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

Working Papers / Documents de travail

Bayesian Inference for TIP curves: An Application to Child Poverty in Germany

Edwin Fourrier-Nicolai
Michel Lubrano

WP 2017 - Nr 10

Bayesian Inference for TIP curves: An Application to Child Poverty in Germany*

Edwin Fourrier-Nicolai† Michel Lubrano‡

March 2017

Abstract

TIP curves are cumulative poverty gap curves used for representing the three different aspects of poverty: incidence, intensity and inequality. The paper provides Bayesian inference for TIP curves, linking their expression to a parametric representation of the income distribution using a mixture of lognormal densities. We treat specifically the question of zero-inflated income data and survey weights, which are two important issues in survey analysis. The advantage of the Bayesian approach is that it takes into account all the information contained in the sample and that it provides small sample confidence intervals and tests for TIP dominance. We apply our methodology to evaluate the evolution of child poverty in Germany after 2002, providing thus an update the portrait of child poverty in Germany given in Corak et al. (2008).

Keywords: Bayesian inference, mixture model, survey weights, zero-inflated model, poverty, inequality.
JEL codes: C11, C46, I32, I38.

*A first version of this paper was presented in May 2016 at the 10th Annual RCEA Bayesian Econometric Workshop in Rimini. Useful comments by participants are gratefully acknowledged. We are especially grateful to Denis de Crombrugghe and to Christian Schluter for a careful reading of a previous version of the paper and for their insightful suggestions. Of course usual disclaimers apply.

†Aix-Marseille University (Aix-Marseille School of Economics), CNRS, EHESS and Centrale Marseille, email: edwin.fourrier@univ-amu.fr.

‡Aix-Marseille University (Aix-Marseille School of Economics), GREQAM-CNRS, EHESS and Centrale Marseille, email: michel.lubrano@univ-amu.fr.

1 Introduction

Poverty is usually measured as the proportion of households having an income below the poverty line. This proportion, equivalently called the headcount ratio or the risk of poverty, is often taken as the unique measure of poverty, ignoring the shape of the income distribution among the poor. This narrow scoped measure equally accounts being poor with a zero income and being poor with an income just below the poverty line. The TIP curve introduced in Jenkins and Lambert (1997) is a cumulative poverty gap curve used for representing the different aspects of poverty usually provided by the Foster et al. (1984)'s poverty indices. It is also a transformation of the generalised Lorenz curve introduced in Shorrocks (1983). Therefore the TIP curve bridges the gap between measures of inequality and measures of poverty. One of the other numerous advantages of the TIP curve is that it comes equipped with a dominance criterion which, under some conditions, is equivalent to restricted second-order stochastic dominance (see Davidson and Duclos 2000). Then the poverty ranking obtained is robust according to the choice of a poverty line and of a poverty measure. Despite their attractiveness, TIP curves have only been used up to now as a descriptive tool of poverty, in a rather small number of articles such as for example Jenkins and Lambert (1997), Del Rio and Ruiz-Castillo (2001) and Kuchler and Goebel (2003). Moreover, none of these papers provide statistical inference for TIP curves.[1] The reason can be found in the difficulties for finding a confidence interval for Lorenz curves. A distribution free estimator for a Generalised Lorenz curve is very easy to derive, once the sample is ordered. However, their asymptotic distribution is quite difficult to establish (see Beach and Davidson 1983 for Lorenz curves and Bishop et al. 1989 for generalised Lorenz curves) and this difficulty increases when one considers survey data (see e.g. Beach and Kaliski 1986 and Binder and Kovacevic 1995). The aim of the present paper is to provide Bayesian inference for TIP curves (and also incidently for Generalised Lorenz curves) in the context of survey data.

A distribution-free approach for the generalised Lorenz curve and more generally for dominance curves might surfer from sensitivity to tails' behaviour and might appear unsatisfactory in small samples (see e.g. Cowell and Flachaire 2015). As our main focus is on TIP curves, the question of tail sensitivity becomes crucial, because we are concerned by the left tail of the income distribution. Moreover, as we shall see later on, a distribution free estimator of a TIP curve throws away all the observations which are above the poverty line, making the question of sample size even more crucial. Bayesian

[1] A notable exception is Thuysbaert (2008).

inference for TIP curves at least partially overcome these difficulties. A distribution-free approach does not make any assumption about the shape of the income distribution. A Bayesian approach has to rely on a parametric or semi-parametric representation of the income distribution. By considering that the income distribution can be represented by a mixture of parametric distributions, we can both have a rather important flexibility obtained simply by letting the sample determine the number of components of the mixture and impose a certain smoothness to the income distribution. The whole sample is then used to make inference on the left tail of the income distribution and no information is lost. We shall propose a mixture representation of the income distribution that takes into account both survey weights and the occurrence of zero value incomes which can be numerous when considering gross income (before taxes and redistribution). Bayesian inference is very well adapted for this purpose just because, as we shall prove it in the text, TIP curves are essentially transformations of the estimated parameters of the income distribution. If we manage to have Monte Carlo draws of the parameters of the income distribution, we shall have automatically draws for the TIP curve (and incidently for the Lorenz curve), opening thus a simple way to statistical inference and dominance tests in a Bayesian framework.

Child poverty in Europe is an important question that has motivated many papers. Several papers by Jenkins and Schluter (Jenkins et al. 2000, Jenkins and Schluter 2003, Corak et al. 2008, Hill and Jenkins 2001, Bradbury et al. 2001, among others) are devoted to measuring and comparing child poverty in the UK and in Germany. More recent data than those used in these papers are now available and many events that had surely an impact on global poverty and on child poverty have occurred since that period: we think about the 2008 financial crisis and more specifically for Germany to the recent social and labour market reforms (the famous Hartz plan). One purpose of the present paper is to show which kind of new information the use of TIP curves can bring in for understanding the evolution of child poverty in Germany, using the recent released data of the GSOEP.

The paper is organised as follows. In section 2, we introduce TIP curves, their relation to the Lorenz curve and define TIP dominance. Section 3 is devoted to Bayesian inference for mixture of log-normal densities in the case of survey weights and zero inflated incomes. In section 4, we derive the analytical formulae for TIP curves when the income distribution is modelled using a mixture of log-normals with sampling weights and zero-incomes. We also propose a test for comparing TIP curves. Section 5 analyses the evolution of child poverty in Germany, making the difference between current and chronic poverty. Section 6 concludes.

2 Measuring poverty using TIP curves

2.1 A formal definition

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:

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

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 $\alpha \in \{0, 1, 2\}$:

$$P^\alpha(z) = \int_0^z (1 - y/z)^\alpha f(y) dy, \tag{2}$$

where $f(y)$ is the income probability density function (pdf). For $\alpha = 1$ and when $P^\alpha(z)$ is considered as a function of z, we have the normalised deficit curve, which is also the second order dominance curve following Atkinson (1987). 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:

$$TIP(p, z) = \int_0^{F^{-1}(p)} (1 - y/z)\mathbb{1}(y \leq z) f(y) dy, \tag{3}$$

$F^{-1}(p)$ 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:

$$\widehat{TIP}(\hat{p}_i = i/n, z) = \frac{1}{n} \sum_{j=1}^{i} (1 - y_{(j)}/z)\mathbb{1}(y_{(j)} < z), \tag{4}$$

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 1. 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

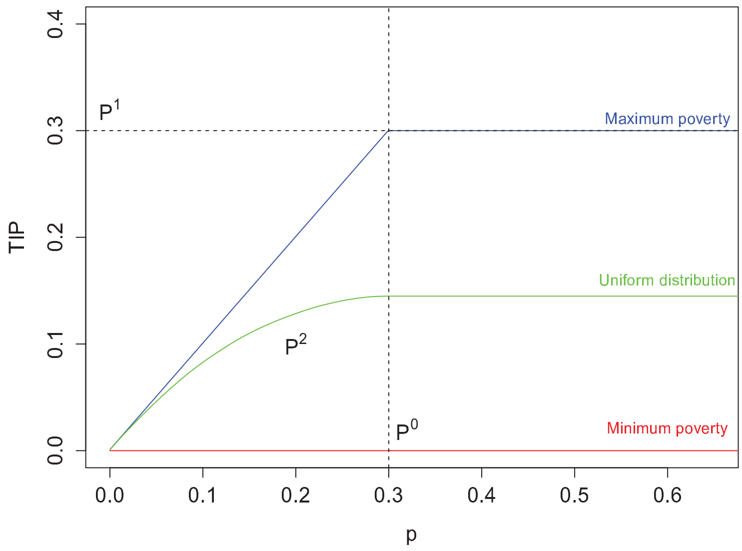

Figure 1: TIP curves from different income distributions

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.

In Figure 1, 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. 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. 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. 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 1 clearly shows that an identical P_0 can hide very different situations, all described adequately by a different TIP curve.

2.2 TIP curves and Lorenz curves

TIP curves have another desirable feature through their close link with the generalised Lorenz curve. The Lorenz curve $L(p)$ is a graphical representation of the cumulative distribution function of income introduced by Lorenz (1905), widely used for representing inequality. It was originally defined by two equations:

$$L(p) = \frac{1}{\bar{y}} \int_0^q t f(t) dt,$$

$$p = F(q) = \int_0^q f(t) dt,$$

where $\bar{y}$ is the mean income. For representing inequality and taking account of the level of income, the generalised Lorenz curve $GL(p)$ has been introduced by Shorrocks (1983). It is simply obtained by multiplying the Lorenz curve by the mean income $\bar{y}$.

