THE UNIVERSITY OF MELBOURNE

DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 960

FEBRUARY 2006

**Bayesian Assessment of Lorenz and
Stochastic Dominance in Income Distributions**

by

Duangkamon Chotikapanich
&
William E. Griffiths

Department of Economics
The University of Melbourne
Melbourne   Victoria   3010
Australia.

Bayesian Assessment of Lorenz and Stochastic Dominance
in Income Distributions

Duangkamon Chotikapanich
Monash University

William E. Griffiths
University of Melbourne

Abstract

Hypothesis tests for dominance in income distributions has received considerable attention in recent literature. See, for example, Barrett and Donald (2003), Davidson and Duclos (2000) and references therein. Such tests are useful for assessing progress towards eliminating poverty and for evaluating the effectiveness of various policy initiatives directed towards welfare improvement. To date the focus in the literature has been on sampling theory tests. Such tests can be set up in various ways, with dominance as the null or alternative hypothesis, and with dominance in either direction (*X* dominates *Y* or *Y* dominates *X*). The result of a test is expressed as rejection of, or failure to reject, a null hypothesis. In this paper we develop and apply Bayesian methods of inference to problems of Lorenz and stochastic dominance. The result from a comparison of two income distributions is reported in terms of the posterior probabilities for each of the three possible outcomes: (a) *X* dominates *Y*, (b) *Y* dominates *X*, and (c) neither *X* nor *Y* is dominant. Reporting results about uncertain outcomes in terms of probabilities has the advantage of being more informative than a simple reject / do-not-reject outcome. Whether a probability is sufficiently high or low for a policy maker to take a particular action is then a decision for that policy maker.

The methodology is applied to data for Canada from the Family Expenditure Survey for the years 1978 and 1986. We assess the likelihood of dominance from one time period to the next. Two alternative assumptions are made about the income distributions –Dagum and Singh-Maddala – and in each case the posterior probability of dominance is given by the proportion of times a relevant parameter inequality is satisfied by the posterior observations generated by Markov chain Monte Carlo.

1.   Introduction

Governments, their policy advisors and society in general are concerned with the notion of becoming "better off" as we progress through time. Assuming that measurement of the state of well-being of an economy can be related to its income distribution, the concept of a social welfare function has been used to assess whether an income distribution at one point in time is preferred to another at a different point in time. Precise specification of a social welfare function requires judgements about the relative utility of particular characteristics of income distributions such as mean income, the variability of income and the proportion of the population categorized as below a poverty line. Because such judgements will rarely be generally accepted, it has been useful to consider broad classes of social welfare functions with a few restrictive assumptions that have more general acceptance. When comparing two income distributions within this context, we say that one income distribution X dominates another Y if X is preferred to Y for all members of a particular class of social welfare functions. Different classes of social welfare functions with varying degrees of restrictive assumptions lead to consideration of alternative dominance relationships. Three types of dominance relationships common in the literature, and those that we consider in this paper, are Lorenz dominance, generalized Lorenz dominance (second order stochastic dominance) and first order stochastic dominance. Also, because ordering of income distributions according to dominance criteria is not always possible, inequality measures such as the Gini coefficient and Atkinson's inequality index, that involve placing more restrictive assumptions on social welfare functions, are frequently used to compare different distributions. For details of these various concepts, and the relationships between them, see, for example, Lambert (1993), Creedy (1996) or Maasoumi (1997).

A comparison of two income distributions, whether it be via a dominance relationship or an inequality measure, typically involves a comparison of income-distribution characteristics from samples of individuals or samples of households taken at the two points in time. Because characteristics calculated from samples are subject to statistical sampling error, comparing them at different points in time does not lead to conclusions that can be made with certainty. Estimates of, for example, mean incomes, Gini coefficients and Lorenz curves can point towards one distribution being preferred to another, but because estimates are indeed estimates, the possibility of incorrect conclusions being drawn always exists. This possibility has led to the development of a large body of literature on sampling-theory properties of estimates of inequality measures, as well as sampling-theory hypothesis-testing procedures for comparing inequality measures and assessing whether one income distribution dominates another in some sense. There has also been a limited number of studies examining the use of Bayesian inference for such purposes. The literature on sampling-theory inference for inequality measures can be accessed through Cowell (1999); for an example of Bayesian inference applied to inequality measures, see Chotikapanich and Creedy (2004). In this paper we are concerned with Bayesian methods for assessing dominance relationships. Although a large effort has been directed towards sampling-theory hypothesis tests for this purpose (see, for example, Anderson (1996), Barrett and Donald (2003a, 2003b), Bishop and Formby (1999), Dardanoni and Forcina (1999), Davidson and Duclos (1997, 2000), Linton, Maasoumi and Whang (2003), Maasoumi (1997), Tse and Zhang (2002) and references therein), there appears to be little or no work on Bayesian methods for such comparisons.

To appreciate the differences between the Bayesian and sampling-theory approaches, suppose that we are comparing two income distributions for randomly

drawn incomes *X* and *Y*, and that we are interested in whether *X* dominates *Y* in some sense ($X \geq_D Y$), or vice versa ($Y \geq_D X$). The results from a Bayesian analysis of this problem are reported as posterior probabilities for each possible outcome, namely, the probability that *X* dominates *Y*, $\Pr(X \geq_D Y)$, the probability that *Y* dominates *X*, $\Pr(Y \geq_D X)$, and the probability that neither dominates, $1 - \Pr(X \geq_D Y) - \Pr(Y \geq_D X)$. Given that probabilities provide a natural framework for describing uncertain information, the reporting of dominance probabilities in this way would seem to be a useful way of summarizing the results of any income distribution comparison. It avoids a number of undesirable characteristics of the sampling-theory approach. Specifically, results reported from a sampling-theory, hypothesis-testing approach will depend on (a) the choice of null and alternative hypotheses (is the null hypothesis an equality or a dominance relation, what happens if the null and alternative hypotheses are reversed), (b) whether the test is based on a distribution function or a quantile function, (c) the test statistic that is chosen, (d) whether an asymptotic or a bootstrapped distribution is used for the test statistic, and (e) the chosen level of significance. The variety of tests that have appeared in the literature can be distinguished according to one or more of these features. One does, however, need to recognize what some may see as disadvantages of the Bayesian approach. In particular, the posterior probabilities for dominance will depend on (a) how the income distribution is modelled through the likelihood function and (b) the prior information that is placed on unknown parameters.

