Bayesian inference for TIP curves: an application to child poverty in Germany

Edwin Fourrier-Nicolaï[1,3] · Michel Lubrano[2,3]

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 log-normal 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 credible 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. (Rev. Income Wealth **54**(4), 547–571, 2008).

Keywords Bayesian inference · Mixture model · Survey weights · Zero-inflated model · Child poverty

1 Introduction

Poverty is usually measured as the proportion of households having an income below a poverty line. This proportion, equivalently called the head-count ratio, the at-risk-of-poverty rate or poverty incidence, is often taken as the unique measure of poverty, ignoring the shape

✉ Michel Lubrano
michel.lubrano@univ-amu.fr

Edwin Fourrier-Nicolaï
edwin.fourrier@univ-amu.fr

1 CNRS, EHESS, Centrale Marseille, AMSE, Aix-Marseille University, Marseille, France

2 School of Economics, Jiangxi University of Finance and Economics and AixMarseille Univ., CNRS, EHESS, Centrale Marseille, AMSE, Marseille, France

3 Amse-Greqam, 5-9 Bd Maurice Bourdet, 13001, Marseille, France

of the income distribution among the poor. Since Sen (1976) and Shorrocks (1995), research in poverty measurement has developed measures which take into account the distributional aspects of poverty. The Three I's of Poverty (TIP) curve, which was formally considered in Jenkins and Lambert (1997), is a cumulative poverty gap curve used for representing the different aspects of poverty usually provided by the first three poverty indices of Foster et al. (1984), measuring Incidence, Intensity and Inequality. This curve is also known in the literature as the Poverty Profile Curve or the Poverty Gap Profile (PGP) curve (Barrett et al. 2016). TIP curves guaranty that poverty comparisons can be done in an unambiguous way (see e.g. Jenkins and Lambert 1998a, b; Davidson and Duclos 2000). This desirable property comes from its close relationship with the generalised Lorenz curve. Specifically, the TIP curve is a transformation of the generalised Lorenz curve, and hence its dominance criterion is equivalent, under some conditions, to restricted second-order stochastic dominance. Consequently poverty ranking obtained by TIP dominance is robust according to the choice of a poverty line and of a set of poverty measures. Despite their attractiveness, TIP curves have only been used as a descriptive tool in few empirical studies, as for example in Del Rio and Ruiz-Castillo (2001) and Kuchler and Goebel (2003). However, a recent growing attention has been devoted on providing statistical inference for TIP curves. Thuysbaert (2008) proposes statistical inference for TIP curves in the presence of stochastic survey weights. More recently, Barrett et al. (2016) provides a non-parametric test for TIP dominance using influence functions. The aim of the present paper is to provide Bayesian inference for TIP curves (and also incidently for generalised Lorenz curves) and Bayesian characterisation of TIP dominance in the context of survey data.

The advantages of using a Bayesian approach are manyfold. First, the statistical test can be performed directly once inference is obtained. Second, distribution-free approaches might suffer 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 the poor, the question of tail sensitivity becomes crucial because we are concerned with the left tail of the income distribution. Particularly, a distribution-free estimator of the TIP curve throws away all the observations which are above the poverty line, making the question of sample size even more crucial when we are concerned with a sub-population like children. Bayesian inference for TIP curves at least partially overcomes these difficulties. While 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 have both a rather important flexibility obtained by letting the sample determine the number of components of the mixture and the advantage of a parametric representation using the whole sample so that no information is lost. By using a Bayesian approach, we increase accuracy and precision for inference and tests.

Using mixtures for modelling the income distribution has a rather rich history, mainly using log-normal densities. For instance, Flachaire and Nunez (2007) use a mixture of normal densities for the UK's log-transformed income distribution, Lubrano and Ndoye (2016) propose a decomposition of inequality indices using Bayesian inference for a mixture of log-normals and Anderson et al. (2014) use a mixture of log-normals to identify the poor. Alternative specifications have been also used, for instance Ndoye and Lubrano (2014) propose Bayesian inference for a mixture of Pareto densities for a model of high wage formation using US data. Chotikapanich and Griffiths (2008) and Lander et al. (2017) use a mixture of gamma densities for comparing income distributions in Canada and in Indonesia. Although the mixture of log-normal densities is particularly appropriated for capturing the left tail of the distribution, we need to take into account both survey weights and the excess of zero value incomes which are two recurrent issues in survey data. This added flexibility makes the analysis more complex, particularly if we have to invert the cumulative distribution function of the mixture using numerical methods in order to compute quantile functions or a Lorenz curve.

We have chosen to illustrate our methods by discussing child poverty in Germany between 2000 and 2012 using the data of the *German Socio-Economic Panel* (GSOEP). Child poverty in Europe is an important question that has motivated many papers (Jenkins et al. 2000; Jenkins and Schluter 2003; Corak et al. 2008; Hill and Jenkins 2001; Bradbury et al. 2001), among others). More recent data than those used in these studies are now available and many events that had surely an impact on poverty have occurred since that period; we think specifically about the recent social and labour market reforms (the well-known Hartz plan, 2003-2005) and the reforms on family policies after 2005. And of course, although not specific to Germany, the 2008 financial crisis. 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 recently released data of the GSOEP in order to update the results found in Corak et al. (2008).

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. Section 6 concludes.

2 Measuring poverty using TIP curves

2.1 A formal definition

Let us consider a population where each household is endowed with an income $y \in \mathbb{R}^+$, the distribution of this income being $F(y)$ and its density $f(y)$. A household is said to be poor if its income is below the poverty line z. Poverty intensity is measured by the relative poverty gap, defined as follows:

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

where $\mathbb{1}(y \leq z)$ is the indicator function which is 1 if $y \leq z$ and 0 otherwise. Integrating this relative poverty gap with respect to the income distribution $f(y)$ provides the TIP curve of Jenkins and Lambert (1997):[1]

$$TIP(p, z) = \int_0^{F^{-1}(p)} (1 - y/z)\mathbb{1}(y \leq z) f(y) dy, \tag{2}$$

where $F^{-1}(p)$ is the quantile function, and p the proportion of individuals.

[1]In this paper we consider the relative TIP curve built on the poverty gaps normalised by the poverty line. This is a well-known variant of the absolute TIP curve based on absolute poverty gaps $\max\{z - y, 0\}$. The advantages of the relative curve are that the FGT indices come directly from the curve and this variant is more appropriated for comparisons of TIP curves when there are different poverty lines.

