A Consistent Nonparametric Test for Lorenz Dominance

Garry F. Barrett
Department of Economics
University of New South Wales
and
Stephen G. Donald*
Department of Economics
Boston University

April 1999
This Version November 1999.

Abstract

This paper proposes a test for Lorenz dominance. The test statistic is based on the standardized largest difference between the empirical Lorenz curves for two independent samples of welfare related variables. The asymptotic properties of the test statistic are derived. Although the limiting distribution of the test statistic is nonstandard, we propose and justify two methods of inference based on simulating the limiting distribution and the bootstrap. The performance of the two methods is examined and compared in the context of a Monte Carlo study and with a comparison of the income based Lorenz curves for the US and Canada over time and across countries.

Keywords: Lorenz dominance, test consistency, simulation, bootstrap.

J.E.L. Classification: C12, C14, C15, D63

---

[*] Corresponding Author: Stephen G. Donald, Department of Economics, 270 Bay State Rd, Boston University, Boston MA, 02215, U.S.A. Stephen Donald gratefully acknowledges financial support from the Afred P. Sloan Foundation.

1 Introduction

A commonly used tool for the empirical analysis of economic inequality is the Lorenz curve which gives the cumulative proportion of total income (or resources) by cumulative proportion of the population when the population is ordered from poorest to richest. When comparing income distributions, a fundamental concept is that of Lorenz dominance. The Lorenz curve associated with an income distribution is said to (weakly) dominate another if it is nowhere below the other. As has been shown by Atkinson (1970), Lorenz domination translates into simple facts concerning the degree of egalitarianism associated with the respective income distributions. The income distribution corresponding to the dominant Lorenz curve is more egalitarian. Moreover, it has been shown by Atkinson (1970) that Lorenz dominance translates into the (partial) ranking of income distributions based on the set of scale-free inequality indices that respect the principle of transfers.[1] 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 the test statistic for the null hypothesis that two separate Lorenz curves were equal. Note that this was a test of Lorenz equality rather than Lorenz dominance. Bishop, Formby and Smith (1991a, 1991b) proposed a test of Lorenz dominance based on the covariance matrix derived by Beach and Davidson (1983). Their approach involved pair-wise multiple comparisons of the empirical Lorenz ordinates for two distributions. Lorenz dominance is inferred when there is at least one positive significant difference and no negative significant difference between the two subsets of Lorenz ordinates.

The aim of the current paper is to develop a flexible yet consistent test for Lorenz dominance. The flexibility of our approach derives from the fact that we use empirical

---

[1] More formally, the set of inequality indices satisfying the principle of transfers is the class of indices that are Schur-Concave (Dasgupta, Sen and Starrett, 1973). The set of scale-free Schur-concave inequality indices includes the Atkinson indices, the Gini coefficient and its extensions, and the coefficient of variation: see Shorrocks and Foster (1987) for a detailed treatment.

Lorenz curves, which can be obtained quite simply from the empirical distributions associated with the two income samples (which are assumed to be independently drawn from two possibly distinct populations). The empirical Lorenz curve is a fully nonparametric estimator of the true underlying Lorenz curve associated with the distribution and was first analysed by Goldie (1977). Therefore, in contrast to the recent empirical work of Basmann, Hayes and Slottje (1993) and Ryu and Slottje (1996), we do not rely on (potentially inconsistent) functional form approximations to the Lorenz curve.

The second feature of our test is that it is consistent in the sense that it can detect any violation of the null hypothesis of weak Lorenz dominance. This is achieved by comparing the empirical Lorenz curves at **all** quantiles. The tests of Lorenz dominance proposed by Bishop, Formby and Smith (1991a, 1991b) are based on estimates of a small number of Lorenz curve ordinates (typically deciles).[2] By restricting attention to a subset of Lorenz curve ordinates, these tests do not use all the information available in a given sample and are therefore potentially inconsistent. The test presented in this paper utilises all the sample information and hence provides a consistent test of Lorenz dominance – our test can be considered as analogous to tests of stochastic dominance proposed in McFadden (1989) and further elaborated and extended by Barrett and Donald (1999b). The fact that the empirical Lorenz curves are piecewise linear and continuous implies that we can compare the Lorenz curves at all points with a finite number of calculations. The main difficulty with our test is that the limiting distribution of the proposed test statistic is nonstandard and will generally depend on the underlying Lorenz curves. We propose and compare two solutions to this difficulty. The first is to use a p-value simulation method that is similar to that used in Hansen (1996). The second is based on the better known bootstrap methods. We show that both methods can be justified theoretically in the sense of providing asymptotically valid inferences concerning the null hypotheses. Additionally, as shown in a Monte Carlo experiment, the tests work well with moderate sample sizes. We also address the issue of computational demands that the two methods impose.[3]

---

[2] The tests of Lorenz dominance implemented in Bishop, Formby and Smith (1991a, 1991b) was based on nine ordinates (corresponding to population deciles) of the Lorenz curve. Althought it is feasible to expand the set of Lorenz curve ordinates used for the test, for their method to be implemented it is necessary to group the samples into a subset of quantiles.

The remainder of the paper is organized as follows. In Section 2 we state our testing problem, review some results pertaining to the properties of empirical Lorenz curves, propose a test statistic 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 we consider the use of a simulation method due to Hansen (1996) for computing p-values for the tests and give a theoretical justification for the method in the present context. In Section 4 the non-parametric bootstrap approach to conducting inference is presented and theoretically justified. Section 5 provides a brief Monte Carlo study that examines how well the asymptotic arguments work in small samples. Finally, in Section 6 we implement the tests by comparing the Lorenz curves for the distribution of family income in the US and Canada over time and across countries. Our empirical results provide strong evidence that inequality has significantly increased in the US over the period 1978-1998. The evidence for Canada is less clearcut but does suggest that the 1990's have seen an unambiguous increase in inequality. When comparing the Lorenz curves across the two countries there is clear evidence to suggest that there is much less inequality in Canada and that, if anything, the degree of dominance of the Canadian Lorenz curve has grown over time.