In this paper we illustrate how the Bayesian approach can be used to find the posterior probability that one income distribution dominates another, and, coincidentally, the probability that neither distribution dominates. We consider two parametric income distributions, the so-called Singh-Maddala and Dagum

distributions and apply them to Canadian income distributions for 1978 and 1986; these data were used by Barrett and Donald (2003a, 2003b) to illustrate their sampling-theory tests. Adopting a parametric approach is less general than desirable, particularly in view of the fact that most sampling-theory tests are nonparametric and, as becomes evident, our results are sensitive to the assumed form of the income distribution. However, our suggested approach is novel and provides a sound base for developing future research into improved techniques for assessing dominance. Future directions that are likely to be productive are the application of our proposals to more flexible income distributions such as mixtures, and the use of our techniques in conjunction with Bayesian nonparametric approaches such as that adopted by Hasegawa and Kozumi (2003). Also, the sensitivity of the results to the nature of the assumed distribution is a finding in itself. Although we consider specific distributions, our analysis is not restricted to within-family comparisons. Our procedures do not preclude comparing a Singh-Maddala distribution with a Dagum distribution or indeed any other distribution that might be considered. In any study where a large number of parametric distributions are considered, our methods could be used to compare best-fitting distributions from each time period or, alternatively, one could work with model-averaged distributions like those derived by Griffiths, Chotikapanich and Rao (2005).

In Section 2 we begin by describing the three kinds of dominance considered in the paper: Lorenz dominance, generalized Lorenz dominance (second order stochastic dominance) and first order stochastic dominance. Although our analysis is confined to these dominance relations, it is straightforward to compare any other measures of interest such as poverty and inequality indices, as long as such indices can be expressed as (analytical or numerical) functions of the parameters of the

income distributions. After describing the alternative forms of dominance in general terms, we describe the Singh-Maddala and Dagum income distributions and relate the dominance conditions to these distributions. In Section 3 we specify prior distributions for the mean, mode and Gini coefficient for the income distributions in each of the two time periods, transform those prior distributions to prior distributions on the parameters of the income distributions, specify the likelihood functions for both individual observations and grouped data, and give expressions for the posterior distributions for the income distribution parameters. The results from applying the methodology to a subset of the Canadian data used by Barrett and Donald (2003a and 2003b) to illustrate their sampling-theory testing procedures are presented in Section 4. Before-tax income distributions in 1978 and 1986 are compared. Some concluding remarks appear in Section 5.

2.   Income Distributions and Dominance Conditions

To introduce Lorenz, generalized Lorenz and first order stochastic dominance consider an income distribution that is described by density and distribution functions $f_X(x)$ and $F_X(x)$, respectively. Also, assume that mean income $\mu_X = E(X)$ is finite. The Lorenz curve that gives the proportion of total income earned by the poorest proportion u of the population is given by

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

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

$$L_X(u) \ge L_Y(u) \qquad\qquad \text{for all } 0 \le u \le 1 \qquad\qquad (2)$$

While this definition is the typical one used in the economics literature (see, for example, Lambert (1993) and Barrett and Donald (2003b)), 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 $X \geq_{LD} Y$. 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 thing equal, less inequality is preferred to more inequality, we refer to this condition as one where X dominates Y.

Because Lorenz dominance considers only the degree of inequality and not the level of income, and higher levels of income are associated with higher levels of welfare, another dominance relation known as generalized Lorenz dominance (GLD) is often considered. We say that X generalized-Lorenz dominates Y, written as $X \geq_{GLD} Y$ if and only if

$$\mu_X L_X(u) \geq \mu_Y L_Y(u) \qquad \text{for all} \quad 0 \leq u \leq 1 \qquad (3)$$

Given the expression for the Lorenz curve in equation (1), the condition in (3) can also be expressed as

$$\int_o^u F_X^{-1}(t)\,dt \geq \int_o^u F_Y^{-1}(t)\,dt \qquad \text{for all} \quad 0 \leq u \leq 1 \qquad (4)$$

Writing the relation for generalized Lorenz dominance (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 (GLD) is that of first-order stochastic dominance (FSD). The distribution for X first-order stochastically dominates Y, written $X \geq_{FSD} Y$, if and only if

$$F_X^{-1}(u) \ \geq \ F_Y^{-1}(u) \qquad\qquad \text{for all} \quad 0 \leq u \leq 1 \qquad (5)$$

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.

We now consider the implications of conditions (2), (3) and (5) for the Dagum and Singh-Maddala income distributions. These distributions were chosen for illustrative purposes because they are popular ones in the literature and their closed-form distribution and inverse distribution functions make them convenient computationally. Many other alternatives could have been chosen. See, for example, McDonald (1984) and McDonald and Xu (1995). Less restrictive distributions that involve more parameters or mixtures may prove to be more realistic empirical representations. The density, distribution, inverse distribution and Lorenz functions for the Dagum and Singh-Maddala distributions are of interest. For the Dagum distribution we have (see, for example, Kleiber and Kotz (2003, p.212)).

$$f_D(x) \ = \ \frac{a\,p\,x^{ap-1}}{b^{ap}\left[1+\left(x/b\right)^a\right]^{p+1}} \qquad\qquad a,\, b,\, p > 0 \qquad (6)$$

$$F_D(x) \ = \ \left[1+\left(x/b\right)^{-a}\right]^{-p} \qquad (7)$$

$$F_D^{-1}(u) \ = \ b\left[u^{-1/p}-1\right]^{-1/a} \qquad (8)$$

$$L_D(u) \ = \ B_{w_1}\left[\left(p+1/a\right),\left(1-1/a\right)\right] \qquad (9)$$

where $w_1 = u^{1/p}$ and $B_w(c,d)$ is the incomplete beta integral

$$B_w(c,d) \ = \ \frac{\displaystyle\int_0^w t^{c-1}(1-t)^{d-1}\,dt}{\displaystyle\int_0^1 t^{c-1}(1-t)^{d-1}\,dt}$$

For the Singh-Maddala distribution these functions are given by (Kleiber and Kotz (2003, p.198)).

$$f_S(x) = \frac{aq\, x^{a-1}}{b^a \left[1 + (x/b)^a\right]^{1+q}} \qquad a, b, q > 0 \qquad (10)$$

$$F_S(x) = 1 - \left[1 + (x/b)^a\right]^{-q} \qquad (11)$$

$$F_S^{-1}(u) = b\left[(1-u)^{-1/q} - 1\right]^{1/a} \qquad (12)$$

$$L_S(u) = B_{w_2}\left[(1 + 1/a), (q - 1/a)\right] \qquad (13)$$

where $w_2 = 1 - (1-u)^{1/q}$. To assess generalized Lorenz dominance using the relation in (3) the means of each distribution are also required. They are given by

$$\mu_D = \frac{b\,\Gamma(p + 1/a)\,\Gamma(1 - 1/a)}{\Gamma(p)} \qquad a > 1$$

and

$$\mu_S = \frac{b\,\Gamma(1 + 1/a)\,\Gamma(q - 1/a)}{\Gamma(q)} \qquad q > \frac{1}{a}$$

Given two distributions, each with <u>known</u> parameter values, one way to assess each form of dominance (LD, GLD or FSD) is to 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. For each set of parameters, in our illustration we use 999 values of u from 0.001 to 0.999 with, as explained in more detailed below, a finer grid in intervals likely to be influential.