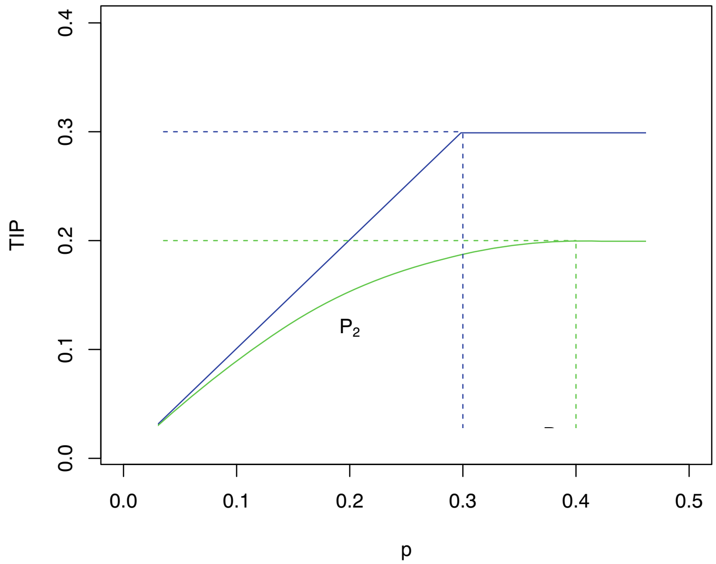

Fig. 1 TIP curves from different income distributions

For values of p greater than the poverty incidence or head-count ratio given by $P_0 = F(z)$, the TIP curve saturates and becomes horizontal. Its 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 introduced in Foster et al. (1984) when $\alpha = 0, 1, 2$ respectively. A TIP curve is thus a convenient device for displaying at the same time the three essential aspects of poverty, justifying its name the Three 'I's of Poverty.

In Fig. 1, we have represented the TIP curves of two simulated income distributions, using the same poverty line z. The top blue TIP curve represents a case of severe poverty where 30% of the individuals have zero income and the rest has an income greater or equal to the poverty line. In this case, the average poverty gap P_1 is the highest (according to the poverty incidence) i.e. $P_1 = P_0$. This is a straight line reflecting the fact that every poor has the same income, hence all poverty gaps are the same. Inequality among the poor P_2 is maximum (maximum slope of 1) and the Gini among the poor is 1.[2] The green curve represents a situation where 40% of the population has an income drawn from a uniform distribution between zero and z and the rest an income greater than z. The average poverty gap is 0.2. 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 represents a situation where the ranking of poverty incidence is the opposite of the ranking of poverty intensity and inequality.

[2]The Gini is 1 among the poor because all the poor have a zero income while one is supposed to have z which is a typical case of maximum inequality.

2.2 TIP curves and generalised Lorenz curves

TIP curves are closely linked to the generalised Lorenz (GL) curve as can be shown now. Decomposing Eq. 2 and substituting q for $F^{-1}(p)$, we obtain:

$$TIP(p,z) = \int_0^q f(y)dy - \frac{1}{z}\int_0^q y\, f(y)dy, \qquad \text{for } p \leq F(z). \tag{3}$$

The first integral is the cumulative distribution function evaluated at q and the second integral is the generalised Lorenz curve $GL(p)$. These integrals can be simplified when an analytical form for the quantile function $q = F^{-1}(p)$ is available and consequently also for the Lorenz curve. In this case, we have access to the direct formulation given in Davidson and Duclos (2000):

$$TIP(p,z) = p - \frac{1}{z}GL(p), \qquad \text{for } p \leq F(z). \tag{4}$$

As an illustration, let us consider the log-normal case with $f_\Lambda(y|\mu,\sigma^2)$. [3] The Lorenz curve of a log-normal is equal to $L_\Lambda(p) = \Phi(\Phi^{-1}(p) - \sigma)$ where Φ is the standard normal cdf. Multiplying $L_\Lambda(p)$ by the mean of the log-normal $\exp(\mu + \sigma^2/2)$ in order to get the generalised Lorenz curve $GL_\Lambda(p)$, we find that the TIP curve associated to log-normal density is:

$$TIP_\Lambda(p,z) = p - \frac{1}{z}\exp(\mu + \sigma^2/2)\,\Phi(\Phi^{-1}(p) - \sigma) \qquad \text{for } p \leq F_\Lambda(z). \tag{5}$$

However, we cannot take advantage of the simplicity of Eq. 4 if we want to model the income distribution using a mixture of log-normals. With mixtures of distributions, the total cdf $F(.)$ is a weighted average of the cdf of each member, but its inverse $q = F^{-1}(p)$ has no analytical form and has to be solved by a numerical procedure. So, we cannot generalise Eq. 5 easily. Instead of considering (4), we have to work with the initial form of the TIP curve, given in (3).

2.3 Stochastic dominance and TIP dominance

The use of a dominance criteria leads to poverty comparisons that are robust to the choice of the poverty line. As TIP curves are related to generalised Lorenz curves, they provide a natural framework for testing restricted second order stochastic dominance. We define a TIP ordering or TIP dominance as follows:[4]

Definition 1 Let us consider two income distributions corresponding to populations A and B and their corresponding TIP curves. Distribution A TIP dominates distribution B for a given poverty line z 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.

Since generalised Lorenz dominance is strictly equivalent to second order stochastic dominance (Atkinson 1987; Foster and Shorrocks 1988), it can be easily shown that TIP

[3] See for instance (Cowell 2011) for the properties of the log-normal density.

[4] 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 presented in Thuysbaert (2008) or Davidson and Duclos (2000).

dominance with a common poverty line z is equivalent to restricted second order stochastic dominance over the interval $[0, z]$ or equivalently for $p \in [0, F(z)]$. This correspond to Theorem 2 of Jenkins and Lambert (1998b) which can be rephrased as:

Theorem 1 *If distribution A TIP dominates distribution B for all* $p \in [0, F(z)]$, *then A stochastically dominates B at the second order over the interval* $[0, z]$. *And this poverty ordering holds for all common poverty lines below z.*

In other words, if $TIP_A(p, z)$ is always below $TIP_B(p, z)$ with a common z, then the two distributions can be ordered with less poverty intensity and inequality in A than in B *for all common poverty lines smaller than or equal to z*.

At this stage, some remarks have to be made. Second order stochastic dominance and TIP ordering imply the ordering of FGT indices corresponding to P_1 and P_2 (intensity and inequality) and all similar indices built around a relative poverty gap measure (see Davidson and Duclos (2000) and Jenkins and Lambert (1998b) among others). However, a robust ordering of poverty incidence is not guaranteed, as, for this, we would have to check for first order stochastic dominance. This situation is illustrated in Fig. 1 as commented before. Finally, and like for Lorenz curves, when TIP curves intersect there is indeterminacy since the poverty ranking can be reversed for some values of p. Various ways to remove this indeterminacy have been proposed: imposing more normative conditions and considering higher-degree dominance criteria as in Sordo and Ramos (2011), Jenkins and Lambert (1998a) or considering a narrower range of values of z as in Atkinson (1987).