---

[3] The authors have written Gauss procedures that allow one to compute the test statistics and to obtain p-values. These are available on request.

2 Asymptotic Properties of Lorenz Dominance Test Statistics

2.1 Hypothesis Formulation

We are interested in comparing the Lorenz curves associated with the distributions of income (or some other measure of welfare) in two different populations. We represent the populations by their respective cumulative distribution functions (c.d.f.'s) F and G. We make the following assumptions regarding these population c.d.f.'s.

Assumption 1 *Assume that* F *and* G are *twice continuously differentiable with associated probability density functions given by* $f(z) = F'(z)$, and $g(z) = G'(z)$ *where* $0 < \inf\ f(z) < \sup f(z) < \infty$ *and* $0 < \inf\ g(z) < \sup g(z) < \infty$.

The Lorenz curves (at ordinate value $p \in [0, 1]$) for the respective populations are defined by,

$$L_F(p) = \frac{\int_{-\infty}^{Q_F(p)} zf(z)dz}{\int_{-\infty}^{\infty} zf(z)dz}$$

and,

$$L_G(p) = \frac{\int_{-\infty}^{Q_G(p)} zg(z)dz}{\int_{-\infty}^{\infty} zg(z)dz}$$

where $Q_F(p) = F^{-1}(p)$ and $Q_G(p) = G^{-1}(p)$ are the respective quantile functions. Convenient alternative representations are,

$$L_F(p) = \frac{\int_0^p Q_F(t)dt}{\mu_F}$$

and,

$$L_G(p) = \frac{\int_0^p Q_G(t)dt}{\mu_G}$$

where μ_F and μ_G are the means for F and G respectively.

The hypotheses that we are interested in testing are:

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

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

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

then the hypotheses H_0^G and H_1^F together imply the strong dominance of L_F over L_G so that in principle one could use the tests to determine whether or not there is strong dominance. Note also that the hypotheses H_0^F and H_0^G together imply that the Lorenz curves are identical.

2.2 Properties of Lorenz Curve Estimators

Our aim is to make inferences regarding the hypotheses described in the previous section based on samples drawn from the respective populations described by F and G. Our assumption regarding the sampling procedure is the following.

Assumption 2: *Assume the following,*

(i) $\{X_i\}_{i=1}^N$ *is a random sample from* F*,*

(ii) $\{Y_i\}_{i=1}^M$ *is a random sample (independent of the sample described in (i)) from* G,

(iii) the sampling scheme is such that as $N, M \rightarrow \infty$,

$$\frac{N}{N+M} \rightarrow \lambda$$

where $0 < \lambda < 1$.

The empirical distributions are respectively,

$$\begin{aligned}
\hat{F}_N(z) &= \frac{1}{N}\sum_{i=1}^{N} 1(X_i \leq z) \\
\hat{G}_M(z) &= \frac{1}{M}\sum_{i=1}^{M} 1(Y_i \leq z)
\end{aligned}$$

We can define the respective quantile functions as,

$$\begin{aligned}
\hat{Q}_F(p) &= \inf\{z : \hat{F}_N(z) \geq p\} \\
\hat{Q}_G(p) &= \inf\{z : \hat{F}_N(z) \geq p\}
\end{aligned}$$

Then the empirical Lorenz curve (hereafter LC), at ordinate value p, can be defined in terms of the quantile function by,

$$\begin{aligned}
\hat{L}_F(p) &= \frac{\int_0^p \hat{Q}_F(t)dt}{\hat{\mu}_F} \\
\hat{L}_G(p) &= \frac{\int_0^p \hat{Q}_G(t)dt}{\hat{\mu}_G}
\end{aligned}$$

where $\hat{\mu}_F = \bar{X}_N$ and $\hat{\mu}_G = \bar{Y}_M$ are the respective sample means. Note that the numerators are the respective generalized Lorenz curves.

As is well known appropriately standardized empirical distribution functions (considered as elements of the function space[4] $D[a, b]$) satisfy the following weak convergence results:

$$\begin{aligned}
\sqrt{N}(\hat{F}_N - F) &\Rightarrow \mathcal{B}_F \circ F \\
\sqrt{M}(\hat{G}_M - G) &\Rightarrow \mathcal{B}_G \circ G
\end{aligned}$$

These imply that for a given value of $z \in [a, b]$,

---

[4] The space $D[a, b]$ is the space of cadlag functions on $[a, b]$.

$$\begin{aligned}\sqrt{N}(\hat{F}_N(z) - F(z)) &\Rightarrow \mathcal{B}_F(F(z)) \sim N(0, F(z)(1 - F(z)))\\ \sqrt{M}(\hat{G}_M(z) - G(z)) &\Rightarrow \mathcal{B}_G(G(z)) \sim N(0, G(z)(1 - G(z)))\end{aligned}$$

Our first result provides a characterization of the limiting properties of the empirical LC's. Note that since the LC is a scaled version of the integral of the quantile function, then the standardized empirical LC's can be considered as members of the function space $C[0,1]$ since they are piecewise linear and continuous. Before providing the result we define some notation. For an arbitrary distribution function H (say) define, $\mathcal{B}_H \circ H$ as the Brownian Bridge process composed of H. That is for a particular value of z $(\mathcal{B}_H \circ H)(z) \equiv \mathcal{B}_H(H(z))$ where $\mathcal{B}_H$ is the usual Brownian bridge for the population H. Also define the Gaussian stochastic process, $\mathcal{G}_H$ on $[0,1]$ to be such that for $p \in [0,1]$,

$$\mathcal{G}_H(p) = -\int_0^p \frac{\mathcal{B}_H(t)}{h(Q_H(t))} dt$$

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

$$\mathcal{L}_H(p) = \frac{\mathcal{G}_H(p)}{\mu_H} - \frac{L_H(p)}{\mu_H}\mathcal{G}_H(1)$$

A process such as $\mathcal{L}_H$ will be referred to as a Lorenz process for the distribution H. The following result shows that the normalized empirical Lorenz Curve processes have limits that are of the form just given.

