case of a direct modelling of the income distribution. Computations to determine posterior probability of pro-poor growth follow the same logic as in section 3.3.

5 Monte Carlo simulation

What is the best parametric model for estimating a GIC among the various propositions we made? The answer to this type of question is usually provided by a Monte Carlo experiment. However, as a Monte Carlo experiment is very dependant on the generating process, we have selected two different generating processes. We first calibrated a GB2 distribution, using the parameter values estimated on UK income data as reported in Jenkins (2009) for 1994 and for 2004.[4] The second

[4] The estimated parameters reported in Jenkins (2009) are for 1994 $a = 2.994$, $b = 227.840$, $p = 1.063$, $q = 1.015$ and for 2004 $a = 4.257$, $b = 341.965$, $p = 0.682$, $q = 0.635$ in the usual notations of the GB2. We used the `R` package `GB2`.

process is a mixture of three log-normals, trying to reproduce the quantiles of the previous GB2.[5] The two corresponding GICs remind clearly the Elephant curve of Alvaredo et al. (2018) as can be seen in Figure 3, but the mixture of three log-normals leads to a slightly more complex shape. We confront the distribution free method, the log-normal model, the mixture of two log-normal

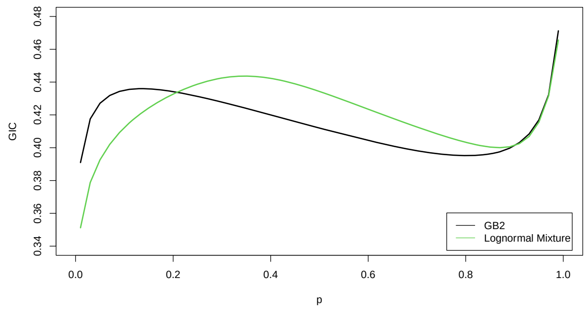


Figure 3: Implied GIC in the Monte Carlo experiment

distributions, and the three functional forms reviewed in section 4, i.e. Kakwani and Podder (1973), Kakwani (1980) and Villasenor and Arnold (1989). Distances to the theoretical values are measured with a root mean square error (RMSE) computed over the $np = 50$ points of a predefined grid for $p \in [0.01, 0.99]$. We used 1000 replications and sample sizes of 1 000, 5 000 and 10 000.

Table 1: Efficiency of different parametric GIC compared to the distribution free GIC

Sampling	NP	LN	MX LN	KP1973	K1980	VA1989
Sample size 1 000						
GB2	0.0488	**0.0370**	0.0498	0.0452	0.0411	0.0394
Mixture 3 LN	0.0526	**0.0406**	0.0474	0.0431	0.0487	0.0584
Sample size 5 000						
GB2	0.0215	0.0220	0.0377	0.0251	**0.0198**	0.0205
Mixture 3 LN	0.0236	0.0259	0.0254	**0.0229**	0.0278	0.0316
Sample size 10 000						
GB2	0.0155	0.0193	0.0322	0.0211	**0.0150**	0.0170
Mixture 3 LN	**0.0168**	0.0231	0.0219	0.0188	0.0242	0.0269

NP stands for the distribution free method, LN is the log-normal model, MX LN a mixture of two log-normals, KP1973 is Kakwani and Podder (1973), K1980 corresponds to Kakwani (1980) and VA1989 to Villasenor and Arnold (1989). Figures indicate average RMSE between the estimated GIC and its theoretical value implied by the generating process on a grid of 50 points between 0.01 and 0.99. Bold figures indicate the best result. 1 000 replications.

[5]The weights are $\eta = (0.25, 0.60, 0.15)$. For the first period $\mu_1 = (5.562412, 4.850735, 5.822487)$, $\sigma_1 = (0.3641002, 0.7003314, 0.6647863)$. For the second period $\mu_2 = (5.946083, 5.367303, 5.924789)$, $\sigma_2 = (0.3338858, 0.7696569, 0.8304167)$.

The first result that comes out of Table 1 is that in relatively small samples the distribution free estimate is always dominated by the simple but however biased log-normal model, illustrating the trade-off between bias and variance. However, this supremacy ceases as soon as the sample size increases. The distribution free estimator becomes the best for the more complex generating process in very large samples. As soon as the sample size increases, Kakwani (1980)'s model starts to be the best model, provided the sample is not too complex. Surprisingly, the mixture of two log-normals is not at ease in this exercise, probably because of the limited number of components and of potential identification issues. We shall see that with real data the mixture model is much more at ease to mimic the distribution free estimate.