3 Bayesian inference for a mixture of log-normals in survey data

Bayesian inference for mixtures of distributions is fairly standard and was detailed for instance in Lubrano and Ndoye (2016) for log-normals. In this section, after recalling the necessary formulae, we extend the analysis to two important features when modelling survey data: the use of sample weights and the modelling of extra zeros. We compare our solutions to the existing literature (Thuysbaert 2008; Si et al. 2015; Savitsky and Toth 2016; Gunawan et al. 2017), for the main references).

3.1 Bayesian inference for a mixture of log-normals

A finite mixture $f(y|\vartheta)$ of log-normals is a linear combination of K log-normal densities $f_\Lambda(y|\mu, \sigma)$ such that:

$$f(y|\vartheta) = \sum_{k=1}^{K} \eta_k f_\Lambda(y|\mu_k, \sigma_k^2), \qquad 0 < \eta_k < 1, \quad \sum_{k=1}^{K} \eta_k = 1, \tag{6}$$

where the η_k are the weights of the mixture and (μ_k, σ_k^2) the parameters of the k^{th} component. $\vartheta = (\eta, \mu, \sigma^2)$ represents the collection of all the parameters of the mixture. The pdf of the log-normal is written as:

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

A mixture model is an incomplete data problem which can be completed by an auxiliary variable z which allocates each observation to a given member of the mixture. Then conditionally on a given sample allocation $[z_i = k]$, each component can be analysed separately. Following Lubrano and Ndoye (2016), each component is equipped with a natural conjugate prior: a conditional normal prior on $\mu_k|\sigma_k^2 \sim f_N(\mu_k|\mu_k^0, \sigma_k^2/n_k^0)$ and an inverted gamma prior on $\sigma_k^2 \sim f_{i\gamma}(\sigma_k^2|v_k^0, s_k^0)$. A Dirichlet prior is used for the vector of 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 which is described as follows:[5]

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 $\vartheta^{(0)} = (\mu^{(0)}, \sigma^{(0)}, \eta^{(0)})$.
2. For $j = 1, \ldots, m + m_0$:
   (a) Generate a classification $z_i^{(j)}$ conditionally on $\vartheta^{(j-1)}$, independently for each observation y_i according to a multinomial process with probabilities:

$$Pr(z_i = k|y, \vartheta^{(j-1)}) = \frac{\eta_k^{(j-1)} f_\Lambda(y_i|\mu_k^{(j-1)}, \sigma_k^{2(j-1)})}{\sum_k \eta_k^{(j-1)} f_\Lambda(y_i|\mu_k^{(j-1)}, \sigma_k^{2(j-1)})}. \tag{7}$$

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

$$n_k = \sum_{i=1}^{n} \mathbb{1}(z_i = k), \tag{8}$$

$$\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^{n} \log(y_i) \mathbb{1}(z_i = k), \tag{9}$$

$$s_k^2 = \frac{1}{n_k} \sum_{i=1}^{n} (\log(y_i) - \bar{y}_k)^2 \mathbb{1}(z_i = k). \tag{10}$$

   (c) Given the classification $z^{(j)}$, generate the parameters $\sigma^{(j)}, \mu^{(j)}, \eta^{(j)}$ from the posterior densities (11), (12) and (13):
      – The conditional posterior density of σ_k^2 is an inverted gamma:

$$p(\sigma_k^2|y, z) = f_{i\gamma}(\sigma_k^2|v_k^*, s_k^*), \tag{11}$$

      – The conditional posterior density of $\mu_k|\sigma_k^2$ is a conditional normal:

$$p(\mu_k|\sigma_k^2, y, z) = f_N(\mu_k|\mu_k^*, \sigma_k^2/n_k^*), \tag{12}$$

      – The conditional posterior density of η is a Dirichlet:

$$p(\eta|y, z) = f_D(\gamma_1^0 + n_1, ..., \gamma_K^0 + n_K) \propto \prod_{k=1}^{K} \eta_k^{\gamma_k^0 + n_k - 1}, \tag{13}$$

[5]Gibbs sampler algorithms in the context of mixture models are presented in Fruhwirth-Schnatter (2001), Fruhwirth-Schnatter (2006) where is discussed, in particular, the question of label switching and its tentative solutions. Lubrano and Ndoye (2016) discuss the Gibbs sampler when applied to mixtures of log-normals and, in particular, how to select prior information.

with:

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

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

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

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

3.2 Introducing survey weights

In population studies, it is common to sample individuals through a complex sample design in which probabilities of inclusion are associated with the variables of interest. This correlation results in observations that are not identically distributed: individuals with different characteristics have different probabilities of being selected. Ignoring the sampling design could bias inference results (Pfeffermann 1996). Survey weights are constructed to correct for discrepancies between the sample and the population. The literature on the use of survey weights is not abundant, although there is a recent growing concern about this issue. Most recently developed methods in a Bayesian context (see e.g. Si et al. (2015) and Savitsky and Toth (2016) for a survey) propose the joint modelling of the weights and of the variable of interest. In this case, weights are random (this is what is considered for instance in Thuysbaert (2008) for classical inference on TIP curves). Savitsky and Toth (2016) consider exogenous weights which serves to modify the likelihood function for a subsequent Bayesian inference. The case of Bayesian inference for mixture of distributions is rather particular, because inference is rarely based on the complete likelihood function. Introducing weights in this context led to a very scare literature. Gunawan et al. (2017) propose a general method based on data augmentation where the representative data are treated as missing and are resampled from the observed sample using the weights. Instead, we propose an approach where the observations are given, the weights are exogenous and are introduced at the level of the conditional likelihood. The conditional likelihood formulated from the completed mixture problem with weights should be:

$$
\ell(y|\vartheta, z) = \prod_{k=1}^{K} \prod_{i;z_i=k} \eta_k^{w_i} f_\Lambda(y_i|\mu_k, \sigma_k^2)^{w_i} \tag{14}
$$

This is an adaptation of the method of Savitsky and Toth (2016) to the case of mixtures of log-normals where the conditional likelihood function is modified by the weights w_i, once the sample separation is made. Because $f_\Lambda(y_i|\mu_k, \sigma_k^2)$ belongs to the exponential family, this is equivalent to computing sufficient statistics using weights.

