# Bayesian Assessment of Lorenz and Stochastic Dominance

David Lander[*] David Gunawan[†]
William Griffiths[‡] Duangkamon Chotikapanich[§]

June 4, 2019

**Abstract**

We introduce a Bayesian approach for assessing Lorenz and stochastic dominance. For two income distributions, say $X$ and $Y$, estimated via Markov chain Monte Carlo, we describe how to compute posterior probabilities for (i) $X$ dominates $Y$, (ii) $Y$ dominates $X$ and (iii) neither $Y$ nor $X$ is dominant. The proposed approach is applied to Indonesian income distributions using mixtures of gamma densities that ensure flexible modelling. Probability curves depicting the probability of dominance at each population proportion are used to explain changes in dominance probabilities over restricted ranges relevant for poverty orderings. They also explain some seemingly contradictory outcomes from the $p$-values of some sampling theory tests.

**JEL codes:** C11, C12, D31.
**Keywords:** dominance probabilities, poverty comparisons, MCMC, gamma mixture.

[*]Affiliation: Peking University HSBC Business School.
[†]Affiliation: University of New South Wales.
[‡]Affiliation: University of Melbourne. Correspondence: wegrif@unimelb.edu.au.
[§]Affiliation: Monash University.

# 1 Introduction

Statistical tests of dominance regularly appear in the economics and finance literature. They employ statistical methods to compare the distributions of two random variables (or one random variable at two points in time) in such a way as to determine if one "dominates" another. These tests have been used to compare income distributions in welfare analysis, distributions of asset returns in portfolio analysis and risk analyses in actuarial science. We focus on Lorenz, stochastic and poverty dominance in income distributions and, in contrast to the existing literature, we propose a Bayesian approach to assessing dominance.[1]

The theoretical foundations for using Lorenz and stochastic dominance to analyse social welfare were laid by Atkinson (1970) and Shorrocks (1983). Among other things, it was shown that, for the class of monotonically increasing concave, but otherwise arbitrary, income utility functions, generalized Lorenz dominance (which is equivalent to second-order stochastic dominance) implies, and is implied by, a social welfare ordering. Robustness of stochastic dominance criteria to changes in the functional form of the social welfare function allows practitioners to avoid specifying a functional form for the utility or poverty function, and provides a compelling reason for its use in analyses of social welfare. Other inequality measures, such as the Gini coefficient and Atkinson's inequality index are useful when income distributions cannot be ordered according to stochastic dominance criteria, but they involve placing more restrictive assumptions on the functional form of the social welfare function or poverty index.[2]

Empirical studies of stochastic, Lorenz and poverty dominance have typically been undertaken in a sampling-theory framework using nonparametric hypothesis tests. The literature is vast. It includes papers concerned with testing Lorenz dominance[3], papers specifically concerned with testing poverty dominance[4], and many others focused on stochastic dominance testing or all three forms of dominance[5]. Extensions to multivariate scenarios have also been

[1] Applications of stochastic dominance to asset returns in portfolio analysis can be found in Wong et al. (2008), Sriboonchitta et al. (2009), and Bai et al. (2011); for applications in actuarial sciences and risk analysis, see Kaas et al. (1994) and Denuit et al. (2005). Maasoumi and Millimet (2005) use stochastic dominance tests to examine trends in environmental quality.

[2] Details of these various concepts, and the relationships between them, can be found in Atkinson and Bourguignon (1982), Chakravarty (2009), Davidson and Duclos (2000), Le Breton and Peluso (2009), Lambert (2001), and Maasoumi (1997).

[3] Beach and Davidson (1983), Bishop et al. (1991), Dardanoni and Forcina (1999), Schluter and Trede (2002), and Barrett et al. (2014).

[4] Tabri (2015) and Barrett et al. (2016).

[5] Bishop et al. (1989), McFadden (1989), Kaur et al. (1994), Bishop et al. (1995), Anderson (1996), Davidson and Duclos (2000, 2013), Maasoumi and Heshmati (2000, 2008), Barrett and Donald (2003), Linton et al. (2005), Horváth et al. (2006), Linton et al. (2010), Berrendero and Cárcamo (2011), Bennett (2013), and Donald and Hsu (2016).

proposed.[6]

We propose a novel Bayesian approach to assessing dominance, and illustrate how it can be applied to income distribution data. There are a number of differences between existing sampling theory approaches and our proposed Bayesian approach. Sampling theory approaches begin with specification of null and alternative hypotheses, with some tests specifying $X$ dominates $Y$ (say) as the null hypothesis, and others using $X$ does not dominate $Y$ as the null. In addition, to exhaust all possibilities, these hypotheses are often reversed, specifying that $Y$ does or does not dominate $X$. A test statistic and its limiting distribution are derived, and results are reported in terms of $p$-values. In our Bayesian approach we begin with flexible parametric specifications for income distributions $X$ and $Y$ and compute posterior probabilities of dominance as the probability that the parameters of the income distributions lie in the constrained region implied by dominance inequalities. We compute the probability that $X$ dominates $Y$, the probability that $Y$ dominates $X$, and the probability that neither distribution is dominant. This general approach—computing posterior probabilities of parameters lying within constrained regions—has been used by Geweke (1988) to assess explosive and oscillatory behavior of real per capita GDP for 19 OECD countries, and by Geweke (1986) in a variety of other contexts.[7]

A major difference between our Bayesian approach and a typical frequentist test is in the interpretation of the results. The Bayesian approach treats each of the three possible outcomes symmetrically, assigning a posterior probability to each of them. A frequentist test can be less conclusive. To see why, consider the following three scenarios. (1) For the first case, suppose a null hypothesis that $X$ dominates $Y$ is rejected; then, the hypotheses are reversed, and a null hypothesis that $Y$ dominates $X$ is also rejected. These outcomes constitute strong evidence that there is no dominance. The Bayesian probability for no dominance would be high and both approaches would be conclusive. (2) Now, suppose that we reject a null hypothesis that $X$ dominates $Y$ but fail to reject a null hypothesis that $Y$ dominates $X$. It is tempting to conclude that these outcomes constitute evidence that $Y$ dominates $X$. However, a large $p$-value, leading to non-rejection of $Y$ dominates $X$, should not be interpreted as evidence that this null hypothesis is necessarily true. It may be true that $Y$ does dominate $X$, but it also may be the case that neither $X$ nor $Y$ is dominant. All we can say is that there is insufficient evidence to establish that $Y$ does not dominate $X$.[8] In our empirical example, we encounter

[6] Duclos et al. (2006), McCaig and Yatchew (2007), and Bennett and Mitra (2013).

[7] Another approach to Bayesian analysis of first and second order stochastic dominance has been proposed by Hasegawa (2013). He develops 95% credible bounds around the difference between Bayesian nonparametric estimates of the relevant curves.

[8] When working with time series, failure to reject a null hypothesis of a unit root is typically used as evidence of nonstationarity. The null hypothesis is assumed to be true. This practice can be justified by the heavy cost associated

a situation where there is a posterior probability of one that neither distribution is dominant when a frequentist test rejects dominance in one direction but not the other. (3) Another example where a frequentist test is likely to be less conclusive is where $X$ dominates $Y$, say, but the curves for $X$ and $Y$ are close together, with the difference between them insufficiently large for a null hypothesis of $Y$ dominates $X$ to be rejected. In this case dominance in both directions would not be rejected by a frequentist test, but the posterior probability that $X$ dominates $Y$ would be greater than 0.5.

Potential disadvantages of our proposal for assessing dominance are the dependence of the posterior probabilities for dominance on how the income distribution is modelled through the likelihood function, and the sensitivity of the results to the prior information that is placed on the unknown parameters. These dependencies can be minimized, however. Subjectivity from specification of prior information can be limited by using relatively uninformative priors and is unlikely to have a major impact given the relatively large sample sizes that are typically used to estimate income distributions. The choice of likelihood function can be more critical. In preliminary work, two of us (Chotikapanich and Griffiths, 2006) found results that were sensitive to a choice between the Dagum and Singh–Maddala income distributions, and concluded that a relatively flexible likelihood function is necessary for robust results. To minimize the impact of the likelihood function, in this paper we choose a more flexible likelihood function: a mixture of gamma densities. Based on an algorithm proposed by Wiper, Insua, and Ruggeri (2001) (hereafter WIR), Chotikapanich and Griffiths (2008) described how to obtain a Bayesian estimate of an income density modelled as a mixture of gamma densities, and its corresponding distribution function and Lorenz curve, but they did not consider dominance. In the current paper we combine the idea in Chotikapanich and Griffiths (2006) with the estimation method in Chotikapanich and Griffiths (2008). In addition to mitigating the inflexibility of the parametric specifications in the first paper, we provide insights into the differences between the Bayesian and frequentist approaches and overcome some computational challenges introduced by using gamma mixtures.

Another important difference between the Bayesian and frequentist approaches is that the Bayesian approach is parametric, whereas the frequentist approach is nonparametric, the latter relying on empirical distribution functions and Lorenz curves. A consequence of this difference is that, conditional on the gamma mixture being a suitable parametric choice, the Bayesian approach is more precise, particularly in the tails of the distribution, often leading to more conclusive outcomes than those from frequentist tests. This advantage of the Bayesian parametric approach is potentially no longer an advantage if the gamma mixture assumption is

with assuming a nonstationary time series is stationary. On the other hand, the cost of assuming dominance when a null hypothesis of dominance cannot be rejected could be high, particularly if the poor people are worse off.

unreasonable. The mixture provides a great deal of flexibility, and it is an excellent fit in our empirical work, but its robustness cannot be guaranteed.

The added flexibility achieved by a mixture makes the analysis more complex than that considered by Chotikapanich and Griffiths (2006). We are confronted with the problem of choosing an arbitrary number of points to assess dominance across the $(0, \infty)$ interval, or having to employ numerical methods to invert the distribution function of the mixture so that the quantile function can be used and the analysis restricted to the $(0, 1)$ interval. We propose numerical methods to resolve this issue, and introduce several refinements to earlier work. The techniques are illustrated using Indonesian household income distributions for the years 1999, 2002, 2005 and 2008.

The remainder of the paper is structured as follows. Conditions for dominance are stated in Section 2; the proposed method for Bayesian assessment is described in Section 3. In Section 4 we introduce a gamma mixture model and describe how parameter draws from the joint posterior density of its parameters can be used to find corresponding values for the quantile function, the generalized Lorenz curve and the Lorenz curve for the mixture of gamma densities. A Markov chain Monte Carlo (MCMC) algorithm for drawing observations from the joint posterior density for the parameters is summarized in Appendix A. Appendix B describes a numerical algorithm for inverting the distribution function of the gamma mixture to obtain its quantiles. Data and estimation results for some Indonesian income distributions are considered in Section 5. Dominance results and a comparison with some sampling theory results are presented in Section 6; some concluding remarks are offered in Section 7.

## 2 Dominance conditions

To introduce the dominance conditions, we consider an income distribution $X$ that is described by density and distribution functions $p_X(x)$ and $F_X(x)$, respectively, with finite mean income $\mu_X = E(X)$. Several expressions have been used in the literature for a Lorenz curve that gives the proportion of total income earned by the poorest proportion $u$ of the population. One that is useful for noting the equivalence of generalized Lorenz dominance with second-order stochastic dominance (see below) is

$$L_X(u) = \frac{1}{\mu_X} \int_0^u F_X^{-1}(t)\, dt, \qquad 0 \le u \le 1. \tag{1}$$

A convenient way to compute ordinates of a Lorenz curve is to begin with a population proportion $u$ $(0 \le u \le 1)$, compute the corresponding value of income $x$ from the quantile function, $x = F_X^{-1}(u)$, and to then compute the corresponding income proportion using the first-moment distribution function for $X$ defined as $F_X^{(1)}(x) = (1/\mu_X) \int_0^x t\, p_X(t)\, dt$. Using this approach, we write the Lorenz curve as

$$L_X(u) = F_X^{(1)}\left[F_X^{-1}(u)\right]. \tag{2}$$

We say that an income distribution for $X$ Lorenz dominates (LD) a distribution for $Y$ (say), expressed as $X \succeq_{\text{LD}} Y$, if and only if

$$L_X(u) \geq L_Y(u) \quad \text{for all } 0 \leq u \leq 1, \quad \text{and} \quad L_X(u) > L_Y(u) \quad \text{for some } 0 < u < 1. \tag{3}$$

While this definition is the typical one used in the economics literature (see, for example, Lambert (2001) and Barrett, Donald, and Bhattacharya (2014); BDD), the definition used in much of the statistics literature follows the opposite convention, with $L_Y(u) \geq L_X(u)$ being the condition for $X \succeq_{\text{LD}} Y$. See, for example, Kleiber and Kotz (2003). Since $L_X(u) \geq L_Y(u)$ implies higher welfare for distribution $X$ in the sense that, other things equal, less inequality is preferred to more inequality, we refer to this condition as one where $X$ dominates $Y$. This is the case when the Lorenz curve of $X$ lies nowhere below that of $Y$ for all population proportions $u$. For two income distributions with the same mean income, Lorenz dominance implies greater utility with respect to all strictly increasing and concave social welfare functions.

Because Lorenz dominance considers only the degree of inequality and not the level of income, generalized Lorenz dominance (GLD) was introduced to recognize that higher levels of income are associated with higher levels of welfare. We say that $X$ generalized-Lorenz dominates $Y$, written as $X \succeq_{\text{GLD}} Y$, if and only if

$$\mu_X L_X(u) \geq \mu_Y L_Y(u) \quad \text{for all } 0 \leq u \leq 1, \tag{4}$$

with strict inequality holding for some $0 < u < 1$. For all strictly increasing and concave social welfare functions, GLD provides an unambiguous ranking of distributions when the generalized Lorenz curves do not intersect. Given the expression for the Lorenz curve in equation (1), the condition in (4) can also be expressed as

$$\int_0^u F_X^{-1}(t)\, dt \geq \int_0^u F_Y^{-1}(t)\, dt \quad \text{for all } 0 \leq u \leq 1, \tag{5}$$

with strict inequality holding for some $0 < u < 1$. Condition (5) can be viewed as the sum of the incomes of the bottom $u$ proportion in $X$ being at least as great as the corresponding sum for $Y$ for any population proportion $u$. Writing the relation for GLD in this way demonstrates its equivalence to second-order stochastic dominance (SSD). See, for example, Maasoumi (1997) or Kleiber and Kotz (2003, p. 25).