Lemma 1: *Given our assumptions on* F and G,

(i) for the F *population we have that,*

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

and in the space $C[0,1]$,

$$\sqrt{N}(\hat{L}_F - L_F) \Rightarrow \frac{\mathcal{G}_F}{\mu_F} - \frac{L_F}{\mu_F}\mathcal{G}_F(1) \equiv \mathcal{L}_F.$$

(ii) *for the* G *population we have that,*

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

and in the space $C[0,1]$*,*

$$\sqrt{M}(\hat{L}_G - L_G) \Rightarrow \frac{\mathcal{G}_G}{\mu_G} - \frac{L_G}{\mu_G}\mathcal{G}_G(1) \equiv \mathcal{L}_G.$$

This result is simply a restatement of Lemma 3 of Barrett and Donald (1999a). The results contained in Lemma 1 are not new and date back to at least Goldie (1977), who presented a full weak convergence result for the LC process under very weak conditions – our assumptions are slightly stronger than required by Goldie (1977). Other results concerning the empirical LC process include Gail and Gastwirth (1978) who derived an asymptotic distribution result for a single ordinate of the normalized LC and Csörgó (1983) who proved that the empirical LC process could be strongly approximated by a sequence of Gaussian processes which are equal in distribution to that given in the result. Beach and Davidson (1983) also presented this result for a vector of ordinates of the Lorenz curve and showed how one could obtain estimates of its variance covariance matrix without imposing distributional assumptions. An important difference in this paper is that we wish to compare Lorenz curves at all points and not at an arbitrary selected (and fixed) set of ordinate values.

2.3 Dominance Test Statistic and Asymptotic Properties

The test statistic that we propose for testing the null hypothesis that distribution F weakly Lorenz dominates distribution G is:

$$\hat{S}_G = \left(\frac{NM}{N+M}\right)^{1/2} \sup_p(\hat{L}_G(p) - \hat{L}_F(p))$$

and the natural decision rule for conducting the test has the form,

$$\text{“reject } H_0^G \text{ if } \hat{S}_G > c_l\text{”}$$

where c_l is some critical value that will be discussed later. The following result characterizes the properties of this test.

Proposition 1: *Given Assumptions 1 and 2 and that c_l is a finite constant, then*

(*i*) *if H_0^G is true,*

$$\lim_{N,M\to\infty} P(\text{reject } H_0^G) \leq P(\sup_p \mathcal{L}_G(p) > c_l) = \alpha(c_l),$$

with equality holding when $L_F(p) = L_G(p)$ for all $p \in [0,1]$,

(ii) *if H_0^G is false,*

$$\lim_{N,M\to\infty} P(\text{reject } H_0^G) = 1.$$

The stumbling block to implementing the test as it stands is that the critical value will generally depend on the LC for G. More particularly the distribution of $\mathcal{L}_G(p)$ depends on L_G so that a critical value corresponding to one's desired significance level cannot generally be found without knowledge of L_G. Note also that the test is consistent in the sense that whenever the null is false the test rejects with probability that approaches one as long as we use a finite critical value. We explore the actual decision rule further in the next section. A similar result that is much simpler to prove is the following.

Corollary 1: *Given Assumptions 1 and 2 and that c_l is a finite constant, then*

(*i*) *if H_0^F is true,*

$$\lim_{N,M\to\infty} P(\text{reject } H_0^F) \leq P(\sup_p \sqrt{\lambda}\mathcal{L}_G(p) - \sqrt{1-\lambda}\mathcal{L}_F(p) > c_l) = \alpha(c_l),$$

with equality holding when $L_F(p) = L_G(p)$ for all $p \in [0,1]$,

(ii) *if H_0^F is false,*

$$\lim_{N,M\to\infty} P(\text{reject } H_0^F) = 1.$$

The difference between the results in (i) of Proposition 1 and Corollary 1 is in the nature of the random variable that appears in the limiting probability of rejection. Although Corollary 1 is simpler to prove its characterization is more complicated, involving the limiting Lorenz processes for both distributions. On the other hand the proof of

Proposition 1 is more involved but results in a characterization that involves only the Lorenz process for the G distribution. An implication of the simpler characterization in Proposition 1 is that it will be computationally much easier to implement. In particular, our approach to inference is computational and involves two methods of simulating the process that appears in the limiting distribution. Because of the cost of the calculations involved the result in Proposition 1 is also of practical importance because it means that one can use roughly half as many calculations to conduct inference – one will only need to simulate the Lorenz process for one population rather than for two as would be required by the characterization in Corollary 1.

3 Simulating P-Values Using the Limiting Distribution

As noted in the previous section the difficulty in implementing the tests in practice arises because the distribution of $\mathcal{L}_G(p)$ will generally depend on $L_G(p)$. In this section we consider the use of simulation or Monte Carlo methods for conducting inference with the test, based on the methods of Hansen (1996), exploiting the fact that we can estimate $L_G(p)$ consistently and that we have a characterization of the process $\mathcal{L}_G(p)$. First we provide a definition of a stochastic process that is essentially linearized version of the empirical Lorenz Curve process corresponding to L_G that has the same limiting behaviour as $\sqrt{M}(\hat{L}_G - L_G)$. The random components of this process are,

$$\begin{aligned}
\tilde{Z}_G &= \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(Y_i - \mu_G) \\
\tilde{B}_G(p) &= \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(1(Y_i \leq Q_G(p)) - p) \\
\tilde{C}_G(p) &= \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(Y_i 1(Y_i \leq Q_G(p)) - E(Y_i 1(Y_i \leq Q_G(p)))
\end{aligned}$$

and we define the linearized version of $\sqrt{M}(\hat{L}_G - L_G)$ as,

$$\tilde{L}_{G,M}(p) = -\frac{1}{\mu_G}(Q_G(p)\tilde{B}_G(p) - \tilde{C}_G(p)) - \frac{L_G(p)}{\mu_G}\tilde{Z}_G$$

The function $\tilde{L}_{G,M}$ can be treated as belonging to $l^{\infty}([0,1])$ – the space of bounded functions on the unit interval, even though the function is actually an element of the much smaller space $D[0,1]$. The following result shows that the stochastic process converges weakly (in $l^{\infty}([0,1])$) to the ($C[0,1]$ function) $\mathcal{L}_G(p)$.