When income distribution data are used to estimate the unknown parameters of the Dagum and/or Singh-Maddala distributions, these parameters are not known with certainty and any conclusion about whether one distribution dominates another cannot be made with certainty. In Bayesian inference uncertainty about whether one distribution dominates another can be expressed in terms of a probability statement. To obtain such a probability statement we begin by generating draws on the parameters from their respective posterior distributions. Computing $F^{-1}(u)$, say, for a given u and for every parameter draw, yields draws from the posterior density function for $F^{-1}(u)$. Given draws from the posterior density functions for two inverse-distribution functions $F_X^{-1}(u)$ and $F_Y^{-1}(u)$, for a fine grid of values for u, an estimate of the probability that $X \geq_{FSD} Y$ is given by the proportion of times (or the proportion of parameter draws) for which $F_X^{-1}(u) \geq F_Y^{-1}(u)$ for all u. Similar probability statements can be made for LD and GLD.

This procedure for finding the posterior probability of dominance can be adopted for any income distribution for which we can compute values for the Lorenz curve and the inverse distribution function at each population proportion. As mentioned, in our application we apply it to the Dagum and Singh-Maddala distributions. For some distributions and some dominance relationships, necessary and sufficient conditions for dominance have been derived in the form of inequalities on the parameters of the distributions. Such inequalities exist for Lorenz dominance when comparing two Dagum distributions or two Singh-Maddala distributions. In these instances an alternative way of computing the posterior probability of dominance is to count the proportion of parameter draws that satisfy the required parameter inequalities. Proceeding in this way is much less demanding

computationally because it does not involve counting a proportion of parameter draws for all population proportions. Although both approaches should yield identical results, in our early calculations we discovered they did not. The source of the problem turned out to be a dramatic decline in the probability of one Lorenz curve exceeding another as the population proportion approached zero. A consequent conclusion that we make is that consideration of the entire Lorenz curve rather than the parameter inequalities may give a more meaningful picture of the probability of dominance because the researcher is able to exclude population proportion values that are too small to be relevant.

The relevant parameter inequalities for Lorenz dominance for the Dagum and Singh-Maddala distributions are as follows. When comparing two Dagum distributions with respective parameters (a_X, b_X, p_X) and (a_Y, b_Y, p_Y), necessary and sufficient conditions for $X \geq_{LD} Y$ are (Kleiber, 1996)

$$a_X \geq a_Y \qquad \text{and} \qquad a_X p_X \geq a_Y p_Y \qquad (14)$$

Similarly, necessary and sufficient conditions for $X \geq_{LD} Y$ when X and Y are Singh-Maddala distributions with respective parameters (a_X, b_X, q_X) and (a_Y, b_Y, q_Y) are (Wilfling and Kramer, 1993)

$$a_X \geq a_Y \qquad \text{and} \qquad a_X q_X \geq a_Y q_Y \qquad (15)$$

Violation of (14) or (15) implies the respective Lorenz curves will cross.

Necessary and sufficient conditions for GLD and FSD in terms of the parameters of the Dagum and Singh-Maddala distributions are not available. However, separate sufficient conditions and separate necessary conditions for FSD have been derived. Klonner (2000) shows that these conditions are:

Dagum sufficient condition:

$$(a_X \leq a_Y, \, a_X p_X \geq a_Y p_Y \quad \text{and} \quad b_X \geq b_Y) \Rightarrow (X \geq_{FSD} Y) \qquad (16)$$

Dagum necessary condition:

$$(X \geq_{FSD} Y) \Rightarrow (a_X \leq a_Y \quad \text{and} \quad a_X p_X \geq a_Y p_Y) \qquad (17)$$

Singh-Maddala sufficient condition:

$$(a_X \geq a_Y, \, a_X q_X \leq a_Y q_Y \quad \text{and} \quad b_X \geq b_Y) \Rightarrow (X \geq_{FSD} Y) \qquad (18)$$

Singh-Maddala necessary condition:

$$(X \geq_{FSD} Y) \Rightarrow (a_X \geq a_Y \quad \text{and} \quad a_X q_X \leq a_Y q_Y) \qquad (19)$$

The posterior probabilities that these inequalities hold can be estimated using the proportion of posterior draws that satisfy the inequalities. Evaluating these probabilities is not as computationally demanding as evaluating the probability of dominance. Since the probability of dominance must be less than the probability that the necessary condition holds, and greater than the probability that the sufficient condition holds, the sufficiency and necessity probabilities provide bounds for the dominance probability. The usefulness of these bounds, in terms of the computations that they save, will depend on how close they are.

Although we focus on Lorenz dominance, generalized Lorenz dominance and first order stochastic dominance, there are many other welfare measures that have appeared in the literature and to which we could apply the methodology described in this paper. For some examples see Lambert (1993), Maasoumi (1997) and Barrett and Donald (2000). As long as these measures can be evaluated as functions of the parameters of the income distributions, we can estimate the probability that a welfare measure from one distribution exceeds its counterpart from another distribution. In

addition to overall LD, GLD and FSD, in our example we consider probabilities for LD, GLD and FSD for the poorest 10% of the population and the poorest 20% of the population. Such measures are important when reduction in poverty is the major policy concern. Also, computing the probabilities for them illustrates the flexibility of our approach.

To implement our proposed methodology we must obtain draws of observations from the posterior density functions of the parameters of the distributions. These posterior density functions are derived from prior density and likelihood functions. Details of these steps are given in the next section.

3.   Priors, Likelihood Functions and Posterior Density Functions

Posterior probability density functions for the parameters of the two income distribution models are obtained by combining prior density functions with the likelihood functions as prescribed by Bayes' theorem. Let θ be a vector containing the unknown parameters of an income distribution. Thus, θ will be a vector of 3 parameters, (a, p, b) for the Dagum distribution and (a, q, b) for the Singh-Maddala distribution. Let $p(\theta \mid x), p(x \mid \theta)$ and $p(\theta)$ denote the posterior density, likelihood and the prior density functions, respectively. The posterior probability density function is obtained as:

$$p(\theta \mid x) \propto p(x \mid \theta)\ p(\theta)$$

In this section we first describe the prior density $p(\theta)$ for parameters of both the Dagum and Singh-Maddala distributions, then the likelihood functions are given for both grouped and individual sample data.

The Prior Specification

Conceptualizing prior information on income distribution parameters θ is likely to be difficult because the parameters of the distributions do not have direct economic meanings. To overcome this problem we begin by considering instead priors for mean income, modal income and the Gini coefficient. It is far easier to elicit prior beliefs and information on quantities of interest such as these, than to find prior information on the parameters of the distribution of income. The additional advantage of specifying priors in this way is that the same prior information is used for both the Dagum and Singh-Maddala distributions.