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 neo-liberal 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), normalized 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 2, 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.[6] 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.

[6] 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.

Table 2: 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.

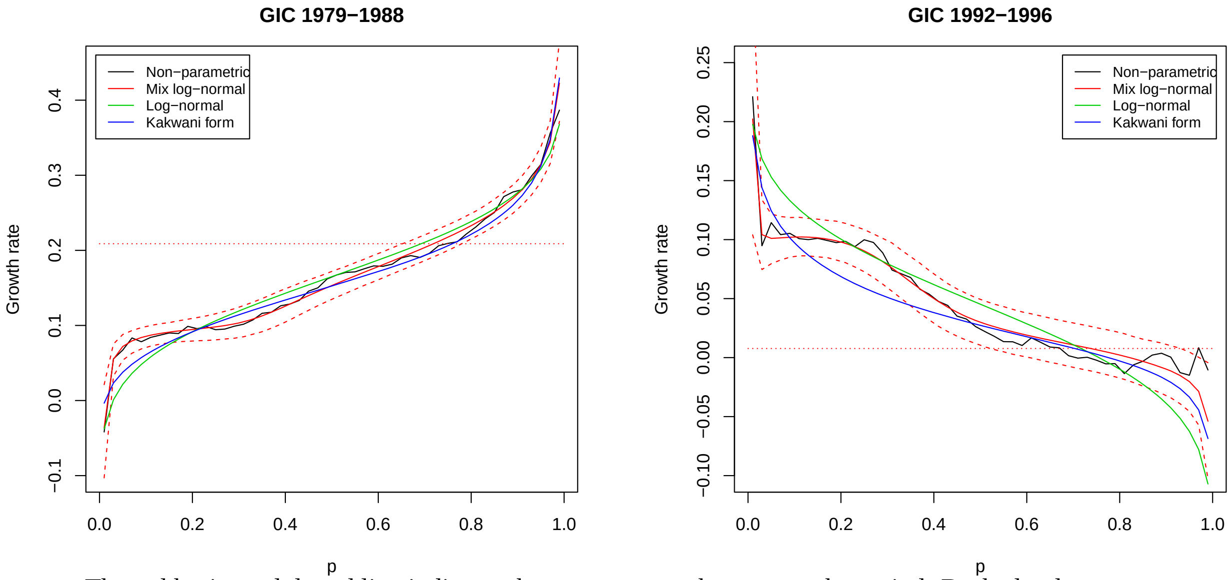
Table 2 provides a partial portray of the evolution of poverty and inequality although these results might be sensitive to the choice of the measure of inequality but also to the choice of the poverty line. The growth incidence curve provides a more complete portray of the evolution of the income distribution over time which is not sensitive to the choice of the measures of poverty and inequality as it is related to first-order stochastic dominance.

6.2 Growth incidence curves for the whole population

Our Monte Carlo experiment has shown that parametric GIC curves were giving better results than a distribution-free approach for moderate to large sample sizes like the ones we have here (slightly more than 6 000 observations). Figure 4 confronts the performance of three of our 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.[7] 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. Our Monte Carlo experiment has shown that this model should have a good precision in large samples. 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. The first period 1979-1988 is clearly characterized by an anti-poor and inequality-increasing growth while the second period 1992-1996 becomes pro-poor and inequality-reducing. For the first period, we must go up to the 0.70^{th} quantile in order to have an average growth rate of income greater than the average

[7] A prior was devised as follows (see Appendix B for details). 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 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 minimize the Bayesian information criterion (BIC).



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 4: Growth incidence curves for 1979-1988 and 1992-1996

growth. During the second period, the situation is reversed. Households up to the 0.70^{th} quantile experienced an income increase greater than average.

Let us now examine which method gives the better fit as compared to the distribution-free estimate. The answer depends on which period is considered. For 1979-1988, the confidence interval of the mixture is quite narrow and it contains whole of the non-parametric estimate. The log-normal and Kakwani (1980) models are within this confidence interval, except for the first decile. But the four curves display the same message where the last two deciles have gained more from growth, while the first decile has gained nothing. The situation is slightly different for the period 1992-1996. The confidence interval for the mixture model is now larger, but still contains the non-parametric estimate. Kakwani (1980)'s functional form is within this confidence interval, except for deciles between 0.20 and 0.30. The log-normal form does not match both ends and is less at ease. But all of those forms again deliver the same message: deciles below the median have gained while deciles over 0.60 have lost in absolute terms. Formal probability tests will deliver a more precise message.