Lemma 2: *Given Assumptions 1 and 2, in the space* $l^{\infty}([0,1])$,

$$\tilde{L}_{G,M}(p) \Rightarrow \mathcal{L}_G(p).$$

The importance of the result lies in the fact that we can easily simulate copies of the three (non-degenerate) random components $\tilde{B}_G, \tilde{C}_G$ and $\tilde{Z}_G$ by replacing unknowns with consistent estimates and by exploiting the multiplier Central Limit Theory exposited in Van der Vaart and Wellner (1996). Moreover the processes can be constructed in such a way as to be independent of the original random components but with identical limiting distributions. To do this let $\{U_i\}_{i=1}^{N}$ denote a sequence of i.i.d. $N(0,1)$ random variables that are independent of the samples. For each value of $p \in [0,1]$ let,

$$\begin{aligned}
\tilde{Z}_G^* &= \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (Y_i - \hat{\mu}_G) U_i \\
\tilde{B}_G^*(p) &= \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (1(Y_i \leq \hat{Q}_G(p)) - p) U_i \\
\tilde{C}_G^*(p) &= \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (Y_i 1(Y_i \leq \hat{Q}_G(p)) - \hat{\mu}_G(p)) U_i
\end{aligned}$$

and finally let,

$$\tilde{L}_G^*(p) = -\frac{1}{\hat{\mu}_G} (\hat{Q}_G(p) \tilde{B}_G^*(p) - \tilde{C}_G^*(p)) - \frac{\hat{L}_G(p)}{\hat{\mu}_G} \tilde{Z}_G^* \tag{1}$$

The following result is fundamental to proving that the approach is valid.

Lemma 3: *Given Assumptions 1 and 2 random process* $\tilde{L}_G^*$ *converges (weakly in probability) to* $\mathcal{L}_G'$ *where* $\mathcal{L}_G'$ *has the same distribution as* $\mathcal{L}_G$ *but is independent of* $\mathcal{L}_G$.

Given the simulated process and its independence from the processes corresponding to the samples we can simulate p-values for the test as,

$$\hat{p}_G = P_U(\sup_p \tilde{L}^*_G(p) > \hat{S}_G) \tag{2}$$

where $P_U(.)$ is the probability function associated with the normal random variables U_i and is conditional on the realized sample(s). Note that these p-values depend on the sample sizes N and M although we have suppressed the dependence for notational convenience. The following result provides a justification for the p-value approach.

Proposition 2: *Given Assumptions 1, 2 and assuming that $\alpha < 1/2$, a test for Lorenz Dominance based on the rule,*

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

satisfies the following,

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

This result is similar to part of the proof of Theorem 2 of Hansen (1996). The main difference is that in our case we must deal with the fact that we have a one sided composite null. The result implies that a test based on the decision rule “reject H_0^G if $\hat{p}_G < \alpha$” will reject a true null hypothesis with probability that is (asymptotically) no larger than α. The probability will be (asymptotically) equal to α when in fact $L_F = L_G$ (in which case the inequalities in the statement of Proposition 1 hold with equality).

In order to compute the p-values in practice we must compute the probability in (2). As in Hansen (1996) we use Monte-Carlo methods to approximate the probability and use a grid to approximate the suprema. Since these are under the control of the statistician one can make the approximations as accurate as one wants given time and computing resources. More specifically let $\{U_i^r\}_{i=1}^N$ denote the rth sample of U_i where we will let $r = 1, ..., R$ where R will denote the number of replications that will be used in the Monte Carlo simulation. Select a grid of values on $[0, 1]$ such as $0 = p_0 < p_1 < ... < p_K = 1$, where K will denote the number of subintervals. Then at the p_j values we let,

$$\begin{aligned}
\tilde{Z}_G^r &= \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (Y_i - \hat{\mu}_G) U_i^r \\
\tilde{B}_G^r(p_j) &= \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (1(Y_i \leq \hat{Q}_G(p_j)) - p) U_i^r \\
\tilde{C}_G^r(p_j) &= \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (Y_i 1(Y_i \leq \hat{Q}_G(p_j)) - \hat{\mu}_G(p_j))) U_i^r
\end{aligned}$$

and we approximate the supremum for the rth sample by,

$$\tilde{S}_r = \max_j \tilde{L}_G^r(p_j)$$

where,

$$\tilde{L}_G^r(p) = -\frac{1}{\hat{\mu}_G} (\hat{Q}_G(p) \tilde{B}_G^r(p) - \tilde{C}_G^r(p)) - \frac{\hat{L}_G(p)}{\hat{\mu}_G} \tilde{Z}_G^r$$

Then the p-values can be approximated by,

$$\hat{p}_G \simeq \frac{1}{R} \sum_{r=1}^{R} 1(\tilde{S}_r > \hat{S}_G).$$

As indicated by Hansen (1996), an appeal to the Central Limit Theorem suggests that the error in approximating $\hat{p}_j$ should have a standard error that is approximately no larger than $(4R)^{-1/2}$ so that if $R = 1000$ (or say 10,000) for instance the standard error in this approximation is roughly 0.015 (or 0.005 when $R = 10000$) and much smaller in cases where $\hat{p}_j$ is close to zero.