A stronger condition for welfare improvement than SSD (or equivalently, GLD) is that of first-order stochastic dominance (FSD). The distribution for $X$ first-order stochastically dominates $Y$, written $X \succeq_{\text{FSD}} Y$, if and only if

$$F_X^{-1}(u) \geq F_Y^{-1}(u) \quad \text{for all } 0 \leq u \leq 1, \tag{6}$$

with strict inequality holding for some $0 < u < 1$. In this case the level of income from distribution $X$ is greater than the level of income from distribution $Y$ for all population proportions $u$. First-order stochastic dominance implies greater utility for all social welfare functions that are strictly increasing. Concavity is no longer required.

While we focus on computing posterior probabilities for Lorenz, first, and second-order stochastic dominance, our computations can be extended to third and higher-order stochastic dominance criteria. Third-order stochastic dominance is relevant for ordering distributions for social welfare functions where the marginal utility function is positive, decreasing, and strictly convex. See, for example, Chakravarty (2009) and Le Breton and Peluso (2009). The higher the degree of dominance, the greater the number of restrictions that need to be imposed on the social welfare function and the weaker the dominance inequality. That is, first-order stochastic dominance implies second-order stochastic dominance, and so on, but the converse is not true. Thus, if $X \succeq_{\text{FSD}} Y$, the generalized Lorenz curve of $X$ will also lie everywhere above $Y$.

Foster and Shorrocks (1988) showed that unidimensional stochastic dominance conditions are equivalent to unidimensional poverty orderings. Let $z$ denote the "poverty line". First-order poverty dominance, in terms of the quantile function of the distribution, is defined as $F_X^{-1}(u) \geq F_Y^{-1}(u)$ for all $0 \leq u \leq F(z)$, with strict inequality for some values of $u$. This is equivalent to the statement that the level of income of individuals is always (weakly) greater in distribution $X$ than in $Y$ up to population proportion given by $F(z)$. Similarly, second-order poverty dominance is defined as $\int_0^u F_X^{-1}(t)\,dt \geq \int_0^u F_Y^{-1}(t)\,dt$ for all $0 \leq u \leq F(z)$, with strict inequality for some values of $u$. Davidson and Duclos (2000) show that poverty dominance at order $s$ and poverty line $z_s$ implies poverty dominance at order $s+1$ for a poverty line $z_{s+1} > z_s$. Thus, there exists a higher order of dominance such that dominance is achieved for any finite $z$. Barrett, Donald, and Hsu (2016) consider testing for poverty dominance in terms of poverty gap profiles—see, for example, Jenkins and Lambert (1997). For a given poverty line, this dominance criterion is equivalent to restricted generalized Lorenz dominance.

## 3 Bayesian assessment of dominance relationships

To assess dominance within a Bayesian framework, we compute the posterior probability that the relevant inequalities, those in equations (3), (4) and (6), hold. This approach, namely, computing posterior probabilities that parameter values satisfy constrained regions, has been used in other contexts by Geweke (1986, 1988). It and other Bayesian approaches to comparing hypotheses are described in Zellner (1971, Ch. 10) and Gill (2015, Ch. 17). Hasegawa (2013) used Bayesian credible intervals around curve differences to assess dominance.

To describe how dominance probabilities are computed, first suppose we have two parametric distributions, each with *known* parameter values. To check each form of dominance (LD, GLD or FSD), we can compute $L(u)$, $\mu L(u)$ and $F^{-1}(u)$ for both distributions for a grid of values for $u$ in the interval $(0, 1)$. If the grid contains a relatively large number of values, and the dominance inequality being considered is satisfied for all those values, then it is reasonable to conclude that the condition is satisfied for all $u$, and hence dominance holds. Now suppose the distributions for $X$ and $Y$ have *unknown* parameter vectors $\theta_X$ and $\theta_Y$ that are estimated using income distribution data. Since these parameters are not known with certainty, any conclusion about whether one distribution dominates another cannot be made with certainty. In Bayesian inference, uncertainty about whether one distribution dominates another, or whether one function is greater than another at a particular point, can be expressed in terms of a probability statement. To obtain or estimate such probability statements, assume we have draws on $\theta_X$ and $\theta_Y$ from the two posterior densities $p(\theta_X \mid \mathbf{x})$ and $p(\theta_Y \mid \mathbf{y})$. Taking FSD as an example, it will be useful to distinguish between $\Pr[F_X^{-1}(u) \geq F_Y^{-1}(u)]$ for a given value of $u$ and $\Pr[F_X^{-1}(u) \geq F_Y^{-1}(u)]$ for all values of $0 < u < 1$. An estimate of the first probability is given by the proportion of draws of $\theta_X$ and $\theta_Y$ for which $F_X^{-1}(u) \geq F_Y^{-1}(u)$ for a given $u$. We will denote this probability by $\Pr[X \succeq_{\text{FSD}} Y \mid u]$. The probability $X \succeq_{\text{FSD}} Y$ for the range $(0, 1)$ is given by the proportion of values $\theta_X$ and $\theta_Y$ for which the inequality holds *for all values of $u$*. We denote this probability by $\Pr[X \succeq_{\text{FSD}} Y]$. In practice we can consider a grid of $u$ values within the interval $(0, 1)$ and count the number of parameter draws where the inequality holds for all $u$ in the grid. Since

$$\Pr[X \succeq_{\text{FSD}} Y] \leq \min_u \Pr[X \succeq_{\text{FSD}} Y \mid u], \tag{7}$$

a finer grid can be taken in the region that counts: those values of $u$ where $\Pr[X \succeq_{\text{FSD}} Y \mid u]$ reaches a minimum. Similar probability statements can be made for LD and GLD.

In our empirical application we used 999 values of $u$ between 0.001 and 0.999 with increments of 0.001. Graphing $\Pr[X \succeq_{\text{FSD}} Y \mid u]$ against $u$—curves that we call "probability curves"—provides useful information about the range(s) of the income distribution that have the largest impact on the probability of dominance. The inequality in (7) means that a probability curve is a powerful tool for visually determining the probability of dominance over the whole range of population proportions as well as within sub-intervals of $u$. Restricted intervals such as poverty dominance can be assessed, along with the sensitivity of poverty dominance to specification of a poverty line.

Having found $\Pr[X \succeq_{\text{FSD}} Y]$ we can reverse the process to find $\Pr[Y \succeq_{\text{FSD}} X]$. The probability that neither $X$ nor $Y$ dominates is given by $1 - \Pr[X \succeq_{\text{FSD}} Y] - \Pr[Y \succeq_{\text{FSD}} X]$.

Suppose that we have a sequence of $M$ draws on $\theta_X$ and a sequence of $M$ draws on $\theta_Y$, and we assess probability using $M$ pairwise comparisons. Estimates of the dominance probabilities may change with different orderings of the draws on $\theta_X$ and $\theta_Y$. To ensure that the probability of dominance is robust to the order of the $M$ parameter draws, we randomly rearrange the order of the draws and repeat the procedure $C$ times. This gives a set of $C$ dominance probabilities, $\Pr(X \succeq_{\text{FSD}} Y)_t$ for $t = 1, 2, \ldots, C$, and a set of $C$ probability curves. We use the mean probability of dominance, $\overline{\Pr(X \succeq_{\text{FSD}} Y)} = \sum_{t=1}^{C} \Pr(X \succeq_{\text{FSD}} Y)_t / C$, as our estimate of the probability of dominance. We will also define the maximum and minimum values of the set of $C$ dominance probabilities as the upper and lower bounds, respectively, for the dominance probability estimate. These bounds give an indication of how sensitive the dominance probability estimate is to the choice of pairwise ordering procedure.

## 4 Estimation and dominance assessment with a gamma mixture

To minimize the dependence of the posterior probabilities for dominance on the assumed parametric functional form for the income distributions, we use a mixture of gamma densities which has the ability to approximate well any functional form. Our specification and estimation procedure follows that proposed by WIR (2001) and used subsequently by Chotikapanich and Griffiths (2008). Other mixtures, such as a log-normal mixture would also provide a good approximation and could be used within the same framework.[9] An income distribution modelled as a gamma mixture with $K$ components can be written as

$$p(x \mid \boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{w}) = \sum_{k=1}^{K} w_k G\left(x \mid v_k, \frac{v_k}{\mu_k}\right), \tag{8}$$

where $x$ is a random income draw from the probability density function (pdf) $p(x \mid \boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{w})$,

[9] One potential disadvantage of a mixture of lognormal densities is that, when the traditional noninformative prior is used for the parameters, the posterior predictive density for a future income observation, obtained by integrating out the parameters, is a mixture of "log-$t$ distributions" which do not have finite moments.

with parameter vectors, $\boldsymbol{w} = (w_1, w_2, \ldots, w_K)'$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)'$, and $\boldsymbol{v} = (v_1, v_2, \ldots, v_K)'$. The pdf $G(x \mid v_k, v_k/\mu_k)$ is a gamma density with mean $\mu_k > 0$ and shape parameter $v_k > 0$,

$$G\left(x \mid v_k, \frac{v_k}{\mu_k}\right) = \frac{(v_k/\mu_k)^{v_k}}{\Gamma(v_k)} x^{v_k - 1} \exp\left(-\frac{v_k}{\mu_k} x\right). \tag{9}$$

Including the mean $\mu_k$ as one of the parameters in the pdf makes the parameterization in (9) different from the standard textbook one, but it is convenient for later analysis. Details of the prior pdfs and the MCMC algorithm proposed by WIR (2001) for drawing observations from the posterior pdf $p(\boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{w} \mid \boldsymbol{x})$ where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ is a random sample from (8), are given in Appendix A.

Having obtained MCMC draws $(\boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$, $m = 1, 2, \ldots, M$ from the posterior distribution $p(\boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{w} \mid \boldsymbol{x})$, we can proceed to set up the machinery to assess various types of dominance. Associated with each draw, we have a gamma mixture density $p(x \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$ that can be obtained from (8). If interest centers on the predictive density for $x$, it can be obtained by averaging $p(x \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$ over all the parameter draws. However, for assessing dominance our interest centers on (i) the Lorenz curves $L_X(u \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$, (ii) the generalized-Lorenz curves

$$GL_X\left(u \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)}\right) = \mu^{(m)} L_X\left(u \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)}\right), \tag{10}$$

with $\mu^{(m)} = \sum_{k=1}^{K} w_k^{(m)} \mu_k^{(m)}$, and (iii) the quantile functions $F^{-1}(u \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$, for each of the parameter draws $m = 1, 2, \ldots, M$.[10]

The first challenge is to obtain a value of the quantile function $F^{-1}(u \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$ for a given proportion $u$, and a given set of parameters $(\boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$. Inverting the distribution function for a mixture of gamma densities is not straightforward. A closed form solution cannot be obtained; solving the required nonlinear equation numerically for a large number of MCMC draws is not only time consuming, but also fraught with dangers of non-convergence. To overcome these problems, we developed the algorithm described in Appendix B.

The quantile function values can be used to assess first-order stochastic dominance. For assessing Lorenz and generalized-Lorenz dominance using equations (2), (3) and (4), we also need to evaluate the first moment distribution function evaluated at quantiles $\widehat{y}$ computed from

[10]There is a large literature on the inadequacies of many MCMC algorithms for mixture models because of the problem of label switching. Geweke (2007) has pointed out that, if interest centres on functions which are invariant to permutations of the components of the mixture, rather than on the parameters of the mixture components, then "simple MCMC works". The three functions $L$, $GL$, and $F^{-1}$ are all invariant with respect to permutations of the components.

the algorithm in Appendix B. Dropping the $(m)$ superscript for convenience, the first-moment distribution function for the mixture of gamma densities is

$$
\begin{aligned}
F_X^{(1)}(\hat{y} \mid \boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{w}) &= \frac{1}{\mu} \int_0^{\hat{y}} x\, p(x \mid \boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{w})\, dx \\
&= \frac{1}{\mu} \sum_{k=1}^{K} w_k \int_0^{\hat{y}} x\, G\left( x \mid v_k, \frac{v_k}{\mu_k} \right) dx \\
&= \frac{1}{\mu} \sum_{k=1}^{K} w_k \int_0^{\hat{y}} \frac{(v_k/\mu_k)^{v_k}}{\Gamma(v_k)} x^{(v_k+1)-1} \exp\left( -\frac{v_k x}{\mu_k} \right) dx \\
&= \frac{1}{\mu} \sum_{k=1}^{K} w_k \int_0^{\hat{y}} \frac{\mu_k (v_k/\mu_k)^{v_k+1}}{\Gamma(v_k+1)} x^{(v_k+1)-1} \exp\left( -\frac{v_k x}{\mu_k} \right) dx \\
&= \frac{1}{\mu} \sum_{k=1}^{K} w_k \mu_k F_k\left( \hat{y} \mid v_k + 1, \frac{v_k}{\mu_k} \right).
\end{aligned}
\tag{11}
$$

where $F_k(\cdot \mid v_k + 1, v_k/\mu_k)$ is the distribution function of a gamma density with parameters $(v_k + 1)$ and $v_k/\mu_k$.

## 5 Data and estimation results

To illustrate the methodology, we use household income data obtained for Indonesian urban regions for the years 1999, 2002, 2005, and 2008. These data were created using household expenditure data obtained from the National Socio-Economic Survey.[11] Appropriate price and equivalence-scale adjustments were made to ensure comparability over the years. Summary statistics for the data are presented in Table 1. The units are thousands of rupiah per month.