Let us now derive the relationship between the GL curve and the TIP curve. Decomposing equation (3) and substituting q for $F^{-1}(p)$, we obtain:

$$TIP(p, z) = \int_0^q f(y) dy - \frac{1}{z} \int_0^q y\, f(y) dy, \qquad \text{for } p \leq F(z). \tag{5}$$

The first integral is the cdf at q, which is p provided that an analytical form for the quantile function is available. Quantiles are directly available for some parametric densities such as the Pareto and the Weibull, or even the log-normal. But there is no direct formula for more complicated densities such as mixtures. The second integral is the generalised Lorenz curve $GL(p)$. Then, the TIP can be expressed as in Davidson and Duclos (2000):

$$TIP(p, z) = p - \frac{1}{z} GL(p), \qquad \text{for } p \leq F(z). \tag{6}$$

This opens the way for parametric modelling and inference for TIP curves when a direct expression for the quantile function is available. For instance, if we assume that the DGP that governs the income distribution is a log-normal $f_\Lambda(y|\mu, \sigma^2)$, the Lorenz curve of a log-normal is known to be $L(p) = \Phi(\Phi^{-1}(p) - \sigma)$, see for instance Cowell (2011). To get the generalised Lorenz curve, we have to multiply it by the mean of the log-normal distribution, that is $\exp(\mu + \sigma^2/2)$. Then, the TIP curve for a log-normal process is:

$$TIP(p, z) = p - \frac{1}{z} \exp(\mu + \sigma^2/2)\, \Phi(\Phi^{-1}(p) - \sigma) \qquad \text{for } p \leq F(z). \tag{7}$$

However, parametric densities have important limitations, particularly due to their lack of flexibility in the tails. Moreover, they are strictly unimodal and thus cannot reflect polarisation for instance. A mixture of parametric densities can overcome these limitations, as advocated in Flachaire and Nunez (2007). But in the case of mixture models $q = F^{-1}(p)$ has no analytical form and thus we cannot generalise equation (7) easily.

2.3 Stochastic dominance and TIP dominance

Robust poverty ranking can be obtained using the poverty deficit curve obtained when (2) is seen as a function of z and letting z vary within a given interval. This corresponds to the notion of restricted stochastic dominance of Atkinson (1987), at the order 2 when $\alpha = 1$. This is the primal approach to stochastic dominance. The dual approach to stochastic dominance consider quantiles and the order 2 corresponds to Generalised Lorenz ordering. As it is related to the Generalised Lorenz curve, the TIP curve provides a natural framework for testing restricted second order stochastic dominance. We have however to show how. We first propose a definition of TIP dominance and then explore what it implies in term of stochastic dominance.[2]

Definition 1. *Let us consider two income distributions corresponding to populations A and B and a common poverty line. Population A TIP dominates B if* $TIP_A(p, z) \leq TIP_B(p, z)$ $\forall p \in [0, 1]$*, and the strict inequality holds at least for one p. The strict TIP dominance requires that this inequality is strict for all p.*

Jenkins and Lambert (1998), with their theorem 1 provide a relation between TIP dominance and poverty ordering. Their theorem could be rephrased as follows (see also Thuysbaert 2008):

Theorem 1. *Let us consider two TIP curves and a common poverty line. The following two conditions are equivalent:*

1. $TIP_A(p, z) \leq TIP_B(p, z)$ *for all* $p \in [0, 1]$
2. $P_A(\lambda z) \leq P_B(\lambda z)$ *for all* $\lambda \in [0, 1]$

This theorem means that TIP dominance is equivalent to restricted stochastic dominance at the order 2 over the range $[0, z]$, which means for all poverty

[2]TIP dominance according to the definition given in the 4th footnote of Jenkins and Lambert (1997) implies that there is more poverty in A than in B if A TIP dominates B. This might appear counter-intuitive when confronted to stochastic dominance. Thus, in our context, TIP dominance will mean less poverty as in Thuysbaert (2008).

lines lower or equal to z. In other words, if $TIP_A(p, z)$ is always below $TIP_B(p, z)$ with a common z, then there is less poverty intensity and less inequality among the poor in A than in B *for all common poverty lines smaller than or equal to* z. However, we cannot say anything about poverty headcount. Furthermore and like for Lorenz curves, when TIP curves intersect there is indeterminacy since the poverty ranking can be reversed for some values of p. So no ranking can be provided in this case.

A situation of TIP dominance is illustrated in Figure 1. The green curve TIP dominates the blue curve *using a common* z. Then the green curve exhibits less poverty than the blue curve *for any poverty line smaller than* z and the ranking cannot be reversed for a smaller value of z. Remark that if $F(z) = 1$ when $z = y_{max}$, we recover the concept of generalised Lorenz dominance or equivalently of (unrestricted) second order stochastic dominance.

The link between TIP dominance and restricted stochastic dominance at the order 2 was established for a common poverty line z. This restriction is not a problem when comparing less developed economies where the famous one-dollar-a-day poverty line is used (or its recent revision). For richer countries, a relative poverty line is the rule, which means different poverty lines. Characterising the relation between TIP dominance and restricted stochastic dominance at the order 2 in this context becomes more difficult. Jenkins and Lambert (1998) have shown that a robust ranking holds for all the possible poverty lines that keep a relation of proportionality between them.

3 Bayesian inference for a log-normal mixture using survey data

As it has been suggested in the previous section, we can use equation (6) to express TIP curves when the income distribution is modelled using a simple parametric density which provides an analytical expression for the Lorenz curve. However, simple parametric densities are too restrictive to describe the richness of the income distribution. They preclude for instance the presence of several modes that arises particularly when looking at gross income. In order to model the income distribution both in a flexible and parsimonious way, mixtures of log-normals have been promoted by Flachaire and Nunez (2007), Lubrano and Ndoye (2016) and Anderson et al. (2014), referring to the Gibrat's law. Mixtures of gamma densities were also considered in Chotikapanich and Griffiths (2008). While still relying on parameters, mixtures are very flexible: it suffices to add a component density to the mixture to increase its flexibility. Moreover, mixtures have desirable features due to the preservation of some component density's properties at the mixture level because of the linearity of the model.

3.1 Finite mixtures of log-normals

A finite mixture $f(y|\vartheta)$ is a linear combination of K parametric densities $f(y|\theta_k)$ such that:

$$f(y|\vartheta) = \sum_{k=1}^{K} \eta_k f(y|\theta_k), \qquad 0 \leq \eta_k < 1, \quad \sum_{k=1}^{K} \eta_k = 1, \tag{8}$$

where $\vartheta = (\eta, \theta)$, and the parameter vectors are $\theta = (\theta_1, ..., \theta_k)$ and $\eta = (\eta_1, ..., \eta_k)$ with η_k and θ_k being, respectively, the weights and the parameters of the k^{th} component. Let us now assume that all the components in (8) are univariate log-normal distributions with $f(y|\theta_k) = f_\Lambda(y|\mu_k, \sigma_k)$. We have thus a mixture of log-normals. With its two parameters (μ, σ), the pdf of the log-normal is written as:

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

3.2 Bayesian inference for a log-normal mixture

A mixture model has to deal with two issues. First, the classification of observations into the K different components with probability η_k. Second, the estimation of the parameters for every component density. The problem would simplify greatly if the classification of the observations were known. This remark led Diebolt and Robert (1994) to consider a mixture as an incomplete data problem. Each observation y_i has to be completed by an unobserved variable z_i taking an integer value in $\{1, ..., K\}$, indicating which member of the mixture each y_i comes from. The model has to explain the pair (y_i, z_i). The EM algorithm in a classical framework and the Gibbs sampler in a Bayesian framework start from an initial hypothetical sample separation $[z_i]$ and conditionally on $[z_i]$ make inference on the parameters ϑ. Once the sample allocation is known, we can treat each component separately meaning that μ_k, σ_k are estimated for all $k = 1, ..., K$ from the observations in group k only, whereas estimation of η is based on the number $n_1(z), ..., n_k(z)$ of observations allocated to each group. Then a new sample separation $[z_i]$ is determined, given the previous values found for μ_k, σ_k and η_k. This approach is particularly well suited in a Bayesian framework because given $[z_i]$ we can manage to find conjugate prior for each sub-model $f_\Lambda(y|\mu_k, \sigma_k)$ and for

η_k. As explained for instance in Lubrano and Ndoye (2016), the natural conjugate priors for each member of a mixture of log-normals are formed by a conditional normal prior on $\mu_k|\sigma_k^2 \sim f_N(\mu_k|\mu_0, \sigma_k^2/n_0)$ and an inverted gamma prior on $\sigma_k^2 \sim f_{i\gamma}(\sigma_k^2|v_0, s_0)$. A Dirichlet prior is used for $\eta \sim f_D(\gamma_1^0, ..., \gamma_K^0)$. The hyperparameters of these priors are $v_0, s_0, \mu_0, n_0, \gamma_k^0$. For a given sample separation $[z_i]$, we get the following sufficient statistics:

$$n_k = \sum_{i=1}^{n} \mathbb{1}(z_i = k), \tag{9}$$

$$\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^{n} \log(y_i) \mathbb{1}(z_i = k), \tag{10}$$

$$s_k^2 = \frac{1}{n_k} \sum_{i=1}^{n} (\log(y_i) - \bar{y}_k)^2 \mathbb{1}(z_i = k). \tag{11}$$