4 Bootstrap Based Inference

A natural alternative to the p-value simulation method is to conduct inferences using a form of the bootstrap. A possible advantage of this is that, although existence of a limiting distribution (for the test statistic) is generally needed, one does not necessarily need to be able to characterize it in the way that we were able to in the previous section.

Therefore the bootstrap more be applicable in more complicated situations. As in the previous section we rely on the result in Proposition 1 which helps greatly in simplifying the number of calculations needed to bootstrap the statistic.

Because of the cost in implementing the bootstrap we allow for the possibility of using only a subset of the observations to estimate the distribution of the object, $\sup_p \mathcal{L}_G(p)$. The sample of observations from the G population was assumed to be a random sample, so without loss of generality we can take the first $k \leq M$ observations as being a random selection from this sample.[5] Let the sample from which we will resample be given by $\mathcal{Y} = \{Y_1, ..., Y_k\}$. For this particular sample we can define the following,

$$\begin{aligned} \hat{G}_k(y) &= \frac{1}{k}\sum_{i=1}^{k} 1(Y_i \leq y) \\ \hat{Q}_{G,k}(p) &= \inf\{z : \hat{G}_k(z) \geq p\} \\ \hat{L}_{G,k}(p) &= \frac{\int_0^p \hat{Q}_{G,k}(t)dt}{\hat{\mu}_{G,k}} \end{aligned}$$

which are the empirical c.d.f., quantile function and Lorenz curve respectively for the sample $\mathcal{Y}$. Note that $\hat{\mu}_{G,k}$ is the mean of the $\mathcal{Y}$ sample. In performing bootstrap computations we fix these objects, or in other words operate conditionally on $\mathcal{Y}$. Let a random sample (drawn with replacement) of size k from $\mathcal{Y}$ be given by $Y_1^*, ..., Y_k^*$. For each random sample we can define,

$$\begin{aligned} \hat{G}_k^*(y) &= \frac{1}{k}\sum_{i=1}^{k} 1(Y_i^* \leq y) \\ \hat{Q}_{G,k}^*(p) &= \inf\{z : \hat{G}_k^*(z) \geq p\} \\ \hat{L}_{G,k}^*(p) &= \frac{\int_0^p \hat{Q}_{G,k}^*(t)dt}{\hat{\mu}_{G,k}^*} \end{aligned}$$

where $\hat{\mu}_{G,k}^*$ is the mean of the randomly drawn sample. These objects are all random variables conditionally on the original sample $\mathcal{Y}$. Through Monte Carlo simulation (by drawing many random samples from $\mathcal{Y}$) we can obtain a good approximation to the distribution of,

---

[5] Of course this assumes that no sorting has been done. If the data has been sorted then for the arguments of this section to be valid we would need to randomly select (without replacement) k of the original data points.

$$\sup_p \sqrt{k}(\hat{L}^*_{G,k}(p) - \hat{L}_{G,k}(p))$$

and can use quantiles as critical values in conducting tests of Lorenz dominance. Equivalently we can conduct inferences by rejecting the null when,

$$\hat{p}_{G,k} = P(\sup_p \sqrt{k}(\hat{L}^*_{G,k}(p) - \hat{L}_{G,k}(p)) > \hat{S}_G|\mathcal{Y}) < \alpha$$

for some prespecified value of α. The p-value can be approximated by Monte Carlo as,

$$\hat{p}_{G,k} \simeq \frac{1}{R}\sum_{r=1}^{R} 1(\sup_p \sqrt{k}(\hat{L}^*_{G,k,r}(p) - \hat{L}_{G,k}(p)) > \hat{S}_F)$$

where $\hat{L}^*_{G,k,r}(p)$ is the rth realization of a random sample from $\mathcal{Y}$. The following result provides a justification for this approach.

Proposition 3: *Let Assumptions 1, 2 hold and assume that $\alpha < 1/2$. Then if $k \to \infty$ as $N, M \to \infty$ then a test for Lorenz Dominance based on the rule,*

"reject H_0^G if $\hat{p}_{G,k} < \alpha$"

satisfies the following,

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

5 Monte Carlo Results

In this section we consider a small scale Monte Carlo experiment in which we gauge the extent to which the preceding asymptotic arguments hold in small samples. Our experimental specifications for the income distributions are in the log-normal family because they are easy to simulate and also because 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 and Y_i as (independent) log-normal random variables using the equations,

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

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 $2.8/1.8 = 1.55$ is similar to that found in actual income data. In Case 1 the Lorenz curves for the two populations are identical and our interest is in the size properties of the testing procedure.

The second case, Case 2, $\mu_1 = 0.85$, and $\sigma_1 = 0.6$ while $\mu_2 = 0.7$ and $\sigma_2 = 0.5$. In this case the Lorenz curve for Y dominates the Lorenz curve for X – indeed the Lorenz curve for Y lies above that for X everywhere except at the endpoints of the interval $[0,1]$. In this case we should expect that we do not reject the hypothesis H_0^G but we should reject H_0^F.[6] We consider tests of both of these hypotheses. Note also that in this case we should expect that the test will reject H_0^G less often than the nominal size of the test because of the inequality in Proposition 1.

In the final case, Case 3, we generate X as before but now generate Y as a mixture of log-normal random variables. In particular,

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

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 Lorenz curves. Neither Lorenz curve dominates the other and we should expect both H_0^G and H_0^F to be rejected.

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

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

---

[6] Recall the distribution function for X_i is denoted F and the distribution function for Y_i is denoted G.

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 $N = M = 50$ and $N = M = 500$. The number of replications was set to 1000 in the case with $N = M = 50$ and 500 in the case where the sample sizes were $N = M = 500$.