If one makes judgments about the ordering of the distributions based on the means or medians, the population becomes better off as it moves from 1999 to 2005, but drops back in 2008, a likely consequence of the global financial crisis. Similarly, using the standard deviations and Gini coefficients to measure inequality suggests inequality increases from 1999 to 2005, but then decreases in 2008. Assuming a welfare function where mean income contributes positively, and the standard deviation or Gini contributes negatively, it is not clear whether a distribution in any one year is preferable. Comparing the mean and median incomes, and checking the means against the maximum values, reveals extremely long right tails; a small number of households have very high incomes. For example, in 2002 the mean income was 432, the maximum income was 24,903, and the proportion of households with incomes greater than 2,000 was only 0.0074.

To choose the number of mixture components, we estimated densities up to a maximum of

[11] The authors are grateful to Ari Handayani for providing the data used in Handayani (2013).

Table 1: Sample statistics (in Rp '000 per month)

| | 1999 | 2002 | 2005 | 2008 |
|---|---|---|---|---|
| Sample mean | 332.63 | 432.36 | 477.04 | 454.08 |
| Median | 270.40 | 337.80 | 357.95 | 355.27 |
| Minimum | 44.12 | 57.65 | 38.32 | 59.84 |
| Maximum | 5,973.92 | 24,902.67 | 30,216.51 | 13,181.99 |
| Standard deviation | 249.32 | 477.29 | 511.19 | 401.23 |
| Gini coefficient | 0.3184 | 0.3509 | 0.3797 | 0.3583 |
| Sample size | 25,175 | 29,280 | 24,687 | 26,648 |
| Proportion more than 2000 | 0.00242 | 0.0074 | 0.0134 | 0.0099 |

**Note:** Strictly speaking, the "income" distributions are consumption distributions based on surveys of household expenditure. These data are urban household expenditure data obtained from the Indonesian National Socio-Economic Survey for the years 1999, 2002, 2005, and 2008. They are adjusted for inflation using the CPI calculated by BPS Statistics Indonesia with 2002 as the base, and converted to per adult equivalent expenditure using the equivalence scale $m = (n_a + \phi n_c)^{\lambda}$, where $n_a$ and $n_c$ are the numbers of adults and children, respectively, and $\phi = \lambda = 0.8$. This equivalence scale was suggested by Banks and Johnson (1994) and Jenkins and Cowell (1994). The Canadian dollar to Indonesian rupiah exchange rate for 2002, supplied by the Penn World Tables, is 5933.34.

five components. Initial MCMC samples of 15,000 were drawn, of which 500 were discarded as burn-in.[12] Maximum likelihood (ML) estimates for the Singh–Maddala and Dagum distributions were also obtained as a benchmark for comparison with the gamma mixtures. For each of the four years, the following three goodness-of-fit criteria were used to compare the estimated distribution functions $\widehat{F}(x_j)$, with the empirical distributions $F_0(x_j) = j/n$, where $j$ refers to the $j$th observation after ordering from lowest to highest. The $\widehat{F}(x_j)$ were computed at the posterior means for the gamma mixture components and at the ML estimates for the Singh–Maddala and Dagum distributions.

$$\begin{aligned} d &= \max_j \left| \widehat{F}(x_j) - F_0(x_j) \right|, \\ \text{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{F}(x_i) - F_0(x_i)\right)^2}, \\ \text{MAE} &= \frac{1}{n}\sum_{i=1}^{n}\left|\widehat{F}(x_i) - F_0(x_i)\right|. \end{aligned} \tag{12}$$

The results are presented in Table 2. They provide unambiguous support for a gamma mixture with four components. It is superior to all other models for every measure of goodness of fit, and for all years. There is also strong support for the choice of a mixture model over the commonly used Singh–Maddala and Dagum distributions.

A possible disadvantage of the three goodness-of-fit criteria that we employed is that they

[12] Settings for the prior parameters are given in Appendix A.

Table 2: Goodness-of-fit comparisons

| | $K=1$ | $K=2$ | $K=3$ | $K=4$ | $K=5$ | Dagum | S–M |
|---|---|---|---|---|---|---|---|
| | | | | 1999 | | | |
| RMSE | 0.0478 | 0.0140 | 0.0059 | 0.0037 | 0.0080 | 0.0084 | 0.0125 |
| MAE | 0.0414 | 0.0120 | 0.0050 | 0.0030 | 0.0078 | 0.0076 | 0.0111 |
| $d$ | 0.0766 | 0.0258 | 0.0199 | 0.0083 | 0.0119 | 0.0130 | 0.0197 |
| | | | | 2002 | | | |
| RMSE | 0.0559 | 0.0189 | 0.0065 | 0.0046 | 0.0086 | 0.0079 | 0.0117 |
| MAE | 0.0491 | 0.0157 | 0.0055 | 0.0037 | 0.0081 | 0.0069 | 0.0103 |
| $d$ | 0.0853 | 0.0333 | 0.0116 | 0.0107 | 0.0131 | 0.0128 | 0.0198 |
| | | | | 2005 | | | |
| RMSE | 0.0556 | 0.0172 | 0.0046 | 0.0023 | 0.0086 | 0.0079 | 0.0122 |
| MAE | 0.0488 | 0.0145 | 0.0040 | 0.0018 | 0.0081 | 0.0067 | 0.0106 |
| $d$ | 0.0864 | 0.0325 | 0.0093 | 0.0062 | 0.0132 | 0.0151 | 0.0209 |
| | | | | 2008 | | | |
| RMSE | 0.0472 | 0.0156 | 0.0047 | 0.0026 | 0.0034 | 0.0112 | 0.0133 |
| MAE | 0.0414 | 0.0134 | 0.0041 | 0.0021 | 0.0030 | 0.0097 | 0.0114 |
| $d$ | 0.0726 | 0.0269 | 0.0086 | 0.0064 | 0.0068 | 0.0191 | 0.0224 |

**Note:** $K$ refers to the number of components in a mixture of gamma distributions. Dagum and S–M refer to the Dagum and Singh–Maddala distributions, respectively. The goodness-of-fit criteria compare the empirical distribution function values with their counterparts estimated from the alternative distributions. RMSE is the root-mean squared error of the differences, MAE is the mean absolute value of the differences, and $d$ is the maximum difference.

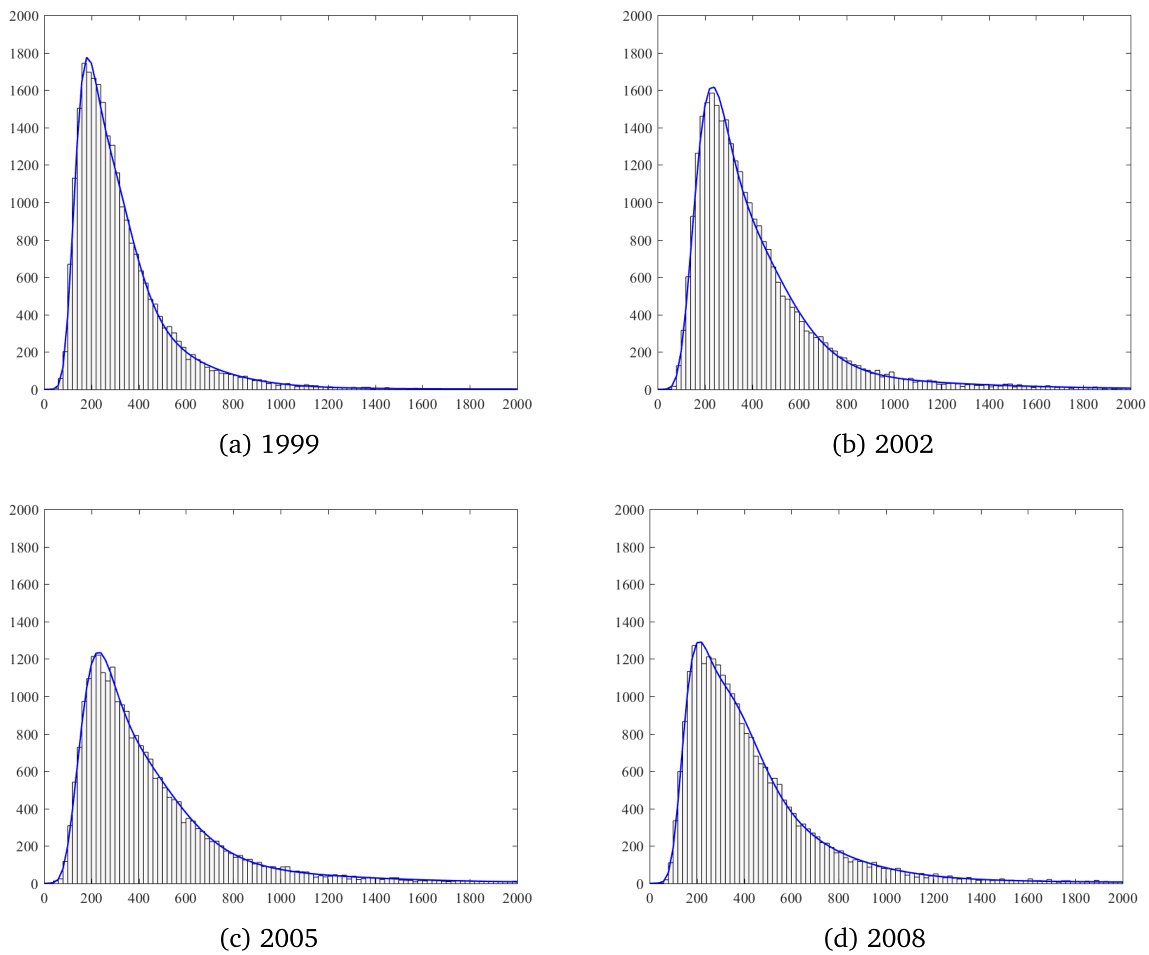
do not include a penalty for extra parameters.[13] To check whether $K = 3$ components might be preferred to $K = 4$ components when we introduce a penalty for the extra parameters, we used the Bayesian Information Criteria (BIC). ML estimates were obtained for all four years for both $K = 3$ and $K = 4$. In all cases reducing the number of components from four to three led to an increase in the BIC, confirming our choice of $K = 4$.

To obtain the draws used for dominance assessment in Section 6, we re-estimated the four-component mixture with 50 different MCMC chains. For each chain, 11,000 draws were generated, the first 1,000 draws were discarded as a burn-in, and only every 50th draw of the subsequent 10,000 draws was retained. With 50 chains and 200 retained observations from each chain, a total of $M = 10,000$ independent draws were retained for subsequent analysis. All parameters showed evidence of convergence. A relatively small number of independent MCMC draws was chosen in preference to a large number of correlated draws to reduce computing time for inverting the distribution functions and carrying out the dominance comparisons. The estimated densities for the four-component mixtures are presented together with corresponding histograms in Figure 1. The shape of each distribution has been captured

[13] See Frühwirth-Schnatter (2006, Ch. 4) for a review of criteria for choosing the number of mixture components.

well by the four-component mixtures, confirming the ability of the mixture distribution to capture the essential features of each distribution. Posterior means and standard deviations for each of the parameters for all components and all years are given in Table 3, alongside ML estimates and their standard errors. As expected with the large sample sizes, the ML estimates and standard errors are very similar to the posterior means and standard deviations. In general, the weight for the last component is very small, but the goodness-of-fit criteria suggest it is important for capturing the upper tails of the distributions.

Figure 1: Histograms and estimated densities for household expenditure




**Note:** The histograms and density estimates are for per adult equivalent household expenditure in thousands of 2002 rupiah. The density estimates are calculated using the posterior means of the parameters of the four-component mixtures of gamma distributions. These posterior means are given in Table 3.

Table 3: Bayesian and maximum likelihood estimates of mixture parameters and weights

| | Component | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | |
| | Bayes | ML | Bayes | ML | Bayes | ML | Bayes | ML |
| 1999 | | | | | | | | |
| $\mu$ | 176.5 | 176.9 | 294.4 | 295.0 | 515.4 | 515.8 | 1296.6 | 1306.3 |
| | (4.8) | (4.7) | (10.3) | (10.09) | (25.6) | (25.0) | (138.7) | (138.0) |
| $w$ | 0.2497 | 0.2501 | 0.4976 | 0.5015 | 0.2361 | 0.2328 | 0.0166 | 0.0156 |
| | (0.0366) | (0.0364) | (0.0336) | (0.0339) | (0.0332) | (0.0328) | (0.0036) | (0.0035) |
| $v$ | 13.989 | 13.763 | 8.155 | 8.004 | 5.041 | 4.918 | 2.193 | 2.118 |
| | (1.194) | (1.124) | (0.534) | (0.504) | (0.430) | (0.397) | (0.350) | (0.317) |
| 2002 | | | | | | | | |
| $\mu$ | 234.9 | 235.0 | 430.5 | 430.6 | 918.7 | 920.7 | 4141.8 | 4243.2 |
| | (3.8) | (3.7) | (8.8) | (8.3) | (40.0) | (39.6) | (637.1) | (645.7) |
| $w$ | 0.3466 | 0.3466 | 0.5383 | 0.5399 | 0.1105 | 0.1094 | 0.0045 | 0.0042 |
| | (0.0256) | (0.0242) | (0.0245) | (0.0232) | (0.0111) | (0.0108) | (0.0009) | (0.0009) |
| $v$ | 9.999 | 9.955 | 6.107 | 6.056 | 3.462 | 3.402 | 1.235 | 1.197 |
| | (0.436) | (0.414) | (0.275) | (0.257) | (0.277) | (0.264) | (0.234) | (0.218) |
| 2005 | | | | | | | | |
| $\mu$ | 232.9 | 233.6 | 441.1 | 442.9 | 902.0 | 912.0 | 2558.0 | 2634.0 |
| | (4.6) | (4.5) | (11.7) | (11.1) | (57.0) | (53.6) | (313.5) | (328.9) |
| $w$ | 0.3073 | 0.3082 | 0.5277 | 0.5327 | 0.1492 | 0.1452 | 0.0157 | 0.0139 |
| | (0.0281) | (0.0273) | (0.0308) | (0.0295) | (0.0200) | (0.0181) | (0.0038) | (0.0034) |
| $v$ | 9.229 | 9.140 | 5.593 | 5.476 | 3.348 | 3.212 | 1.556 | 1.510 |
| | (0.486) | (0.472) | (0.374) | (0.318) | (0.337) | (0.299) | (0.229) | (0.223) |
| 2008 | | | | | | | | |
| $\mu$ | 198.4 | 198.2 | 361.0 | 360.8 | 625.5 | 620.4 | 1411.7 | 1410.4 |
| | (5.7) | (5.3) | (15.7) | (14.1) | (41.2) | (33.5) | (95.1) | (82.3) |
| $w$ | 0.2014 | 0.2007 | 0.4674 | 0.4678 | 0.2800 | 0.2810 | 0.0512 | 0.0505 |
| | (0.0264) | (0.0256) | (0.0424) | (0.0404) | (0.0468) | (0.0428) | (0.0079) | (0.0071) |
| $v$ | 12.665 | 12.546 | 7.420 | 7.345 | 5.225 | 5.112 | 2.210 | 2.194 |
| | (0.989) | (0.922) | (0.583) | (0.524) | (0.525) | (0.458) | (0.169) | (0.164) |

