Consistent Nonparametric Tests for Lorenz Dominance

Garry F. Barrett
School of Economics
University of Sydney
Sydney, NSW, 2006

Stephen G. Donald
Department of Economics
University of Texas at Austin
Austin, TX, 78712

Debopam Bhattacharya
Department of Economics
University of Oxford
Oxford, OX1 3UQ, UK

Version: April 2013.

Abstract

Consistent nonparametric methods for testing the null hypothesis of Lorenz dominance are proposed. The methods are based on a class of statistical functionals defined over the difference between the Lorenz curves for two samples of welfare related variables. Two specific test statistics belonging to the general class are presented and their asymptotic properties derived. As the limiting distributions of the test statistics are non-standard, we propose and justify bootstrap methods of inference. We provide methods appropriate for case where the two samples are independent as well as the case where the two samples represent different measures of welfare for one set of individuals. The small sample performance of the two tests is examined and compared in the context of a Monte Carlo study and an empirical analysis of income and consumption inequality.

Keywords: Lorenz dominance, test consistency, bootstrap simulation.
J.E.L. Classification: C12, C14, C15, D63

1. INTRODUCTION

A fundamental tool for the analysis of economic inequality is the Lorenz curve which graphs the cumulative proportion of total income, or other measure of individual welfare, by cumulative proportion of the population after ordering from poorest to richest. The related concept of Lorenz dominance provides a partial ordering of income distributions based on minimal normative criteria. Distribution A weakly Lorenz dominates distribution B if the Lorenz curve for A is nowhere below that for B. As shown by Atkinson (1970), Lorenz dominance translates into simple facts concerning the degree of egalitarianism in the respective income distributions. Lorenz dominance is equivalent to the ranking of income distributions based on the class of scale-free inequality indices that respect the 'principle of transfers' - whereby a progressive transfer is associated with a decrease in inequality - while avoiding the imposition of stronger additional normative criteria embodied in a specific scalar index of inequality. An empirical method for directly inferring Lorenz dominance is therefore very desirable.

The work of Beach and Davidson (1983) represented a key development in the use of Lorenz curves for statistical inference in economics. They derived the sampling properties of a subset of ordinates from the empirical Lorenz curve and presented a test for the null hypothesis that two independent Lorenz curves are equal. Note that this was a test of Lorenz equality, rather than dominance, at a fixed set of population proportions. Bishop, Formby and Smith (1991a, 1991b) proposed a test of Lorenz dominance based on multiple pair-wise comparisons of empirical Lorenz ordinates. Davies, Green and Paarsch (1998), Dardanoni and Forcina (1999) and Davidson and Duclos (2000) presented tests for Lorenz dominance based on a predetermined grid of points. The null hypothesis of dominance across those fixed points imply a series of inequality restrictions which can be tested using the methods of Wolak (1989). Although these tests use information on the covariances among the set of estimated Lorenz ordinates, making them more powerful than the Bishop et al. (1991a, 1992b) tests, these methods are also potentially inconsistent. By limiting attention to a small fixed set of grid points, the tests do not take account of the full set of restrictions implied by Lorenz dominance.

The aim of the current article is to develop consistent tests for Lorenz dominance. Our approach to testing is based on a class of statistical functionals defined over the difference between two Lorenz curves. A test of Lorenz dominance may be considered as a scalar measure of the extent to which one Lorenz curve (hereafter LC) is everywhere above the other. Two test statistics based on specific functionals from the general class are examined in detail. The first test statistic is based on the largest difference between the two LCs - a supremum or Kolmogorov-Smirnov (KS) type test, while the second is a Cramer von-Mises (CVM) type test based on the integral of the difference between the curves over the range of ordinates for which one lies above the other. This second test statistic was first presented in Bhattacharya (2007) in the context of analysing inequality using stratified and clustered survey data. Both measures will be zero when one curve weakly dominates another, and both will be strictly positive when this is not the case. The tests are nonparametric and based on normalized estimates of quantities involving the empirical LCs. The empirical LC is a fully nonparametric, $\sqrt{n}$-consistent estimator of the true underlying LC. The empirical LC does not share the disadvantages associated with other nonparametric estimators such as for density and regression models. Our estimation problem is analogous to the estimation of a cumulative distribution function for which nonparametric estimation via the empirical distribution (or smoothed empirical distribution) function is known to be $\sqrt{n}$-consistent and asymptotically normal with Brownian Bridge limit processes. Representing the empirical LC as a smooth functional of the empirical distribution function permits the application of the functional delta method to obtain the limit processes.

The second feature of our tests is that they are consistent in that they detect any violation of the null hypothesis of weak Lorenz dominance. This is achieved by comparing the empirical LCs at all quantiles. The tests presented in this article utilize all the sample information and provide a consistent test of Lorenz dominance. Our tests are analogous to tests of stochastic dominance (SD) proposed in McFadden (1989) and elaborated and extended by Barrett and Donald (2003). SD relations are based on a comparisons of CDFs (or partial integrals of CDFs) and provide partial orderings in terms of welfare levels or poverty. In contrast, Lorenz dominance is based on a comparisons of (mean independent) LCs which provides a partial ordering in terms of relative inequality, as articulated in Atkinson (1970; 1987) and Deaton (1997: 157-169). Further, as the empirical LC is given by the partial integral of the empirical quantile function normalised by the mean, LD testing must address the issue of small denominators in studying convergence, which is an issue that does not arise in SD testing. The main difficulty with our tests of Lorenz dominance is that the limiting distributions of the test statistics are nonstandard and generally depend on the underlying LCs. We propose and justify the use of the bootstrap for conducting inference. The application of the bootstrap in approximating the asymptotic distribution of a test statistic has been used for similar problems in Andrews (1997), Barrett and Donald (2003) and Linton, Maasoumi and Whang (2005). Our main results are obtained for two possible sampling schemes for estimating the LCs. The first is that we have two independent samples of comparable variables, for differing numbers of individuals. The second is that we have one sample of individuals and two measures of welfare (e.g.: before and after tax income or in panel contexts), which we refer to as "matched pair" sampling. The difference between the two sampling schemes is that in the latter case the estimated LCs will be correlated, whereas in the former case they will not. This has important implications for how we use the bootstrap in each case. One could also justify inference using the bootstrap for more elaborate sampling schemes, such as those considered in Bhattacharya (2005).

The remainder of the article is organized as follows. In Section 2 we state our testing problem, review key results on the properties of empirical LCs, propose two test statistics and provide a characterization of the limiting distributions of the test statistics under the null hypothesis in terms of well known stochastic processes. In Section 3 the non-parametric bootstrap approach to conducting inference is presented and theoretically justified. Section 4 provides a brief Monte Carlo study that examines how well the asymptotic arguments work in small samples. In Section 5 we implement the tests by comparing the LCs for the distribution of income and consumption in Australia from

1984 to 2009/10. In Section 6 concluding comments are presented.

2. ASYMPTOTIC PROPERTIES OF LORENZ DOMINANCE TEST STATISTICS

2.1 Preliminaries

We are interested in comparing the LCs associated with the distributions of income (or some other measure of welfare) for variables X^1 and X^2. These could either be corresponding variables from two different populations for which we have independent random samples or else these could be two measures of welfare for a specific individual from a single population. We let F_1 and F_2 denote the respective marginal cumulative distribution functions (CDFs). We make the following assumptions regarding these CDFs.

Assumption 1 *Assume that the population described by $F_j : [0, \infty) \to [0, 1]$ (for $j = 1, 2$) has finite first two moments and is continuously differentiable with associated probability density function given by $f_j(z) = F_j'(z)$ such that $f_j(z)$ is strictly positive everywhere on $[0, \infty)$ and for some $\gamma \in (0, 1)$ the following tail condition is satisfied,*

$$\lim_{x\to\infty} \frac{\{1 - F_j(x)\}^{1+\gamma}}{f_j(x)} = 0 = \lim_{x\to 0} \frac{\{F_j(x)\}^{\gamma}}{f_j(x)} \tag{1}$$

The existence of two moments is sufficient for us to define the LCs (at ordinate value $p \in [0, 1]$) for the respective populations by,

$$L_j(p) = \frac{\int_0^{Q_j(p)} z f_j(z)dz}{\int_0^{\infty} z f_j(z)dz} = \frac{\int_0^p Q_j(t)dt}{\mu_j}$$

where $Q_j(p) = F_j^{-1}(p)$ are the respective quantile functions and μ_j is the mean of the distribution. The tail condition on the distributions will allow us to derive weak convergence results for the empirical LC as shown in the next subsection.