In using the p-value simulation method we set the number of gridpoints at $K = 100$ and we used $R = 100$ replications in computing the p-value for each statistic. For the bootstrap method we used bootstrap sample sizes equal to $k = 50$ in the $N = M = 50$ case and $k = 100, 200$ and 500 in the case where $N = M = 500$. As in the p-value simulation method we used $R = 100$ replications to approximate the p-value in each Monte Carlo simulation. The results for the Monte Carlo simulations are reported in Tables I-III. The tables all report the proportion of times that the respective null hypothesis was rejected for three different nominal significance levels α.

		H_0^G			H_0^F		
		Nominal Size			Nominal Size		
Sample Size	Method	0.1	0.05	0.01	0.1	0.05	0.01
50	S	0.149	0.097	0.051	0.162	0.116	0.041
	B50	0.158	0.099	0.048	0.130	0.089	0.038
500	S	0.118	0.072	0.016	0.120	0.062	0.018
	B100	0.130	0.084	0.032	0.124	0.062	0.026
	B200	0.134	0.076	0.028	0.128	0.068	0.016
	B500	0.132	0.078	0.016	0.146	0.072	0.020

Table I: Monte Carlo Results:Case 1

		H_0^G			H_0^F		
		Nominal Size			Nominal Size		
Sample Size	Method	0.1	0.05	0.01	0.1	0.05	0.01
50	S	0.471	0.364	0187	0.028	0.014	0.006
	B50	0.472	0.374	0.222	0.021	0.013	0.002
500	S	0.990	0.972	0.836	0.000	0.000	0.000
	B100	0.982	0.956	0.832	0.000	0.000	0.000
	B200	0.984	0.962	0.852	0.000	0.000	0.000
	B500	0.988	0.968	0.842	0.000	0.000	0.000

Table II: Monte Carlo Results:Case 2

		H_0^G			H_0^F		
		Nominal Size			Nominal Size		
Sample Size	Method	0.1	0.05	0.01	0.1	0.05	0.01
50	S	0.074	0.030	0.004	0.489	0.384	0.212
	B50	0.062	0.017	0.004	0.480	0.367	0.198
500	S	0.872	0.602	0.118	0.988	0.984	0.896
	B100	0.798	0.430	0.060	0.990	0.984	0.904
	B200	0.834	0.474	0.064	0.992	0.984	0.922
	B500	0.952	0.710	0.170	0.996	0.980	0.884

Table III: Monte Carlo Results:Case 3

Several features of the tests are of note. The tests tend to over-reject in small samples although the degree of over-rejection is not too severe. As would be expected, there is some improvement as one moves to the larger sample size – this is most evident for the p-value simulation method. For the bootstrap method the improvement is not as obvious and it is suspected that because of the large number of pseudo random number used, some cycling may have occurred. This type of problem would be most likely to occur in the case where the bootstrap sample size was 500. In terms of size the p-value simulation test appears to be slightly better than the bootstrap, although one should bear in mind the potential cycling in the case of the bootstrap. It should be noted that the sample sizes considered are rather small compared to many empirical applications so that the fact that the actual sizes of the test are close to the nominal size is encouraging.

In terms of power the tests appear to be quite similar. In case 2 the tests detect the fact that the Lorenz curve for Y (which has distribution G) dominates that for X. As a result the hypothesis H_0^G is rejected with high probability. It is interesting to note that the hypothesis H_0^F is never rejected (in the large sample) 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 curve is dominant and so the tests reject both nulls with high probability in the large sample case. In small samples the tests have a hard time detecting the failure of the hypothesis H_0^G because the violation of this hypothesis is quite small.

In terms of the bootstrap method one is interested in the impact of the choice of bootstrap sample size on the performance of the tests. The results in Tables I-III indicate that the choice of k does not seem to matter so that the impact of using only a subsample is minimal – the computational savings from using only a subsample come at little or no cost. This is a nice feature, especially given the large sizes of samples typically used in income distribution studies.

6 Empirical Example

In this section we implement the methods for the testing for Lorenz dominance relations by examining the distribution of family income in the United States and Canada. The data for Canada are from the Family Expenditure Survey (FAMEX) for years 1974, 1978, 1982, 1990 and 1996. We consider annual total family income, both before and after taxes. The data for the United States are drawn from the March Demographic files of the Current Population Survey (CPS) which record total family income before taxes for the years 1978, 1988 and 1998.[7] Sample sizes and summary statistics are reported in Tables II and III for Canada and the US, respectively.[8] The reported dollar amounts are in 1998 US dollars. For Canada/US, the GDP/GNP implicit price deflator for consumption is used as the inflation index, and the Canadian dollar values are converted to US dollars using the 1990 OECD estimate of the purchasing power parity for final private consumption goods which was 1.29 (OECD, 1993).[9]

Year	before/after	Sample Size	Mean	Median	Std. Dev
1974	before	6408	33074	28894	22900
1978	before	8526	35535	32423	22098
1982	before	9999	37881	33914	24526
1990	before	4268	40326	35554	26926
1996	before	9739	38188	32714	26841
1974	after	6408	27945	25110	17471
1978	after	8526	29840	27813	16873
1982	after	9999	31696	28937	18573
1990	after	4268	31979	28912	19103
1996	after	9739	30708	27209	19393

Table IV: Canadian Summary Statistics

---

[7] For each CPS a random sample of approximately 8000 family records were selected for the analysis.

[8] The sample frame of the FAMEX varied across the surveys. To ensure the same population was sampled in each year we restricted the analysis to households consisting of a single 'economic family' (individuals related by blood, marriage or adoption). The Canadian definition of an economic family is equivalent to the family concept used in the CPS. Additionally, it is noted that observations with zero or negative reported incomes were dropped from the analysis: these accounted for a very small number of observations.

[9] The conversion of nominal values to constant US dollars is simply for descriptive purposes as income shares and hence the Lorenz curve are unaffected by a common scaling of all incomes.