Let $\delta = (\mu, m_0, \gamma)'$ be a 3-dimensional vector containing mean income μ, modal income m_0 and the Gini coefficient γ. We refer to δ as the "economic quantities of interest" because its elements are more likely to be of interest than those of θ and because these are quantities for which some prior information is likely to be available. Following Griffiths, Chotikapanich and Rao (2005), independent gamma distributions were chosen for μ and m_0 and a beta distribution was chosen for γ. The general form of these priors can be written as

$$p(\delta) = p(\mu, m_0, \gamma) = p(\mu) p(m_0) p(\gamma) \tag{20}$$

where

$$p(\mu) = \frac{1}{d^c \Gamma(c)} \mu^{c-1} e^{-\mu/d} \tag{21}$$

$$p(m_o) = \frac{1}{\beta^\alpha \Gamma(\alpha)} m_o^{\alpha-1} e^{-m_o/\beta} \tag{22}$$

and

$$p(\gamma) \;=\; \frac{1}{B(v,w)}\gamma^{v-1}(1-\gamma)^{w-1} \tag{23}$$

Gamma densities were chosen for μ and m_o because they are relatively flexible distributions defined over the range $(0,\ \infty)$; by making suitable choices for the prior parameters we can accommodate a wide range of prior views about mean and modal incomes. Similarly, the beta distribution defined on $(0,\ 1)$ can represent a variety of prior views about the Gini coefficient. Our settings for the prior parameters c, d, α, β, v and w, and the reasons for them, are discussed in Section 4.

The prior distribution on the parameters $p(\theta)$ is obtained by transforming the prior density on the economic quantities of interest, $p(\delta)$. That is,

$$p(\theta) = p(\delta)\left|\frac{\partial\delta}{\partial\theta}\right| \tag{24}$$

Griffiths, Chotikapanich and Rao (2005) provide expressions for the economic quantities of interest as functions of the income distribution parameters, $\delta = g(\theta)$, as well as useful derivatives for computing the Jacobian term $\left|\partial\delta/\partial\theta\right|$.

The Likelihood Functions

The likelihood function, $p(x\,|\,\theta)$ depends on the form of the data. For individual observations it is defined as the density function of the income distribution. That is, if the income distribution follows a Dagum distribution, then $p(x\,|\,\theta)$ is defined by equation (6) and for the Singh-Maddala distribution it is defined by equation (10). If the data are grouped as the number of sampled income units in each of a set of income classes, the likelihood function is the density function for a potential sample of numbers of income units in each of the groups,

$n = \left(n_1, n_2, \ldots, n_M \right)'$, where M is the number of income classes. Let (z_{j-1}, z_j) be the income class limits, with $j = 1, 2, \ldots, M$. The likelihood function is given by the multinomial distribution

$$p(x \mid \theta) \propto \prod_{j=1}^{M} \left[F\left(z_j; \theta \right) - F\left(z_{j-1}; \theta \right) \right]^{n_j} \tag{25}$$

where $F(.)$ is the cumulative distribution function for the income distribution. It is given by equations (7) and (11) for the Dagum and Singh-Maddala distributions, respectively. In our empirical example, we have individual observations and hence work with the Singh-Maddala and Dagum density functions.

In our application we compare dominance probabilities obtained assuming the two income distributions being compared are (i) both Dagum, (ii) both Singh-Maddala, (iii) the first is Dagum and the second is Singh-Maddala, and (iv) the first is Singh-Maddala and the second is Dagum. Thus, in each case we are assuming the chosen distributions are valid descriptions of the population income distributions. One way to relax this assumption within the framework of Bayesian inference is to model average the Singh-Maddala and Dagum distributions (and other distributions if more are considered) and to consider dominance conditions for the averaged distributions.

The Posterior Densities

Combining the prior $p(\theta)$ and the likelihood function $p(x \mid \theta)$, the posterior densities for the parameters of each income distribution are given by

$$\begin{aligned} p(\theta \mid x) &\propto p(x \mid \theta) \; p(\theta) \\ &= p(x \mid \theta) \; p(\delta) \left| \frac{\partial \delta}{\partial \theta} \right| \end{aligned} \tag{26}$$

The posterior densities $p(\theta \,|\, x)$ for both the Dagum and Singh-Maddala parameters are not sufficiently tractable for derivation of the moments of the elements in each θ or for deriving marginal posterior densities. However, it is straightforward to use a Metropolis-Hastings algorithm to draw observations $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(N)}$ from each of the posterior densities. We used a random-walk Metropolis-Hastings algorithm, similar to that employed by Griffiths and Chotikapanich (1997).

4.   Application

The data used to illustrate the methodology are the pre-tax income data obtained from the Canadian Family Expenditure Surveys for the years 1978 and 1986; it was provided by Garry Barrett who used it to illustrate various sampling theory tests (Barrett and Donald 2003a). The sample sizes are 8526 and 9470 for 1978 and 1986, respectively. The sample means and standard deviations are 35,535 and 22,098 for 1978 and 36,975 and 24,767 for 1986. Histograms of the observations with income expressed in thousand dollar units are presented in Figure 1. The means and standard deviations provide preliminary evidence about which income distribution might be preferred. If per capita income is our only criterion for comparison, then 1986 is preferred. The increase in per capita income comes at the expense of an increase in the standard deviation, however. Criteria that favour income distributions with smaller variation may lead to 1978 as preferable.

The first step in our Bayesian approach is to choose prior parameters for the prior distributions on the means, modes and Gini coefficients of the populations from which these samples were drawn. Our objective is to choose settings that yield priors that are relatively noninformative in the sense that they do not conflict with a wide range of prior opinions of applied researchers in the income distribution area. Setting

values in this way means that our results are dominated by the sample information and that we are not open to criticism for employing excessive subjectivity. The same priors were used for both 1978 and 1986. The prior parameter settings for the gamma distributions for the mean and mode were ($c = 1.8$, $d = 30$) and ($\alpha = 1.2$, $\beta = 30$), respectively. Some prior probability intervals from these settings are

$$\Pr(\mu \leq 20) = 0.19 \qquad\qquad \Pr(m_0 \leq 20) = 0.39$$

$$\Pr(20 \leq \mu \leq 60) = 0.46 \qquad\qquad \Pr(20 \leq m_0 \leq 60) = 0.43$$

When viewed in relation to the histograms in Figure 1 and the sample means of 35.5 and 37.0, these probability intervals show that the priors have a relatively large spread and will not conflict with a range of prior views that might be more precise.

The parameter settings chosen for the prior beta distribution for the Gini coefficient were $v = 1.1$ and $w = 2$. Two prior probability intervals from this choice are

$$\Pr(\gamma \leq 0.3) = 0.47 \qquad\qquad \Pr(0.1 \leq \gamma \leq 0.7) = 0.74$$

Again, these intervals demonstrate the relatively noninformative nature of our prior. The sample Gini coefficients for the two years were 0.336 and 0.356.

The techniques described in Sections 2 and 3 were applied, with 35,000 observations being drawn using a random-walk Metropolis-Hastings algorithm, and 5,000 of these being discarded as a burn-in. Plots of the observations were taken to confirm the convergence of the Markov chain. Posterior means and standard deviations for the parameters of the income distributions and the estimated mean incomes (in $1000 units), obtained using the expressions for μ_D and μ_S in Section 2, are presented in Table 1, along with the corresponding maximum likelihood estimates

and their standard errors. The similarity of the Bayesian point estimates to those from maximum likelihood estimation confirms that the prior information has been relatively mild. Also, the Bayesian posterior means for mean incomes μ_D and μ_S are very close to the sample values. The Singh-Maddala and Dagum income distributions obtained by setting the parameters equal to their posterior means are graphed against the histograms in Figure 1. They appear to capture the essential characteristics of these distributions.