More precisely, let us consider n individuals that are sampled from the whole population with sampling weights $w_i = 1/\pi_i$ constructed to be the inverse of the inclusion probability π_i that individual i belongs to the survey. Let $\tilde{w_i}$ be the normalised weights such that $\tilde{w_i} = n w_i / \sum w_i$, so that they sum to the sample size n. Conditional inference using (14) is equivalent to using sampling weights when evaluating the sufficient statistics in (8-10), such that:

$$n_k = \sum_{i=1}^{n} \tilde{w}_i \mathbb{1}(z_i = k), \tag{15}$$

$$\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^{n} \tilde{w}_i \log(y_i) \mathbb{1}(z_i = k), \tag{16}$$

$$s_k^2 = \frac{n_k}{n_k^2 - \sum_{i=1}^{n} \tilde{w}_i{}^2 \mathbb{1}(z_i = k)} \sum_{i=1}^{n} \tilde{w}_i (\log(y_i) - \bar{y}_k)^2 \mathbb{1}(z_i = k). \tag{17}$$

Given these weighted sufficient statistics, the η_k are drawn using (13) together with a new sample separation obtained using (7). So a mixture estimate of the income distribution, representative of the original population, is obtained simply by introducing sampling weights at the level of the conditional sufficient statistics.

3.3 Modelling zero-inflated data

Another recurrent 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, such as elderly persons or unemployed workers, has no market income. As the log-normal is defined with strictly positive support, we propose to add an extra component for modelling the zero incomes:

$$f(y|\vartheta) = \begin{cases} \bar{\omega} & \text{if } y = 0, \\ (1-\bar{\omega}) \sum_{k=1}^{K} \eta_k f(y|\theta_k) & \text{if } y > 0, \end{cases} \tag{18}$$

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. The cdf corresponding to this zero-inflated mixture is:

$$F(y|\vartheta) = \begin{cases} \bar{\omega} & \text{if } y = 0, \\ \bar{\omega} + (1-\bar{\omega}) \sum_{k=1}^{K} \eta_k F(y|\theta_k) & \text{if } y > 0. \end{cases} \tag{19}$$

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

This modelling of zero-inflated data 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 sample. Figure 2 represents estimates of the German disposable income distribution for the whole population in 2009 using the GSOEP data.[6] The kernel density estimate smooths out the zero-excess observations. Worse, still 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.

[6]The GSOEP will be described in Section 5 and disposable as well as market income defined in the same section.

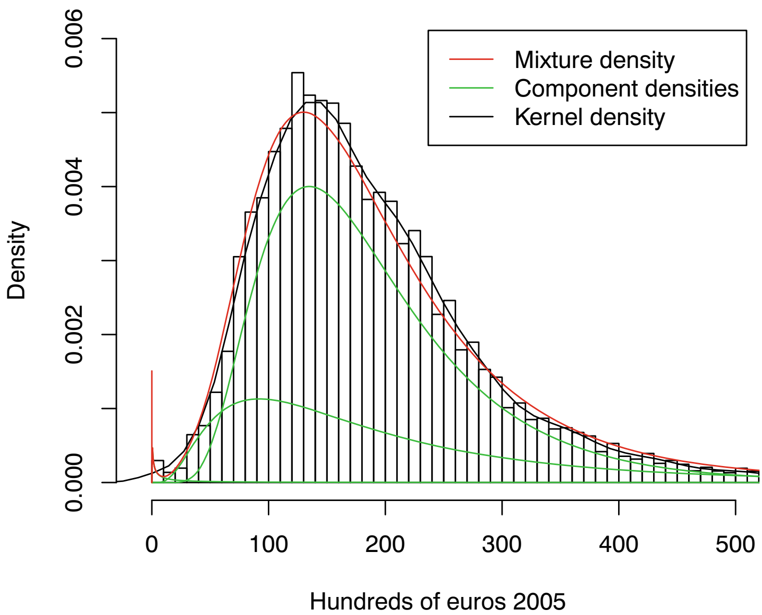

Fig. 2 The distribution of disposable household income in 2011 using weights and modelling zero incomes. Source: authors calculations using the 2011 wave of the GSOEP

4 Bayesian inference for TIP curves

In order to provide inference for TIP curves, we have to express them as a function of the model's parameters. Since the quantile function of a mixture has no closed form, we cannot generalise Eq. 4 to the case of mixtures of log-normals. We have to start from the original definition of the TIP curve given in (3), where the expression of the quantile q is left unspecified.

4.1 An alternative TIP curve formula

Assuming that the data generating process of $f(y)$ is a mixture of log-normal as given in (6), 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 (5), since q here represents the value of the p quantile of the complete mixture and the quantile of a mixture is, of course, not a linear function of the quantile of each component and has to be evaluated numerically. We can write the following expression of the TIP curve for a mixture of log-normal densities:

$$\begin{aligned} TIP_\Lambda(p, z) = \sum_{k=1}^{K} \eta_k \Bigg(\Phi \left(\frac{\ln q - \mu_k}{\sigma_k} \right) \\ - \frac{1}{z} \exp(\mu_k + \sigma_k^2/2) \Phi \left(\frac{\ln q - \mu_k - \sigma_k^2}{\sigma_k} \right) \Bigg), \end{aligned} \tag{20}$$

for $p \leq F(z)$. It turns out that the TIP curve of a mixture of log-normal densities is the weighted sum of the components' TIP curves, as given in (5), evaluated at the mixture's quantiles. As 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 numerically in q the equation $p = F(q)$. The expression of $F(y)$ is $\sum_{k=1}^{K} \eta_k F(y|\theta_k)$. We consider a grid of points p_s for $p \in [0, 1]$ and $s = 1, \ldots, S$. For each point p_s and for each value of ϑ, we solve numerically in q the equation:

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

This can be done using the `uniroot` function in `R`. This function is based on the Brent (1973)[Chapter 4]'s method, a root-finding algorithm combining the bisection method, the secant method and the inverse quadratic interpolation. Convergence is always guaranteed provided that the predefined interval contains the solution. The execution time of the Brent's algorithm is negligible.[7]

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, applying (20) 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% credible interval of the TIP curve.

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 function is changed into that given in equation (19). Moreover, we have to solve

[7] Note also the method proposed in Appendix D of Gunawan et al. (2017) which is a stochastic algorithm for computing the quantiles of a Gamma mixture.