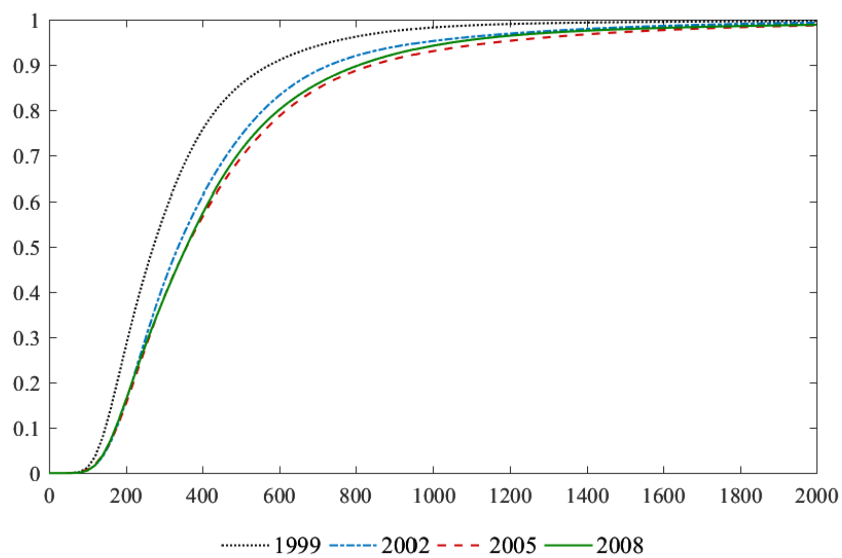
**Note:** The Bayesian estimates are the posterior means estimated from 10,000 independent MCMC draws. To achieve independence, 50 chains, each with 10,000 post burn-in draws, were run; every 50th draw was selected from each chain. Estimates of the posterior standard deviations, obtained from the same draws, are in parentheses below the posterior means. The maximum likelihood estimates were obtained using EViews' Berndt–Hall–Hall–Hausman outer-product gradient algorithm with Marquardt steps. Standard errors in parentheses below the estimates were computed using the observed Hessian.

# 6 Dominance results

## 6.1 Visual inspection

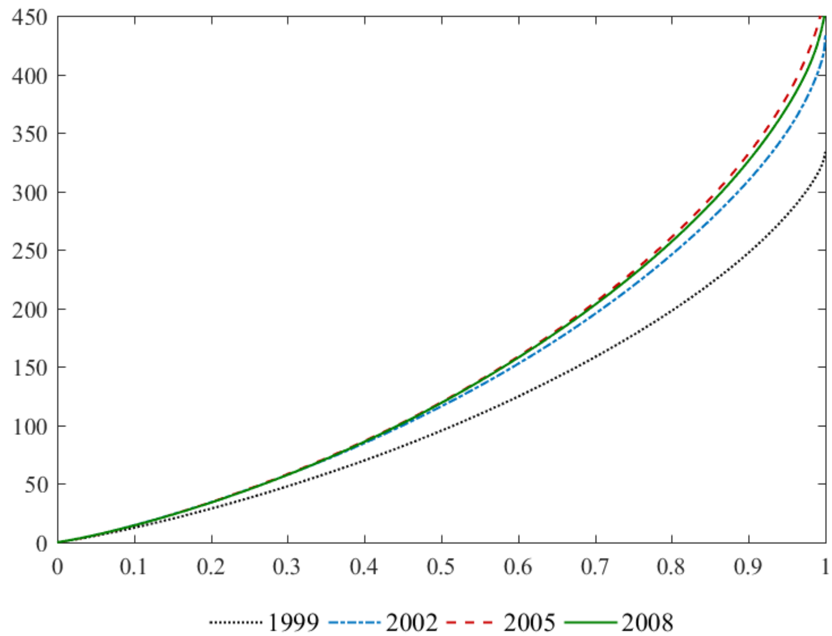
We begin a discussion of the dominance results by visually comparing the estimated distribution functions (Figure 2), the generalized-Lorenz curves (Figure 3), and the Lorenz curves (Figure 4) for each of the four years. In Figure 2, it appears that the distributions in 2002, 2005 and 2008 all FSD the 1999 distribution. However, it seems that there is no clear ranking between 2002, 2005, and 2008. In Figure 3 the generalized-Lorenz curves for 2002, 2005, and 2008 are also everywhere above the 1999 distribution, consistent with the fact that FSD implies GLD. A similar but less clear remark can be made about the Lorenz curves in Figure 4. Here the 1999 distribution Lorenz dominates 2002, 2005, and 2008, except for grid points where $u$ is close to zero or one.

Figure 2: Estimated distribution functions



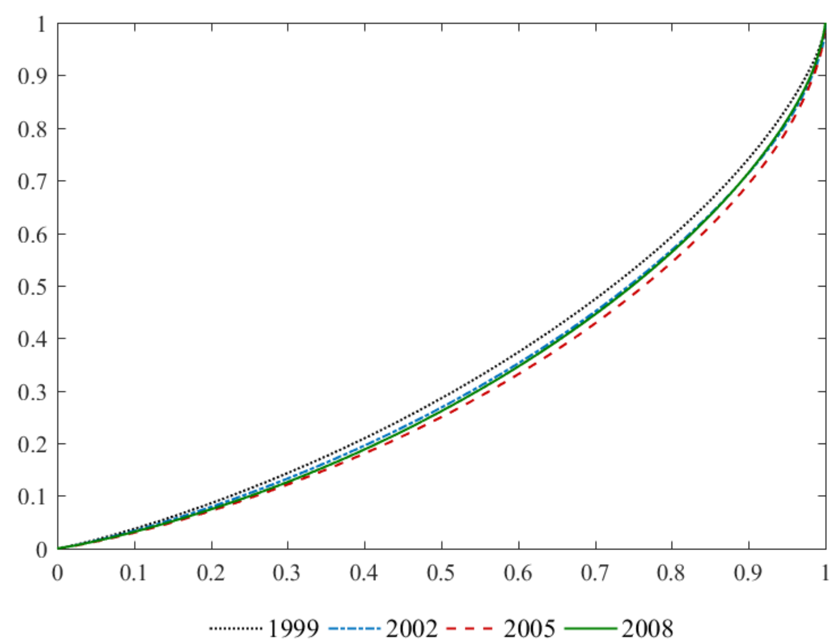

**Note:** Distribution functions for four-component gamma mixtures with parameters set at their posterior means. The distribution for 1999 appears to be first-order stochastically dominated by the distributions in the other years. The relationships between the other distributions is less clear.

Figure 3: Estimated generalized-Lorenz curves




**Note:** Generalized Lorenz curves for four-component gamma mixtures with parameters set at their posterior means. The distribution for 1999 appears to be second-order stochastically dominated by the distributions in the other years. The relationships between the other distributions is less clear.

Figure 4: Estimated Lorenz curves




**Note:** Lorenz curves for four-component gamma mixtures with parameters set at their posterior means. There is some evidence that the distribution for 1999 Lorenz dominates the distributions in other years.

## 6.2 Dominance probabilities

When comparing the four years of income distributions, there are six possible pairwise comparisons that can be made. For each comparison, we calculate the probability that $X$ dominates $Y$, the probability that $Y$ dominates $X$, and the corresponding upper and lower bounds for the estimated probabilities. These probabilities are reported in Table 4, along with bounds computed using $C = 1000$. The values in the table can be used to determine the probability that neither $X$ nor $Y$ dominates, given by $1 - \Pr(X \succeq_{\mathrm{D}} Y) - \Pr(Y \succeq_{\mathrm{D}} X)$. In each of the three segments in the table, the last two columns provide evidence of how the income distribution has changed from one period to the next. We consider each of these pairs of years in turn.

Table 4: Dominance probabilities, with bounds in parentheses

| | 08 D 05 | 05 D 08 | | | | |
|---|---|---|---|---|---|---|
| FSD | 0.0000<br>(0.0000,0.0000) | 0.0001<br>(0.0000,0.0004) | | | | |
| GLD | 0.0000<br>(0.0000,0.0000) | 0.0001<br>(0.0000,0.0005) | | | | |
| LD | 1.0000<br>(0.9997,1.0000) | 0.0000<br>(0.0000,0.0000) | | | | |
| | 08 D 02 | 02 D 08 | 05 D 02 | 02 D 05 | | |
| FSD | 0.0000<br>(0.0000,0.0001) | 0.0000<br>(0.0000,0.0000) | 0.0001<br>(0.0000,0.0004) | 0.0000<br>(0.0000,0.0000) | | |
| GLD | 0.0148<br>(0.0117,0.0183) | 0.0000<br>(0.0000,0.0000) | 0.0002<br>(0.0000,0.0006) | 0.0000<br>(0.0000,0.0000) | | |
| LD | 0.0000<br>(0.0000,0.0000) | 0.0000<br>(0.0000,0.0000) | 0.0000<br>(0.0000,0.0000) | 0.0037<br>(0.0021,0.0056) | | |
| | 08 D 99 | 99 D 08 | 05 D 99 | 99 D 05 | 02 D 99 | 99 D 02 |
| FSD | 0.9980<br>(0.9966,0.9991) | 0.0000<br>(0.0000,0.0000) | 0.2529<br>(0.2432,0.2624) | 0.0000<br>(0.0000,0.0000) | 0.9943<br>(0.9923,0.9962) | 0.0000<br>(0.0000,0.0000) |
| GLD | 0.9997<br>(0.9989,1.0000) | 0.0000<br>(0.0000,0.0000) | 0.4676<br>(0.4595,0.4772) | 0.0000<br>(0.0000,0.0000) | 0.9990<br>(0.9979,0.9998) | 0.0000<br>(0.0000,0.0000) |
| LD | 0.0000<br>(0.0000,0.0000) | 0.3331<br>(0.3247,0.3421) | 0.0000<br>(0.0000,0.0000) | 0.1915<br>(0.1829,0.2020) | 0.0000<br>(0.0000,0.0000) | 0.9996<br>(0.9988,1.0000) |

**Note:** Dominance probabilities are calculated as the proportion of MCMC draws for which the relevant inequalities are satisfied for all population proportions between 0.001 and 0.999. The bounds are the minimum and maximum values obtained from 1000 random rearrangements of the MCMC draws. FSD, GLD and LD refer to first-order stochastic dominance, generalized Lorenz dominance and Lorenz dominance, respectively. For each pairwise comparison, dominance probabilities for both directions are reported. These probabilities can be used to find the probability that neither distribution is dominant. For example, when comparing 2002 with 2008 for FSD, there is a probability of one that neither dominates. When comparing 1999 with 2005 for FSD, there is a probability of $(1 - 0.2529) = 0.7471$ that neither dominates.