2.2 Hypothesis Formulation

The hypotheses that we are interested in testing are:

$$H_0^1 \; : \; L_2(p) \leq L_1(p) \text{ for all } p \in [0, 1]$$

$$H_1^1 \; : \; L_2(p) > L_1(p) \text{ for some } p \in [0, 1]$$

The null hypothesis is that the LC for population F_1 is everywhere at least as large as that for the population F_2. This will be referred to as weak Lorenz Dominance of L_1 over L_2. This formulation of the hypotheses is consistent with much of the literature on testing stochastic dominance (McFadden 1989). Note that the null hypothesis also includes the case where the LCs coincide. As has been shown in Lambert (1993), this can only occur if $F_1(z) = F_2(\alpha z)$ for some non-negative value of α. That is, multiplying all incomes in a population by the same constant does not affect the LC associated with the distribution. The alternative hypothesis is true whenever the LC for F_2 is above that for F_1 at some point. Note that we can reverse the roles of F_1 and F_2 and test similar hypotheses. This would allow one to determine whether a LC dominated another in a stronger sense. In particular, if one considered the hypotheses

$$
\begin{aligned}
H_0^2 &: L_1(p) \leq L_2(p) \text{ for all } p \in [0,1] \\
H_1^2 &: L_1(p) > L_2(p) \text{ for some } p \in [0,1]
\end{aligned}
$$

then the hypotheses H_0^1 and H_1^2 together imply the strong dominance of L_1 over L_2 so that in principle one could use the tests to determine whether or not there is strong Lorenz dominance. In addition, the hypotheses H_0^1 and H_0^2 together imply that the LCs are identical. The Bonferroni inequality provides a bound for the p-value for the union of the two LD tests. Alternatively, a direct test of the null of LC equality, $H_0^{eq} : L_2(p) = L_1(p)$ for all $p \in [0,1]$, can be constructed based on the standard KS test applied to LCs rather than CDFs.

We consider the approach to testing based on a functional of the difference between the two LCs which gives a scalar result that indicates which of the hypotheses is correct. In order to justify a bootstrap approach to inference we impose additional regularity conditions on the functional. For this purpose we define $\phi(p) = L_2(p) - L_1(p)$ and note that under our assumptions ϕ is a continuous function on $[0,1]$. Thus we can write, $\phi \in C[0,1]$. Also let $\|.\|$ denote the sup norm on $C[0,1]$. We develop our theory of testing and inference for a general functional $\mathcal{F} : C[0,1] \to R$, which we can normalize such that $\mathcal{F}(0) = 0$ and $\mathcal{F}$ satisfies the following properties:

Property 1: *For any* $\phi^*, \phi' \in C[0,1]$:

(i) If $\phi'(p) \leq 0$ *for all* $p \in [0,1]$ *then* $\mathcal{F}(\phi^*) \leq \mathcal{F}(\phi^* - \phi')$,

(ii) If $\phi^*(p) > 0$ *for some* $p \in (0,1)$ *then* $\mathcal{F}(\phi^*) > 0$,

(iii) $|\mathcal{F}(\phi^*) - \mathcal{F}(\phi')| \leq \|\phi^* - \phi'\|$,

(iv) any scalar constant $c > 0$, $c\mathcal{F}(\phi^*) = \mathcal{F}(c\phi^*)$

(v) $\mathcal{F}$ *is convex*

Properties 1(i) (ii) and the normalization are sufficient to show that the functional can be used to distinguish between the null and alternative hypothesis based on the scalar value of the functional. The latter properties are continuity conditions that allow one to derive weak convergence properties for the test statistics based on the functional and also allow easy justification of the bootstrap method. The condition 1(v) is a convexity condition that will allow us to show that the distribution of the test statistic is absolutely continuous. This condition is not the only one that will guarantee that this result holds but is satisfied for the two functionals considered in this article (see Davydov, Lifshits and Smorodina (1998) for methods and assumptions for establishing absolute continuity of distributions of functionals of random processes). Our first result shows that Property 1(i) and (ii) allow one to distinguish between the null and alternative based on the functional.

Lemma 1: *If* $\mathcal{F}$ *satisfies Property 1(i) and (ii) then* H_0^1 (H_1^1) *is equivalent to* $\mathcal{F}(\phi) \leq 0$ $(\mathcal{F}(\phi) > 0)$.

The two specific functionals considered in this article are

$$\begin{aligned} \mathcal{S}(\phi) &= \sup_{p \in [0,1]} (\phi(p)) \\ \mathcal{I}(\phi) &= \int_0^1 \phi(p) 1(\phi(p) > 0) dp \end{aligned}$$

where $1(A)$ represents the indicator function which is equal to 1 when A is true (and 0 otherwise). The next Lemma establishes that these functionals satisfy all parts of Property 1. Therefore these two functional are capable of distinguishing between the two hypotheses plus they satisfy the regularity conditions for weak convergence and justification of the bootstrap approach to inference considered in subsequent sections.

Lemma 2: *Each of the functionals $\mathcal{S}$ and $\mathcal{I}$ satisfy Property 1.*

2.3 Properties of the Empirical Lorenz Curve and Test Statistics

Our aim is to make inferences regarding Lorenz dominance based on samples drawn under two possible sampling situations. The first is classical independent random sampling from two populations.

Assumption 2 (IS): *Assume that:*

(i) $\{X_i^j\}_{i=1}^{n_j}$ is a random sample from F_j and the sample for $j = 1$ is independent from the sample for $j = 2$.

(ii) the sampling scheme is such that as $n_1 \to \infty$

$$\lim_{n_1 \to \infty} \frac{n_1 n_2}{n_1 + n_2} \to \infty$$

The first part is the standard independent random samples assumption that would be appropriate in situations where we have two separate random samples from non-overlapping populations such as countries or regions and would also generally be a plausible assumption if the two samples are random samples at two different points in time for the same population. Note we allow for differing sample sizes. The requirement in (ii) is that, as far as the asymptotic analysis is concerned, the number of observations in each sample is not fixed as the other grows. We do allow for the possibility that one sample size grows at a faster rate than the other. This condition is key for the consistency properties of the test under the random sampling assumption. For this case we define the following,

$$\lim_{n_1 \to \infty} \frac{n_1}{n_1 + n_2} \to \lambda \in [0, 1]$$

and λ can take on one of the endpoints when one sample size grows faster than the other. Note that the random sampling assumption could be relaxed in ways that are discussed in Bhattacharya (2005).

We also consider an alternate sampling scheme whereby X^1 and X^2 represent different random variables for the same individual, referred to as the matched pairs case. We have in mind that X^j could represent measures of the same welfare variable at different points in time, such as with panel data, or where they represent different measures of welfare for an individual at a single point in time, such as income and expenditure. In the former case one is then considering LD based on panel data while in the latter case one is interested in relative inequality between two notions of welfare. For these types of situations we use the following assumption, where MP is shorthand for matched pairs.

Assumption 2 (MP): *Assume that $\{(X_i^1, X_i^2)\}_{i=1}^n$ is a random sample from a joint distribution $F(x_1, x_2)$ whose marginals are given by F_1 and F_2.*

In this case there is only one sample size n so in what follows, except where indicated, the notation n_j refers to this common n for this sampling assumption. Also, unlike the independent random sampling case, while it makes sense to assume that (X_i^1, X_i^2) is independent of (X_j^1, X_j^2) (for $i \neq j$) it is implausible to assume that X_i^1 is independent of X_i^2. As we see below this will imply that the estimated LC's for the two variable will be dependent and this will need to be taken into account in the inference procedure. Provided the pair (X_i^1, X_i^2) are iid a simple adjustment of the bootstrap can be performed so that valid inference is possible even without knowing the nature of the dependence between the two variables.

The empirical distributions are given by

$$\hat{F}_j(z) = \frac{1}{n_j} \sum_{i=1}^{n_j} 1(X_i^j \leq z)$$

and the quantile functions as

$$\hat{Q}_j(p) = \inf\{z : \hat{F}_j(z) \geq p\}$$

Then the empirical LC at ordinate value p, can be defined in terms of the quantile function by

$$\hat{L}_j(p) = \frac{\int_0^p \hat{Q}_j(t)dt}{\hat{\mu}_j}$$

where $\hat{\mu}_j = \bar{X}_N^j$ are the sample means. Since the quantile process is a step function (right continuous) then the empirical LC is a piecewise linear function starting at the origin and reaching the value 1 when $p = 1$. For a given sample $\{X_i^j\}_{i=1}^{n_j}$, denote the unique values by $x_1 < x_2 < \ldots < x_{n_j^*}$ (with $n_j^* \leq n_j$), the sample mean by $\hat{\mu}$, and denote the proportion of observations in the sample that take on each of these values as $\hat{\pi}_r$ then the empirical LC is obtained by connecting the points

$$\left(\sum_{l=1}^{r} \hat{\pi}_l, \sum_{l=1}^{r} \frac{\hat{\pi}_l x_l}{\hat{\mu}} \right) : r = 1, ..., n_j$$

with straight lines. Thus the empirical LC is easily computed and, like the population LC, is continuous and convex.