Eq. 2 separately for $0 \le p \le \bar{\omega}$ and then for $\bar{\omega} \le p \le 1$. For the range $0 \le p \le \bar{\omega}$, we know that $y = 0$, so equation (2) becomes:

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

For the range $\bar{\omega} \le p \le 1$, we make use of Eq. 19, so the expression of the TIP curve (2) becomes:

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

and after integration:

$$\begin{aligned} TIP(p,z) \; = & \; \bar{\omega} + (1-\bar{\omega}) \sum_{k=1}^{K} \eta_k \Phi \left(\frac{\ln q - \mu_k}{\sigma_k} \right) \\ & - \frac{(1-\bar{\omega})}{z} \sum_{k=1}^{K} \eta_k \exp(\mu_k + \sigma_k^2/2) \Phi \left(\frac{\ln q - \mu_k - \sigma_k^2}{\sigma_k} \right). \end{aligned} \tag{21}$$

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

$$TIP(p,z) = \begin{cases} p & 0 \le p \le \bar{\omega} \\ \bar{\omega} + (1-\bar{\omega}) TIP_{\Lambda}(p,z) & \bar{\omega} \le p \le 1, \end{cases} \tag{22}$$

where $TIP_{\Lambda}(p,z)$ was defined in (20). For evaluating (22), the value of q has to be determined using (19).

4.3 Testing for TIP dominance

Let us consider two dominance curves $D_A^s(y)$ and $D_B^s(y)$ at the order s for the two states A and B, respectively. We say that A dominates B if $D_A^s(y) \le D_B^s(y)$. When testing for stochastic dominance (or TIP dominance), we have to compare two curves at a given number of points. Most statistical tests of stochastic dominance consider a null of dominance, see Cowell and Flachaire (2015) for a survey of the classical approach. This means that the null hypothesis $H_0 : D_A^s(y) \le D_B^s(y)$ has to be verified for all values of y while the alternative is chosen if it is valid for some values of y. It is worth mentioning that non-rejection of the null does not imply a situation of non-dominance.

Davidson and Duclos (2000, 2013) consider a test where the null is non-dominance and the alternative is dominance. In this case, we say that A does not dominate B and the null is $H_0 : D_A^s(y) \ge D_B^s(y)$ which has to be verified for some y. If the null is rejected, the alternative is dominance and is valid for all values of y. Because the test is based on finding the least favourable case, it relies on the minimum over y of the distance between the two curves. Translated into the domain of TIP dominance, this type of test would consider the minimum distance between two TIP curves computed over a grid for p.

Hypothesis testing is the domain where there is the greatest difference between the classical and the Bayesian approach. Essentially in a Bayesian framework there is no privileged hypothesis. The two hypotheses H_0 and H_1 are compared by means of the ratio between their posterior probability, the famous Bayes factor $\Pr(H_0|y)/\Pr(H_1|y)$. The main question for evaluating a Bayes factor is to find a convenient way to compute the posterior probability of a hypothesis. In our case, the probability we want to compute is:

$$\Pr[\delta(x, p|\theta) \leq 0], \qquad \forall p \in [0, 1],$$

where

$$\delta(x, p|\theta) = TIP_A(p, z|\theta_A) - TIP_B(p, z|\theta_B),$$

and θ_A and θ_B are the parameters associated to TIP curves A and B, respectively. The probability that $\delta(x, p|\theta) \leq 0$ is the probability that A TIP dominates B. We evaluate each TIP curve on a grid of fixed points, conditionally on a draw of the parameters, say $\theta_A^{(j)}$ and $\theta_B^{(j)}$. The condition $\delta(x, p|\theta) \leq 0$ defines a logical vector of zeros and ones of dimension S, the dimension of the grid. It is equivalent to check any of the three following conditions:

$$\prod_{i=1}^{S} \mathbb{1}(\delta(p_i|\vartheta) < 0) = 1, \quad \max_i \mathbb{1}(\delta(p_i|\theta) < 0) = 1, \quad \min_i \mathbb{1}(-\delta(p_i|\theta) > 0) = 1.$$

The posterior probability of this event can be evaluated easily once we have obtained m posterior draws of the parameters $\vartheta^{(j)}$ from the Gibbs sampler. More formally:

$$\begin{aligned} \Pr\left(\max_p d(p|y) < 0\right. &= \int_\vartheta \mathbb{1}\left[\max_p d(p|\vartheta) < 0\right. \; \mu(\vartheta|y)d\vartheta \\ &\simeq \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}\left[\max_p d(p|\vartheta^{(j)}) < 0\right. , \end{aligned} \tag{23}$$

where $\mu(\vartheta|y)$ is the posterior density of the parameters.[8] 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.

Because TIP dominance is equivalent to restricted second order stochastic dominance, there is no guarantee that in the case of TIP dominance there is also less poverty incidence as shown for instance in Fig. 1. We thus have to complement our TIP dominance test by a test comparing two poverty incidence measures. When evaluating $TIP_A(p, z|\theta_A^{(j)})$ and $TIP_B(p, z|\theta_B^{(j)})$, we have to compute $P_A^0 = h(z, \theta_A^{(j)})$ and $P_B^0 = h(z, \theta_B^{(j)})$ in order to determine at which point the graph of the TIP curve becomes horizontal. Testing for a lower poverty incidence in A than in B is equivalent to computing the proportion of cases where $h(z, \theta_A^{(j)}) < h(z, \theta_B^{(j)})$ for $j = 1, \ldots, m$.

Finally, when can we say that the situation in A is not statistically different from the situation in B? Equality is rejected if, for at least one value of p_s, $\delta(p_s|\vartheta)$ is statistically different from zero. This means that we have to compute a credible interval for $\delta(p_s|\vartheta)$ and see if zero is included in this interval. If we find a single p_s for which zero does not belong to a say 90% credible interval for $\delta(p_s|\vartheta)$, then we can reject at the 90% level that the two TIP curves are equal.

How do our method is related to the existing literature? Chotikapanich and Griffiths (2006) relate Lorenz dominance and stochastic dominance to parametric restrictions when the income distribution is modelled with either a Dagum or a Singh-Maddala distribution.

[8] 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 $p \in [0.01, 0.99]$, values adopted in e.g. Davidson and Duclos (2013).

On one side, they compute the posterior probability that these parametric restrictions are satisfied. On the other side, they compute the probability that two dominance curves are ordered for all values of y or all values of p when they compare Lorenz curves. In a MCMC framework, these probabilities are evaluated by counting the number of times the inequality condition is satisfied. Lander et al. (2017) take into account the necessity to have a flexible formulation for modelling the income distribution using a mixture of gamma distributions. Numerous papers have proposed statistical tests based on stochastic dominance in a classical approach, see Xu and Osberg (1998) and Berihuete et al. (2018) among others.