Table 3 provides probabilities of pro-poor growth computed using (17) for each decile for the periods 1979-1988 and 1992-1996, using three different parametric models. We keep in mind that the mixture models should be the closest one to the distribution free approach as it corresponds to a semi-parametric approach and thus represent the case with minimum bias. The three models deliver the same probability of anti-poor growth up to the 0.60^{th} quantile and the same probability of pro-rich growth from the 0.90^{th} quantile for 1979-1988. There are however differences for the 0.70^{th} quantile where the log-normal model over-evaluates the probability that this quantile has benefited more from growth, nearly crossing the mythical value of 0.50. When we consider the second period (1992-1996), Kakwani (1980)'s form is fairly

Table 3: Probability of pro-poor growth, whole sample

p	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
					1979-1988				
Log-normal	0.00	0.00	0.00	0.00	0.00	0.04	0.48	0.97	1.00
Kakwani	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.75	1.00
Mixture	0.00	0.00	0.00	0.00	0.00	0.01	0.29	0.93	1.00
					1992-1996				
Log-normal	1.00	1.00	1.00	1.00	0.99	0.93	0.64	0.18	0.00
Kakwani	1.00	1.00	1.00	0.97	0.91	0.76	0.53	0.28	0.06
Mixture	1.00	1.00	1.00	0.99	0.93	0.77	0.58	0.39	0.19

well in accordance with the mixture model. The log-normal model provides slightly different probabilities for quantiles greater than 0.60, but do not contradict the essential message when 0.50 is taken as the reference to decide if growth is pro-poor or not. In summary, the extra parameter of the Kakwani (1980)'s model allows to provide more reliable conclusions than the simple log-normal model.

Table 3 can also be read in more restrictive way. Pro-poor growth is defined in (3) as the probability that the lower quantiles increase more than average up to the quantile defined by the poverty rate in the first period. Poverty rates for 1979 and 1992 are respectively 0.135 and 0.196, using the poverty line defined as 60% of the median income. This implies that Table 3 can be limited to the first two columns to decide if growth was pro-poor or not. The answer becomes unambiguous for all models.

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: households where the head is retired, households where both heads are working, households where both heads are unemployed. In addition, we also consider lone parenthood, that is single headed families with children.

Figure 5 displays 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.

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.

The second period is characterized by a strong equalization. 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

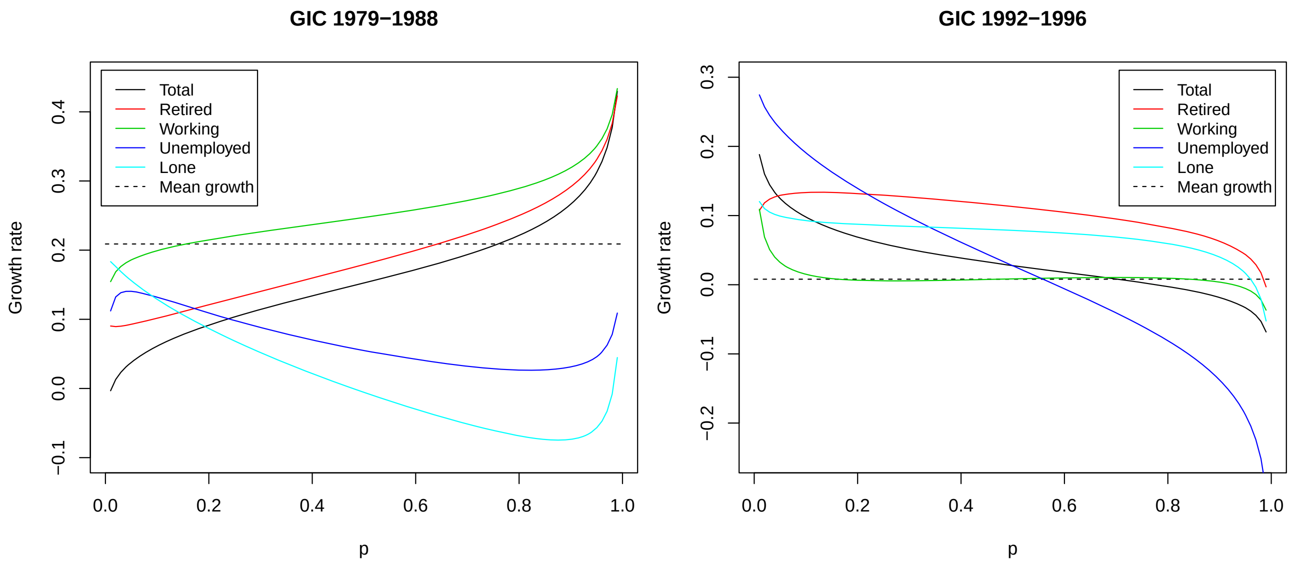


Figure 5: Subgroups GIC for 1979-1988 and 1992-1996

of the unemployed is more contrasted. The lower deciles gained more than the mean while the upper deciles experienced a negative growth.