To set notation, for an arbitrary distribution function F_j define $\mathcal{B}_j \circ F_j$ as the Brownian Bridge process composed of F_j. As is well known, appropriately standardized empirical distribution functions (considered as elements of the space of cadlag functions $D[a, b]$ on $[a, b]$) satisfy the following weak convergence results:

$$\sqrt{n_j}(\hat{F}_j - F_j) \Rightarrow \mathcal{B}_j \circ F_j$$

Note that in the case of Assumption 2(IS) it follows that since the two samples are independent then $\mathcal{B}_1 \circ F_1$ is also independent of $\mathcal{B}_2 \circ F_2$. In the case of Assumption 2(MP) we have that,

$$\sqrt{n} \begin{pmatrix} \hat{F}_1 - F_1 \\ \hat{F}_2 - F_2 \end{pmatrix} \Rightarrow \begin{pmatrix} \mathcal{B}_1 \circ F_1 \\ \mathcal{B}_2 \circ F_2 \end{pmatrix}$$

where the limit is a bivariate correlated Brownian Bridge with covariance function (at the point (x_1, x_2)) given by,

$$\begin{pmatrix} F_1(x_1)(1 - F_1(x_1)) & F(x_1, x_2) - F_1(x_1)F_2(x_2) \\ F(x_1, x_2) - F_1(x_1)F_2(x_2) & F_2(x_2)(1 - F_2(x_2)) \end{pmatrix} \quad (2)$$

Such a result follows from marginal weak convergence using arguments in van der Vaart and Wellner (1996, Sections 1.1 and 1.4). Since in this case X_i^1 and $X_i^2$are from the same unit of observation, it is unreasonable to assume that the off diagonals are zero.

Our first result provides a characterization of the limiting properties of the empirical LCs. Since the LC is a scaled version of the integral of the quantile function the standardized empirical LCs can be considered as members of the function space $C[0,1]$ since they are piecewise linear and continuous. Define the Gaussian stochastic process, $\mathcal{G}_j$ on $[0,1]$ to be such that for $p \in [0,1]$,

$$\mathcal{G}_j(p) = -\int_0^p \frac{\mathcal{B}_j(t)}{f_j(Q_j(t))} dt$$

and finally the process $\mathcal{L}_j$ to be such that for $p \in [0,1]$,

$$\mathcal{L}_j(p) = \frac{\mathcal{G}_j(p)}{\mu_j} - \frac{L_j(p)}{\mu_j}\mathcal{G}_j(1)$$

Under Assumption 2(IS) these $\mathcal{L}_1$ and $\mathcal{L}_2$ will be independent since $\mathcal{B}_1$ and $\mathcal{B}_2$ are independent. On the other hand, under Assumption 2(MP) since the Brownian Bridge processes $\mathcal{B}_1$ and $\mathcal{B}_2$ are correlated then the Lorenz processes $\mathcal{L}_1$ and $\mathcal{L}_2$will also be correlated. The following result concerning the asymptotic behavior of the empirical Lorenz processes is is stated for completeness and will be the basis for inference methods based on the functionals satisfying Property 1.

Lemma 3: *Given Assumption 1 and either 2(IS) or 2(MP),*

(i) for each j,

$$\sup |\hat{L}_j(p) - L_j(p)| \overset{a.s.}{\to} 0$$

and in the space $C[0,1]$, $\sqrt{n_j}(\hat{L}_j - L_j) \Rightarrow \mathcal{L}_j$

(ii) letting $\hat{\phi} = \hat{L}_2 - \hat{L}_1$ under 2(IS) with $T_n = n_1 n_2/(n_1+n_2)$ we have,

$$\sqrt{T_n}(\hat{\phi} - \phi) \Rightarrow \bar{\mathcal{L}} = \sqrt{\lambda}\mathcal{L}_2 - \sqrt{1-\lambda}\mathcal{L}_1$$

and under 2(MP) we have $T_n = n$,

$$\sqrt{T_n}(\hat{\phi} - \phi) \Rightarrow \bar{\mathcal{L}} = \mathcal{L}_2 - \mathcal{L}_1$$

Results such as in (i) for the single Lorenz process date back to Goldie (1977) under slightly different conditions. The weak convergence result in (i) can be derived using functional delta methods described in van der Vaart and Wellner (1996). This requires showing that the LC is a Hadamard differentiable function of the CDF for which the tail condition in Assumption 1 is sufficient (using a result shown in Bhattacharya (2007)). Beach and Davidson (1983) also presented results for a vector of LC ordinates and importantly showed how to do inference, by providing estimates of the variance covariance, matrix without imposing distributional assumptions. Here we consider inference on the entire LC.

The second result follows immediately from the first part and assumptions concerning the sample sizes that are explicit in Assumption 2(IS)(ii) or implicit in Assumption 2(MP). This result is stated formally so as to define the process $\bar{\mathcal{L}}$ which appears in the limiting distributions of the test statistics considered in the next section. Note that this differs in terms of its properties depending on whether we are using Assumption 2(IS), in which case λ appears and $\mathcal{L}_1$and $\mathcal{L}_2$ are independent, or Assumption 2(MP) in which case $\mathcal{L}_1$and $\mathcal{L}_2$ are correlated. Our inference methods are designed to deal with these differences in behavior.

This result allows one to obtain the properties of the test statistic for general functional $\mathcal{F}$ in a straightforward fashion. As in Lemma 3 we allow the normalizing factor for each sampling Assumptions 2(IS) and 2(MP), $\sqrt{T_n}$, to differ as stated in Lemma 3.

Lemma 4: *Under Assumptions 1 and 2(IS) or 2(MP) and assuming that $\mathcal{F}$ satisfies Property 1 then,*

(i) Under H_0^1, $T_n\mathcal{F}(\hat{\phi}) \leq T_n\mathcal{F}(\hat{\phi} - \phi) \Rightarrow \mathcal{F}(\bar{\mathcal{L}})$ where $\bar{\mathcal{L}}$ is as given in Lemma 3 and for $\alpha < 1/2$ the $1 - \alpha$ quantile of the distribution of $\mathcal{F}(\bar{\mathcal{L}})$ is strictly positive, finite and unique,

(ii) Under H_1^1 $T_n\mathcal{F}(\hat{\phi}) \xrightarrow{p} \infty$

This result shows that the test statistic can be used to test between the null and alternative in much the same way as one would test a one sided hypothesis on a single parameter. The test statistic is dominated under the null hypothesis by a statistic that is asymptotically distributed as $\mathcal{F}(\bar{\mathcal{L}})$. The inequality in (i) is an equality when the LCs are identical with $\phi = 0$. One rejects the null for large values of the test statistic $T_n\mathcal{F}(\hat{\phi})$ and one would require a critical value with the property that $P(\mathcal{F}(\bar{\mathcal{L}}) > c_\alpha | H_0^1) = \alpha$ so that the test will have significance level equal to α. The result in (i) guarantees that the critical value is finite so that the divergence of the test statistic under the alternative guarantees that the test will be consistent. An alternative and equivalent way to test the hypotheses is using G (say), the distribution of $\mathcal{F}(\bar{\mathcal{L}})$. One would reject the null if the p-value $\hat{p}(\mathcal{F}) = 1 - G(T_n\mathcal{F}(\hat{\phi}))$ is less than α. In this particular situation because the distribution G is both nonstandard and population dependent (i.e. it depends on both F_1 and F_2 as well as the covariance between the associated Brownian Bridges under Assumption 2(MP)) we require a data based bootstrap approach to inference.

3. BOOTSTRAP BASED INFERENCE

In order to conduct the tests in such a way that they have known asymptotic significance levels we propose using the bootstrap to estimate asymptotic p-values. In the case of Assumption 2(IS) we treat the original samples independently and for this purpose let $\mathcal{X}^j = \{X_i^j\}_{i=1}^{n_j}$ for $j = 1, 2$ be the two original samples. In this case one can bootstrap by independently drawing (with replacement) samples of size n_j from each of $\mathcal{X}^1$ and $\mathcal{X}^2$. Denote these samples by $X_1^{j*}, ..., X_{n_j}^{j*}$ for $j = 1, 2$. In this case one will have bootstrap estimates of empirical distributions given by,

$$\hat{F}_j^*(x) = \frac{1}{n}\sum_{i=1}^{n_j} 1(X_i^{j*} \leq x)$$

where X_i^{1*} is randomly drawn from $\hat{F}_1$ and X_i^{2*} is randomly drawn from $\hat{F}_2$. Under Assumption 2(MP) $\hat{F}_1^*$ and $\hat{F}_2^*$ are obtained by sampling from the n matched pairs $\{(X_i^{1*}, X_i^{2*})\}_{i=1}^n$with replacement from the observed sample X= $\{(X_i^1, X_i^2)\}_{i=1}^n$. That is, we randomly select observational units (with replacement) so that (X_i^{1*}, X_i^{2*}) is the set of both measures for the ith randomly chosen unit. This adjustment will allow one to capture the dependence across the two dependent Lorenz processes.