Year	before/after	Sample Size	Mean	Median	Std. Dev
1978	before	7896	32009	26586	24616
1988	before	7852	37064	28870	31175
1998	before	7879	44006	31000	48469

Table V: US Summary Statistics

The first series of comparisons examine changes in the Canadian distribution of family income over time. Tables IV and V give results based on the p-value simulation method and the bootstrap method respectively. In the case of the p-value simulation method we chose a grid of 100 points on $[0, 1]$ and performed 1000 replications to compute the p-value. For the bootstrap method we chose $k = 1000$ and conducted inferences using a random subsample drawn from the appropriate sample in order to perform the bootstrap simulations. We also used 1000 draws to compute the approximate p-value. In terms of the results the two methods gave essentially the same results, suggesting that either method could be used in practice.[10] Moreover, the bootstrap method was tried using several different randomly selected subsamples with identical results.[11]

Because of the similarity of the two sets of results we focus on Table IV. The first two columns of results in Tables IV are for the distribution of after-tax income, and the last two columns relate to the before-tax income distribution. The test statistics reject the hypothesis that the 1974 after-tax income distribution weakly Lorenz dominated the 1982 distribution, whereas the converse is not rejected. The evidence therefore shows that between 1974 and 1982 the Canadian distribution of after-tax family income became unambiguously more equal. Comparing the distributions for 1982 and 1990 indicates that weak dominance is not rejected for either distribution, which together implies that the Lorenz curves for these two years are not significantly different. This result was replicated using the bootstrap method. Considering more recent trends, the tests reveal that the 1990 after-tax income distribution strongly Lorenz dominated the 1996 distribution. Overall, the tests of Lorenz dominance reveal a clear trend toward equality in the Canadian distribution of after-tax family income during the later part of the 1970's, there was a period of relative stability during the course of the 1980's while the 1990's have witnessed a clear increase in family after-tax income inequality.

---

[10] Given that the simulation and bootstrap methods gave very similar results, if one method is to be adopted then an important practical consideration may be the demand on computer cpu time. The empirical application presented in this section was performed on a 300 MHz Pentium II machine with 128MB RAM running Gauss for Windows. Taking the comparison of the Canadian before-tax income distributions for 1982-1990 as an example (the 1982 distribution has the largest sample size with 9999 observations and hence this is the most time consuming run), the test statistics took under 3 seconds to compute. In constructing the p-value, the simulation method took 25:34 (minutes:seconds) using all the observations and a grid of 100 values. The bootstrap method using a subsample of 1000 (2000) observations took 4:20 (11:38). The boostrap based on a subsample of 5000 observations took 35:01, while the bootstrap using the full sample took 1:15:38.

[11] The bootstrap simulations were also performed for $k = 2000$ as well as with the full sample. The bootstrapped p-values were numerically very similar, and the inferences identical, to those reported for $k = 1000$ in Table VII .

The Lorenz dominance tests were replicated using the Canadian before-tax family income. The inferences drawn from these tests generally coincide to those discussed above for the after-tax distribution. The exception was that the 1982 before-tax income distribution was found to dominate the 1990 before tax distribution at conventional levels of significance using both methods.[12] The difference in the test results for the after-tax and before-tax comparisons for 1982-1990 suggest that the Canadian tax system effectively redistributed income toward the bottom of the distribution over this period, counteracting the trend toward greater inequality evident in the before-tax income distribution.

The next series of results relate to changes in the United States distribution of before-tax family income. The p-values (for both methods) clearly show that there was a trend toward greater inequality in family income between 1978 and 1988, and that this trend continued throughout the decade from 1988 to 1998.

The final set of tests relate to a comparison of the Canadian and US family income distribution at different points in time. In 1978 the Canadian distribution of family income strongly Lorenz dominated the US distribution. Given the findings reported above for trends in inequality within each country over the 1980's, it is unsurprising that the Canadian family income distribution in 1990 strongly Lorenz dominated the US distribution for 1988. Although both countries experienced growing income inequality over the 1990s, the 1996 Canadian distribution clearly remained more equal than the 1998 US family income distribution. Indeed, although not reported in the tables, the magnitude of the test statistics indicate that the extent of the dominance of the Canadian distribution over the US distribution has grown over time.

---

[12] With p-values of 0.072 and 0.60 for the simulation and bootstrap methods, respectively.

		after tax		before tax	
G	F	H_0^G	H_0^F	H_0^G	H_0^F
C1974	C1982	0.000	0.900	0.000	0.904
C1982	C1990	0.716	0.149	0.926	0.072
C1990	C1996	0.911	0.000	0.863	0.000
U1978	U1988			0.927	0.000
U1988	U1998			1.000	0.000
C1978	U1978			0.898	0.000
C1990	U1988			0.912	0.000
C1996	U1998			0.916	0.000

Table VI: p-values for Lorenz Dominance Tests
Simulation Method

		after tax		before tax	
G	F	H_0^G	H_0^F	H_0^G	H_0^F
C1974	C1982	0.000	0.986	0.000	0.973
C1982	C1990	0.752	0.159	1.000	0.060
C1990	C1996	1.000	0.001	0.946	0.000
U1978	U1988			1.000	0.000
U1988	U1998			1.000	0.000
C1978	U1978			1.000	0.000
C1990	U1988			1.000	0.000
C1996	U1998			0.990	0.000

Table VII: p-values for Lorenz Dominance Tests
Bootstrap Method:$k = 1000$

7 Conclusion

In this paper we have proposed a method of testing for Lorenz dominance based on independent samples from two populations. The test is fully non–parametric and consistent being based on global comparisons of the empirical Lorenz curves. Although the proposed test statistic has a non-standard and case specific limiting distribution we were able to show that asymptotically valid inferences could be drawn using simulation and the bootstrap. The tests were shown to have a fairly good performance in quite small samples and were illustrated in the context of an empirical example comparing Lorenz curves for Canada and the US over a number of time periods. Although Lorenz dominance relations only provide a partial ordering of distributions, the empirical example illustrates that it is possible to make many meaningful inferences regarding trends in inequality over time, and across countries at a point in time.