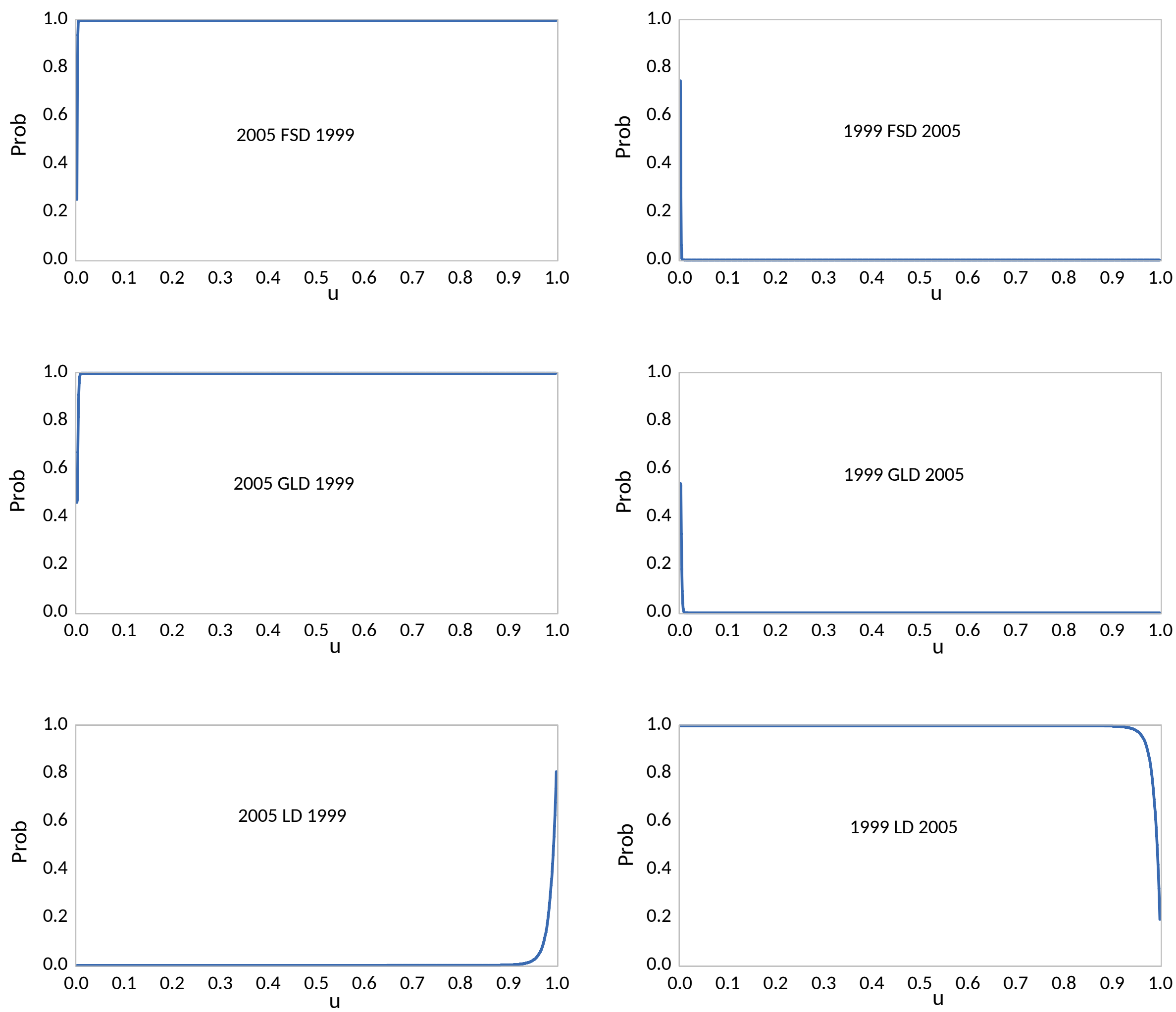
For 1999 and 2002, the dominance probabilities in the table make clear predictions. In terms of FSD and GLD, the probability that 2002 dominates 1999 is greater than 0.99. While for LD, the probability that 1999 dominates 2002 is 0.9996. The level of welfare measured by FSD and the level of relative inequality described by LD have increased significantly from 1999 to 2002. With respect to GLD, we can conclude that growth over the period has been sufficiently large to compensate for the increase in inequality, making social welfare unambiguously greater in 2002. In contrast, when comparing 2002 and 2005, we find the probability that neither year is dominant is very close to one for FSD, GLD and LD. There is only a very low probability (0.0037) that 2002 dominates 2005 in terms of LD. We obtain a similar result for FSD and GLD when comparing 2005 to 2008; the probability that neither year is dominant is very close to one. However, in terms of LD, the probability that 2008 dominates 2005 is close to one. There are also positive probabilities (0.9996, 0.1915, and 0.3331) that 1999 Lorenz dominates 2002, 2005, and 2008, respectively. These results lead us to conclude that the level of inequality has increased from 1999 to 2005, and decreased from 2005 to 2008. The only clear improvement in terms of FSD and GLD is from 1999 to 2002 (0.9990), and 1999 to 2008 (0.9997). It is somewhat surprising that the FSD and GLD probabilities of 2005 over 1999 (0.2529 and 0.4676, respectively) are far less than those for 2002 over 1999 (both over 0.99). Since both the mean and median for 2005 are greater than those for 2002, we might expect the FSD and GLD probabilities for 2005 over 1999 to be greater than those for 2002 over 1999. Another surprising result comes from comparing the Lorenz dominance probabilities and the estimated Gini coefficients for 2002, 2005 and 2008. The Gini coefficient for 2005 (0.3797) is higher than those for both 2002 (0.3509) and 2008 (0.3583). Hence, we may expect similar LD probabilities for 2008 over 2005, and 2002 over 2005. However, the probability 2008 Lorenz dominates 2005 is one, while the probability 2002 Lorenz dominates 2005 is only 0.0037. These examples illustrate the usefulness of using posterior probabilities in addition to summary statistics from the distributions. They provide us with some evidence that there may be subtle changes occurring in different parts of the distribution over time that are not captured by the summary statistics. These aspects can be investigated further by using probability curves to assess the probability of dominance at different sub-intervals of the population proportions.

### 6.3 Using probability curves

Recall that the probability curves provide us with the probability of dominance at each population proportion $u$. Thus, we can use these curves to see which population proportions have the greatest effect on the probability of dominance, or lack of it. Furthermore, we can show how the probability of dominance will change if we restrict our focus to a particular segment of the population, such as the poorest 10% or 20%. The probability of dominance within the restricted range will necessarily be less than or equal to the minimum value of the probability curve for that restricted range. Additionally, the probability curve for dominance in one direction is the mirror image of the probability curve of dominance in the other direction. For example, $\Pr(2002 \succeq_{\text{FSD}} 1999 \mid u)$ is the mirror image of $\Pr(1999 \succeq_{\text{FSD}} 2002 \mid u)$.

The probability curves for 2005 over 1999 are given in Figure 5. They show that the

Figure 5: Estimated probability curves for FSD, GLD, and LD comparing 1999 and 2005

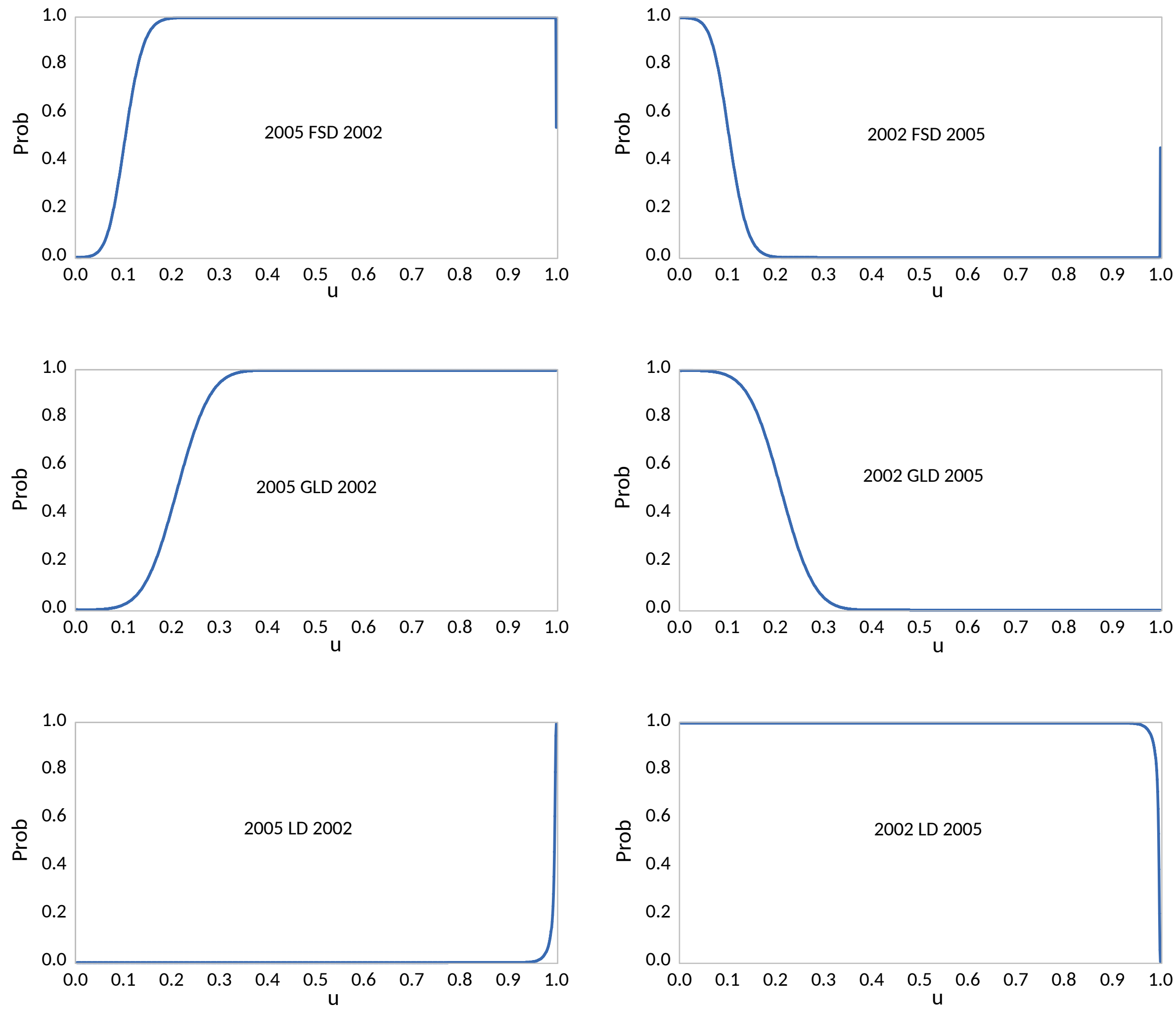



**Note:** The graphs show the probability that one distribution dominates another at a given population proportion $u$. The probability of (unconditional) dominance is less than or equal to the minimum point on each graph. $\Pr(2005 \succeq_{\text{FSD}} 1999) = 1$ if the poorest 0.4% of the population are ignored. For $\Pr(2005 \succeq_{\text{GLD}} 1999) = 1$ the poorest 0.9% of the population need to be ignored. For $\Pr(1999 \succeq_{\text{LD}} 2005) = 1$, the richest 12% of the population have to be ignored.

FSD and GLD probabilities for 2005 over 1999 are very sensitive to the starting point of the population proportion $u$. If we restrict attention to $u \geq 0.005$, then, instead of being only 0.25 and 0.47, both dominance probabilities are greater than 0.95, bringing them much in line with the FSD and GLD probabilities for 2002 over 1999. Thus, the probability curves enable us to isolate the cause of significant differences between the dominance probabilities. The results suggest that, between 2002 and 2005, the incomes of the poorest 0.5% of the sample must have declined sufficiently such that we are unable to establish either a FSD or GLD relationship for

2005 over 1999. Such a precise statement cannot be made using only the summary statistics.

Figure 6: Estimated probability curves for FSD, GLD, and LD comparing 2002 and 2005

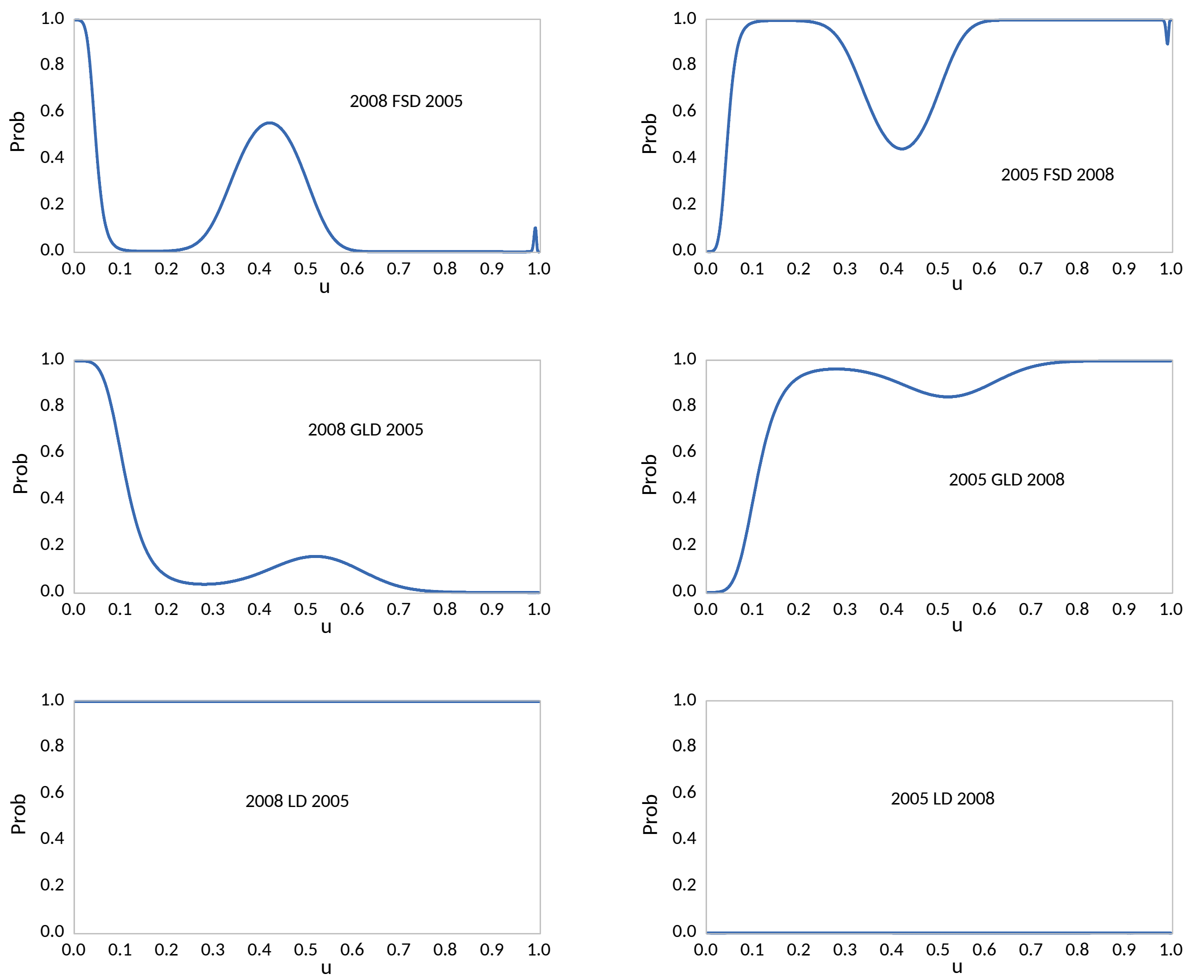



**Note:** The graphs show the probability that one distribution dominates another at a given population proportion $u$. The probability of (unconditional) dominance is less than or equal to the minimum point on each graph. For both FSD and GLD, the probability of one of the distributions dominating the other is zero. However, if we focus on only the poorest 10%, we find $\Pr(2002 \succeq_{\text{FSD}} 2005) = 0.533$ and $\Pr(2002 \succeq_{\text{GLD}} 2005) = 0.975$. For LD, $\Pr(2002 \succeq_{\text{LD}} 2005) = 1$ if we ignore the richest 0.6% of the population.