For each bootstrap sample define

$$\begin{aligned} \hat{Q}_j^*(p) &= \inf\{z : \hat{F}_j^*(z) \geq p\} \\ \hat{L}_j^*(p) &= \frac{\int_0^p \hat{Q}_j^*(t)dt}{\hat{\mu}_j^*} \end{aligned}$$

where $\hat{\mu}_j^*$ is the mean of the bootstrap samples (either independent samples or matched pairs). Then we define $\hat{\phi}^*(p) = \hat{L}_2^*(p) - \hat{L}_1^*(p)$. In order to obtain a valid approximation to the distribution of the test statistic under the null hypothesis we need to subtract $\hat{\phi}(p)$ so that the object $T_n\mathcal{F}(\hat{\phi}^*(p) - \hat{\phi}(p))$ will have the same limiting distribution as $\mathcal{F}(\bar{\mathcal{L}})$. Using this, our bootstrap p-values can be computed by finding (under Assumption 2(IS)),

$$\hat{p}(\mathcal{F}) = P(T_n\mathcal{F}(\hat{\phi}^*(p) - \hat{\phi}(p)) > T_n\mathcal{F}(\hat{\phi}(p))|\mathcal{X}^1, \mathcal{X}^2)$$

or (under Assumption 2(MP)),

$$\hat{p}(\mathcal{F}) = P(T_n\mathcal{F}(\hat{\phi}^*(p) - \hat{\phi}(p)) > T_n\mathcal{F}(\hat{\phi}(p))|\mathcal{X})$$

Equivalently, one can find the probability that the random variable $T_n\mathcal{F}(\hat{\phi}^*(p)-\hat{\phi}(p))$ lies above the test statistic conditional on the sample(s). This p-value can be approximated by Monte Carlo simulation as

$$\hat{p}(\mathcal{F}) \simeq \frac{1}{R}\sum_{r=1}^{R} 1(T_n\mathcal{F}(\hat{\phi}_r^*(p) - \hat{\phi}(p)) > T_n\mathcal{F}(\hat{\phi}(p)))$$

where $\hat{\phi}_r^*(p)$ is the rth resampled difference of LCs. The test is then based on the decision rule,

$$\text{“reject } H_0^1 \text{ if } \hat{p}(\mathcal{F}) < \alpha\text{”} \tag{3}$$

Proposition 1: *Under Assumptions 1 and 2 and given that $\mathcal{F}$ satisfies Property 1 then the test based on the decision rule (3) has the following properties,*

$$\begin{aligned} \lim P(\text{reject } H_0^1) &\leq \alpha \text{ if } H_0^1 \text{ is true} \\ \lim P(\text{reject } H_0^1) &= 1 \text{ if } H_0^1 \text{ is false} \end{aligned}$$

An immediate implication of this result is that the bootstrap approach will work for the test statistics based on the functionals $\mathcal{S}$ and $\mathcal{I}$:

$$\begin{aligned}\hat{M}_S &= \sqrt{T_n}\mathcal{S}(\hat{\phi}) \\ \hat{M}_I &= \sqrt{T_n}\mathcal{I}(\hat{\phi})\end{aligned}$$

For completeness, we also present the KS test of LC equality: $H_0^{eq} : L_2(p) = L_1(p)$ for all $p \in [0,1]$ against $H_1^{eq} : L_2(p) \neq L_1(p)$ for some $p \in [0,1]$. A test based on the statistical functional $\mathcal{M}_E(\phi) = \sup_{p\in[0,1]}(|\phi(p)|)$ with associated test statistic $\hat{M}_E = \sqrt{T_n}\mathcal{S}(|\hat{\phi}|)$ is readily constructed. The asymptotic distribution of this statistic under the null can be approximated using an analogous bootstrap procedure with the p-value given by $\hat{p}(\hat{M}_E) \simeq \frac{1}{R}\sum_{r=1}^{R} 1(\hat{M}_{Er}^* > \hat{M}_E)$ where $\hat{M}_{Er}^* = \sqrt{T_n}\mathcal{S}(|(\hat{\phi}_r^*(p) - \hat{\phi}(p)|)$. It is straightforward to show the validity of this approximation.

4. MONTE CARLO RESULTS

4.1 Independent Sampling

In this section we consider a small scale Monte Carlo experiment to gauge the extent to which the preceding asymptotic properties hold in small samples. The initial experiments examine the properties of the tests under independent random sampling. In the first set of experiments, our specifications for the distributions are in the log-normal family because they are easy to simulate and they have been used in empirical work on income distributions. We generate two sets of samples from two possibly different distributions. In the first two cases we generate X_i^1 and X_i^2 as independent log-normal random variables using the equations,

$$\begin{aligned}X_i^1 &= \exp(\sigma_1 Z_{1i} + \mu_1) \\ X_j^2 &= \exp(\sigma_2 Z_{2j} + \mu_2)\end{aligned} \qquad (4)$$

where the Z_{1i} and Z_{2j} are independent $N(0,1)$. In Case 1, $\mu_1 = \mu_2 = 0.85$ and $\sigma_1 = \sigma_2 = 0.6$. With this choice of parameters the two populations have the same distribution with means equal to 2.8 and standard deviations equal to 1.8 – the ratio of the mean to the standard deviation of 1.55 is similar to that found in actual income data. In Case 1 the LCs for the two populations are identical and our interest is in the size properties of the testing procedure.

For Case 2 $\mu_1 = 0.85$, and $\sigma_1 = 0.6$ while $\mu_2 = 0.85$ and $\sigma_2 = 0.55$. In this case the LC for X^2 dominates the LC for X^1 – indeed the LC for X^2 lies above that for X^1 everywhere except at the endpoints of the interval $[0, 1]$. In this case we should expect to to reject the hypotheses H_0^1 and H_0^{eq} but not H_0^2. Note that in this case we expect that the test will reject H_0^2 less often than the nominal size of the test because of the inequality in Proposition 1.

In Case 3, we generate X^1 as before but now generate X^2 as a mixture of log-normal random variables. In particular,

$$X_i^2 = 1(U_i \geq 0.2)\exp(\sigma_2 Z_{2j} + \mu_2) + 1(U_i < 0.2)\exp(\sigma_3 Z_{2j} + \mu_3)$$

where U_i is a uniform $[0, 1]$ random variable, Z_{2j} and Z_{3j} are independent standard normal random variables and where $\mu_2 = 0.6$, and $\sigma_2 = 0.2$ while $\mu_3 = 1.8$ and $\sigma_3 = 0.3$. In this case we have crossing LCs. Neither LC dominates the other, nor are the LCs equal, and we expect H_0^1, H_0^2 and H_0^{eq} to be rejected.

In a second set of experiments, we simulate distributions based on the Singh-Maddala (SM) specification. This family of distributions has been popular in empirical and experimental work and, unlike the log-normal, the SM distribution is "heavy-tailed". The CDF of the SM distribution is given by $F(X) = 1 - \frac{1}{[1+x^a]^q}$ where a and q are shape parameters. In designing these experiments we exploit a theoretical result of Wilfling and Kramer (1993): for two SM distributions, denoted $SM(a_1, q_1)$ and $SM(a_2, q_2)$ respectively, with $a_1 \leq a_2$, $SM(a_2, q_2)$ will Lorenz dominate $SM(a_1, q_1)$ iff $a_1 q_1 \leq a_2 q_2$.

We generate X_i^1 and X_i^2 as SM random variables using the equations for the inverse SM CDF:

$$\begin{aligned} X_i^1 &= ((1 - U_{1i})^{-1/q_1} - 1)^{1/a_1} \\ X_i^2 &= ((1 - U_{2i})^{-1/q_2} - 1)^{1/a_2} \end{aligned} \quad (5)$$

where the U_{1i} and U_{2i} are independent uniform $[0, 1]$ random variables.

In Case 4, we set $a_1 = a_2 = 1.6$ and $q_1 = q_2 = 2.265$. These parameter values were obtained by fitting the SM distribution to the United States individual-equivalent gross income distribution data from the 1998 March Current Population Survey. Like Case 1, the LCs for the two distributions are equal and we consider the size properties of the tests (but, here, simulating from a heavy-tailed distribution). In Case 5, we generate X_i^1 as in Case 4 but set $a_2 = 1.7$ and $q_2 = q_1$. For this case the X^2 distribution Lorenz dominates that for X^1,though by only a relatively small amount. We should expect to reject the hypotheses H_0^1 and H_0^{eq} but not the hypothesis H_0^2. In Case 6, X_i^1 is generated as before, while $a_2 = 1.8$ and $q_2 = q_1$. The distribution for X^2 Lorenz dominates that for X^1 by a greater amount than in Case 5 and consequently we expect a stronger rejection of H_0^1 and H_0^{eq} (and less rejection of H_0^2) in Case 6. For Case 7, the final experiment, X_i^1 is generated as before and $a_2 = 3.8$ and $q_2 = 0.47$. This specification leads to a single crossing of the LCs, violating H_0^1 over the bottom three quintiles of the distributions, and therefore we expect rejection of H_0^1, H_0^2 and H_0^{eq}.