If one is interested only in point estimates, and not the probability of dominance, we can use the parameter values in Table 1 to assess whether dominance occurs at the posterior means of the parameters. Checking the necessary and sufficient conditions for LD given in equations (14) and (15), we find that $1978 \geq_{LD} 1986$ when comparing two Singh-Maddala distributions, but not for the two Dagum distributions. In the latter case, however, the estimated Lorenz curves cross only at a population proportion of 0.01, with 1978 being dominant for population proportions greater than 0.01. Comparing the complete 1978 estimated Singh-Maddala Lorenz curve with its 1986 Dagum counterpart, yields $1978 \geq_{LD} 1986$. A 1978 Dagum versus 1986 Singh-Maddala comparison does not yield dominance, however, with the Lorenz curves crossing at population proportions 0.13 and 0.97. A similar exercise can be carried out for GLD and FSD by examining the generalized Lorenz curves and inverse distribution functions, respectively, evaluated at the posterior means of the parameters. In all of these comparisons the relevant curves crossed at least once, indicating that dominance does not occur at these parameter values.

Our remaining results are obtained from relevant pairwise comparisons of the 30,000 draws from each of the four posterior density functions, for the parameters of

the Singh-Maddala and Dagum distributions, for the years 1978 and 1986. In each case probabilities are estimated as the proportion of draws that satisfies an inequality or a dominance relation. We begin by considering the probabilities for the necessary, sufficient and necessary and sufficient conditions for first order stochastic dominance presented in Table 2. The separate sufficient and separate necessary probabilities are obtained using equations (16) to (19); for the necessary and sufficient probabilities the complete inverse distribution function in equation (5) is used.

Table2: First Order Stochastic Dominance

| | 86 over 78 | | 78 over 86 | |
	D vs D	S vs S	D vs D	S vs S
Sufficient	0.0059	0.0000	0.0000	0.0000
Necessary and sufficient	0.1851	0.1058	0.0000	0.0000
Necessary	0.7618	0.1612	0.0000	0.0023

As expected, in each case the probability that the necessary condition is satisfied is greater than or equal to the probability that the necessary and sufficient condition is satisfied which in turn is greater than the probability of satisfying the sufficient condition. There is some probability that the income distribution in 1986 is preferred to that in 1978 in terms of FSD, and zero probability that the converse is true. In the case where two Dagum distributions are compared, the computationally-convenient strategy of looking only at the probability for the necessary condition gives a misleading picture of the actual probability of dominance. The fact that we have some non-zero probabilities for dominance implies there are some parameter draws for which the inverse distribution functions do not cross although, when evaluated at the posterior means, the functions do cross.

A more comprehensive comparison involving not only a comparison of like functions, and not only FSD, is presented in Table 3 and Figures 2 and 3. First, consider Figure 2. The graphs in the FSD, LD and GLD columns were obtained by computing the proportion of draws satisfying, respectively, the dominance relations in equations (5), (2) and (3), for each of 999 values of u from 0.001 to 0.999. Thus, each graph gives the probability of the dominance inequality holding at each population proportion u. Since dominance occurs only when an inequality is satisfied for all u, each dominance probability in Table 3 will be less than or equal to the corresponding minimum value of the "probability graphs" that appear in Figures 2 and 3. To illustrate this fact consider the upper left graph in Figure 2 where two Dagum distributions are compared to see if $1986 \geq_{FSD} 1978$. The minimum point on this graph is approximately 0.21 at the point $u = 0.29$. The probability of dominance, given by the proportion of parameter draws for which the relevant inequality is satisfied for all u (and given in Table 3) is 0.1851. As a check to see if we were considering a sufficiently fine grid of points for u to accurately estimate the probabilities in each case we computed proportions for a finer grid of values for u in the region of the minimum.

The probability graphs in Figures 2 and 3 convey a great deal of information. They show the population proportion(s) that have the greatest effect on dominance or lack of it, and they show how dominance is likely to change if we restrict our focus to a particular segment of the population such as the poorest 10% or 20%. Note that the graphs in Figure 3, concerned with dominance of 1978 over 1986, are a mirror image of the corresponding comparisons in Figure 2 where dominance of 1986 over 1978 is considered.

From the first four columns in Table 3 we find there is some evidence that $1986 \geq_{FSD} 1978$ and $1986 \geq_{GLD} 1978$, but no evidence that the converse is true. On the other hand, there is also evidence that $1978 \geq_{LD} 1986$, but not the converse. Zero probabilities, suggesting no evidence of dominance, correspond to probability curves (Figures 2 and 3) that go to zero for some values of the population proportion u. The fact that there is some preference for 1986 over 1978 in terms of FSD and GLD, but not LD, is because increasing income has been accompanied by increasing inequality.

As expected, in Table 3 the probability of GLD is always as least as great as that for FSD, reflecting the fact that FSD is a stronger condition. In the Figures, this result is conveyed by having GLD curves with minimums greater than or equal to the minimums of the FSD curves.

In the last eight columns of Table 3 we consider what might be viewed as a restricted form of dominance. We examine the probability of one curve exceeding another when only the poorest 20% and then only the poorest 10% of the population are considered. Considering a much smaller subset of the population leads to quite large increases in the probabilities, and the possibility that dominance could occur in either direction. For example, in the lowest 10%, Dagum versus Dagum FSD comparison, we have $\Pr(1986 \geq_{FSD} 1978) = 0.413$, $\Pr(1978 \geq_{FSD} 1986) = 0.264$ and $\Pr(\text{neither dominates}) = 0.323$. This example illustrates the depth of information conveyed by the Figures. They show the effect of considering a restricted form of dominance defined over a subset of population proportions, and we can evaluate the probability of dominance in either direction as well as the probability that there is no dominance.

Although the general message about when dominance might occur is a consistent one, and consistent with the sampling-theory results in Barrett and Donald (2003a, 2003b), some of the probabilities we present are quite sensitive to the distributions used to make the comparisons. For example, the four different comparisons for $1978 \geq_{LD} 1986$ yield probabilities 0.24, 084, 0.00 and 1.00. Insights into these differences can be obtained by examining Figure 3 and recalling our discussion about the behaviour of the Lorenz curves evaluated at the posterior means. Consider the first probability curve in the middle column of Figure 3. Although its lowest point is where it appears to cut the vertical axis at a probability of approximately 0.47, the probability of dominance that we report in Table 3 is 0.24. There is actually a dramatic decline in the probability curve as it approaches zero. How we discovered this fact is discussed shortly. For the moment, note that we previously found that the two Dagum posterior-mean Lorenz curves crossed at a population proportion of 0.01. Thus, there is strong evidence of dominance at most population proportions, but the critical population proportion that determines the probability of overall dominance is close to zero, before the posterior-mean Lorenz curve cross. A similar dramatic fall at zero arises with the second probability curve where two Singh-Maddala distributions are compared. The graph suggests a dominance probability of approximately 0.95 when the actual probability is 0.84. In this case, because the two posterior-mean Lorenz curves do not cross, the dominance probability remains high. In the third probability curve in the second column of Figure 3, the probability of dominance is zero, but, if we restrict the population proportion to lie between approximately 0.25 and 0.9, the dominance probability becomes one. Outside the range (0.25, 0.9) the probability curve drops sharply at both ends, reaching zero at about 0.05 and 0.99. This behaviour can also be explained in terms of

where the posterior-mean Lorenz curves cross, in this case at population proportions of 0.13 and 0.97. Finally, in the last case where the probability of dominance is one, the posterior-mean Lorenz curves do not cross.