Combining these sufficient statistics with the prior hyperparameters, we get:

$$\begin{aligned} n_{*k} &= n_0 + n_k, \\ \mu_{*k} &= (n_0\mu_0 + n_k\bar{y}_k)/n_{*k}, \\ v_{*k} &= v_0 + n_k, \\ s_{*k} &= s_0 + n_k s_k^2 + \frac{n_0 n_k}{n_0 + n_k}(\mu_0 - \bar{y}_k)^2, \end{aligned}$$

which are used to index the conditional posterior densities of first σ_k^2 which is an inverted gamma:

$$p(\sigma_k^2|y, z) = f_{i\gamma}(\sigma_k^2|v_{*k}, s_{*k}), \tag{12}$$

and second of $\mu_k|\sigma_k^2$, which is a conditional normal:

$$p(\mu_k|\sigma_k^2, y, z) = f_N(\mu_k|\mu_{*k}, \sigma_k^2/n_{*k}). \tag{13}$$

The conditional posterior distribution of η_k is a Dirichlet with:

$$p(\eta|y, z) = f_D(\gamma_1^0 + n_1, ..., \gamma_K^0 + n_K) \propto \prod_{k=1}^{K} \eta_k^{\gamma_k^0 + n_k - 1}. \tag{14}$$

We can then determine the posterior probability that the i-th observation comes from the k-th component $(z_i = k)$, conditionally on the value of ϑ and the value of y_i:

$$Pr(z_i = k|y, \vartheta) = \frac{\eta_k f_\Lambda(y_i|\mu_k, \sigma_k^2)}{\sum_k \eta_k f_\Lambda(y_i|\mu_k, \sigma_k^2)}. \tag{15}$$

A Gibbs sampler algorithm is easy to implement, given those conditional posterior densities. Examples can be found in Lubrano and Ndoye (2016) and Fruhwirth-Schnatter (2001) together with a discussion concerning label switching and its remedies (see also Fruhwirth-Schnatter 2006, p. 78).

3.3 Introducing survey weights

In population studies, it is common to sample individuals through complex sampling designs in which some individuals or groups are over or under-represented corresponding to specific sampling weights. Analysing data from such designs can be tricky, since the collected sample is not representative of the overall population. Survey weights are constructed to correct for discrepancies between sample and population. The literature on the use of survey weights in mixture models is not abundant and concerns mainly stratification (see e.g. Kunihama et al. 2016). However, most of socio-economic panels (PSID, BHPS, GSOEP, ...) provide information only on sampling weights, while information concerning stratification is limited. We propose a simple method, easy to implement within a Gibbs sampler, to take account of the sampling weights.

Let us consider n individuals that are sampled from the whole population with survey weights $w_i = c/\pi_i$, where c is a positive constant and π_i is the inclusion probability that individual i belongs to the survey. A mixture estimate of the income distribution representative of the genuine population can be obtained by introducing weights when evaluating the sufficient statistics in (9-11), such that:

$$n_k = \sum_{i=1}^{n} w_i \mathbb{1}(z_i = k), \tag{16}$$

$$\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^{n} w_i \log(y_i) \mathbb{1}(z_i = k), \tag{17}$$

$$s_k^2 = \frac{n_k}{n_k^2 - \sum_{i=1}^{n} w_i^2 \mathbb{1}(z_i = k)} \sum_{i=1}^{n} w_i (\log(y_i) - \bar{y}_k)^2 \mathbb{1}(z_i = k). \tag{18}$$

The other formulae of the Gibbs sampler are left unchanged: re-weighting the conditional sufficient statistics is enough to modify the sample allocation with probabilities (15).

Introducing survey weights directly in the estimation of the mixture's parameters is one way to consider weights for TIP curves. If we consider the distribution free approach, one way to modify the estimator (4) is as follows:

$$\widehat{TIP}\left(\hat{p}_i = \frac{\sum_{j=1}^{i} w_{(j)}}{\sum_{j=1}^{n} w_{(j)}}, z\right) = \frac{1}{\sum_{j=1}^{n} w_{(j)}} \sum_{j=1}^{i} w_{(j)} (1 - y_{(j)}/z) \mathbb{1}(y_{(j)} < z), \tag{19}$$

where $w_{(i)}$ is the weight attached to i^{th} order statistics. The weights are used for determining which observation corresponds to p and for computing the partial summation. Alternatives ways can be found for instance in

Thuysbaert (2008) in the case of random weights. From our experience the distribution-free estimator of the TIP curve using weights needs a large number of observations whereas our semi-parametric Bayesian method performs better in small samples.

3.4 Modelling zero-inflated data

Another redundant feature of survey data is the excess number of zeros (greater than expected under the distributional assumption). Particularly in income studies, zero incomes are numerous when measured before taxes and transfers. Actually, a large part of the population has no market income such as elderly persons or unemployed workers. This is a problem when estimating the income distribution in both a parametric approach and a non-parametric approach using smoothing techniques. As the log-normal is defined on the strict positive support, we propose to add an extra-component for modelling the zero incomes:

$$f(y|\vartheta) = \begin{cases} \bar{\omega} & \\ (1-\bar{\omega})\sum_{k=1}^{K}\eta_k f(y|\theta_k) & \text{if } y > 0, \end{cases} \tag{20}$$

where $\bar{\omega} = \Pr(y = 0) \simeq (\sum_i \mathbb{1}(y_i = 0)w_i)/\sum w_i$. This is a zero-inflated mixture model where $\bar{\omega}$ is estimated as the (weighted) proportion of zeros in the sample, while inference on the other parameters is performed on the sample excluding the zeros. Hence, zeros are not a problem for inference. The CDF corresponding to this zero-inflated mixture is:

$$F(y|\vartheta) = \begin{cases} \bar{\omega} & \text{if } y = 0, \\ \bar{\omega} + (1-\bar{\omega})\sum_{k=1}^{K}\eta_k F(y|\theta_k) & \text{if } y > 0. \end{cases} \tag{21}$$

It will be used in subsection 4.2 for making Bayesian inference on TIP curves.

Such modelling of the zero-inflated data problem is simple and particularly useful within a parametric approach, knowing that values close to 0 cause problems for many parametric families. Even more, non-parametric estimators might fail to represent this feature of the distribution. Figure 2 represents estimates of the German market income distribution for the whole population in 2009 using the GSOEP data. The kernel density estimate smoothes out the zero-excess observations. Worst, the kernel density estimate seems to be less flexible than the 3 component mixture at the beginning of the distribution, which is important for analysing poverty. Even if the proportion of zeros seems negligible in this sample of disposable income with 0.15%, it is still important to take it into account. If we had considered market income, that proportion would have gone up to 5% for the same year.

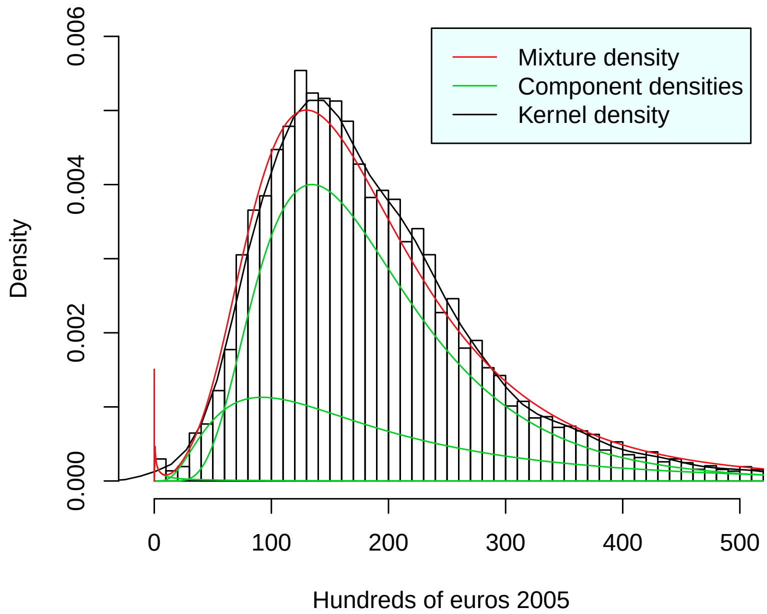

Figure 2: The distribution of disposable household income in 2011

4 Bayesian inference for TIP curves

Up to now, we have provided inference for a mixture model of log-normals taking into account both the survey weights and the zero-inflated observations. In order to provide inference for TIP curves, we have to express these as a function of the model's parameters. Since the quantile function of a mixture has no analytical form, we cannot generalise equation (6) to the case of mixtures of lognormals. We have to start from the original definition of the TIP curve given in equation (5) where the expression of the quantile q is left unspecified.