To identify the cause of the other surprising result, we can view the probability curves for 2002 over 2005 and 2008 over 2005, given in Figures 6 and 7, respectively. If we restrict the analysis to approximately $u < 0.98$, instead of being only 0.0037, the probability that 2002 Lorenz dominates 2005 is greater than 0.9, bringing it more in line with the probability of one for 2008 Lorenz dominating 2005. We can use this information to infer several interesting things. The richest 2% of the sample in 2002 must be sufficiently richer than the richest 2% in

2008 to prevent us from making a definitive statement about how income inequality changed from 2002 to 2005. In other words, we cannot establish a Lorenz dominance relationship for 2002 over 2005 *only* because of the level of incomes of the richest 2%. This captures an important aspect of the distributions that we could not establish simply through the use of Gini coefficients as summary statistics. It is, however, consistent with another summary statistic; the maximum income for 2002 (24,903) is almost double that of 2008 (13,182).

Figure 7: Estimated probability curves for FSD, GLD, and LD comparing 2005 and 2008




**Note:** The graphs show the probability that one distribution dominates another at a given population proportion $u$. The probability of (unconditional) dominance is less than or equal to the minimum point on each graph. For both FSD and GLD, the probability of one of the distributions dominating the other is zero. However, if we focus on only the poorest 10%, we find $\Pr(2008 \succeq_{\text{FSD}} 2005) = 0.011$ and $\Pr(2008 \succeq_{\text{GLD}} 2005) = 0.591$. For Lorenz dominance, we have $\Pr(2008 \succeq_{\text{LD}} 2005) = 1$.

### 6.4 Tail behavior

In the last subsection, we encountered two examples where results for dominance probabilities depended critically on outcomes in the tails of the distributions. The range of population proportions considered was 0.001 to 0.999. Changing this range, in one case to 0.01 to 0.99, and in the other to 0.02 to 0.98, had a dramatic effect on the probability of dominance. This raises two questions. (1) Do we regard the outcomes in the tails as important? (2) If so, are we estimating the tails of the curves accurately? The answer to the first question will depend on the research objectives. If we are prepared to base conclusions on the middle 99% or 98% of the population, then, at least for these two examples, accurate estimation in the tails is of limited importance. For situations where the tails are important, it is useful to consider whether they have been estimated accurately, and whether results for those tails are real or imagined.

The $(0, 1)$ endpoints for the distribution functions and the Lorenz curves will be "estimated" perfectly. As different curves approach these endpoints, they converge. Variances of estimates at these points (or posterior standard deviations in our context), will approach zero. However, if the curves converge more quickly than the standard deviations go to zero, both positive and negative differences between the curves will arise from the MCMC parameter draws. An outcome such as this can explain why a dominance probability of one, obtained by ignoring the tails, can change considerably when the tails are included. Also, it is likely that the nature of the outcome will be sensitive to the parametric specifications. By using mixture distributions, we hope to minimise the sensitivity, but we cannot guarantee robustness. If there is concern about the legitimacy of results in the extreme tails, it may be better to base conclusions on a central, but still substantial, portion of the population.

To check goodness of fit in the tails of the four-component gamma mixtures, in Table 5 we report *RMSE*, *MAE* and $d$ for distribution function estimates for the poorest 1%, 5%, 10% and 20%, and the richest 10%, 5% and 2% of the population. The poorest 1% and richest 2% relate to the two examples considered in the previous subsection. The larger percentages relate to regions of restricted dominance considered in the next subsection, and to apparent inconsistencies with frequentist tests considered in Subsection 6.6. Except for *MAE* for the bottom 20% in 2005, all fits in the tails are at least as good as the overall fits, and, in most cases, much better.

Table 5: Goodness-of-fit comparisons for selected proportions of the population

| | Bottom 1% | Bottom 5% | Bottom 10% | Bottom 20% | Top 10% | Top 5% | Top 2% | All obs. |
|---|---|---|---|---|---|---|---|---|
| | | | | 1999 | | | | |
| RMSE | 0.0011 | 0.0012 | 0.0021 | 0.0026 | 0.0022 | 0.0006 | 0.0004 | 0.0037 |
| MAE | 0.0009 | 0.0010 | 0.0018 | 0.0023 | 0.0017 | 0.0004 | 0.0004 | 0.0030 |
| $d$ | 0.0017 | 0.0026 | 0.0044 | 0.0050 | 0.0040 | 0.0018 | 0.0009 | 0.0083 |
| | | | | 2002 | | | | |
| RMSE | 0.0010 | 0.0008 | 0.0011 | 0.0026 | 0.0012 | 0.0014 | 0.0003 | 0.0046 |
| MAE | 0.0010 | 0.0007 | 0.0010 | 0.0022 | 0.0010 | 0.0012 | 0.0003 | 0.0037 |
| $d$ | 0.0013 | 0.0017 | 0.0027 | 0.0046 | 0.0024 | 0.0024 | 0.0010 | 0.0107 |
| | | | | 2005 | | | | |
| RMSE | 0.0010 | 0.0010 | 0.0016 | 0.0023 | 0.0016 | 0.0015 | 0.0005 | 0.0023 |
| MAE | 0.0010 | 0.0009 | 0.0014 | 0.0021 | 0.0012 | 0.0013 | 0.0005 | 0.0018 |
| $d$ | 0.0014 | 0.0021 | 0.0035 | 0.0041 | 0.0034 | 0.0034 | 0.0010 | 0.0062 |
| | | | | 2008 | | | | |
| RMSE | 0.0005 | 0.0010 | 0.0011 | 0.0017 | 0.0015 | 0.0009 | 0.0012 | 0.0026 |
| MAE | 0.0005 | 0.0009 | 0.0010 | 0.0016 | 0.0012 | 0.0007 | 0.0010 | 0.0021 |
| $d$ | 0.0009 | 0.0018 | 0.0018 | 0.0032 | 0.0036 | 0.0020 | 0.0020 | 0.0064 |

**Note:** The goodness-of-fit criteria compare the empirical distribution function values in the left and right tails with their counterparts estimated from the gamma mixture with $K = 4$ components. RMSE is the root-mean squared error of the differences, MAE is the mean absolute value of the differences, and $d$ is the maximum difference. "All obs" refers to all observations.

### 6.5 Restricted dominance

The probability curves are not only useful for isolating population proportions that are critical for dominance assessment. They can also be used to investigate dominance over restricted regions. Poverty orderings, for example, are concerned with dominance below a poverty line. A dominance probability within a restricted range is only likely to differ from a corresponding unrestricted dominance probability if the minimum of the probability curve occurs outside the restricted region. Moreover, the probability curve shows how the restricted dominance probability is likely to change as the restricted interval (for example, the poverty line) changes.

Table 6 contains dominance results for the lowest 10% of the population. Comparing the unrestricted and restricted probabilities of dominance for 2002 over 2005 illustrates how striking the difference can be. Probabilities for FSD, GLD and LD for 2002 over 2005 were 0.0000, 0.0000, and 0.0037, respectively. After restricting the range, however, these probabilities increased to 0.5330, 0.9754, and 1.0000, respectively. Examining the probability curves for FSD and GLD in Figure 6 reveals that restricting the range to less than 10% would further increase the probability of dominance. Similar remarks can be made for GLD dominance of 2008 over 2005 where the probability increased from 0.000 to 0.591; see Figure 7.

Table 6: Restricted dominance probabilities (for the lowest 10% of incomes), with bounds in parentheses

| | 08 D 05 | 05 D 08 | | | | |
|---|---|---|---|---|---|---|
| FSD | 0.0108<br>(0.0081,0.0136) | 0.0001<br>(0.0000,0.0005) | | | | |
| GLD | 0.5910<br>(0.5796,0.6000) | 0.0001<br>(0.0000,0.0005) | | | | |
| LD | 1.0000<br>(0.9999,1.0000) | 0.0000<br>(0.0000,0.0000) | | | | |
| | 08 D 02 | 02 D 08 | 05 D 02 | 02 D 05 | | |
| FSD | 0.0027<br>(0.0014,0.0042) | 0.1378<br>(0.1306,0.1447) | 0.0002<br>(0.0000,0.0006) | 0.5330<br>(0.5250,0.5419) | | |
| GLD | 0.0288<br>(0.0240,0.0330) | 0.1297<br>(0.1229,0.1360) | 0.0002<br>(0.0000,0.0006) | 0.9754<br>(0.9710,0.9788) | | |
| LD | 0.0000<br>(0.0000,0.0000) | 0.8426<br>(0.8357,0.8506) | 0.0000<br>(0.0000,0.0000) | 1.0000<br>(1.0000,1.0000) | | |
| | 08 D 99 | 99 D 08 | 05 D 99 | 99 D 05 | 02 D 99 | 99 D 02 |
| FSD | 0.9980<br>(0.9966,0.9991) | 0.0000<br>(0.0000,0.0000) | 0.2529<br>(0.2432,0.2624) | 0.0000<br>(0.0000,0.0000) | 0.9943<br>(0.9923,0.9962) | 0.0000<br>(0.0000,0.0000) |
| GLD | 0.9997<br>(0.9989,1.0000) | 0.0000<br>(0.0000,0.0000) | 0.4676<br>(0.4595,0.4772) | 0.0000<br>(0.0000,0.0000) | 0.9999<br>(0.9979,0.9998) | 0.0000<br>(0.0000,0.0000) |
| LD | 0.0000<br>(0.0000,0.0000) | 1.0000<br>(1.0000,1.0000) | 0.0000<br>(0.0000,0.0000) | 1.0000<br>(1.0000,1.0000) | 0.0000<br>(0.0000,0.0000) | 1.0000<br>(1.0000,1.0000) |

**Note:** Dominance probabilities are calculated as the proportion of MCMC draws for which the relevant inequalities are satisfied for all population proportions between 0.001 and 0.1. The bounds are the minimum and maximum values obtained from 1000 random rearrangements of the MCMC draws. FSD, GLD and LD refer to first-order stochastic dominance, generalized Lorenz dominance and Lorenz dominance, respectively. For each pairwise comparison, dominance probabilities for both directions are reported. These probabilities can be used to find the probability that neither distribution is dominant. For example, when comparing 2002 with 2008 for FSD, there is a probability of $(1 - 0.0027 - 0.1378) = 0.8595$ that neither dominates.

## 6.6 Comparison of Bayesian and sampling theory results

It is instructive to compare the Bayesian posterior probabilities with some sampling theory results and highlight any differences that may emerge. In Table 7 we report the probabilities of dominance alongside the $p$-values obtained using the Barrett and Donald (2003) tests for FSD and GLD and the BDD (2014) test for LD. The $p$-values are those obtained assuming a null hypothesis of dominance is true. In terms of the general conclusions about dominance that are likely to be made from each approach, there is agreement in most cases. However, there are instances where the Bayesian results lead to more conclusive outcomes. These are instances where a Bayesian probability of dominance is close to zero, and the frequentist $p$-value is relatively large.

For example, consider a FSD comparison of 2005 and 2008. Both approaches suggest 2008 does not dominate 2005; the Bayesian probability is zero and the $p$-value is 0.0011. However, when we investigate whether 2005 might dominate 2008, the Bayesian probability of dominance is 0.0001, and the $p$-value is 0.8907. In other words, the Bayesian probability that neither distribution dominates is 0.9999, but the sampling theory test has been unable to disprove that 2005 dominates 2008. From a frequentist standpoint, 2005 dominating 2008, and neither distribution dominating, are both possible outcomes.

Table 7: A comparison of sampling theory $p$-values with stochastic dominance probabilities

| | | 08 D 05 | 05 D 08 | | | | |
|---|---|---|---|---|---|---|---|
| FSD | probability | 0.0000 | 0.0001 | | | | |
| | $p$-value | 0.0011 | 0.8907 | | | | |
| GLD | probability | 0.0000 | 0.0001 | | | | |
| | $p$-value | 0.0000 | 0.7610 | | | | |
| LD | probability | 1.0000 | 0.0000 | | | | |
| | $p$-value | 1.0000 | 0.0000 | | | | |
| | | 08 D 02 | 02 D 08 | 05 D 02 | 02 D 05 | | |
| FSD | probability | 0.0000 | 0.0000 | 0.0001 | 0.0000 | | |
| | $p$-value | 0.4264 | 0.0000 | 0.4625 | 0.0000 | | |
| GLD | probability | 0.0148 | 0.0000 | 0.0002 | 0.0000 | | |
| | $p$-value | 0.6100 | 0.0000 | 0.6480 | 0.0000 | | |
| LD | probability | 0.0000 | 0.0000 | 0.0000 | 0.0037 | | |
| | $p$-value | 0.0930 | 0.0090 | 0.0000 | 0.2590 | | |
| | | 08 D 99 | 99 D 08 | 05 D 99 | 99 D 05 | 02 D 99 | 99 D 02 |
| FSD | probability | 0.9980 | 0.0000 | 0.2529 | 0.0000 | 0.9943 | 0.0000 |
| | $p$-value | 0.9713 | 0.0000 | 0.9988 | 0.0000 | 0.9580 | 0.0000 |
| GLD | probability | 0.9997 | 0.0000 | 0.4676 | 0.0000 | 0.9990 | 0.0000 |
| | $p$-value | 1.0000 | 0.0000 | 0.8590 | 0.0000 | 0.9990 | 0.0000 |
| LD | probability | 0.0000 | 0.3331 | 0.0000 | 0.1915 | 0.0000 | 0.9996 |
| | $p$-value | 0.0000 | 0.9970 | 0.0000 | 1.0000 | 0.0000 | 0.9940 |

**Note:** The probability values are the Bayesian posterior probabilities of dominance. The $p$-values for FSD and GLD are for the test statistics proposed by Barrett and Donald (2003). The $p$-value for Lorenz dominance is for the test statistic proposed by BDD (2014). These values are for a null hypothesis of dominance and an alternative hypothesis of no dominance. A large $p$-value should not be treated as a sampling theory probability of dominance. It means that the test has been unable to disprove dominance.