Table 4: Probability of pro-poor growth for 1979-1988 using the Kakwani (1980)'s GIC

p	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
					1979-1988				
Retired	0.00	0.00	0.00	0.00	0.03	0.26	0.78	0.99	1.00
Working	0.24	0.66	0.89	0.97	0.99	1.00	1.00	1.00	1.00
Unemployed	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lone parents	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
					1992-1996				
Retired	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Working	0.67	0.47	0.45	0.47	0.51	0.54	0.56	0.54	0.43
Unemployed	1.00	1.00	0.99	0.95	0.74	0.35	0.09	0.01	0.00
Lone parents	1.00	0.99	0.99	0.99	0.99	0.98	0.97	0.95	0.86

We complete this picture by computing pro-poor growth probabilities as reported in Table 4. During the first period, retired people gained more than average after the 0.70^{th} decile while the working group gained more than average starting from the 0.20^{th} decile. Unemployed and lone parents gained less than average for all deciles. During the second period, all the deciles gained more than average for retired and lone parents. Unemployed gained more than average up to the median. The probabilities for the working group are un-conclusive.

7 Summary and discussion

The growth incidence curve of Ravallion and Chen (2003) has been introduced to illustrate how economic growth was distributed over the quantiles of an income distribution. In doing so, it allows to gauge whether growth has been favourable to the poor (the lower quantiles) or whether growth has been inequality-increasing, favouring the upper quantiles. With a Bayesian approach, we were able to provide simple confidence intervals for the GIC and dominance tests which are valid in small samples. This is particularly useful when analysing subgroups, where by definition the sample size is reduced. Bayesian inference relies most of the time on parametric models. This can be seen as a restriction, compared to a distribution-free approach. We have managed to show that this was not the case here. Parametric models were able to reduce the variability of the distribution-free GIC, without introducing a major distortion. Indeed, results of the Monte Carlo experiment suggest that even very crude parametric models, when applied for the GIC were more efficient than the distribution-free estimator, except when the sample size was very large (more than 10 000 observations). The empirical application demonstrated the ability of mixtures to reproduce the peculiarities contained in real data, thus minimizing bias while keeping a low variance.

We have examined here the simple case of the anonymous GIC, which does not follow an individual over time. This means that an individual in a given initial quantile will not be necessarily in the same quantile in the next period. Grimm (2007) introduces income mobility for analysing if growth is pro-poor. By removing the anonymity axioms, he reduces the analysis to those who were initially poor, excluding thus those who become poor over time. He defines a naGIC (non-anonymous GIC) from inference on the bivariate distribution $F(y_{i,t-1}, y_{i,t})$. Van Kerm (2009) also considers the initial ranking of y_{t-1} and demonstrates the flexibility of a non-parametric approach. A quantile function can be approximated by regressing a variable over its rank. The transition between $y_{i,t-1}$ and $y_{i,t}$ can be measured by the difference $\log(y_{i,t}) - \log(y_{i,t-1})$ while the naGIC is obtained by a non-parametric regression of this difference over the initial rank of $y_{i,t-1}$. These two views rely on a particular definition of what is a quantile in a bivariate context (see also Jenkins and Van Kerm 2006 or Bourguignon 2011). We know from Barnett (1976) that a multivariate quantile is not an evident concept, even if there exist recent developments with e.g. Hallin et al. (2010).[8] Building on these considerations, a Bayesian approach to naGIC is on our research agenda.

References

Aitchison, J. and Brown, J. A. (1957). The lognormal distribution with special reference to its uses in economics. *The American Economic Review*, 48(4):690–692.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2018). The elephant curve of global inequality and growth. *AEA Papers and Proceedings*, 108:103–08.

[8]We are grateful to Karim Abadir for this reference, together with useful discussions.