4.1 An alternative TIP curve formula

Assuming that the DGP of $f(y)$ is a mixture of log-normal as given in equation (8), we can decompose the general TIP formula, using the linearity property of the integral:

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

for $p \leq F(z)$. The first part of the right hand side is the cdf of the mixture of log-normals:

$$\sum_{k=1}^{K} \eta_k F_\Lambda(q|\mu_k,\sigma_k) = \sum_{k=1}^{K} \eta_k \Phi\left(\frac{\ln q - \mu_k}{\sigma_k}\right),$$

knowing that Φ is the standard normal cdf and that the mixture's cdf is the weighted sum of the components' cdf. In the second part of the right hand side, we find the weighted sum of components' generalised Lorenz curve for the log-normal:

$$\sum_{k=1}^{K} \eta_k \int_0^q y f_\Lambda(y|\mu_k,\sigma_k)\,dy = \sum_{k=1}^{K} \eta_k \exp(\mu_k + \sigma_k^2/2)\Phi\left(\frac{\ln q - \mu_k - \sigma_k^2}{\sigma_k}\right).$$

As a matter of fact, q cannot be substituted within each component as in equation (7), since q here represents the value of the p quantile of the complete mixture. And by no way is the quantile of a mixture a linear function of the quantile of each component. A quantile of a mixture has to be evaluated numerically. Leaving aside for the while the question of determining the value of q, we can regroup the previous results in order to propose an expression of the TIP curve for a mixture of log-normal densities:

$$\begin{aligned} TIP_\Lambda(p,z) &= \sum_{k=1}^{K} \eta_k \left(\Phi\left(\frac{\ln q - \mu_k}{\sigma_k}\right)\right. \\ &- \left.\frac{1}{z}\exp(\mu_k + \sigma_k^2/2)\Phi\left(\frac{\ln q - \mu_k - \sigma_k^2}{\sigma_k}\right)\right), \end{aligned} \tag{22}$$

for $p \leq F(z)$. Note that the left hand side is a function of p, while the right hand side is a function of q. We have to complete this equation by a relation between p and q, solving $q = F^{-1}(p)$. We know that the cdf of a mixture $F(y|\vartheta)$ is the weighted sum of the cdf of each component: $F(y|\vartheta) = \sum_{k=1}^{K} \eta_k F(y|\mu_k,\sigma_k)$. As there is no expression for $F^{-1}(p)$, $F(y|\vartheta)$ has to be inverted numerically. We consider a grid of points p_s for p between 0 and 1. For each point p_s and for each value of ϑ, we solve numerically in q the equation:[3]

$$F(q|\vartheta) = p_s.$$

[3]For instance in `R`, this is programmed as `uniroot(function(x){Fln(x,mu,s) - p},c(0,100000))$root`, with for instance `Fln = function(x,mu,s){plnorm(x,mu,s)}`. The execution time is negligible.

Let us now suppose that we have obtained m posterior draws of the parameters, $\vartheta^{(j)}$ from the Gibbs sampler. For each point p_s of a predefined grid, we can obtain m draws for the TIP curve (or the Lorenz curve), applying (22) and solving numerically for q the equation $F(q|\vartheta^{(j)}) = p_s$. The posterior mean of each point of the TIP curve corresponding to a value of p_s is obtained as the mean of all these m posterior draws for each value of p_s. The 0.05 and 0.95 quantiles of these m draws provide an evaluation of a 90% confidence interval of the TIP curve.[4]

4.2 TIP curves for the zero-inflated model

When we take into account the excess of zero incomes, the expression of the cumulative distribution is changed in to (21). Moreover, we have to solve equation (3) separately for $0 \leq p \leq \bar{\omega}$ and then for $\bar{\omega} \leq p \leq 1$. For the range $0 \leq p \leq \bar{\omega}$, we know that $y = 0$, so equation (3) becomes:

$$TIP(p,z) = \int_0^{F^{-1}(p)} f(y)dy = p.$$

For the range $\bar{\omega} \leq p \leq 1$, we make use of equation (21), so the expression of the TIP curve (3) becomes:

$$\begin{aligned} TIP(p,z) &= \int_0^{F^{-1}(p)} (1-\bar{\omega}) \sum_{k=1}^{K} \eta_k f(y|\theta_k) dy \\ &\quad - \frac{1}{z} \int_0^{F^{-1}(p)} y(1-\bar{\omega}) \sum_{k=1}^{K} \eta_k f(y|\theta_k) dy, \end{aligned}$$

and after integration:

$$\begin{aligned} TIP(p,z) &= \bar{\omega} + (1-\bar{\omega}) \sum_{k=1}^{K} \eta_k \Phi\left(\frac{\ln q - \mu_k}{\sigma_k}\right) \\ &\quad - \frac{(1-\bar{\omega})}{z} \sum_{k=1}^{K} \eta_k \exp(\mu_k + \sigma_k^2/2) \Phi\left(\frac{\ln q - \mu_k - \sigma_k^2}{\sigma_k}\right). \quad (23) \end{aligned}$$

[4] There are the same numerical difficulties for deriving the posterior density of a TIP curve than for deriving the posterior density of a Lorenz curve. Chotikapanich and Griffiths (2008) have proposed Bayesian inference for a mixture of Gamma densities and have derived the corresponding Lorenz curve. They propose however an ad hoc procedure to compute the quantiles that does not take a full account of parameter uncertainty. This explains why their posterior confidence intervals are so narrow and consequently unconvincing.

The TIP curve for an inflated-zero model can be written as:

$$TIP(p,z) = \begin{cases} p & 0 \leq p \leq \bar{\omega} \\ \bar{\omega} + (1-\bar{\omega})TIP_{\Lambda}(p,z) & \bar{\omega} \leq p \leq 1. \end{cases} \tag{24}$$

where $TIP_{\Lambda}(p,z)$ was defined in (22). For evaluating (23), the value of q has to be determined using (21).

4.3 Testing for TIP dominance

Testing for TIP dominance means comparing two curves. If the income distribution is modelled as a simple lognormal distribution, then the TIP curve receives a parametric form which is given in (7). Sufficient conditions for TIP dominance in the case of identical poverty lines can be inferred from the results of Levy and Kroll (1976) or Yitzhaki (1982) for second order stochastic dominance. $TIP_A(p,z)$ is always lower than $TIP_B(p,z)$ if we have:

$$\begin{aligned} \mu_A &\geq \mu_B, \\ \sigma_A &\leq \sigma_B, \\ \mu_A + \sigma_A^2 &\geq 2\mu_B + \sigma_B^2. \end{aligned}$$

It is rather easy to check for these conditions in a Bayesian framework and to compute the posterior probability that these three conditions hold simultaneously. This is a simple MCMC exercise. However, in the more realistic case when the income distribution is modelled as a mixture of lognormal distributions, it becomes impossible to find parametric restrictions implying TIP dominance.

We are thus back to the question of comparing directly two curves. This question has a long history in a classical framework, as surveyed for instance in Davidson and Duclos (2000). Among the various possibilities, Davidson and Duclos (2000), Davidson and Duclos (2013) promote the use of a test based on a previous work by Kaur et al. (1994). This procedure aims at testing H_0 of non-dominance against the alternative of dominance. More formally the hypothesis are as follows for TIP dominance:

$$H_0 : \min_p (TIP_A(z,p) - TIP_B(z,p)) \geq 0, \tag{25}$$

$$H_1 : \min_p (TIP_A(z,p) - TIP_B(z,p)) < 0. \tag{26}$$

The test statistics is a simple t which has asymptotically a standardised Gaussian distribution. However, this test has been criticised for having low power (see for instance Thuysbaert 2008). But Davidson and Duclos (2013) show that the performance of this type of test can be greatly improved by bootstrapping. Because of the relation between Bayesian inference and bootstrapping (see Efron 2012), we are confident in the development of a similar test in a Bayesian framework.

Consider two populations A and B and the associated posterior draws of the parameters $\vartheta = (\vartheta_A, \vartheta_B)$ of their respective distributions modelled as mixtures. Testing for TIP dominance of A over B can be reframed as determining whether the conditional difference between the associated two TIP curves

$$d(p|\vartheta) = TIP_A(p, z) - TIP_B(p, z)$$

is significantly negative for every value of p on a given grid. This is equivalent to consider the probability of the event:

$$\min_p(d(p|\vartheta)) < 0,$$

for all values of ϑ. The posterior density of this event can be evaluated easily once we have obtained m posterior draws of the parameters $\vartheta^{(j)}$ from the Gibbs sampler. For each draw $\vartheta^{(j)}$, we compute the posterior density of the minimum distance between the two TIP curves $d(p|\vartheta_A, \vartheta_B)$ and see if the above event is true or not. More formally:

$$\begin{aligned} \Pr\left(\min_p d(p|y) < 0\right) &= \int_\vartheta \mathbb{1}\left[\min_p d(p|\vartheta) < 0\right] \mu(\vartheta|y) d\vartheta \\ &\simeq \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}\left[\min_p d(p|\vartheta^{(j)}) < 0\right], \end{aligned} \tag{27}$$

where $\mu(\vartheta|y)$ is the posterior density of the parameters. Note here that the range of p has to be slightly restricted because all TIP curves are zero at $p = 0$ and solving numerically for the $p = 1$ quantile can be troublesome. So the practical range for the test should be something like $p \in [0.01, 0.99]$. This is what is commonly done in a classical framework, see e.g. Davidson and Duclos (2013). This procedure is general and allows us to obtain the probability of stochastic dominance, restricted stochastic dominance, Lorenz dominance and TIP dominance depending on the grid defined and the poverty lines chosen.