In performing the test of Lorenz Dominance we use the decision rule,

$$\text{“reject } H_0^j \text{ if } \hat{p}_G < \alpha\text{”}$$

where $\hat{p}_G$ is the simulated p-value for the test statistic $\hat{T}_G$. For all of the experiments we used sample sizes of $N = M = 500$. The number of bootstrap replications was set to 500 to approximate the p-value in each Monte Carlo iteration, and 1000 iterations were performed for each experiment. The results for the Monte Carlo simulations are reported in Table 1. The table reports the proportion of times that the respective null hypothesis was rejected for three different nominal significance levels α.

Table 1. Monte Carlo Rejection Rates: Independent Sampling

	Test	H_0^1			H_0^2			*Test*	H_0^{eq}		
		Nominal Size			Nominal Size				Nominal Size		
		0.10	0.05	0.01	0.10	0.05	0.01		0.10	0.05	0.01
Log-normal distributions											
Case 1	M_S	0.102	0.049	0.013	0.081	0.039	0.012	M_E	0.100	0.051	0.008
	M_I	0.108	0.052	0.011	0.084	0.037	0.014				
Case 2	M_S	0.609	0.450	0.204	0.003	0.000	0.000	M_E	0.960	0.907	0.723
	M_I	0.709	0.573	0.293	0.000	0.000	0.000				
Case 3	M_S	0.923	0.700	0.178	0.994	0.988	0.944	M_E	0.997	0.995	0.900
	M_I	0.801	0.501	0.095	0.620	0.401	0.114				
Singh-Maddala distributions											
Case 4	M_S	0.114	0.059	0.017	0.107	0.054	0.009	M_E	0.110	0.048	0.005
	M_I	0.111	0.057	0.017	0.106	0.044	0.009				
Case 5	M_S	0.329	0.197	0.069	0.032	0.009	0.001	M_E	0.213	0.124	0.035
	M_I	0.414	0.284	0.112	0.011	0.006	0.002				
Case 6	M_S	0.645	0.491	0.230	0.004	0.001	0.000	M_E	0.499	0.367	0.165
	M_I	0.765	0.651	0.385	0.001	0.000	0.000				
Case 7	M_S	0.166	0.081	0.020	0.717	0.549	0.244	M_E	0.652	0.477	0.153
	M_I	0.257	0.163	0.062	0.214	0.112	0.030				

A number of features of the tests are of note. The first series of experiments were based on the log-normal distribution, and in the first case the size properties of the tests were examined Each of the M_S, M_I and M_E test procedures led to the rejection of the true null hypotheses at rates similar to the nominal size. There was a slight under rejection for H_0^2; however the under rejection was not severe, and the actual size of the tests were close to their nominal size. In terms of power, the test procedures appear to be quite similar where there is strong dominance. In Case 2 the tests detect the fact that the LC for X^2 dominates that for X^1. The hypotheses H_0^1 and H_0^{eq} are rejected with high probability. Note that the hypothesis H_0^2 is rarely rejected in this case – this feature of the test is related to the one sided composite nature of the null hypothesis and is similar to the behavior of tests of one sided restrictions on parameters. In Case 3, neither LC is dominant and the tests reject each null considered with very high probability, although the rejection is stronger for the M_S test compared to M_I.

The second series of experiments were based on the SM distribution. Case 4 provides a further comparison of the size properties of the tests. Again, the actual and nominal size of the tests were very similar. Although there was a slight over rejection of H_0^1, the discrepancy between actual and nominal size was minor. It is useful to note that the sample sizes considered in the experiments were relatively small compared to many empirical applications and the fact that the actual sizes of the tests in these experiments are close to the nominal size is encouraging. In terms of power, the tests are able to detect the violation of H_0^1 and H_0^{eq} in Case 5, with the strong rejection of these nulls and, conversely, the rejection of true null of H_0^2 is well below the nominal size. In Case 6, where the Lorenz dominance of X^2 over X^1 is stronger, the rejection rates for H_0^1 and H_0^{eq} are greater, and the rejection of the true null H_0^2 is further below the nominal size. In Case 7, with crossing LCs, the M_S, M_I and M_E tests detect the violation of the null hypotheses. The null H_0^1 is violated by a small amount over the bottom three quintiles, which the M_I is relatively better at detecting than M_S. Conversely, the null H_0^2 is sharply violated over the top quintiles which the M_S test is relatively superior at detecting. Even so, both tests detect the violation of the false null sufficiently well to reject at rates well in excess of the nominal size.

4.2 Matched Pair Sampling

The Monte Carlo experiments were repeated with the simulated samples drawn from dependent distributions to reflect matched pair sampling. Each case was repeated with identical specifications for the marginal distributions and pre-determined correlation. The method proposed by Cario and Nelson (1997) for generating correlated random samples was adopted, which involved generating bivariate standard normal random variables $(X_{1i}, \tilde{X}_{2i})$ with correlation $\tilde{\rho}$ using the algorithm $\tilde{X}_{2i} = \tilde{\rho}.X_{1i} + (1-\tilde{\rho})^2.X_{2i}$ where (X_{1i}, X_{2i}) are independent, as in the initial series of experiments. For the log-normal simulations the variates $(X_{1i}, \tilde{X}_{2i})$ are demeaned and transformed as in (4), and for the SM simulations the variates are demeaned, converted to uniform variates by applying the normal CDF then transformed to SM variates using the quantile function in (5). A numerical search over values of $\tilde{\rho}$ was performed to obtain the desired correlation ρ of the simulated log-normal and SM variates. The Monte Carlo experiments were performed for values of the correlation coefficient $\rho = \{0.3, 0.7, 0.9\}$. These values are comparable to the correlation between family income and food expenditure, family income and non-durable expenditures, and pre-tax and post-tax income, respectively.

Results of the Monte Carlo simulation for the cases with $\rho = 0.7$ are reported in Table 2. Rejection rates for the simulations involving different values of correlation coefficients were very similar to those in Table 2 and hence are not reported. As is evident from Table 2, the tests under matched pair sampling continue to exhibit very good size characteristics. In terms of power performance, the tests tend to reject more strongly the false null hypotheses in Cases 2-3 and 5-7 under dependent sampling. Overall, series of small scale Monte Carlo experiments indicate that each of the test procedures, under both independent and matched-pair sampling, exhibits good size and power properties. Further, when the sample size for the Monte Carlo experiments is increased slightly, the asymptotic properties are clearly reflected in enhanced size and power characteristics.

Table 2. Monte Carlo Rejection Rates: Matched Pair Sampling (ρ=0.7)

	Test	H_0^1			H_0^2			*Test*	H_0^{eq}		
		Nominal Size			Nominal Size				Nominal Size		
		0.10	0.05	0.01	0.10	0.05	0.01		0.10	0.05	0.01
Log-normal distributions											
Case 1	M_S	0.088	0.047	0.013	0.097	0.054	0.014	M_E	0.102	0.052	0.007
	M_I	0.090	0.052	0.014	0.104	0.064	0.013				
Case 2	M_S	0.697	0.555	0.259	0.002	0.001	0.000	M_E	0.588	0.432	0.201
	M_I	0.832	0.735	0.468	0.001	0.001	0.001				
Case 3	M_S	0.956	0.765	0.193	0.995	0.985	0.939	M_E	1.000	0.995	0.910
	M_I	0.871	0.593	0.128	0.662	0.429	0.131				
Singh-Maddala distributions											
Case 4	M_S	0.096	0.042	0.007	0.116	0.060	0.015	M_E	0.109	0.052	0.009
	M_I	0.100	0.047	0.010	0.123	0.069	0.019				
Case 5	M_S	0.425	0.270	0.092	0.021	0.005	0.000	M_E	0.333	0.210	0.069
	M_I	0.608	0.464	0.203	0.007	0.003	0.000				
Case 6	M_S	0.861	0.748	0.438	0.003	0.000	0.000	M_E	0.773	0.643	0.366
	M_I	0.966	0.923	0.752	0.000	0.000	0.000				
Case 7	M_S	0.249	0.137	0.039	0.865	0.722	0.363	M_E	0.829	0.663	0.306
	M_I	0.371	0.240	0.088	0.297	0.137	0.042				

5. EMPIRICAL EXAMPLE

The methods for testing Lorenz dominance relations are illustrated with an analysis of the distribution of income and consumption in Australia. The data are from the Australia Bureau of Statistics Household Expenditure Survey (HES) conducted in 1984, 1988/89, 1993/94, 1998/99, 2003/04 and 2009/10 (hereafter referenced by the first year of the survey period). The income measure is gross annual family income. The consumption measure is expenditure on non-durables, consisting of food, alcohol and tobacco, fuel, clothing, personal care, medical care, transport, recreation, utilities and current housing services. Current housing services for renters is equal to rent paid while for home-owners it is imputed from a regression of rent payments on a series of indicator variables for number of bedrooms and location of residence by survey year for the subsample of renters. The sample is restricted to families where the household reference person is between 25 and 60 years of age.