Anderson, G., Pittau, M. G., and Zelli, R. (2014). Poverty status probability: a new approach to measuring poverty and the progress of the poor. *The Journal of Economic Inequality*, 12(4):469–488.

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.

Barnett, V. (1976). The ordering of multivariate data. *Journal of the Royal Statistical Society. Series A (General)*, 139(3):318–355.

Bauwens, L., Lubrano, M., and Richard, J.-F. (1999). *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, Oxford.

Bourguignon, F. (2011). Non-anonymous growth incidence curves, income mobility and social welfare dominance. *The Journal of Economic Inequality*, 9(4):605–627.

Brent, R. P. (1973). An algorithm with guaranteed convergence for finding a zero of a function. In Brent, R. P., editor, *Algorithms for Minimization without Derivatives*, chapter 4. Prentice-Hall, Englewood Cliffs, NJ.

Chancel, L. and Piketty, T. (2019). Indian income inequality, 1922-2014: From British Raj to billionaire Raj? *The Review of Income and Wealth*, 65(S1):S33–S62.

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.

Darvas, Z. (2016). Some are more equal than others: New estimates of global and regional inequality. Discussion Paper 8, Bruegel, Brussels, Belgium.

Datt, G. and Ravallion, M. (1992). Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980. *Journal of Development Economics*, 38:275–295.

Davidson, R. and Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, 68(6):1435–1464.

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

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

Fruhwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer Series in Statistics. Springer-Verlag New York.

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

Grimm, M. (2007). Removing the anonymity axiom in assessing pro-poor growth. *The Journal of Economic Inequality*, 5(2):179–197.

Hallin, M., Paindaveine, D., and Siman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: from L1 optimization to halfspace depth. *Annals of Statistics*, 38:635–669.

Hyndman, R. J. and Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician*, 50(4):361–365.

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

Jenkins, S. P. (2009). Distributionally-sensitive inequality indices and the GB2 income distribution. *Review of Income and Wealth*, 55(2):392–398.

Jenkins, S. P. and Van Kerm, P. (2006). Trends in income inequality, pro-poor income growth, and income mobility. *Oxford Economic Papers*, 58(3):531–548.

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.

Krause, M. (2014). Parametric Lorenz curves and the modality of the income density function. *Review of Income and Wealth*, 60(4):905–929.

Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review*, 45(1):1–28.

Lakner, C. and Milanovic, B. (2016). Global income distribution: From the fall of the Berlin wall to the great recession. *World Bank Economic Review*, 30(2):203–232.

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.

Novokmet, F., Piketty, T., and Zucman, G. (2018). From Soviets to oligarchs: Inequality and property in Russia, 1905-2016. *The Journal of Economic Inequality*, 16:189–223.

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

Sarabia, J. M. (2008). Parametric Lorenz curves: Models and applications. In Chotikapanich, D., editor, *Modeling Income Distribuions and Lorenz Curves*, volume 5 of *Economic Studies in Equality, Social Exclusion and Well-Being*, chapter 9, pages 167–190. Springer, New-York.

Sarabia, J.-M., Castillo, E., and Slottje, D. J. (1999). An ordered family of Lorenz curves. *Journal of Econometrics*, 91(1):43 – 60.

Son, H. H. (2004). A note on pro-poor growth. *Economics Letters*, 82(3):307 – 314.

Van Kerm, P. (2009). Income mobility profiles. *Economics Letters*, 102:93–95.

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

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

Appendix

A Bayesian inference for the log-linear regression model

Bayesian inference for the log-linear regression model is similar to that of the usual linear regression model that can be found in textbooks such as Bauwens et al. (1999). Here we provide a summary of these results. Suppose that we have n observations of the random variable y_i generated by a log-normal model. The vector of the n observations is noted y and the matrix of observed covariates being X. The likelihood function is:

$$\begin{aligned} L(y;\beta,\sigma^2) &= \left(\prod_{i=1}^{n}(y_i)^{-1/2}\right)(2\pi)^{-n/2}\sigma^{-n} \\ &\quad \times \exp -\frac{1}{2\sigma^2}(\log(y)-X'\beta)'(\log(y)-X'\beta). \end{aligned}$$

Under a non-informative prior on all the parameters:

$$p(\beta,\sigma^2) \propto 1/\sigma^2,$$

the posterior density of β and σ^2 are, respectively, an inverted gamma 2 and a Student:[9]

$$p(\sigma^2|y) = f_{i\gamma}(\sigma^2|n, s_*^2) \tag{25}$$

$$p(\beta|y) = f_t(\beta|\beta_*, M_*, s_*^2, n), \tag{26}$$

with:

$$M_* = X'X, \tag{27}$$

$$\beta_* = M_*^{-1}X'\log(y), \tag{28}$$

$$s_* = (\log(y)-X\beta_*)'(\log(y)-X\beta_*). \tag{29}$$

For the simple log-normal process, X contains just a constant term.