5 A new portrait of child poverty in Germany

Corak et al. (2008) provides some of the most recent results concerning the evolution of child poverty in Germany, using the data of the GSOEP. However, their reporting period ends in 2004. As new data are now available, we propose to investigate the period 2000-2012 in order to enlighten differences and changes in evolution. We shall use the tool of TIP curves in order to investigate the following points: Provide a better description of the different dimensions of child poverty over the period, illustrate the evolution of chronic child poverty and its differences with adult chronic poverty, illustrate the differences between East and West Germany and measure the impact of the redistributive system.

5.1 Preliminary points

The first step is to determine a poverty line. We consider 50% of the median disposable income as was made in Corak et al. (2008), taking into account all households (those having children and those without children, but eliminating the households which were given a zero cross-section weight in the GSOEP data set), normalised by the new OECD equivalence scale.[5] The usual practice, as recommended by Eurostat for instance, is to take either 60% of median disposable income or 50% of mean disposable income. We justify our choice in order to keep coherency with the paper of Corak et al. (2008).

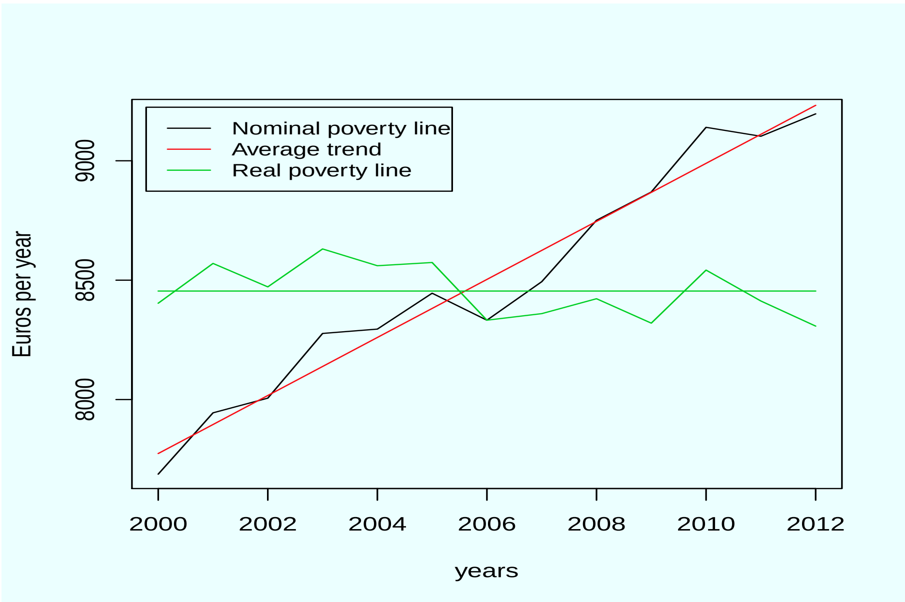

Figure 3: Nominal and real poverty lines

Figure 3 shows that the nominal poverty line is evolving steadily around

[5]We used the cross-section household weights for computing the median.

its trend. However, there is a drop in 2006 and the poverty line takes time to recover its trend. As the poverty line is defined as a fraction of the disposable income, this drop might reflect the delayed effect of the Hartz plan. When we divide disposable income by the 2005 CPI, the real poverty line fluctuates above a constant mean before 2006 and below it after that date. We might eventually consider a unique real poverty line for the period, equal to its average of 8 455 euros per year and per equivalent adult. Another practice could also lead us to consider two different poverty lines, one for East and one West Germany. We opted for a single poverty line valid for all Germany, because we would like to investigate if there has been a convergence between the poverty rates of the East and West parts of Germany.

Investigating child poverty means that we have to consider a sample where only the children are present, which means having possibly several children coming from the same household. This is the usual practice followed by Hill and Jenkins (2001), Jenkins and Schluter (2003), Corak et al. (2008) and many others. An alternative way of doing would be to consider households with children, so having a single observation per household, independently of the number of children in the household. This would lead to an under-evaluation of child poverty. We define a child as a person under 18 years old (wether it is considered as a child (95% of the sample) or a simple relative or friend, in a given household). We report in Figure 4 the evolution of poverty headcounts, distinguishing between East and West Germany. Child poverty headcounts are following a rising trend till 2006. This corroborates the findings of Corak et al. (2008) for the period 1999-2004. This rising evolution stops in 2006 as from that date child poverty seems to fluctuate around its mean in East and decrease in West Germany. We conclude that there is a clear break around that date, which justifies considering separately the two periods.

A TIP curve is a graphical representation of the left tail of an income distribution, below a poverty line. When current income is used, we have a portrait of current or instant poverty. However, a household might be currently in a state of poverty and escape from poverty in the next period. If households can transfer income from one period to the other, poverty has to be portrayed with respect to this smoothed or permanent income. This point of view was introduced in Rodgers and Rodgers (1993) and was later developed in Hill and Jenkins (2001) and Kuchler and Goebel (2003). For this, we must consider longitudinal data and more precisely a balanced panel. We have decided to separate the period in two subperiods, leading to two balanced data sets and two series of smoothed income (obtained as individual weighted means using longitudinal weights). For the sake of comparability, we consider 2002-2006 and 2007-2011. Because of attrition, we got 2 991

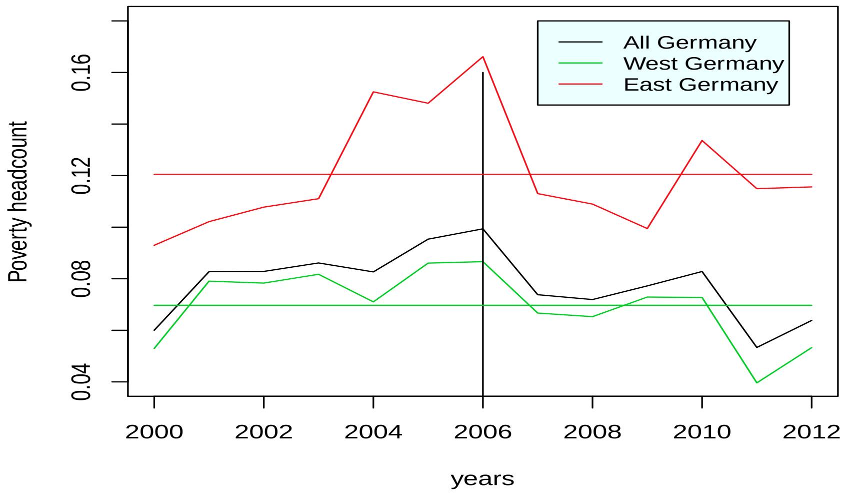

Figure 4: Evolution of poverty headcounts in Germany

The two horizontal lines represent the average poverty headcount over the period for East and West Germany. The poverty line is 50% of median disposable income. A constant poverty line equal to 8 455 euros would simply emphasise some of the fluctuations, without changing the general shape of the graphs.

observations for the first period and 2 236 for the second period. We were obliged to discard the year 2012 because including it in our second balanced panel would have entailed too much attrition. When analysing smoothed income, the corresponding poverty line is formed by the mean of the poverty lines corresponding to the smoothing period under consideration.

The last thing we have to do is to run as many Gibbs samplers as we have to produce graphs of TIP curves. We proceed as follows. We have used each time 10 000 draws, dropping the first 4 000 for warming the chain. On average, we used a three member mixture for disposable income and for smoothed income while four components were needed for market income. The number of components was selected using a BIC. We took equal prior prior probabilities for η_k with $\gamma_0 = 5$. We took identical prior expectation of μ for each component, setting it equal to the weighted sample mean of $\log y$ and choosing prior precision $n_0 = 1.0$. For each $\mathrm{E}(\sigma^2)$, we took an increasing fraction of the weighted sample variance of $\log y$, correspond to the sequence $(0.25, 0.5, 1.0)$ for 3 components with $\nu_0 = 50$. This prior is coherent with the Gibbs algorithm given in Appendix A where an ordering constraint is imposed on σ_k^2 to cope with label switching.[6] Standardised CUMSUM graphs were used to check for convergence. We do not report posterior results which are available on request.