Family income and consumption were divided by the adult equivalent scale (AES) equal to the square-root of family size. To minimise reporting errors only multiple-family households are excluded. The HES is a stratified random sample and for each observation there is an associated weight representing the inverse probability of selection into the survey. The observational weights were multiplied by the number of family members in order to make the sample representative of individuals; the adjusted weights were used throughout the analysis.

Summary statistics are reported in Table 3. Nominal prices are inflated to 2010 real values using the CPI. The summary statistics show that the mean budget share of the non-durable commodity bundle was 68 percent in 1984. Over the sample period non-durable consumption grew at an average annual rate of 2.36 percent while income grew at an average annual rate of 2.53 percent. Point estimates for the Gini coefficient suggest a substantial increase in income inequality, and a minor change in consumption inequality, over the 25 year period.

Table 3: HES 1984-2009 Summary Statistics

Year	Sample Size	Income			Consumption		
		Mean	Std. Dev.	Gini	Mean	Std. Dev.	Gini
1984	2895	254.90	157.54	0.322	173.09	84.06	0.251
1988	4654	263.26	180.79	0.323	173.02	84.04	0.247
1993	5396	265.49	197.15	0.346	186.50	94.08	0.247
1998	4645	296.11	198.85	0.342	201.40	96.89	0.248
2003	4583	325.97	227.25	0.330	215.14	104.75	0.249
2009	5009	408.75	364.59	0.361	251.42	131.73	0.260

The first comparisons examine changes in the distribution of individual equivalent income over time. Table 4 presents the p-values for the test statistics of the null hypothesis that distribution 1 weakly Lorenz dominated distribution 2, against the alternative that the null is false. To calculate the p-values, 2000 bootstrap repetitions were used to simulate the distribution of the test statistics. The first two rows of the table are for the test with distribution 1 corresponding to 1984 income and distribution 2 corresponding to 1988 income. The results show that neither null of dominance, H_0^{84} or H_0^{88}, can be rejected at the 5% level of significance. The p-value for the null of LC equality of 0.159, which is similar to the bound based on the Bonferroni inequality for M_S of 0.146, does not lead to rejection at conventional levels of significance. The tests indicate that the two income distributions were 'equally unequal.' The following two rows show the null hypothesis that the 1988 income distribution Lorenz dominated the 1993 distribution cannot be rejected, while the converse null, that the 1993 distribution dominated 1988, can be rejected at the 5% level of significance. Strong Lorenz dominance of the 1988 income distribution over the 1993 distribution can be inferred at conventional levels of significance. Comparison of the 1993 and 1998 income distributions shows that the two LCs are equal, while the 2003 distribution is found to strongly Lorenz dominate both the 1998 and 2009 distributions at the 10% level of significance. Across the full observation period, the 1984 income distribution strongly Lorenz dominates the 2009 income distribution. The increase in income inequality over the 1984-2009 period was concentrated in the 1988-93 and 2003-09 subperiods, where the former coincided with the severe recession of 1990/1991 while the latter includes the global finance crisis which began in 2007.

Consumers with access to credit facilities may smooth transitory fluctuates in current income. It is therefore of interest to examine changes in consumption inequality over time. In comparing across surveys, relative inequality in the distribution of consumption shows much greater stability. The consumption LCs between adjacent surveys in the 1984-2003 period were found to coincide. The 2003 consumption distribution weakly Lorenz dominated the 2009 distribution, although the null of LC equality is not rejected at conventional levels of significance (the Bonferroni bound on the p-value from the sequential application of M_S is 0.126). The income and consumption LCs for the 2003 and 2009 samples are plotted in Figures 1 and 2, respectively, showing the greater increase in income inequality between the two survey years. The findings suggest that the increases in income inequality coinciding with the 1990/1991 recession and the onset of the global financial crisis in 2007 had a transitory component which households were largely able to smooth.

The lower panels of Table 4 present several additional comparisons. The standard life-cycle model of consumption implies that the distribution of non-durable consumption will be more equal than the distribution of current income at a point in time. This hypothesis was tested with the bootstrap procedure adapted for matched pair sampling to replicate the dependence in the data. The comparison of the empirical income and consumption distributions for each survey year strongly supports this hypothesis. Further, comparing income and consumption at different points in time, the consumption distribution in 2009 strongly Lorenz dominated the distribution of income in 1984 (the most equal income distribution). Less surprising, the 1984 consumption distribution strongly Lorenz dominated the 2009 income distribution.

Overall, the Lorenz dominance tests show a rise in income inequality in Australia between 1984 and 2009, though consumption inequality remained stable. The empirical results suggest that households were generally insured against shocks to the income process over the observation period. The test results show that the distribution of consumption was more equal than the distribution of income at each point in time, and over the study period. In terms of the performance of the two tests of Lorenz dominance, both gave essentially the same result which suggests that either may be used in practice.

Table 4. P-Values for Lorenz Dominance Tests

F_1	F_2	Test	H_0^1	H_0^2	H_0^{eq}
Y1984	Y1988	M_S	0.611	0.079	0.159
		M_I	0.486	0.640	
Y1988	Y1993	M_S	0.942	0.002	0.001
		M_I	0.913	0.001	
Y1993	Y1998	M_S	0.129	0.699	0.266
		M_I	0.221	0.601	
Y1998	Y2003	M_S	0.030	0.181	0.066
		M_I	0.030	0.542	
Y2003	Y2009	M_S	0.976	0.011	0.016
		M_I	0.972	0.001	
Y1984	Y2009	M_S	0.846	0.000	0.000
		M_I	0.857	0.000	
C1984	C1988	M_S	0.188	0.550	0.385
		M_I	0.233	0.630	
C1988	C1993	M_S	0.451	0.306	0.619
		M_I	0.442	0.405	
C1993	C1998	M_S	0.216	0.515	0.443
		M_I	0.498	0.399	
C1998	C2003	M_S	0.431	0.415	0.807
		M_I	0.463	0.368	
C2003	C2009	M_S	0.886	0.063	0.127
		M_I	0.907	0.050	
C1984	C2009	M_S	0.965	0.193	0.358
		M_I	0.969	0.093	
C1984	Y1984	M_S	0.974	0.000	0.000
		M_I	0.974	0.000	
C1988	Y1988	M_S	0.991	0.000	0.000
		M_I	0.991	0.000	
C1993	Y1993	M_S	0.974	0.000	0.000
		M_I	0.974	0.000	
C1998	Y1998	M_S	0.995	0.000	0.000
		M_I	0.995	0.000	
C2003	Y2003	M_S	0.949	0.000	0.000
		M_I	0.974	0.000	
C2009	Y2009	M_S	0.989	0.000	0.000
		M_I	0.989	0.000	
Y1984	C2009	M_S	0.000	0.686	0.000
		M_I	0.000	0.824	
C1984	Y2009	M_S	0.991	0.000	0.000
		M_I	0.991	0.000	

6. CONCLUSION

In this article we proposed two methods for testing for Lorenz dominance, along with a test of LC equality, based on samples from two, potentially dependent, populations. The tests presented are fully non–parametric and consistent being based on global comparisons of the empirical LCs. Although the proposed test statistics have non-standard and case specific limiting distributions we were able to show that asymptotically valid inferences could be drawn using the bootstrap. Each of the tests were shown to have a good performance in quite small samples and were illustrated in the context of an empirical example comparing income and consumption LCs for Australia over the period 1984-2009/10.

APPENDIX: PROOFS OF RESULTS

Proof of Lemma 1: Suppose that H_0^1 holds then $\phi \leq 0$,

$$\mathcal{F}(\phi) \leq \mathcal{F}(\phi - \phi) = \mathcal{F}(0) = 0$$

On the other hand $\mathcal{F}(\phi) \leq 0$ implies that $\phi \leq 0$ by Property 1(ii). Clearly under H_1^1 we have $\mathcal{F}(\phi) > 0$ by Property 1(ii). The converse follows easily since if $\mathcal{F}(\phi) > 0$ then it cannot be the case that H_0^1 is true since if it were true, i.e.$\phi \leq 0$ then using Property 1(i),

$$0 < \mathcal{F}(\phi) \leq \mathcal{F}(\phi - \phi) = \mathcal{F}(0) = 0$$

which is false. Consequently H_1^1 must be true. **Q.E.D.**

Proof of Lemma 2: For Property 1(i) we have that,

$$\phi^*(p) \leq \phi^*(p) - \phi'(p)\ \forall p$$

so that $\mathcal{S}$ is easily seen to satisfy the property while $\mathcal{I}$ does by properties of the integral and since, $\phi^*(p) > 0 \Longrightarrow \phi^*(p) - \phi'(p) > 0$ so that,

$$\phi^*(p)1(\phi^*(p) > 0) \leq (\phi^*(p) - \phi'(p))\, 1\,(\phi^*(p) - \phi'(p) > 0)$$