Some insights into this result are revealed by examining Figure 8. Here we have graphed the empirical distribution function difference $d = F_{2005}(x) - F_{2008}(x)$ and Bayesian probability $\Pr(2005 \succeq_{\text{FSD}} 2008 \mid x)$ for values of income up to $x = 450$. First, notice that the scale on the left axis is very small; in line with Figure 2, there is little difference between the two distribution functions. We should not be surprised to find a high probability that neither distribution is dominant. The positive values of $d$ up to an income of 147 have been sufficient for the Bayesian approach to register a probability of zero for $\Pr(2005 \succeq_{\text{FSD}} 2008)$. For a frequentist, the positive values of $d$ contradict a null hypothesis that 2005 dominates 2008, but they are not sufficiently large to reject this null hypothesis.

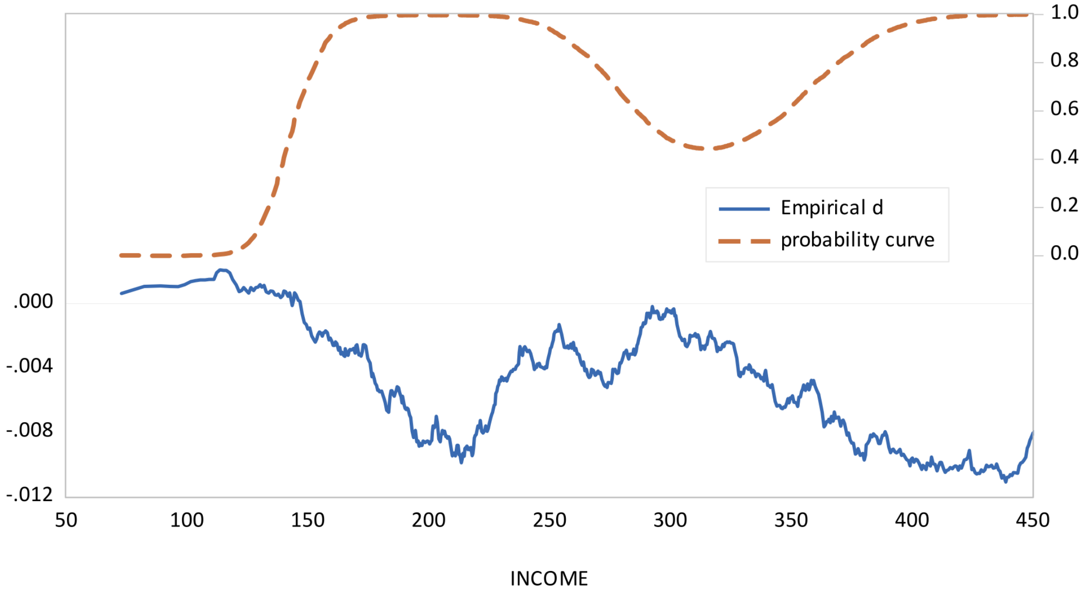



Figure 8: Comparison of distribution function difference with probability curve for 2005 $\succeq_{\text{FSD}}$ 2008

**Note:** The empirical distribution function difference $d = F_{2005}(x) - F_{2008}(x)$ is the jagged curve. The smooth probability curve is reproduced from the top right panel of Figure 7, with the horizontal axis converted to equivalent quantiles. The maximum income in the graph of 450 is the 0.641 quantile. Large negative values for $d$ correspond to high values for $\Pr(2005 \succeq_{\text{FSD}} 2008 \mid x)$.

When we consider values of income greater than 147, $d$ becomes negative, and $\Pr(2005 \succeq_{\text{FSD}} 2008 \mid x)$ gradually increases, becoming 0.998 at an income of $x = 162$. The dip in the probability curve from $x = 250$ to $x = 375$ corresponds to a region where $d$ becomes less negative, almost reaching zero, such that $\Pr(2005 \succeq_{\text{FSD}} 2008 \mid x)$ is no longer close to one. The largest negative value for $d$ is $-0.0163$ at $x = 792$ (beyond the range of Figure 8). This value has been sufficiently large for the frequentist test to reject the null hypothesis that 2008 dominates 2005.

The other cases where the Bayesian probability that neither distribution dominates is close to one, but there is a relatively large $p$-value for one null hypothesis of dominance, are similar. The difference between the two curves is positive for some population proportions and negative for others. With a frequentist test, the maximum value of one of the signed differences—either positive or negative—is sufficiently large to be "significant", but the one in the other direction is not. In terms of Bayesian posterior probabilities, the differences in both directions are "significant". The different conclusions can be attributed to the parametric assumption—the mixture of gamma densities—made by the Bayesian approach.

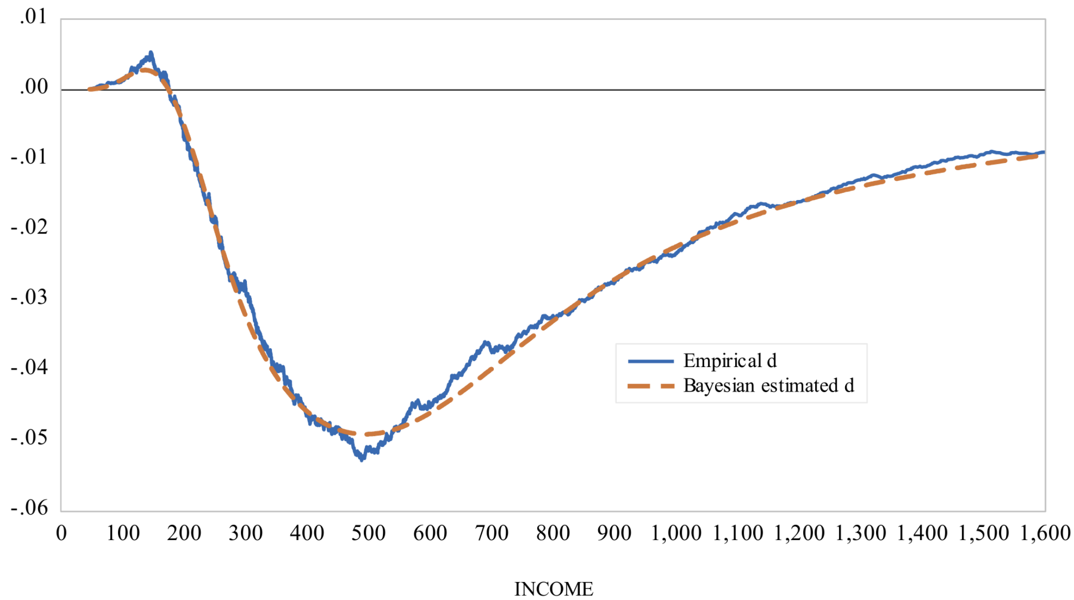



Figure 9: Differences between 2002 and 2005 distribution functions

**Note:** The distribution function differences $d = F_{2005}(x) - F_{2002}(x)$ are evaluated at 9999 quantiles of the 2005 distribution. The smooth curve is the difference in distribution functions evaluated at the Bayesian posterior means. The jagged curve is the difference in empirical distribution functions. Negative values of $d$ suggest 2005 is preferable; positive values of $d$ suggest 2002 is preferable.

To illustrate, and explain further, consider Figure 9 where we have graphed both the empirical and Bayesian estimated differences $d = F_{2005}(x) - F_{2002}(x)$. This is a case where the Bayesian conclusion is that neither distribution is first-order dominant (probability = 0.9999). The frequentist conclusions are 2002 does not dominate 2005 ($p$-value $= 0.0000$), and we fail to reject that 2005 dominates 2002 ($p$-value $= 0.4625$). Rejection of $H_0 : 2002 \succeq_{FSD} 2005$ occurs because the largest negative value of $d$ (at $x = 488$) is a sufficiently large negative value. The failure to reject $H_0 : 2005 \succeq_{FSD} 2002$ occurs because the largest positive value of $d$ (at $x = 145$) is not large enough; the variance of the test statistic is not small enough to rule out negative values for the "true" $d$. However, evidence from the Bayesian parametric approach suggests a negative underlying $d$ is extremely unlikely at $x = 145$. This will occur if the tail of the posterior density for $d$ at this point is barely in the negative region. Thus, the posterior standard deviation of $d$ at $x = 145$ must be small relative to the standard error of the frequentist test statistic. Assuming a gamma mixture leads to a small posterior standard deviation in the tail, meaning a more definite conclusion can be reached.

What we can take away from this investigation is that an advantage of the Bayesian parametric approach is that it leads to more precise estimation in the tails of the distributions. A disadvantage is that it is dependent on a parametric assumption. The gamma mixture approximation does appear to be a good one, however. What is also clear is that one should be very reluctant to conclude that dominance exists when a null hypothesis of dominance is not rejected. Both the empirical $d$ and the Bayesian estimated $d$ are positive for incomes between 46 and 175 (the poorest 10% of the population); it would be brave to conclude that the difference between the two distribution functions is negative over the whole range. Using probability curves, or, alternatively, graphing the function differences that appear in the test statistics, provides valuable insights not revealed by considering only the $p$-value of a test statistic.

## 7 Conclusions

The development of statistical inference for assessing how income distributions have changed over time in what might be considered a desirable way has attracted a great deal of attention within the sampling theory framework. Hypothesis testing procedures have been developed for, among other things, Lorenz dominance, generalized-Lorenz dominance and first-order stochastic dominance. This paper provides an alternative to the existing hypothesis tests, by defining a novel approach to assessing dominance relationships within a framework of Bayesian inference.

A desirable feature of this framework is the reporting of results in terms of probabilities—a natural way to express our uncertainty. Furthermore, these probabilities can be provided for dominance in either direction, as well as the probability that dominance does not occur. They overcome the problem of giving favorable treatment to what is chosen as the null hypothesis in sampling theory inference. By employing a flexible gamma mixture model, we minimize the sensitivity of dominance results to the chosen distribution.

The methodology is applied to data for Indonesia for the years 1999, 2002, 2005, and 2008. In general, the dominance test results led us to conclude that the level of inequality increased from 1999 to 2005, and decreased from 2005 to 2008. There was also strong evidence of an increase in welfare from 1999 to 2002, but little evidence of improvements from 2002 to 2005 and from 2005 to 2008.

The introduction of probability curves that trace how the probability of dominance at a particular population proportion changes as the population proportion changes enabled us to isolate segments of the population having the greatest impact on overall dominance and to explain seemingly contradictory outcomes from the Bayesian and sampling theory approaches. Future research will extend the framework to multivariate settings and to orderings of more than two distributions.

## References


ANDERSON, G. (1996): "Nonparametric Tests for Stochastic Dominance," *Econometrica*, 64, 1183–1193.

ATKINSON, A. B. (1970): "On the Measurement of Inequality," *Journal of Economic Theory*, 2, 244–263.

ATKINSON, A. B. AND F. BOURGUIGNON (1982): "The Comparison of Multi-Dimensioned Distributions of Economic Status," *Review of Economic Studies*, 49, 183–201.

BAI, Z., L. HUA, H. LIU, AND W.-K. WONG (2011): "Test Statistics for Prospect and Markowitz Stochastic Dominances with Applications," *Econometrics Journal*, 14, 278–303.

BANKS, J. AND P. JOHNSON (1994): "Equivalence Scale Relativities Revisited," *Economic Journal*, 104, 883–890.

BARRETT, G. F. AND S. G. DONALD (2003): "Consistent Tests for Stochastic Dominance," *Econometrica*, 71, 71–104.

BARRETT, G. F., S. G. DONALD, AND D. BHATTACHARYA (2014): "Consistent Non-parametric Tests for Lorenz Dominance," *Journal of Business & Economic Statistics*, 32, 1–13.

BARRETT, G. F., S. G. DONALD, AND Y.-C. HSU (2016): "Consistent Tests for Poverty Dominance Relations," *Journal of Econometrics*, 191, 360–373.

BEACH, C. M. AND R. DAVIDSON (1983): "Distribution-Free Statistical Inferences with Lorenz Curves and Income Shares," *Review of Economic Studies*, 50, 723–735.

BENNETT, C. J. (2013): "Inference for Dominance Relations," *International Economic Review*, 54, 1309–1328.

BENNETT, C. J. AND S. S. MITRA (2013): "Multidimensional Poverty: Measurement, Estimation and Inference," *Econometric Reviews*, 32, 57–83.

BERRENDERO, J. R. AND J. CÁRCAMO (2011): "Tests for Second Order Stochastic Dominance Based on $L$-Statistics," *Journal of Business & Economic Statistics*, 29, 260–270.

BISHOP, J. A., S. CHAKRABORTI, AND P. D. THISTLE (1989): "Asymptotically Distribution-Free Statistical Inferences for Generalized Lorenz Curves," *Review of Economics and Statistics*, 71, 725–727.

BISHOP, J. A., J. P. FORMBY, AND R. SAKANO (1995): "Lorenz and Stochastic Dominance Comparisons of European Income Distributions," *Research on Economic Inequality*, 6, 77–92.

BISHOP, J. A., J. P. FORMBY, AND W. J. SMITH (1991): "Lorenz Dominance and Welfare: Changes in the U.S. Distribution of Income, 1967–1986," *Review of Economics and Statistics*, 73, 134–139.

CHAKRAVARTY, S. R. (2009): *Inequality, Polarization and Poverty: Advances in Distributional Analysis*, New York: Springer.

CHOTIKAPANICH, D. AND W. E. GRIFFITHS (2006): "Bayesian Assessment of Lorenz and Stochastic Dominance in Income Distributions," *Research on Economic Inequality*, 13, 297–321.

——— (2008): "Estimating Income Distributions Using a Mixture of Gamma Densities," in *Modeling Income Distributions and Lorenz Curves*, ed. by D. Chotikapanich, New York: Springer, 285–302.

DARDANONI, V. AND A. FORCINA (1999): "Inference for Lorenz Curve Orderings," *Econometrics Journal*, 2, 49–75.