5.2 The evolution of global child poverty

The left part of Figure 5 clearly shows that there is a significant increase in global child poverty over the first period as TIP curves do not intersect as well as 10% confidence intervals. This increase is thus apparent in the three dimensions of poverty. Chronic poverty is of course lower than current poverty and represents 47% of current poverty in 2002, 30% in 2006, so that the distance between the two increases over time during the first period. The behaviour of poverty during the second period is totaly different. As

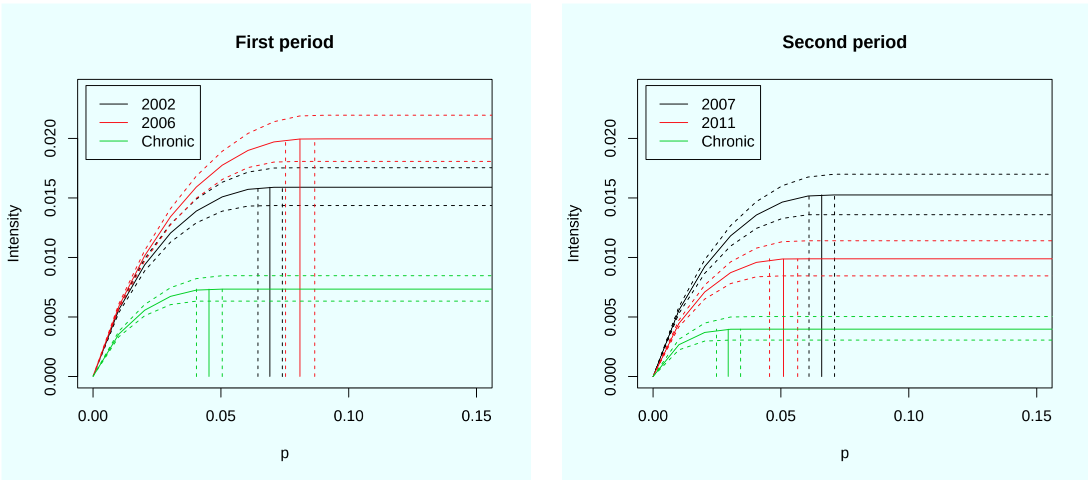

Figure 5: The evolution of global child poverty

the graphs have the same scale, visual comparison is made easy. Current poverty has decreased significantly between the beginning and the end of the second period as again confidence intervals do not intersect. The fraction of chronic poverty is increasing as it represents 34% of current poverty in 2007 and 38% in 2011. But chronic poverty has decreased significantly between the two periods. Formal dominance tests confirm these impressions. The probability that 2007-2011 TIP dominates 2002-2006 is equal to 0.997. For current child poverty, Table 1, shows that 2011 dominates all the other years, which means that child poverty is lowest for this year. Child poverty has significantly increased between 2002 and 2006 as 2002 TIP dominates 2006

[6] See e.g. Fruhwirth-Schnatter (2006, Chap. 3) for more details on label switching.

Table 1: Probability of TIP dominance for current child poverty

Year	2002	2006	2007	2011
2002	0.000	0.907	0.260	0.000
2006	0.003	0.000	0.002	0.000
2007	0.525	0.939	0.000	0.000
2011	1.000	1.000	1.000	0.000

Each line represents the probability that the corresponding year TIP dominates the year given in column.

at 91% while it decreases after that date because 2007 TIP dominates 2006 at 94%.

5.3 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).[7] In Figure 6, we 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), contrary to what happened to child poverty as reported from Figure 5. So the evolution of poverty over the two subperiods concerned mainly children with an increase and a decrease, while there was no significant effect on the population of adults. This is confirmed by TIP dominance tests reported in Table 2. There is no convincing probability of Tip dominance of any year.

[7]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 using the two balanced panels.

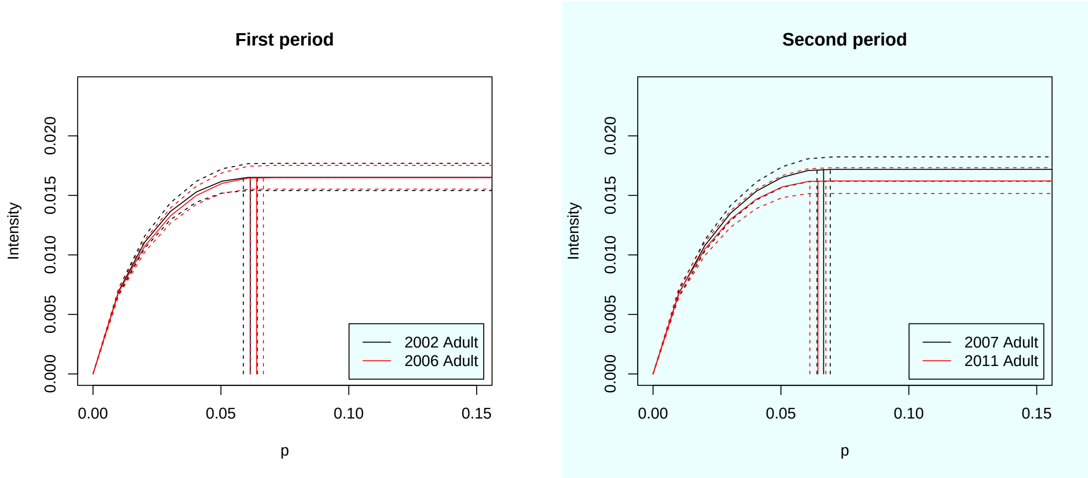

Figure 6: The constancy of current adult poverty over the period

Table 2: Probability of TIP dominance for current adult poverty

Year	2002	2006	2007	2011
2002	0.000	0.241	0.228	0.215
2006	0.496	0.000	0.435	0.370
2007	0.194	0.150	0.000	0.135
2011	0.493	0.403	0.424	0.000

Each line represents the probability that the corresponding year TIP dominates the year given in column.

5.4 Child poverty dynamics and the redistributive system

Up to now, we have been working with disposable income, which means income after taxes and redistribution. So poverty is compensated and we know that the German redistributive system is supposed to be generous with children (see Konigs 2014). But how far? Figure 7 contrast smoothed income before and after redistribution and taxes. Before any intervention of the redistributive system, chronic poverty has slightly decreased between the two periods over all dimensions, but this decrease is not significant as the probability of TIP dominance between 2007-2011 and 2002-2006 is only of 0.253. However poverty incidence is important with a poverty rate around 13%. Introducing redistribution first reduces poverty in a very important extent as shown in Figure 7. And second, chronic poverty is strongly reduced

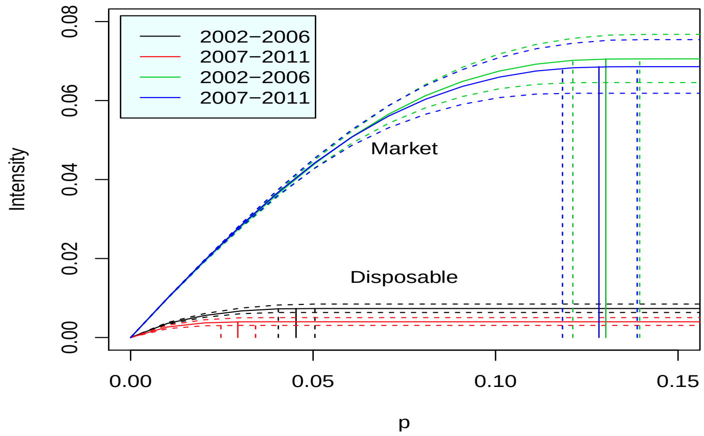

Figure 7: The importance of the redistributive system
on chronic child poverty

between the two periods, passing from 4.5% to 2.9%. This reduction is significant because the probability of TIP dominance of the second period over the first period is equal to 0.997 and that confidence interval for poverty rates do not overlap.

5.5 The East-West contrast

East and West Germany have been reunified in 1990. However, the convergence between these two regions is slow and the economic differences are still important. At the level of the whole country, we concluded that over the period 2002-2011 adult poverty has not changed significantly, while child poverty followed an up and down pattern. Do we observe the same pattern when we observe the two regions separately? In order to make these comparisons, we restrict our attention to chronic poverty.

The first striking fact is that chronic poverty is much more important in East Germany than in West Germany during the first period, as seen when looking at the lower left panel of Figure 8. Poverty incidence and poverty intensity are significantly greater in the East than in the West. However, a test of TIP dominance is not significant, because the TIP curves of the West and the East intersect during the first period. This explain the low values of 0.016 and 0.212 reported in Table 3 below.