For Property 1(ii) we have that if there is a p such that $\phi^*(p) > 0$ then,

$$\mathcal{S}(\phi^*) \geq \phi^*(p) > 0$$

while continuity of ϕ^* implies that there is a neighborhood of p on which $\phi^*(p') > \varepsilon > 0$ for all p' such that $|p' - p| < \delta$ so that,

$$\begin{aligned}
\mathcal{I}(\phi^*) &\geq \int_{p-\delta}^{p+\delta} \phi^*(p')1(\phi^*(p') > 0)dp \\
&> \int_{p-\delta}^{p+\delta} \varepsilon dp = 2\delta\varepsilon > 0
\end{aligned}$$

For Property 1(iii) for $\mathcal{S}$ we have,

$$\mathcal{S}(\phi^*) \leq \mathcal{S}(\phi' + \phi^* - \phi') \leq \mathcal{S}(\phi') + \mathcal{S}(\phi^* - \phi')$$

so,

$$\mathcal{S}(\phi^*) - \mathcal{S}(\phi') \leq \mathcal{S}(\phi^* - \phi') \leq \|\phi^* - \phi'\|$$

reversing ϕ' and ϕ^* we have,

$$\mathcal{S}(\phi') - \mathcal{S}(\phi^*) \leq \|\phi^* - \phi'\|$$

so,

$$-\|\phi^* - \phi'\| \leq \mathcal{S}(\phi^*) - \mathcal{S}(\phi')$$

and Property 1(iii) follows. For $\mathcal{I}$ we have,we have that,

$$|\phi^*(p)1(\phi^*(p) > 0)\, dp - \phi'(p)1(\phi'(p) > 0)| \leq |\phi^*(p) - \phi'(p)|$$

which is obvious when $\phi^*(p) > 0$ and $\phi'(p) > 0$ and also when both are negative. When, $\phi^*(p) > 0$ and $\phi'(p) \leq 0$ we have,

$$\begin{aligned}
|\phi^*(p)1(\phi^*(p) > 0)\, dp - \phi'(p)1(\phi'(p) > 0)| &= |\phi^*(p)| \\
&\leq |\phi^*(p) - \phi'(p)|
\end{aligned}$$

and similarly for the other case. Hence,

$$|\mathcal{I}(\phi^*) - \mathcal{I}(\phi')| \leq \int_0^1 |\phi^*(p) - \phi'(p)|\, dp \leq \|\phi^* - \phi'\|$$

Property 1(iv) is obvious for $\mathcal{S}$ and follows for $\mathcal{I}$ by linearity of the integral operator and the fact that,

$$c\phi^*(p) > 0 \Longleftrightarrow \phi^*(p) > 0$$

For Property 1(v) let ϕ' and ϕ^* be continuous functions and let $\beta \in (0,1)$. Then for $\mathcal{S}$ the result follows by properties of supremum since,

$$\begin{aligned} \mathcal{S}(\beta\phi' + (1-\beta)\phi^*) &\leq \mathcal{S}(\beta\phi') + \mathcal{S}((1-\beta)\phi^*) \\ &= \beta\mathcal{S}(\phi') + (1-\beta)\mathcal{S}(\phi^*). \end{aligned}$$

For the functional $\mathcal{I}$ we have that,

$$\begin{gathered} \mathcal{I}(\beta\phi' + (1-\beta)\phi^*) = \int_0^1 \beta\phi'(p)1(\beta\phi'(p) + (1-\beta)\phi^*(p) > 0)dp \\ + \int_0^1 (1-\beta)\phi'(p)1(\beta\phi'(p) + (1-\beta)\phi^*(p) > 0)dp \\ \leq \beta\int_0^1 \phi'(p)1(\beta\phi'(p) > 0) + \int_0^1 (1-\beta)\phi'(p)1((1-\beta)\phi^*(p) > 0) \\ = \beta\mathcal{I}(\phi') + (1-\beta)\mathcal{I}(\phi^*) \end{gathered}$$

using the following facts,

$$\begin{aligned} \beta\phi'(p)1(\beta\phi'(p) + (1-\beta)\phi^*(p) &> 0) \leq \beta\phi'(p)1(\beta\phi'(p) > 0) \\ (1-\beta)\phi'(p)1(\beta\phi'(p) + (1-\beta)\phi^*(p) &> 0) \leq (1-\beta)\phi'(p)1((1-\beta)\phi^*(p) > 0) \end{aligned}$$

To see that these hold consider the first expression. There are two possible ways in which $\beta\phi'(p) + (1-\beta)\phi^*(p) > 0$ holds. First it could be that $\beta\phi'(p) > 0$, in which case the expression on the left is equal to the expression on the right. Second, it could be that $\beta\phi'(p) < 0$, in which case the left hand side is negative while the right hand side is positive. Thus the inequality holds, and the same argument can be applied to the second expression. **Q.E.D.**

Proof of Lemma 4: (i) Under the null hypothesis $\phi(p) = L_2(p) - L_1(p) \leq 0$ for all $p \in (0,1)$. By Property 1(i) and (iv) we then have that,

$$\begin{aligned} T_n\mathcal{F}(\hat{\phi}) &\leq T_n\mathcal{F}(\hat{\phi} - \phi) = \mathcal{F}(T_n(\hat{\phi} - \phi)) \\ &\Longrightarrow \mathcal{F}(\bar{\mathcal{L}}) \end{aligned}$$

with the weak convergence following from Lemma 3 (ii) and the continuous mapping theorem which applies by Property 1(iii). Note that the $1-\alpha$ quantile is positive by the fact that, $\mathcal{F}(\bar{\mathcal{L}}) \leq 0$ is equivalent to $\sup \bar{\mathcal{L}}(p) \leq 0$ using Property 1(ii) and,

$$P(\sup \bar{\mathcal{L}}(p) \leq 0) < 1/2$$

using the fact that $\bar{\mathcal{L}}$ is a separable mean zero Gaussian process. The quantile is finite for any $1/2 > \alpha > 0$ using Borell's inequality (stated as Proposition A.2.1 of van der Vaart and Wellner, 1996). Finally, uniqueness of the quantile follows from the fact that $\mathcal{F}$ is convex using Proposition 11.1 of Davydov, Lifshits and Smorodina (1998) $(0, \infty)$. For (ii) by Lemma 3(i) and using Property 1(i) and (iii),

$$\mathcal{F}(\hat{\phi}) \xrightarrow{p} \mathcal{F}(\phi) > 0$$

so that $T_n \mathcal{F}(\hat{\phi}) \xrightarrow{p} \infty$. **Q.E.D.**

Proof of Proposition 1: The LC is a Hadamard differentiable functional of the empirical distribution function following the results in Bhattacharya (2005). We must establish that the bootstrap applied to the empirical distributions yields processes with covariance properties corresponding to those for the empirical distributions of X^1 and X^2 under Assumptions 2(IS) or 2(MP). In the case of Assumption 2(IS) bootstrap empirical processes are respectively (see van der Vaart and Wellner (1996, 3.6),

$$\begin{aligned}
\mathbb{G}_{1n}(x_1) &= \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} (1\{X_{1i}^* \leq x_1\} - \hat{F}_1(x_1)) = \frac{1}{\sqrt{n_1}} \sum_{i=1}^{n_1} (M_{1i} - 1) 1\{X_{1i} \leq x_1\} \\
\mathbb{G}_{2n}(x_2) &= \frac{1}{\sqrt{n_2}} \sum_{i=1}^{n_2} (1\{X_{2i}^* \leq x_2\} - \hat{F}_1(x_2)) = \frac{1}{\sqrt{n_2}} \sum_{i=1}^{n_2} (M_{2i} - 1) 1\{X_{2i} \leq x_1\}
\end{aligned}$$

where M_{1i} and M_{2i} are multinomial random variables (with parameters n_1, n_2 and probabilities $1/n_1$ and $1/n_2$ respectively) independent of the sample and also independent of each other. It is easy to verify that, conditional on the sample these are independent mean zero processes with covariance kernels given by,

$$E(\mathbb{G}_{jn}(x_{j0})\mathbb{G}_{jn}(x_{j0})|\mathcal{X}_j) = \hat{F}_j(x_{j0}) - \hat{F}_j(x_{j0})\hat{F}_j(x_{j1}) \tag{6}$$

for $x_{j0} \leq x_{j1}$ and that this converges to the covariance kernel of the limiting process corresponding to the empirical process based on the empirical distribution.

On the other hand under Assumption 2(MP) we have bootstrap empirical processes,

$$\begin{aligned} \mathbb{G}_{1n}(x_1) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (M_i - 1) 1\{X_{1i} \leq x_1\} \\ \mathbb{G}_{2n}(x_2) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (M_i - 1) 1\{X_{2i} \leq x_1\} \end{aligned}$$

using the same multinomial variable M_i (with parameter n and probabilities $1/n$). In this case the covariance kernel of each process has the same form as (6) but the processes are correlated since,

$$E(\mathbb{G}_{1n}(x_1)\mathbb{G}_{2n}(x_2)|\mathcal{X}) = \hat{F}(x_1, x_2) - \hat{F}_1(x_1)\hat{F}_2(x_2)$$

Thus the bootstrap processes, in the limit, have a correlation structure corresponding to (2).