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 data from the *German Socio-Economic Panel* (GSOEP). They provide information on the evolution of poverty *incidence* both for children and for adults (adults are defined as those in households without children). They analyse the East-West contrast, the impact of family composition and of citizenship status. They provide information on poverty dynamics by examining the average length of poverty spells, computing probabilities of entry into poverty and exit of poverty. Finally, they compare the impact of the redistributive system between Germany and a group of English speaking countries.

However, their reporting period ends in 2004. As new data are now available, we propose to investigate the period 2000-2012 in order to check if child poverty follows the same trends as the ones depicted before or if there has been a structural break. TIP curves will also complement the portrait of poverty incidence given in Corak et al. (2008) with information about poverty *intensity* and poverty *inequality*. They will also provide information on poverty dynamics when applied to smoothed income as in Kuchler and Goebel (2003) in order to depict *chronic poverty*, as explained below.

5.1 Data and methodological issues

The GSOEP is a socio-economic panel provided by the German Institute for Economic Research in Berlin (DIW). It is a representative sample of households living in Germany since 1983 and including former East German households after the reunification. For the period 2000-2012, each wave covers on average 11 623 households. Among these, on average 3 154 households have children and their average number of children is 1.48 so that on average, the sample contains 4 680 children in each wave.

The survey is a stratified sample with weights. Cross sectional weights are introduced in order to match individual and household profiles to those of the population. Longitudinal weights are computed to account for the probability of a household to stay in the survey next year. New households are regularly added to the panel to compensate for attrition.

Two types of income can be reconstructed: an annual *market income* which represents labour income, capital income, in fact all incomes coming from a market activity; a *disposable income* which is the market income minus taxes and plus redistribution including unemployment benefits, social security pensions, family allowances and all remaining forms of social redistribution. We consider the real disposable income obtained by dividing the current income by the Consumer Price Index (2005) provided in the GSOEP.

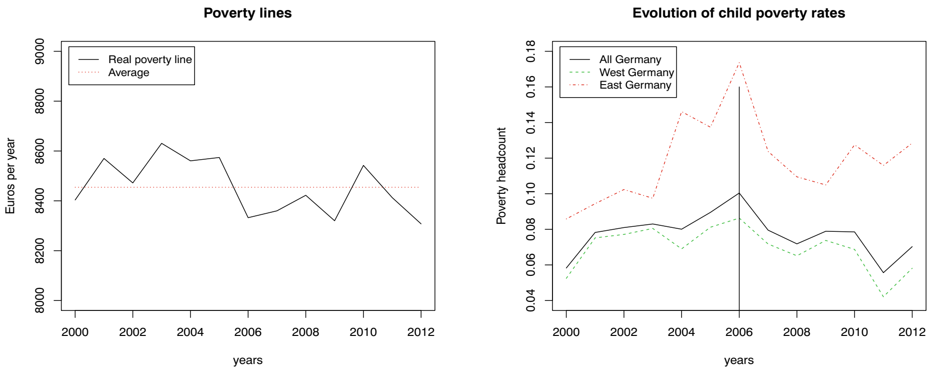

Fig. 3 How to define a poverty line and the corresponding poverty rates. (The left panel displays yearly poverty lines based on 50% of the sample median. The horizontal line is the average poverty line which is equal to 8 455 euros. It is used for computing distribution-free poverty head-counts rates in the right panel.)

In order to keep coherency with the paper of Corak et al. (2008), we define the poverty line as 50% of the sample median disposable income (taking into account all households, those having children and those without children). Remark that the disposable income is normalised by the new OECD equivalence scale and expressed in real terms.[9] Computed poverty lines are provided in the left panel of Fig. 3. The average poverty line is 8 455 euros per year and per equivalent adult over the whole period. The annual poverty line is slightly greater than this value before 2006 and slightly below after that date. But the fluctuation is less than 2%, so we have decided to keep the same poverty line over the entire period, which is more convenient to implement tests of TIP dominance.

Analysing child poverty means that we consider households with children having an equivalent disposable income below the poverty line. We define a child as a person under 18 years old. It would be misleading to consider a household as the unit of observation, ignoring the number of children in the household. This would certainly under-estimate the number of poor children in the country. A more realistic approach, followed by Hill and Jenkins (2001), Jenkins and Schluter (2003), Corak et al. (2008) and many others, consists in considering a child (and not a household) as the unit of observation, which means having possibly several observations coming from the same household. The child poverty incidence rate corresponds thus to the ratio between the number of children living in poor households over the total number of children.

We report in the right panel of Fig. 3 the evolution of child poverty incidence, distinguishing between East and West Germany. The child poverty head-count ratio follows a rising trend until 2006. This corroborates the findings of Corak et al. (2008) for the period 1999-2004. But this rising evolution stops in 2006 as from that date child poverty incidence decreases, except in East Germany after the financial crisis. This change might reflect the delayed effect of the Hartz reforms which introduced large changes in social assistance.[10]

[9]We used the cross-section household weights for computing the median. We eliminated the households which were given a zero cross-section weight in the data set.

[10]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. See the online Appendix B for more details.

We conclude that there is a clear break around 2006, which justifies considering separately two periods for dynamic analysis.

Analysing poverty dynamics has a long history starting with Bane and Ellwood (1986). They propose to study the length of poverty spells, a poverty spell; that is, the period in which a household is below the poverty line. Corak et al. (2008) analyse the length of poverty spells of households with children. They also compute the probability of entering into poverty and the probability of exiting from poverty. However, this approach has been criticised because of the bias introduced by censored spells and also because it considers only one dimension of poverty, namely poverty incidence. Rodgers and Rodgers (1993) promoted the idea that households can transfer income from one year to the other, then poverty has to be portrayed with respect to the smoothed or permanent income. It is computed at the household level over a given period, usually taking the mean income over the period as in Hill and Jenkins (2001). Then, *chronic poverty* is when the household smoothed income (as evaluated over a given period) is below the poverty line. Kuchler and Goebel (2003) pushed the analysis one step further, considering TIP curves using smoothed income to depict chronic poverty.

Analysing the dynamics of poverty requires panel data. We have to be able to follow the same household over a given period, which means that the panel has to be balanced over that period. Because the year 2012 led to huge attrition, we preferred to discard this end-of-sample year. Second, we have seen in Fig. 3 that the dynamics of poverty incidence has changed in 2006. So there is a strong interest in considering two subperiods and thus providing a separate analysis of chronic poverty for each of them. The balanced panel for the subperiod 2007-2011 covers five years and contains 2 236 children. In order to have the same number of years in the first period, we drop 2000 and 2001 so that the period 2002-2006 contains also five years and 2 991 children. We then compute two series of smoothed income obtained as individual weighted means using longitudinal weights.