[9] Notations for these two densities are provided in the appendix of Bauwens et al. (1999), together with procedures to draw random numbers from them.

B Bayesian inference for a mixture of log-normals

Bayesian inference for a mixture of log-normals does not differ fundamentally from Bayesian inference in mixture of normal densities that can be found in textbooks like Fruhwirth-Schnatter (2006). Specific issues concerning prior specification to avoid label switching are discussed in Lubrano and Ndoye (2016).

The model and its notations are defined in (12) and $[y_i]$ is a vector of n observations. The usual way of doing inference is to consider a mixture model as an incomplete data problem. An auxiliary integer variable z allocates each observation y_i to a member of the mixture, identified by its label so that conditionally on a given sample allocation $[z_i = k]$, each component of the mixture can be analysed separately in a natural conjugate framework. We define a conditional normal prior on μ_k:

$$\mu_k|\sigma_k^2 \sim f_N(\mu_k|\mu_k^0, \sigma_k^2/n_k^0),$$

and an inverted gamma prior on σ_k^2:

$$\sigma_k^2 \sim f_{i\gamma}(\sigma_k^2|v_k^0, s_k^0),$$

for each of the K members, completed by a Dirichlet prior for the weights:

$$\eta \sim f_D(\gamma_1^0, ..., \gamma_K^0).$$

The hyperparameters of these priors are the $K-$vectors $v^0, s^0, \mu^0, n^0, \gamma^0$. The conditional posterior densities belong to the same families and serve to draw a new sample allocation, resulting in a Gibbs sampler algorithm:

1. Set K, the number of components, m the number of draws, m_0 the number of warming draws and initial values of the parameters $\theta^{(0)} = (\mu^{(0)}, \sigma^{(0)}, \eta^{(0)})$.

2. For $j = 1, \ldots, m + m_0$:

   (a) Generate a classification $z_i^{(j)}$ conditionally on $\theta^{(j-1)}$, independently for each observation y_i according to a multinomial process with probabilities:

$$Pr(z_i = k|y, \theta^{(j-1)}) = \frac{\eta_k^{(j-1)} f_\Lambda(y_i|\mu_k^{(j-1)}, \sigma_k^{2(j-1)})}{\sum_k \eta_k^{(j-1)} f_\Lambda(y_i|\mu_k^{(j-1)}, \sigma_k^{2(j-1)})}. \tag{30}$$

   (b) Compute the conditional sufficient statistics $n_k, \bar{y}_k, s_k^2$:

$$n_k = \sum_{i=1}^n \mathbb{I}(z_i = k), \tag{31}$$

$$\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^n \log(y_i) \mathbb{I}(z_i = k), \tag{32}$$

$$s_k^2 = \frac{1}{n_k} \sum_{i=1}^n (\log(y_i) - \bar{y}_k)^2 \mathbb{I}(z_i = k). \tag{33}$$

(c) Given the classification $z^{(j)}$, generate:

$\sigma^{(j)} \sim p(\sigma_k^2|y, z) = f_{i\gamma}(\sigma_k^2|v_k^*, s_k^*)$,
$\mu^{(j)} \sim p(\mu_k|\sigma_k^2, y, z) = f_N(\mu_k|\mu_k^*, \sigma_k^2/n_k^*)$,
$\eta^{(j)} \sim p(\eta|y, z) = f_D(\gamma_1^0 + n_1, ..., \gamma_K^0 + n_K) \propto \prod_{k=1}^K \eta_k^{\gamma_k^0+n_k-1}$,

with:

$n_k^* = n_k^0 + n_k$,
$\mu_k^* = (n_k^0\mu_k^0 + n_k\bar{y}_k)/n_k^*$,
$v_k^* = v_k^0 + n_k$,
$s_k^* = s_k^0 + n_k s_k^2 + \frac{n_k^0 n_k}{n_k^0+n_k}(\mu_k^0 - \bar{y}_k)^2$.

(d) Increase j by one and return to step (a).

3. Discard the first m_0 stored draws and compute posterior moments and marginal densities using the remaining draws.