During the second period, there is a massive reduction of chronic child poverty, both in East and West Germany as seen from the two upper pan-

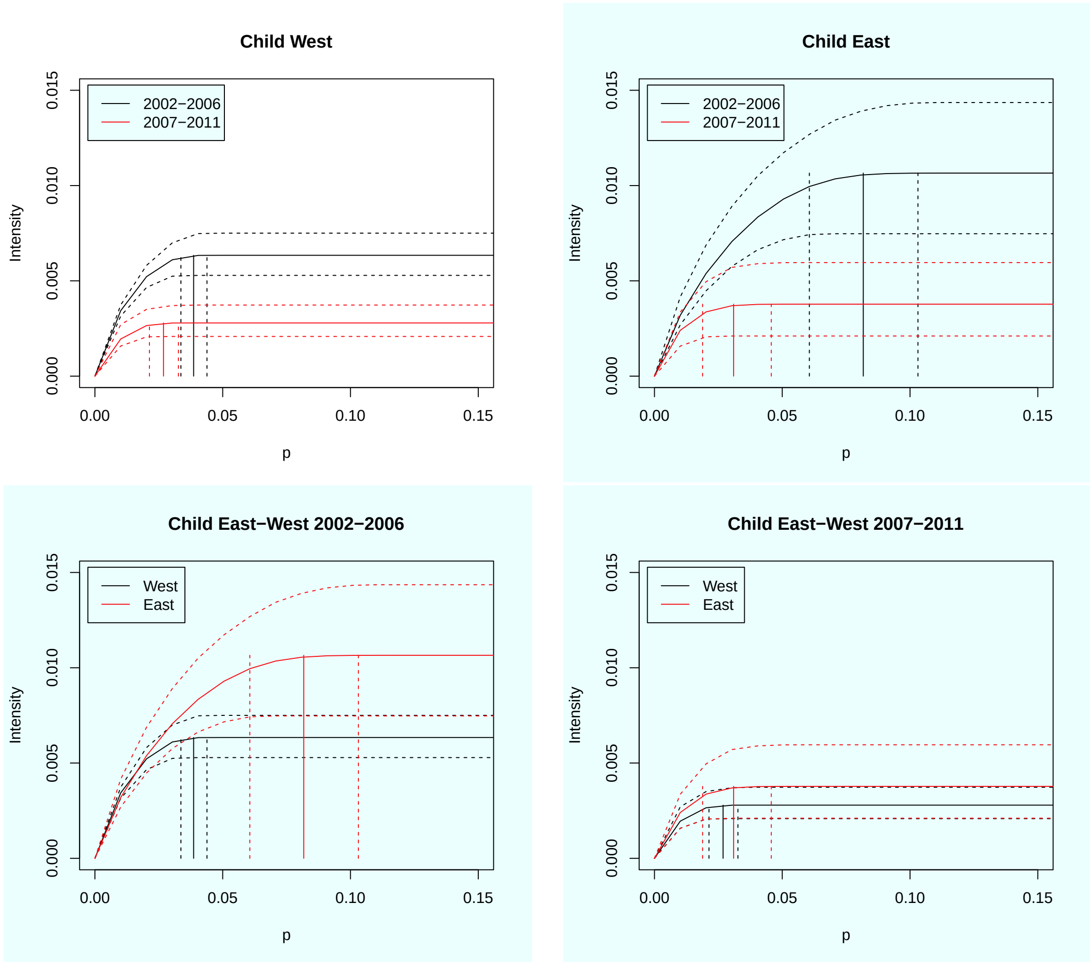

Figure 8: The West-East contrast of chronic child poverty

els of Figure 8. The results of Table 3 confirm this global and significant reduction of chronic poverty in the West between the two periods with a probability of TIP dominance of 0.998. This reduction is also important in the East, but the significance of TIP dominance is less marked with a probability of 0.872, certainly due to the relatively small number of observations (around 500 observations for children in the East). This massive reduction of chronic child poverty has erased the differences between the two regions as confidence intervals for poverty incidence and poverty intensity overlap for the two regions in the second period as seen when looking at the lower right panel of Figure 8. And there is no TIP dominance in one way or the other with reported probabilities of 0.673 and 0.169 in Table 3. We conclude the redistributive system has been very efficient during the second period for fighting against child chronic poverty, after the absorption of the effects of

Table 3: TIP dominance test for child chronic poverty between West and East Germany

		West		East	
		2002-2006	2007-2011	2002-2006	2007-2011
West	2002-2006	0.000	0.001	0.212	0.028
	2007-2011	0.998	0.000	0.991	0.673
East	2002-2006	0.016	0.000	0.000	0.002
	2007-2011	0.941	0.169	0.872	0.000

the reforms introduced in the Hartz plan.

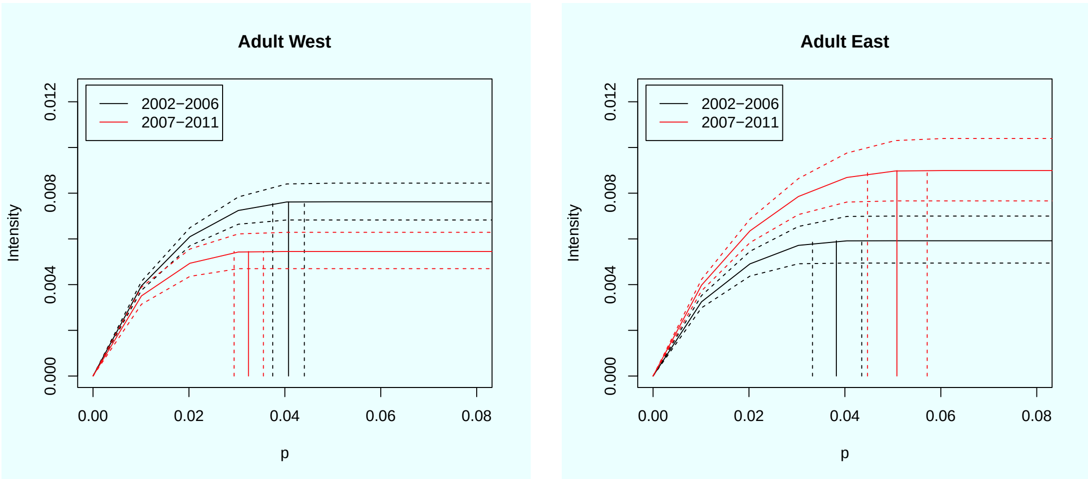

Figure 9: The West-East contrast of chronic adult poverty

If we now turn to adult chronic poverty, the situation is totaly different. During the first period, the rate of adult chronic poverty is lower in the East and the East TIP dominates the West with a probability equal to 0.979 (see Table 4). During the second period, adult chronic poverty has significantly decreased in West Germany. The probability of TIP dominance of the second period over the first is equal to 0.956. While it has significantly increased in East Germany. The probability of TIP dominance of the first period over the second is equal to 0.999. There is definitively more chronic poverty in the East part of Germany in the second period (TIP dominance of the West over the East with a probability equal to 0.954). We can conclude that there are still strong differences between the West and the East part of Germany. If the break in the data is due to the Hartz plan, the effect of the latter was very different in the two regions. The question of chronic poverty was very well treated for children in both regions of Germany by the redistributive

Table 4: TIP dominance test for adult chronic poverty between West and East Germany

		West		East	
		2002-2006	2007-2011	2002-2006	2007-2011
West	2002-2006	0.000	0.002	0.001	0.555
	2007-2011	0.956	0.000	0.175	0.954
East	2002-2006	0.979	0.279	0.000	0.999
	2007-2011	0.073	0.000	0.001	0.000

system and convergence was reached. But the effect were devastating for adult chronic poverty in East Germany.

6 Conclusion and summary

Thanks to their numerous properties, TIP curves are called to play an important role in the measurement of poverty. Because they are a transformation of the general Lorenz curve, TIP dominance is linked to second order stochastic dominance as shown in Davidson and Duclos (2000). We have provided tools for Bayesian inference for TIP curves. In this attempt, we have proposed a parametric modelling of the income distribution, using a mixture of lognormals. For this, we had to solve two questions raised by the use of survey data: incorporating survey weights and taking into account explicitly zero-inflated samples. Once we have obtained random draws from the posterior distribution of the parameters of the modelled income distribution, the TIP curves are a (not so) simple transformation of these draws, which means that we have a direct access to statistical inference (both confidence intervals and testing). This approach can be found to be more powerful than the traditional distribution-free approach for two reasons. First, we take into account of the whole sample when modelling the income distribution, both the observations which are below the poverty line and those which are above. The distribution free approach neglect all the observations which are above the poverty line. Second, the rate of convergence of non-parametric estimators is $n^{-1/5}$ while that of parametric estimators is $n^{-1/2}$. Of course there is always the risk of misspecification, but with a mixture approach which is a semi-parametric approach, we are on the safe side. We have illustrated the applicability and usefulness of our methods to the question of child poverty in Germany. We used the German SOEP between 2001 and 2011. Our results are complementary to those of Corak et al. (2008) which stop in 2004. Child poverty continued to follow the upward trend pointed out in Corak et al. (2008) for the period 2000-2004. We might see there the impact of the Hartz plan (2003-2005). However, after 2006 the portrait of child poverty in Germany become totaly different with a decrease in both total and chronic poverty and a total reduction of the gap between adult and child poverty. However, remains a large difference in adult chronic poverty between the East and West part of Germany.

All this being said, we did not answer all the questions concerning child poverty in Germany. We point out in particular the modelling of poverty dynamics (entry into poverty probability, and exit from poverty) and the measurement of the exact impact of the Hartz plan. Are the effects we measured due to the Hartz plan or have they another origin? There is little chance for the while that simple TIP curves could answer those questions. More research is needed.

References

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.

Atkinson, A. B. (1987). On the measurement of poverty. *Econometrica*, 55(4):749–764.

Beach, C. M. and Davidson, R. (1983). Distribution-free statistical inference with Lorenz curves and income shares. *The Review of Economic Studies*, 50(4):723–735.

Beach, C. M. and Kaliski, S. F. (1986). Lorenz curve inference with sample weights: An application to the distribution of unemployment experience. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 35(1):38–45.

Binder, D. A. and Kovacevic, M. S. (1995). Estimating some measures of income inequality from survey data: An application of the estimating equations approach. *Survey Methodology*, 21(2):137–145.