5.2 The evolution of child poverty

We shall first compare TIP curves computed at the beginning of a sub-period and at the end of the same subperiod to measure the evolution of the three aspects of current poverty. If there is TIP dominance, this will give a clear indication on the evolution of at least two dimensions of poverty, intensity and inequality. Investigating the first dimension, incidence, requires a specific formal test.

The left part of Fig. 4 clearly shows that there is a significant increase in the three aspects of child poverty over the first period as a) TIP curves do not intersect, b) their 90% credible intervals do not overlap, c) the 90% credible intervals of poverty incidence also do not overlap. We have a confirmation of the results of Corak et al. (2008) who concluded to an increase of the incidence of child poverty in Germany between 1999 and 2004. But this result is extended a) to the three dimensions of child poverty and b) till the year 2006, and for all poverty lines lower than 8 455 euros.

The behaviour of child poverty during the second period (2007-2011) is just the opposite. As the graphs have the same scale, visual comparison is made easy. Current poverty has decreased significantly between the beginning and the end of that period. There is a significant reduction of child poverty in all its dimensions as all credible intervals do not overlap.

Formal dominance tests confirm these results. For current child poverty, the last line of Table 1 shows that 2011 dominates all the other reported years and that poverty incidence

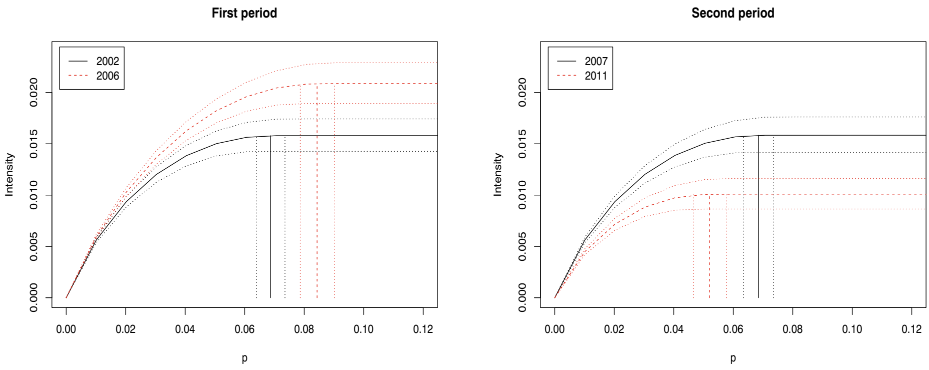

Fig. 4 The three I's of child poverty in Germany. (90% credible intervals are represented by dotted lines. In black solid line is represented the TIP curve at the beginning of each sub-period. The red dashed line corresponds to the TIP curve of the end of each subperiod.)

is the lowest in 2011. The other lines of Table 1 show that child poverty has significantly increased between 2002 and 2006 (2002 TIP dominates 2006 with probability 0.952) while it decreases after that date (2007 TIP dominates 2006 with probability 0.947 and 2011 dominates all the other years with probability 1.000). The ordering of poverty incidence is in accordance, so that we can order simultaneously the three aspects of poverty at similar degree of probability. We have also tested and found that the TIP curves in 2002 and 2007 were not statistically different. The global portrait of child poverty depicted in Corak et al. (2008) has thus completely changed.

5.3 Chronic poverty and the East-West contrast

It is much worse for a household to be in a state of chronic poverty over a long period (here five years) than being temporarily in poverty. So we shall concentrate from now on chronic poverty, computing TIP curves with smoothed income instead of current income,

Table 1 Probability of TIP dominance for current child poverty

	TIP dominance				Lower incidence			
Year	2002	2006	2007	2011	2002	2006	2007	2011
2002	–	0.952	0.401	0.000	–	0.999	0.481	0.000
2006	0.000	–	0.001	0.000	0.001	–	0.001	0.000
2007	0.384	0.947	–	0.000	0.519	0.999	–	0.000
2011	1.000	1.000	1.000	–	1.000	1.000	1.000	–

Each line represents the probability that there is less poverty in the corresponding year than in the year given in column. The first panel corresponds to TIP dominance (intensity and inequality) while the right panel indicates the probability of lower incidence

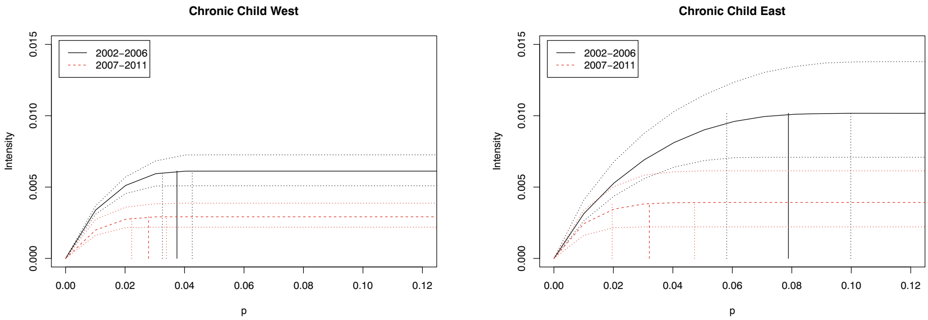

Fig. 5 The West-East contrast of chronic child poverty

following the approach of Kuchler and Goebel (2003), in order to explore 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. Corak et al. (2008) concluded that child poverty incidence was much more important in the East part of Germany.

This is confirmed for child chronic poverty during period I (2002-2006) as seen when comparing the two panels of Fig. 5. However because the credible intervals are large, there is no TIP dominance of the West over the East for the first period (the probability of TIP dominance is 0.212, but incidence is lower with a probability of 1.000 as reported from Table 2).

There is a massive reduction of chronic child poverty during period II (2007-2011), both in East and West Germany as seen from the two panels of Fig. 5. As a matter of fact, we observe TIP dominance of period II over period I for the two regions (left part of Table 2) and also a significant lower child poverty incidence (right part of Table 2). This massive reduction of chronic child poverty has erased the differences between the two regions concerning chronic child poverty. The two TIP curves were tested not to be statistically different between the two regions (the two red dashed curves in the two panels of Fig. 5). We conclude that the redistributive system has been very efficient during the second period for fighting against chronic child poverty and after the Hartz reforms came into force.

Table 2 TIP dominance test for child chronic poverty between West and East Germany

		West		East		West		East	
		I	II	I	II	I	II	I	II
West	I	–	0.001	0.212	0.042	–	0.017	1.000	0.267
	II	0.997	–	0.989	0.668	0.982	–	1.000	0.647
East	I	0.019	0.000	–	0.003	0.000	0.000	–	0.001
	II	0.917	0.166	0.842	–	0.733	0.353	0.999	–

Each column represents the probability that there is less poverty for the category indicated in line. The first panel corresponds to TIP dominance (intensity and inequality) while the right panel indicates the probability of lower incidence. Period I corresponds to 2002-2006 and Period II to 2007-2011.