It is not surprising that the relative locations of the posterior densities for $L(u)$, $\mu L(u)$ and $F^{-1}(u)$, at a few values of u, have a big impact on the probability of dominance. Dominance requires an inequality to hold for all values of u. The value of u that matters the most will be the one for which the probability of the inequality holding is smallest. It is important, therefore, that the density functions chosen to represent the income distributions are sufficiently flexible to model the data well over the whole income range. Otherwise, unrealistic behaviour at particular population proportions could have a large unrealistic effect on the dominance probabilities. The sensitivity of our results suggests that at least one of the Dagum or Singh-Maddala densities is too restrictive for our data. Kleiber (1996) has pointed out that the Dagum density is likely to be better fitting distribution for incomes around zero, while the Singh-Maddala density is likely to be preferable for modelling the right-tail of the income distribution. In this regard future research that applies our methodology to more flexible income distributions, including mixtures, or within a Bayesian nonparametric framework, is likely to be productive. In any event, the probability curves that we introduced in Figures 2 and 3 are a good device for assessing the degree of sensitivity, and the population proportions that contribute most to that sensitivity.

The remaining issue to resolve is the behaviour around zero for the Dagum vs. Dagum and Singh-Maddala vs. Singh-Maddala Lorenz dominance probabilities. Using the inequalities in equations (14) and (15), we find that

Dagum vs Dagum: $\qquad$ $\Pr(1978 \geq_{LD} 1986) = 0.238$

Singh-Maddala vs Singh-Maddala: $\quad \Pr(1978 \geq_{LD} 1986) = 0.837$

However, in our calculations using a grid of u values from 0.001 to 0.999 these probabilities were 0.474 and 0.953, respectively. On further investigation, we discovered the discrepancy occurred because we did not (initially) consider values of u sufficiently close to zero. The sensitivity of the results to the minimum value of u is given in Table 4 for these two cases. The differences are dramatic. These results also show that quite different Lorenz dominance probabilities can be obtained if one is prepared to ignore a small proportion of the population. Considering only the inequalities that are necessary and sufficient, without also considering the complete probability curve, can lead to a large loss of information.

5. Concluding Remarks

The development of statistical inference for assessing whether 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. The purpose of this paper was to illustrate how such dominance relationships can be assessed within a framework of Bayesian inference. Bayesian inference has the advantage of reporting results in terms of probabilities - a natural way to express our uncertainty. Because it enables us to give probabilities for dominance in either direction, as well as the probability that dominance does not occur, it overcomes the problem of giving favourable treatment to what is chosen as the null hypothesis in sampling theory inference. We introduced the concept of a probability curve that describes the

posterior probability of a dominance inequality being satisfied at every population proportion. These curves are useful for examining what population proportions have the biggest effect on dominance and for assessing how conclusions change if a restricted range of population proportions is considered. The sensitivity of our results to the assumed family of income distributions suggests that application of our techniques to more flexible families of distributions is likely to be a productive avenue for future research.

References

Anderson, G. (1996), "Nonparametric Tests of Stochastic Dominance in Income Distributions" *Econometrica*, 64(5), 1183-1194.

Barrett, G.F. and S.G. Donald (2000), "Statistical Inference with Generalized Gini Indices of Inequality and Poverty", Working Papers, School of Economics, University of New South Wales.

Barrett, G.F. and S.G. Donald (2003a), "Consistent Tests for Stochastic Dominance", *Econometrica*, 71(1), 71-104.

Barrett, G.F. and S.G. Donald (2003b), "Consistent Nonparametric Tests for Lorenz Dominance", Working Papers, School of Economics, University of New South Wales.

Bishop J. A. and J.P. Formby (1999), "Tests of Significance for Lorenz Partial Orders", in J. Silber, editor, *Handbook on Income Inequality Measurement*, London: Kluwer Academic Publishers.

Chotikapanich, D. and J. Creedy (2004), "The Atkinson Inequality Measure and its Sampling Properties: Bayesian and Classical Approaches", *Australian Economic Papers*, 43, 3, 288-301.

Cowell, F.A. (1999), " Estimation of Inequality Indices", in J. Silber, editor, *Handbook on Income Inequality Measurement*, London: Kluwer Academic Publishers.

Creedy, J (1996), *Fiscal Policy and Social Welfare*, Cheltenham: Edward Elgar Publishing.

Dardanoni, V. and A. Forcina (1999), "Inference for Lorenz Orderings", *Econometrics Journal*, 2, 49-75.

Davidson, R. and J.-Y. Duclos (1997), "Statistical Inference for the Measurement of the Incidence of Taxes and Transfers", *Econometrica*, 65(6), 1453-1466.

Davidson, R. and J.-Y. Duclos (2000), "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality", *Econometrica*, 68(6), 1435-1464.

Griffiths, W.E., D. Chotikapanich (1997), "Bayesian Methodology for Imposing Inequality Constraints on a Linear Expenditure Function with Demographic Factors", *Australian Economic Papers*, 36, 321-341.

Griffiths, W.E., D. Chotikapanich and D.S.P. Rao (2005), "Averaging Income Distributions", *Bulletin of Economic Research*, 57, 347-367.

Hasegawa, H and H. Kozumi (2003), "Estimation of Lorenz Curves: A Bayesian Nonparametric Approach", *Journal of Econometrics*, 115, 277-291.

Kleiber, C (1996), "Dagum vs. Singh-Maddala Income Distributions", *Economics Letters*, 53, 265-268.

Kleiber, C. and S. Kotz (2003), *Statistical Size Distributions in Economics and Acturial Sciences*, New Jersey: Weiley & Sons.

Klonner, S. (2000), "The First-order Stochastic Dominance Ordering of the Singh-Maddala Distribution", *Economics Letters*, 69, 123-128.

Lambert, P. J. (1993), *The Distribution and Redistribution of Income*, Manchester: Manchester University Press.

Linton, O., E. Maasoumi, and Y.J. Whang (2005), "Consistent Testing for Stochastic Dominance under General Sampling Schemes", *Review of Economic Studies*, 72(3), 735-766.

Maasoumi, E. (1997), "Empirical Analyses of Inequality and Welfare", in M. H. Pesaran and P. Schmidt, editors, Handbook of Applied Econometrics, Volume II: Microeconomics, Malden, MA: Blackwell.