Bishop, J. A., Chakraborti, S., and Thistle, P. D. (1989). Asymptotically distribution-free statistical inference for generalized Lorenz curves. *The Review of Economics and Statistics*, 71(4):725–27.

Bradbury, B., Jenkins, S., and Micklewright, J. (2001). The dynamics of child poverty in seven industrialised nations. In Bradbury, B., Jenkins,

S. P., and Micklewright, J., editors, *The Dynamics of Child Poverty in Industrialised Countries*, chapter 4, pages 92–132. Cambridge University Press, Cambridge, UK.

Chotikapanich, D. and Griffiths, W. (2008). Estimating income distributions using a mixture of gamma densities. In *Modeling Income Distributions and Lorenz Curves*, volume 5 of *Economic Studies in Equality, Social Exclusion and Well-Being*, pages 285–302. Springer, New York.

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

Cowell, F. (2011). *Measuring Inequality*. Oxford University Press, 3rd edition.

Cowell, F. A. and Flachaire, E. (2015). Statistical methods for distributional analysis. In Atkinson, A. B. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 2, chapter 6, pages 359–465. Elsevier, Amsterdam.

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.

Davidson, R. and Duclos, J.-Y. (2013). Testing for restricted stochastic dominance. *Econometric Reviews*, 32(1):84–125.

Del Rio, C. and Ruiz-Castillo, J. (2001). TIPs for poverty analysis. the case of Spain, 1980-81 to 1990-91. *Investigaciones Economicas*, 25(1):63–91.

Diebolt, J. and Robert, C. (1994). Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistical Society. Series B (Methodological)*, 56(2):363–375.

Efron, B. (2012). Bayesian inference and the parametric bootstrap. *The Annals of Applied Statistics*, 6(4):1971–1997.

Flachaire, E. and Nunez, O. (2007). Estimation of the income distribution and detection of subpopulations: An explanatory model. *Computational Statistics and Data Analysis*, 51(7):3368–3380.

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 52(3):761–766.

Fruhwirth-Schnatter, S. (2001). Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association*, 96(453):194–209.

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

Hill, M. and Jenkins, S. (2001). Poverty among British children: chronic or transitory? In Bradbury, B., Jenkins, S. P., and Micklewright, J., editors, *The Dynamics of Child Poverty in Industrialised Countries*, chapter 7, pages 174–195. Cambridge University Press, Cambridge, UK.

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

Jenkins, S. and Lambert, P. (1998). Three I's of poverty curves and poverty dominance: TIPs for poverty analysis. In Slottje, D., editor, *Research on Economic Inequality*, volume 8, pages 39–56. JAI Press Inc.

Jenkins, S., Schluter, C., and Wagner, G. (2000). The dynamics of child poverty: Britain and Germany compared. Discussion paper 233, DIW.

Jenkins, S. P. and Schluter, C. (2003). Why are child poverty rates higher in Britain than in Germany? a longitudinal perspective. *The Journal of Human Resources*, 38(2):441–465.

Kaur, A., Rao, B. P., and Singh, H. (1994). Testing for second-order stochastic dominance of two distributions. *Econometric Theory*, 10:849–866.

Konigs, S. (2014). The dynamics of social assistance benefit receipt in Germany - state dependence before and after the 'Hartz reforms'. Technical Report 628, SOEP Papers on Multidisciplinary Panel Data Research at DIW Berlin.

Kuchler, B. and Goebel, J. (2003). Incidence and intensity of smoothed income poverty in European countries. *Journal of European Social Policy*, 13(4):357–369.

Kunihama, T., Herring, A. H., Halpern, C. T., and Dunson, D. B. (2016). Nonparametric Bayes modelling with sample survey weights. *Statistics & Probability Letters*, 113:41–48.

Levy, H. and Kroll, Y. (1976). Stochastic dominance with riskless assets. *The Journal of Financial and Quantitative Analysis*, 11(5):743–777.

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70):209–219.

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.

Rodgers, J. and Rodgers, J. (1993). Chronic poverty in the United States. *The Journal of Human Resources*, 28(1):25–54.

Shorrocks, A. F. (1983). Ranking income distributions. *Economica*, 50(197):3–17.

Thuysbaert, B. (2008). Inference for the measurement of poverty in the presence of a stochastic weighting variable. *The Journal of Economic Inequality*, 6(1):33–55.

Yitzhaki, S. (1982). Stochastic dominance, mean variance, and Gini's mean difference. *The American Economic Review*, 72(1):178–185.

APPENDIX

A Gibbs sampler algorithm

We present a simple Gibbs sampler algorithm which implements a traditional ordering solution for label switching (see Fruhwirth-Schnatter 2001 and Fruhwirth-Schnatter 2006, p. 78):

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 $\vartheta^{(0)} = (\mu^{(0)}, \sigma^{(0)}, \eta^{(0)})$ for $j = 0$.

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

   (a) Generate a classification $z_i^{(j)}$ independently for each observation y_i according to a multinomial process with probabilities given by equation (15), using the value of $\vartheta^{(j-1)}$.

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

   (c) Generate the parameters $\sigma^{(j)}, \mu^{(j)}, \eta^{(j)}$ from the posterior distributions given in equations (12), (13) and (14) respectively, conditionally on the classification $z^{(j)}$.

(d) Order $\sigma^{(j)}$ such that $\sigma_1^{(j)} < ... < \sigma_K^{(j)}$ and sort $\mu^{(j)}$, $\eta^{(j)}$ and $z^{(j)}$ accordingly.

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

3. Finally discard the first m_0 stored draws to compute posterior moments and marginals.

B The Hartz reforms in a nutshell

The Hartz reforms, started in 2003 and ended at the beginning of 2005, have fundamentally changed the labour market, the social assistance and insurance systems. They have triggered a lot of political and social protests.

The income-support for working-age individuals has been the most affected by the Hartz reforms, see Konigs (2014) for an extensive review. Until 2005, the individual's income after a job loss was partially replaced by the unemployment insurance benefits (UI, Arbeitslosengeld) for a limited amount of time (12-31 months), with eligibility being conditional on contribution records. The level of the benefit was independent of individual means and it was greater for individuals with children. When UI expired, individuals could claim unemployment assistance benefits (UA, Arbeitslosenhilfe) for an unlimited amount of time, they are also earnings-related but less generous than UI and means-tested on family income. Finally, social assistance (SA, Sozialhilfe) was the last resort even if it was initially been targeted at individuals with limited employability, but a gradual tightening of eligibility criteria for UI and UA over time resulted in a growing numbers of individuals had shifted into SA, as explained in Konigs (2014). After the introduction of the Hartz reforms, the UI was replaced by the unemployment benefit I (UBI, Arbeitslosengeld I) with an initially unchanged maximum benefit duration and replacement rate. In 2006, the maximum duration was lowered to 18 months but raised again to 24 months in 2008. The UA was replaced by the unemployment benefit II (UBII, Arbeitslosengeld II) which was not earnings-related. Social assistance was henceforth restricted to individuals incapable of work.

Before and after the reform, an income-tested Housing Benefit (HB, Wohngeld) is targeted at low-income households (except those entitled in SA). Since 2005, as recipients of SA, recipients of UBII cannot be eligible for HB but they can receive support for eligible housing expenses (HE).

A large part of the family support policy in Germany comes from the child benefits (Kindergeld), about 1.6% of GNI in 2009, even if they only aim at compensating for the financial burden of raising children. They have not changed deeply since 1996, and benefits only depend on the number of children. They are monthly paid to every legal guardian of children (under 18 years old, exceptions exist until 25) as a cash benefit or as a tax deduction (Kinderfreibetrag), the latter being rather rare, about 4.4% of total child benefit in 2009. If children live with persons in need of social assistance, they are entitled to social assistance too. As well, if children live with persons with very low incomes, they can perceive the means-tested supplementary child benefit (Kinderzuschlag). This was introduced in 2005 along with the Hartz reforms and aims at targeting households that fall below the needs thresholds of the new unemployment benefit II only because they have children. Finally, the parental allowance (Elterngeld) is a benefit for parents who would like to look after their child themselves after their birth and therefore are not full-time employed or not working at all. Since 2007, parents can apply for parental leave (Erziehungsgeld) and receive 67% of their net income as a parental allowance from the government for a duration of up to 14 months. Parental leave offers parents the opportunity of looking after their child whilst allowing them to maintain contact with working life. Employees can be entitled to parental leave until the child's third birthday. Their job is kept for them, and their contract cannot be ended by their employer. Parental leave can be taken by the mother and the father individually or jointly.

The insight behind the Hartz reforms was to push former unemployed individuals into the labour market. Although, the total number of employed persons has risen, the social situation of low-paid earners and unemployed persons has deteriorated, as suggested in Konigs (2014). In 2005, the benefit receipt rate of UBI should decrease as the benefit receipt average of UBII, but the benefit receipt rate of UBII should increase. This suggests a decrease of income for unemployed individuals. In 2006, the benefit receipt rate of UBI should decrease (due to the reduction of maximum duration of UBI) and that of UBII should increase as a consequence. This suggest a decrease of income too for the unemployed. And, in 2008, this last pattern should be reversed.