DAVIDSON, R. AND J.-Y. DUCLOS (2000): “Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality,” *Econometrica*, 68, 1435–1464.

——— (2013): “Testing for Restricted Stochastic Dominance,” *Econometric Reviews*, 32, 84–125.

DENUIT, M., J. DHAENE, M. J. GOOVAERTS, AND R. KAAS (2005): *Actuarial Theory for Dependent Risks: Measures, Orders, and Models*, New York: Wiley.

DONALD, S. G. AND Y.-C. HSU (2016): “Improving the Power of Tests of Stochastic Dominance,” *Econometric Reviews*, 35, 553–585.

DUCLOS, J.-Y., D. E. SAHN, AND S. D. YOUNGER (2006): “Robust Multidimensional Poverty Comparisons,” *Economic Journal*, 116, 943–968.

FOSTER, J. E. AND A. F. SHORROCKS (1988): “Poverty Orderings and Welfare Dominance,” *Social Choice and Welfare*, 5, 179–198.

FRÜHWIRTH-SCHNATTER, S. (2006): *Finite Mixture and Markov Switching Models*, New York: Springer.

GEWEKE, J. (1986): “Exact Inference in the Inequality Constrained Normal Linear Regression Model,” *Journal of Applied Econometrics*, 1, 127–142.

——— (1988): “The Secular and Cyclical Behavior of Real GDP in 19 OECD Countries, 1957–1983,” *Journal of Business & Economic Statistics*, 6, 479–486.

——— (2007): “Interpretation and Inference in Mixture Models: Simple MCMC Works,” *Computational Statistics & Data Analysis*, 51, 3529–3550.

GILL, J. (2015): *Bayesian Methods: A Social and Behavioural Sciences Approach*, New York: CRC Press, 3 ed.

HANDAYANI, A. (2013): “An Analysis of Changes in Indonesia’s Income Distribution 1996–2008,” Ph.D. thesis, Monash University, Melbourne.

HASEGAWA, H. (2013): “Bayesian Analysis of Stochastic Dominance,” Discussion Paper Series A 2013-253, Graduate School of Economics and Business Administration, Hokkaido University.

HORVÁTH, L., P. KOKOSZKA, AND R. ZITIKIS (2006): “Testing for Stochastic Dominance Using the Weighted McFadden-Type Statistic,” *Journal of Econometrics*, 133, 191–205.

JENKINS, S. P. AND F. A. COWELL (1994): “Parametric Equivalence Scales and Scale Relativities,” *Economic Journal*, 104, 891–900.

JENKINS, S. P. AND P. J. LAMBERT (1997): “Three “Ls” of Poverty Curves, with an Analysis of UK Poverty Trends,” *Oxford Economic Papers*, 49, 317–327.

KAAS, R., A. E. VAN HEERWAARDEN, AND M. J. GOOVAERT (1994): *Orderings of Actuarial Risks*, Brussels: Caire.

KAUR, A., B. L. S. P. RAO, AND H. SINGH (1994): “Testing for Second-Order Stochastic Dominance of Two Distributions,” *Econometric Theory*, 10, 849–866.

KLEIBER, C. AND S. KOTZ (2003): *Statistical Size Distributions in Economics and Actuarial Sciences*, New Jersey: Wiley.

LAMBERT, P. J. (2001): *The Distribution and Redistribution of Income: A Mathematical Analysis*, Manchester: Manchester University Press, 3 ed.

LE BRETON, M. AND E. PELUSO (2009): "Third-Degree Stochastic Dominance and Inequality Measurement," *Journal of Economic Inequality*, 7, 249–268.

LINTON, O., E. MAASOUMI, AND Y.-J. WHANG (2005): "Consistent Testing for Stochastic Dominance under General Sampling Schemes," *Review of Economic Studies*, 72, 735–765.

LINTON, O., K. SONG, AND Y.-J. WHANG (2010): "An Improved Bootstrap Test for Stochastic Dominance," *Journal of Econometrics*, 154, 186–202.

MAASOUMI, E. (1997): "Empirical Analyses of Inequality and Welfare," in *Handbook of Applied Econometrics: Volume II Microeconomics*, ed. by M. H. Pesaran and P. Schmidt, Malden, MA: Blackwell, 202–245.

MAASOUMI, E. AND A. HESHMATI (2000): "Stochastic Dominance among Swedish Income Distributions," *Econometric Reviews*, 19, 287–320.

——— (2008): "Evaluating Dominance Ranking of PSID Incomes by Various Household Attributes," in *Advances on Income Inequality and Concentration Measures*, ed. by G. Betti and A. Lemmi, Oxford: Routledge, 47–69.

MAASOUMI, E. AND D. L. MILLIMET (2005): "Robust Inference Concerning Recent Trends in US Environmental Quality," *Journal of Applied Econometrics*, 20, 55–78.

MCCAIG, B. AND A. YATCHEW (2007): "International Welfare Comparisons and Nonparametric Testing of Multivariate Stochastic Dominance," *Journal of Applied Econometrics*, 22, 951–969.

MCFADDEN, D. (1989): "Testing for Stochastic Dominance," in *Studies in the Economics of Uncertainty: In Honor of Josef Hadar*, ed. by T. B. Fomby and T. K. Seo, New York: Springer, 113–134.

SCHLUTER, C. AND M. TREDE (2002): "Tails of Lorenz Curves," *Journal of Econometrics*, 109, 151–166.

SHORROCKS, A. F. (1983): "Ranking Income Distributions," *Economica*, 50, 3–17.

SRIBOONCHITTA, S., W.-K. WONG, S. DHOMPONGSA, AND H. T. NGUYEN (2009): *Stochastic Dominance and Applications to Finance, Risk, and Economics*, Boca Raton, FL: Chapman & Hall.

TABRI, R. V. (2015): "Empirical Likelihood for Robust Poverty Comparisons," Working Paper 2015-2, Department of Economics, University of Sydney.

WIPER, M., D. R. INSUA, AND F. RUGGERI (2001): "Mixtures of Gamma Distributions with Applications," *Journal of Computational and Graphical Statistics*, 10, 440–454.

WONG, W.-K., K. F. PHOON, AND H. H. LEAN (2008): "Stochastic Dominance Analysis of Asian Hedge Funds," *Pacific-Basin Finance Journal*, 16, 204–223.

ZELLNER, A. (1971): *An Introduction to Bayesian Inference in Econometrics*, New York: John Wiley.

# APPENDIX

## Appendix A Gibbs sampler for parameters of gamma mixture

The gamma mixture density function is given by

$$p(x \mid \boldsymbol{\mu}, \boldsymbol{v}, \boldsymbol{w}) = \sum_{k=1}^{K} w_k G\left(x \mid v_k, \frac{v_k}{\mu_k}\right).$$

Let $Z_1, Z_2, \ldots, Z_n$ be indicator variables such that $Z_i = k$ when the $i$th observation comes from the $k$th component in the mixture. Then, $P(Z_i = k \mid \boldsymbol{w}) = w_k$ for $k = 1, 2, \ldots, K$, with $w_k > 0$ and $\sum_{k=1}^{K} w_k = 1$. Conditional on $Z_i = k$, the distribution of $x_i$ is $G(x_i \mid v_k, v_k/\mu_k)$. Following Wiper et al. (2001), we use the following prior distributions for each of the parameters:

$$\begin{aligned}
&p(\boldsymbol{w}) = D(\boldsymbol{\varphi}) \propto w_1^{\varphi_1 - 1} w_2^{\varphi_2 - 1} \cdots w_K^{\varphi_K - 1} && \text{(Dirichlet)},\\
&p(v_k) \propto \exp(-\theta v_k), \qquad k = 1, 2, \ldots, K && \text{(exponential)},\\
&p(\mu_k) = GI(\alpha_k, \beta_k) \propto \mu_k^{-(\alpha_k + 1)} \exp\left(-\frac{\beta_k}{\mu_k}\right), \quad k = 1, 2, \ldots, K && \text{(inverted gamma)}.
\end{aligned}$$

These prior densities combine nicely with the likelihood function and are sufficiently flexible to represent vague prior information which can be dominated by the sample data. We set $\varphi_k = 1$ for all $k$, implying a flat prior for the mixture weights on each component. For the exponential prior on the shape parameters $v_k$, we set $\theta = 0.01$ for all $k$. A 95% probability interval from this prior is $(2.54, 202.86)$ implying a large range of values are possible. For the $\mu_k$ we set $\alpha_k = 2.2$ for all $k$, and $\beta_1 = 60$, $\beta_2 = 100$, $\beta_3 = 150$, $\beta_4 = 200$ and $\beta_5 = 250$. From these prior settings, the 95% probability intervals for $\mu_1, \mu_2, \ldots, \mu_5$ are $(10.31, 202.86)$, $(17.06, 320.84)$, $(24.93, 488.68)$, $(33.85, 626.13)$, and $(42.47, 813.53)$, respectively. All these intervals suggest that the priors are relatively uninformative.

As described in Wiper et al. (2001), the conditional posterior densities for use in a Gibbs sampler are:

1)

$$P(Z_i = k \mid \boldsymbol{x}, \boldsymbol{w}, \boldsymbol{v}, \boldsymbol{\mu}) = \frac{p_{ik}}{p_{i1} + p_{i2} + \cdots + p_{iK}},$$

$$p_{ik} = w_k \frac{(v_k/\mu_k)^{v_k}}{\Gamma(v_k)} x_i^{v_k - 1} \exp\left(-\frac{v_k x_i}{\mu_k}\right).$$

2)

$$p(\boldsymbol{w} \mid \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{v}, \boldsymbol{\mu}) = D(\boldsymbol{\varphi} + \boldsymbol{n}),$$

where $\boldsymbol{z}' = (k_1, k_2, \ldots, k_n)$ are the realized components for each of the observations, and $\boldsymbol{n}' = (n_1, n_2, \ldots, n_K)$, with $n_k$ being the number of observations for which $Z_i = k$. Also, $\sum_{k=1}^{K} n_k = n$.

3)
$$p(\mu_k \mid \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{w}, \boldsymbol{v}) = GI(\alpha_k + n_k v_k,\ \beta_k + S_k v_k), \qquad S_k = \sum_{i:Z_i=k} x_i.$$

4)
$$p(v_k \mid \boldsymbol{x}, \boldsymbol{z}, \boldsymbol{w}, \boldsymbol{\mu}) \propto \frac{v_k^{n_k v_k}}{[\Gamma(v_k)]^{n_k}} \exp\left\{ -v_k \left( \theta + \frac{S_k}{\mu_k} + n_k \log \mu_k - \log P_k \right) \right\},$$
$$P_k = \prod_{i:Z_i=k} x_i.$$

This density is not a recognizable form and requires a Metropolis step. A candidate $\tilde{v}_k^{(t+1)}$ is drawn from a gamma density $G(r, r/v_k^{(t)})$ with mean equal to the previous draw $v_k^{(t)}$ and is accepted with probability

$$\min\left\{ 1, \begin{array}{c} \dfrac{p(\tilde{v}_k^{(t+1)} \mid \boldsymbol{x}, \boldsymbol{z}^{(t+1)}, \boldsymbol{w}^{(t+1)}, \boldsymbol{\mu}^{(t+1)})}{p(v_k^{(t)} \mid \boldsymbol{x}, \boldsymbol{z}^{(t+1)}, \boldsymbol{w}^{(t+1)}, \boldsymbol{\mu}^{(t+1)})} \\ \times \dfrac{p(\tilde{v}_k^{(t+1)}, v_k^{(t)})}{p(v_k^{(t)}, \tilde{v}_k^{(t+1)})} \end{array} \right\}.$$

Here $p(v_k^{(t)}, \tilde{v}_k^{(t+1)})$ is the gamma density used to generate $\tilde{v}_k^{(t+1)}$. The value of $r$ is chosen by experimentation to give an acceptance rate of approximately 0.25 to 0.4.

5) Following Wiper et al. (2001), we ordered the components of $\boldsymbol{\mu}$ according to $\mu_1 < \mu_2 < \mu_3 < \mu_4$, and sorted $\boldsymbol{v}$ and $\boldsymbol{w}$ accordingly.

This last step is not needed for estimation of the mixture distribution (Geweke, 2007), but it aids identification of the individual components, should that be an objective.

## Appendix B Algorithm for obtaining gamma mixture quantiles

1. Implement for each parameter draw $m = 1, 2, \ldots, M$.
2. Generate a large number of draws from $p(x \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$, say 100,000, and sort them from lowest to highest. Let the vector of these draws be denoted by $\boldsymbol{x}$ and the $j$th ordered value by $x_j$.
3. Find the 100,000 cumulative proportions $\boldsymbol{F}(\boldsymbol{x} \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$.
4. Implement for proportions $u_i$, $i = 1, 2, \ldots, n$.
5. For a given $u_i$, find the smallest $j$, call it $j_i$, for which $F(x_j \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)}) \geq u_i$.
6. Set a tolerable error $\varepsilon$. We used $\varepsilon = 10^{-4}$ in our empirical work.
7. Find an initial value $\widehat{y}_1 = F^{-1}(u_i \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$ as a draw from the uniform distribution $U(x_{j_i-1}, x_{j_i})$.
8. Compute $F_X(\widehat{y}_1 \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)})$. If $|F_X(\widehat{y}_1 \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)}) - u_i| > \varepsilon$, go to step 9; otherwise, go back to 4. (Step 9 was seldom needed in our empirical work.)
9. We improve on the initial value $\widehat{y}_1$ as follows. If $F_X(\widehat{y}_1 \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)}) - u_i > 0$, generate a random draw $\widehat{y}_2 = U(x_{j_i-1}, \widehat{y}_1)$. If $F_X(\widehat{y}_1 \mid \boldsymbol{\mu}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{w}^{(m)}) - u_i < 0$, generate $\widehat{y}_2 = U(\widehat{y}_1, x_{j_i})$.
10. Repeat step 8 with the new $\widehat{y}$.
11. Go back to step 1.