McDonald, J.B. (1984), "Some Generalized Functions of the Size Distribution of Income," *Econometrica*, 52, 647-663.

McDonald, J.B. and Y.J. Xu (1995), "A Generalization of the Beta Distribution with Applications", *Journal of Econometrics* 66, 133-152. Errata 69 (1995), 427-428.

Wilfing, B and W. Kramer (1993), "The Lorenz Ordering of Singh-Maddala Income Distributions", *Economics Letters*, 43, 53-57.

Tse, Y.K. and X. Zhang (2002), "A Monte Carlo Investigation of Some Tests for Stochastic Dominance", Working Papers, Department of Econometrics and Business Statistics, Monash University.

Table 1: ML estimates and standard errors and posterior means
and standard deviations (based on income in $1000units)

| | ML | | Bayesian | |
	1978	**1986**	**1978**	**1986**
Dagum				
a	4.3153	3.6071	4.3176	3.6072
	(0.1030)	(0.0874)	(0.1051)	(0.0879)
p	0.3566	0.4343	0.3566	0.4350
	(0.0130)	(0.0173)	(0.0132)	(0.0174)
b	49.399	47.793	49.410	47.770
	(0.6740)	(0.8742)	(0.7060)	(0.8747)
mean	36.056	37.893	36.058	37.897
			(0.2507)	(0.2788)
Singh-Maddala				
a	1.8727	1.8327	1.8663	1.8309
	(0.0254)	(0.0252)	(0.0251)	(0.0254)
q	6.1431	4.2298	6.3912	4.2711
	(0.7420)	(0.3831)	(0.8116)	(0.4121)
b	98.418	82.111	101.01	82.674
	(8.0960)	(5.4296)	(8.6691)	(5.8103)
mean	35.571	37.032	35.567	37.040
			(0.2363)	(0.2512)

Table 3: Probabilities for Lorenz, Generalized Lorenz and First Order Stochastic Dominance

86 over 78

	Over all dominance				Lowest 20%				Lowest 10%			
	D vs D	S vs S	D vs S	S vs D	D vs D	S vs S	D vs S	S vs D	D vs D	S vs S	D vs S	S vs D
FSD	0.1851	0.1058	0.0000	0.0000	0.2577	0.1056	0.0000	0.0051	0.4130	0.1096	0.0000	0.0051
Lorenz	0.0000	0.0000	0.0000	0.0000	0.0043	0.0002	0.0000	0.0534	0.0567	0.0018	0.0000	0.0534
G Lorenz	0.2735	0.1081	0.0000	0.0138	0.4065	0.1096	0.0000	0.4391	0.5160	0.1145	0.0000	0.4391

78 over 86

	Over all dominance				Lowest 20%				Lowest 10%			
	D vs D	S vs S	D vs S	S vs D	D vs D	S vs S	D vs S	S vs D	D vs D	S vs S	D vs S	S vs D
FSD	0.0000	0.0000	0.0000	0.0000	0.2643	0.7204	0.0000	0.1041	0.2643	0.8074	0.0000	0.8284
Lorenz	0.2382	0.8365	0.0000	1.0000	0.3243	0.9014	0.0000	1.0000	0.3243	0.9014	0.0000	1.0000
G Lorenz	0.0000	0.0000	0.0000	0.0000	0.2632	0.7979	0.0000	0.8352	0.2632	0.8300	0.0000	0.9928