6 Conclusion

We have provided Bayesian inference for TIP curves. Specifically, we have proposed parametric modelling of the income distribution using a mixture of log-normal densities. We also solved two questions raised by the use of survey data: incorporating survey weights and taking into account explicitly zero-income data. Once we have obtained random draws of the income distribution parameters from the posterior distributions, the TIP curves are a (not so) simple transformation of these draws, which means that we have direct access to statistical inference (both credible intervals and testing). In doing so, we propose a solution for computing the quantiles of a mixture model.

Using TIP curves, we gave an empirical portrait of child poverty in Germany, which is complementary to that of Corak et al. (2008), both for the period considered and for the tools we used. We showed that poverty head-counts rates have to be completed by other dimensions and that TIP curves are very convenient for this issue. In some cases, the poverty incidence ordering is consistent with the poverty ordering in the other dimensions, but in other cases, this is not so clear. We provided an example with adult poverty reported in Table 3 of the online Appendix C.

Child poverty continued to follow the upward trend found in Corak et al. (2008) and that up to 2006. However, after 2006, the portrait of child poverty in Germany has totally changed with a decrease in both current and chronic poverty and a reduction of the gap between adult and child poverty. The gap between East and West chronic child poverty was also much reduced, becoming not significant. There has thus been a convergence on this point between the two parts of Germany. Meanwhile, adult chronic poverty (as documented in online Appendix C) has increased a lot in East Germany while it was decreasing in the West. We might see there a consequence of the structural changes introduced by the Hartz plan (2003-2005) and the related reforms.

Acknowledgements 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, Christian Schluter, Luc Bauwens, the editor and an anonymous referee for a careful reading of a previous version of the paper and for their insightful suggestions. Of course usual disclaimers apply. This work has been carried out thanks to the support of the A*MIDEX project (No ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR).

References

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

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

Bane, M.J., Ellwood, D.T.: Slipping into and out of poverty: the dynamics of spells. J. Hum. Resour. **21**(1), 1–23 (1986)

Barrett, G.F., Donald, S.G., Hsu, Y.-C.: Consistent tests for poverty dominance relations. J. Econ. **191**(2), 360–373 (2016)

Berihuete, A., Ramos, C., Sordo, M.: Welfare, inequality and poverty analysis with rtip: An approach based on stochastic dominance. R J. **10**, 328–341 (2018)

Bradbury, B., Jenkins, S., Micklewright, J.: The dynamics of child poverty in seven industrialised nations. In: Bradbury, B., Jenkins, S.P., Micklewright, J. (eds.) The Dynamics of Child Poverty in Industrialised Countries, chapter 4, pp. 92–132. Cambridge University Press, Cambridge (2001)

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

Bruckmeier, K., Wiemers, J.: A new targeting: a new take-up? Empir. Econ. **43**(2), 565–580 (2012)

Chotikapanich, D., Griffiths, W.: Bayesian assessment of Lorenz and stochastic dominance in income distributions. In: Creedy, J., Kalb, G. (eds.) Dynamics of Inequality and Poverty, volume 13 of Research on Economic Inequality, pages 297–321. Elsevier (2006)

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

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

Cowell, F.: Measuring Inequality. Oxford University Press, 3rd edn (2011)

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

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

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

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

Flachaire, E., Nunez, O.: Estimation of the income distribution and detection of subpopulations: an explanatory model. Comput. Stat. Data Anal. **51**(7), 3368–3380 (2007)

Foster, J., Shorrocks, A.: Poverty orderings. Econometrica **56**(1), 173–177 (1988)

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

Fruhwirth-Schnatter, S.: Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. J. Am. Stat. Assoc. **96**(453), 194–209 (2001)

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

Gunawan, D., Panagiotelis, A., Griffiths, W., Chotikapanich, D.: Bayesian Weighted Inference from Surveys, Technical report, University of Melbourne (2017)

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

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

Jenkins, S., Lambert, P.: Ranking poverty gap distributions: Further TIPs for poverty analysis. In: Slottje, D. (ed.) Research on Economic Inequality, volume 8, pages 31–38. JAI Press Inc (1998a)

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

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

Jenkins, S.P., Schluter, C.: Why are child poverty rates higher in Britain than in Germany? A longitudinal perspective. J. Hum. Resour. **38**(2), 441–465 (2003)

Konigs, S.: 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 (2014)

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

Lander, D., Gunawan, D., Griffiths, W., Chotikapanich, D.: Bayesian Assessment of Lorenz and Stochastic Dominance. Technical report, University of Melbourne (2017)

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

Ndoye, A.A.J., Lubrano, M.: Tournaments and superstar models: A mixture of two Pareto distributions. In: Bishop, J.A., Rodriguez, J.G. (eds.) Economic Well-Being and Inequality: Papers from the Fifth ECINEQ Meeting, chapter 21, pp. 449–479. Emerald Group Publishing Limited (2014)

Nygård, M., Campbell-Barr, V., Krüger, N.: The financial crisis and recent family policy reforms in finland, Germany and the United Kingdom : Is there a connection? Journal of Comparative Social Work **8**(2), 1–33 (2015)

Pfeffermann, D.: The use of sampling weights for survey data analysis. Stat. Methods Med. Res. **5**(3), 239–261 (1996)
Rodgers, J., Rodgers, J.: Chronic poverty in the United States. J. Hum. Resour. **28**(1), 25–54 (1993)
Savitsky, T.D., Toth, D.: Bayesian estimation under informative sampling. Electronic Journal of Statistics **10**, 1677–1708 (2016)
Sen, A.: Poverty: an ordinal approach to measurement. Econometrica **44**(2), 219–231 (1976)
Shorrocks, A.: Revisiting the Sen poverty index. Econometrica **63**(5), 1225–1230 (1995)
Si, Y., Pillai, N.S., Gelman, A.: Bayesian nonparametric weighted sampling inference. Bayesian Anal. **10**(3), 605–625 (2015)
Sordo, M., Ramos, C.: Poverty comparisons when TIP curves intersect. Stat. Oper. Res. Trans. **35**(1), 65–80 (2011)
Thuysbaert, B.: Inference for the measurement of poverty in the presence of a stochastic weighting variable. J. Econ. Inequal. **6**(1), 33–55 (2008)
Xu, K., Osberg, L.: A distribution free tests for deprivation dominance. Econ. Rev. **17**, 415–429 (1998)