The result then follows using the delta method for the bootstrap (van der Vaart and Wellner (1996) 3.9.11) and the continuous mapping theorem. In particular note that the decision rule is equivalent to the rule that $T_n \mathcal{F}(\hat{\phi}) > \hat{c}_n(\alpha)$ where,

$$\hat{c}_n(\alpha) = \inf\{t : P(T_n \mathcal{F}(\hat{\phi}^*(p) - \hat{\phi}(p)) > t | \mathcal{X}) \leq \alpha)$$

where we condition on the sample(s) in computing the probability. The Hadamard differentiability of the LC and Property 1(iii) and (iv) of the map $\mathcal{F}$ we have that,

$$T_n \mathcal{F}(\hat{\phi}^*(p) - \hat{\phi}(p)) \Longrightarrow \mathcal{F}(\bar{\mathcal{L}})$$

in probability given $\mathcal{X}$ so that

$$\hat{c}_n(\alpha) \xrightarrow{p} c(\alpha) = \inf\{t : P(\mathcal{F}(\bar{\mathcal{L}}) > t) \leq \alpha)\}$$

where the latter is strictly positive, finite and unique given Lemma 4. The result then follows using Lemma 4. **Q.E.D.**

ACKNOWLEDGMENTS

The authors thank Keisuke Hirano, the associate editor and two referees for useful suggestions. The research was partially supported by the Australian Research Council and NSF grant SES-0196372. STATA and GAUSS programs for implementing the tests in this article are available on request from the authors.

REFERENCES

Abadie, A. (2002), "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97(457), 284-92.

Atkinson, A.B. (1970), "On the measurement of inequality," *Journal of Economic Theory*, 2, 244-63.

Atkinson, A.B. (1987), "On the measurement of poverty," *Econometrica*, 55, 749-764.

Andrews, D.W.K. (1997), "A Conditional Kolmogorov Test," *Econometrica*, 65(5), 1097-1128.

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

Bhattacharya, D. (2005), "Asymptotic inference from multi-stage samples," *Journal of Econometrics*, 126, 145-171.

Bhattacharya, D. (2007), "Inferences on inequality from household surveys," *Journal of Econometrics*, 137, 674-707.

Beach, C.M., and R.Davidson (1983), "Distribution-free statistical inference with Lorenz curves and income shares," *Review of Economic Studies*, 50, 723-35.

Bishop, J.A., J.P. Formby and W.J. Smith (1991a), "Lorenz Dominance and Welfare: Changes in the U.S. Distribution of Income, 1967-1986," *Review of Economics and Statistics*, 73, 134-39.

Bishop, J.A., J.P. Formby and W.J. Smith (1991b), "International Comparisons of Income Inequality: Tests for Lorenz Dominance across Nine Countries," *Economica,* 58, 461-77.

Cario, M.C. and B.L. Nelson (1997), "Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix," *Department of Industrial Engineering and Management Sciences Technical Report 50*, Northwestern University, Evanston, Illinois.

Csörgó, M. (1983), *Quantile Processes with Applications*, (Regional Conference Series on Applied Mathematics), S.I.A.M. Philadelphia.

Dardanoni, V. and A. Forcina (1999), "Inference for Lorenz curve orderings," *Econometrics Journal*, 2(1), 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.

Davies, J.B., D.A. Green and H.J. Paarsch (1998), "Economic Statistics and Social Welfare Comparisons: A Review" in *Handbook of Applied Economic Statistics,* Ullah, A. and Giles, D.E.A. eds. New York, CRC Press, chapter 1.

Davydov, Y.A., M. A. Lifshits and N.V. Smorodina (1998), *Local Properties of Distributions of Stochastic Functionals.* American Mathematical Society, Providence, Rhode Island.

Deaton, A.S. (1997), *The analysis of household surveys: microeconometric analysis for development policy.* Baltimore, Johns Hopkins University Press for The World Bank.

Goldie, C. M. (1977), "Convergence Theorems for Empirical Lorenz Curves and Their Inverses," *Advances in Applied Probability*, 9, 765-791.

Linton, O., E. Maasoumi and Y.J. Whang (2005), "Consistent Testing for Stochastic Dominance: A Subsampling Approach," *Review of Economic Studies*, 72, 735-765.

Lambert, P.J. (1993), *The Distribution and Redistribution of Income: A Mathematical Analysis*, Manchester University Press, Manchester, 2nd edition.

McFadden, D.(1989), "Testing for Stochastic Dominance" in *Studies in the economics of uncertainty: In honor of Josef Hadar,* Fomby,T-B, and Seo,T-K, eds. New York; Berlin; London and Tokyo: Springer, 113-34.

Van der Vaart, A. W. and J. A. Wellner (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer-Verlag: New York.

Wilfling, B. and W. Kramer (1993), "The Lorenz-ordering of Singh-Maddala income distributions," *Economics Letters*, 43(1), 53-57.

Wolak, F. A. (1989), "Testing Inequality Constraints in Linear Econometric Models," *Journal of Econometrics*, 41, 205-235.

FIGURE 1 -- HES 2003 and 2009 Income LC

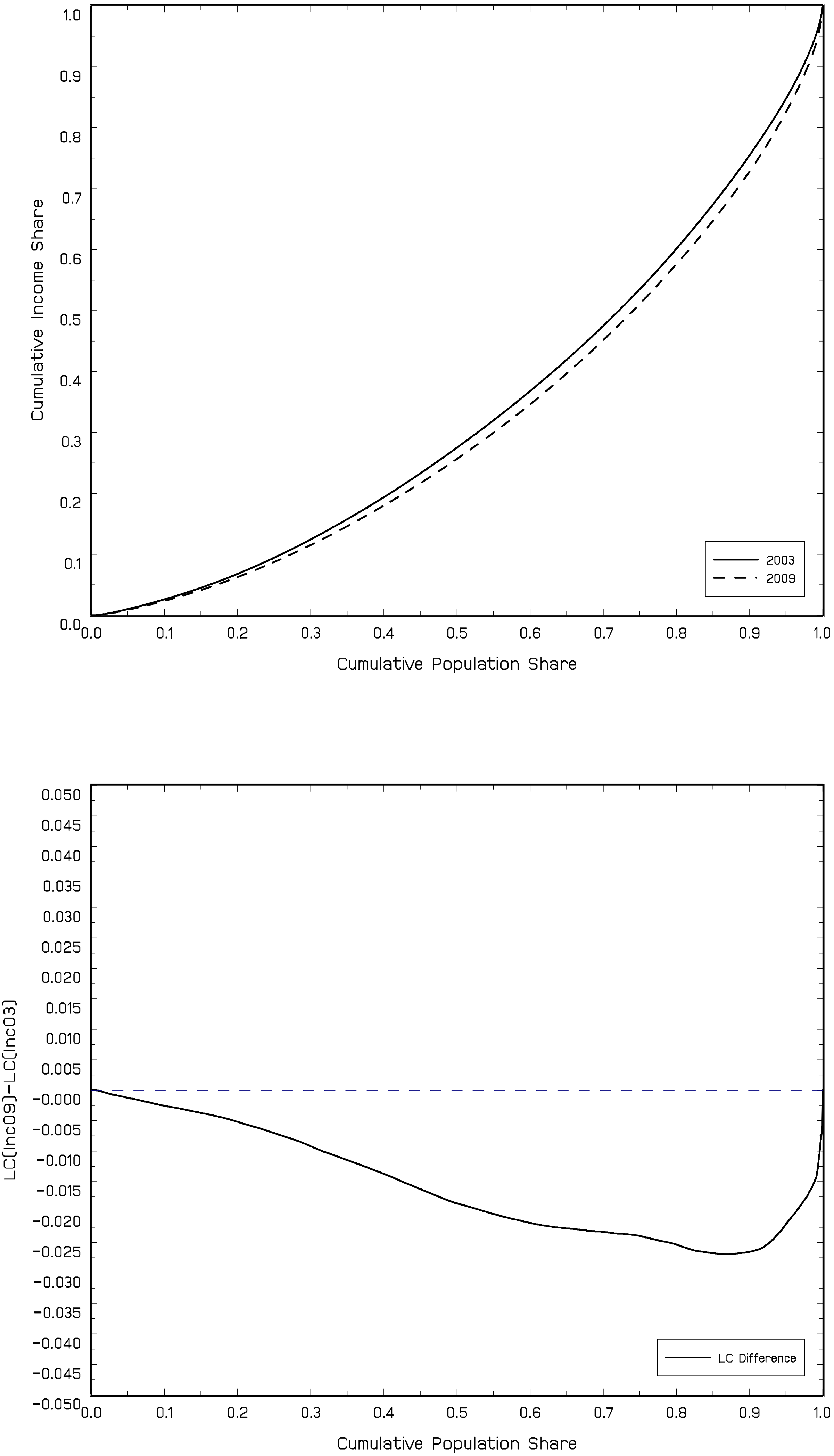

FIGURE 2 -- HES 2003 and 2009 Expenditure LC