Figure 1: Histogram and density functions
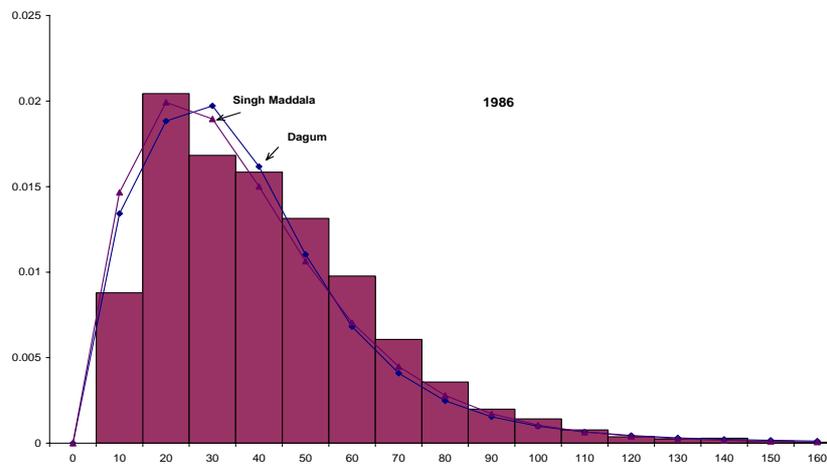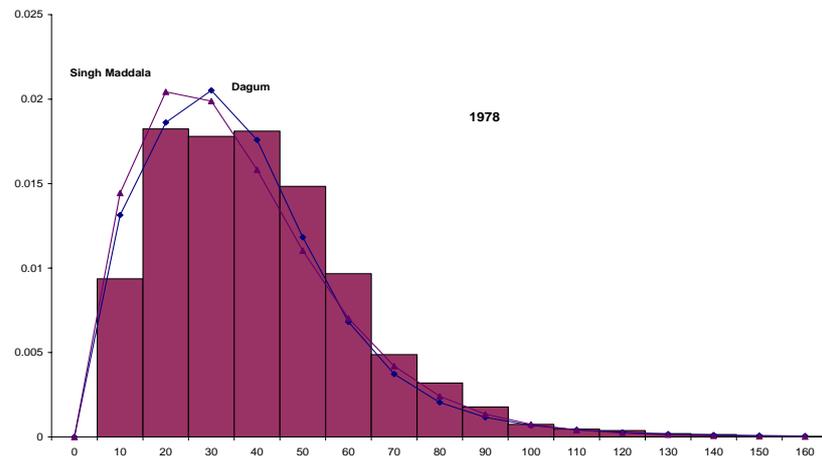
Figure 2: Probability Curves for Dominance of 86 over 78
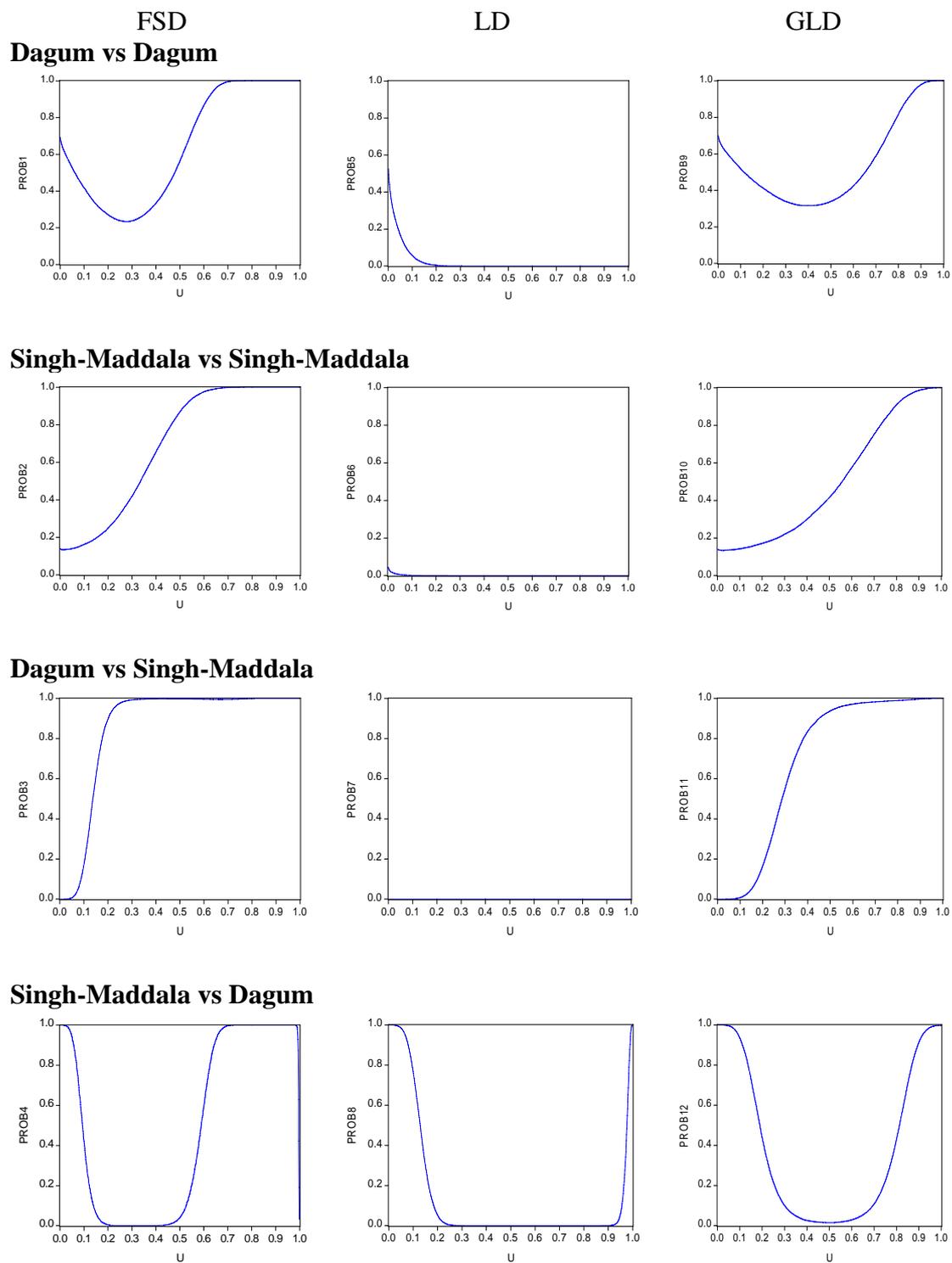
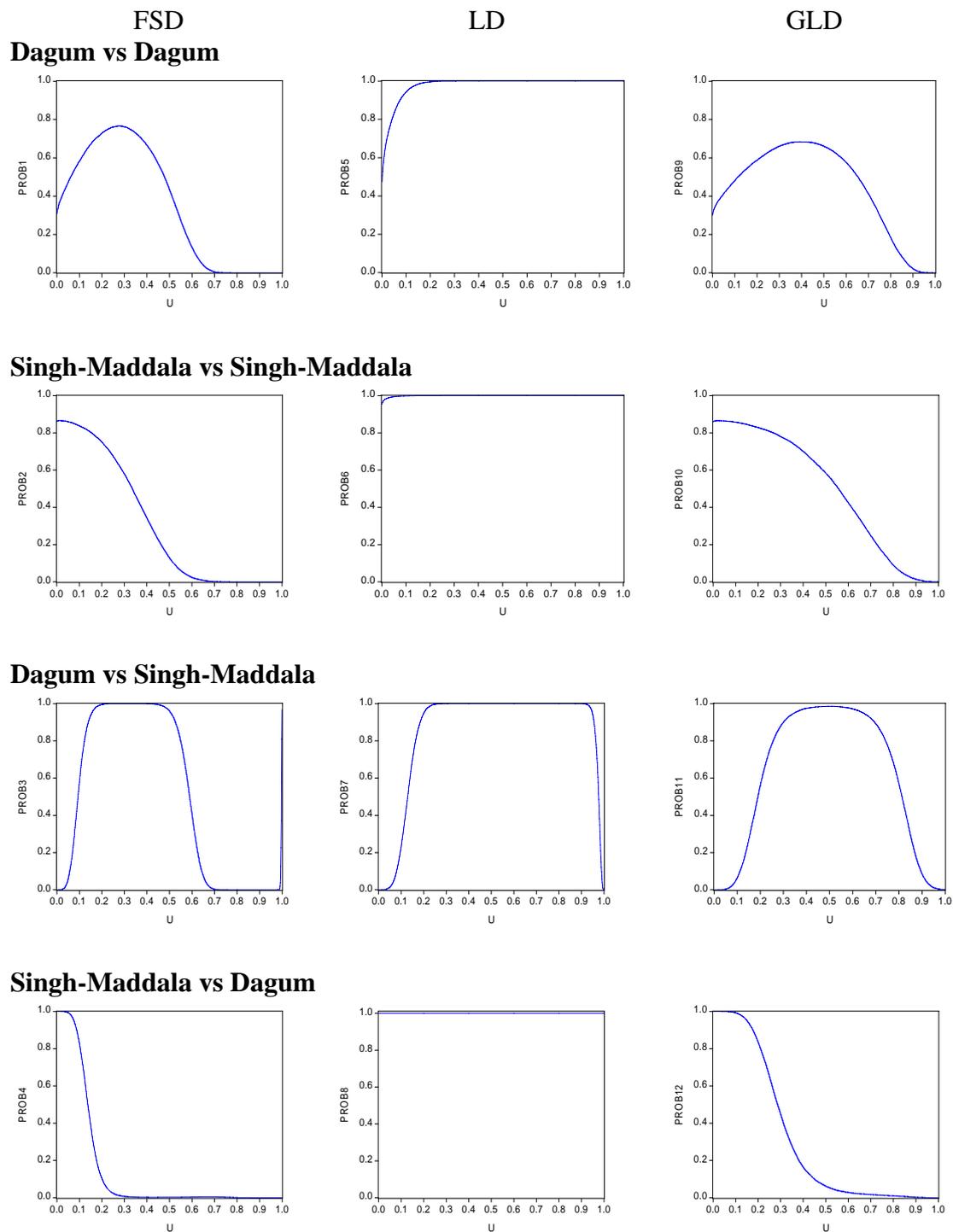**Figure 3**: Probability Curves for Dominance of 78 over 86

Table 4: Lorenz Dominance for 1978 over 1986

Min population proportion	D vs D	S vs S
0.001	0.4744	0.9533
0.0001	0.4044	0.9336
0.00001	0.3661	0.9195
0.000001	0.3417	0.9096
0.0000001	0.3243	0.9014
0	0.2382	0.8365