Appendix A: Proofs of Results

Proof of Proposition 1: The proof is based on a characterization for the limiting distribution and the application of an inequality. The result in Lemma 1 implies that,

$$\sup_z |(\hat{L}_G(p) - \hat{L}_F(p)) - (L_G(p) - L_F(p))| \overset{a.s.}{\to} 0 \tag{3}$$

and,

$$\begin{aligned} \hat{T} &= \sqrt{\frac{NM}{N+M}}(\hat{L}_G - L_G) - \sqrt{\frac{NM}{N+M}}(\hat{L}_F - L_F) \\ &= \sqrt{\frac{N/M}{N/M+1}}\sqrt{M}(\hat{L}_G - L_G) - \sqrt{\frac{1}{N/M+1}}\sqrt{M}(\hat{L}_F - L_F) \\ &\Rightarrow \lambda^{1/2}\mathcal{L}_G - (1-\lambda)^{1/2}\mathcal{L}_F \\ &\equiv \bar{T} \end{aligned} \tag{4}$$

Use the notation $\hat{T}(p)$ for $\hat{T}$ evaluated at the specific point $p \in P = [0,1]$. An implication of the weak convergence result is that for any $\gamma, \varepsilon > 0$ that there exists a $\delta > 0$ such that the following stochastic equicontinuity condition holds,

$$\lim P(\sup_{|p_1 - p_2| < \delta} |\hat{T}(p_1) - \hat{T}(p_2)| > \varepsilon) < \gamma \tag{5}$$

We first prove (i) and assume that $L_G(p) \leq L_F(p)$ for all p. Let the set of values for which there is equality be equal to P^* so that for any $p \in P^*$ we have that

$$\hat{T}(p) = \left(\frac{NM}{N+M}\right)^{1/2} (\hat{L}_G(p) - \hat{L}_F(p))$$

because for such p, $L_G(p) - L_F(p) = 0$. It is easily seen that P^* is a compact set because of Assumption 1. We aim to show that for $c > 0$,

$$P(\hat{S}_G > c) \to P(\sup_{p \in P^*} \bar{T}(p) > c) \tag{6}$$

To show this we first note that,

$$\begin{aligned} \hat{S}_G &= \left(\frac{NM}{N+M}\right)^{1/2} \sup_p (\hat{L}_G(p) - \hat{L}_F(p)) \\ &\geq \sup_{p \in P^*} \hat{T}(p) \\ &\Rightarrow \sup_{p \in P^*} \bar{T}(p) \end{aligned}$$

because of the fact that $P^* \subset P$ and using Continuous Mapping Theorem (CMT). Consequently,

$$\limsup P(\hat{S}_G \leq c) \leq P(\sup_{p \in P^*} \bar{T}(p) \leq c) \tag{7}$$

Let $\hat{p}$ denote any value of p that solves the problem,

$$\sup_p (\hat{L}_G(p) - \hat{L}_F(p))$$

and note that $\hat{p} \in P$. We suppress the dependence of $\hat{p}$ on N and M for ease of notation. Then, for any non-empty $P^+ \subset P^*$ we have that,

$$\begin{aligned} \hat{S}_G &= \left(\frac{NM}{N+M}\right)^{1/2} (\hat{L}_G(\hat{p}) - \hat{L}_F(\hat{p})) \\ &\leq \sup_{p \in P^*} \hat{T}(p) + \left(\frac{NM}{N+M}\right)^{1/2} (L_G(\hat{p}) - L_F(\hat{p})) + \hat{T}(\hat{p}) - \inf_{p \in P^+} \hat{T}(p) \\ &\leq \sup_{p \in P^*} \hat{T}(p) + \sup_{p \in P^+} (\hat{T}(\hat{p}) - \hat{T}(p)) \\ &\leq \sup_{p \in P^*} \hat{T}(p) + \sup_{p \in P^+} |\hat{T}(\hat{p}) - \hat{T}(p)| \end{aligned} \tag{8}$$

where the second line follows from the fact that,

$$\left(\frac{NM}{N+M}\right)^{1/2} \inf_{p \in P^+} (\hat{L}_G(p) - \hat{L}_F(p)) \leq \left(\frac{NM}{N+M}\right)^{1/2} \sup_{p \in P^*} (\hat{L}_G(p) - \hat{L}_F(p))$$

the third line follows from the fact that under the null hypothesis,

$$(L_G(p) - L_F(p)) \leq 0$$

Now pick any $\varepsilon^* > 0$. Let c' and c'' be such that $c' < c < c''$,

$$P(\sup_{p \in P^*} \bar{T}(p) \leq c) - P(\sup_{p \in P^*} \bar{T}(p) \leq c') < \varepsilon^* \tag{9}$$

$$P(\sup_{p \in P^*} \bar{T}(p) \leq c'') - P(\sup_{p \in P^*} \bar{T}(p) \leq c) < \varepsilon^* \tag{10}$$

Let ε_1 be a positive number such that $0 < \varepsilon_1 < \max\{c'' - c, c - c'\}$ and then pick a $\delta > 0$ such that (5) holds with $\varepsilon = \varepsilon_1$ and $\gamma = \varepsilon^*$. Define the set $P^+ = P^* \cap B(\hat{p}, \delta)$ and let $A_{N,M}$ denote the event that P^+ is nonempty. We first demonstrate that $P(A_{N,M}) \rightarrow 1$. Let $\bar{P}^*_\delta = \{p \in P : d(p, P^*) \geq \delta\}$ where $d(p, P^*) = \inf_{p' \in P^*} |p - p'|$ is a measure of the distance of the point p from the compact set P^*. It is only necessary to consider the case that $\bar{P}^*_\delta$ is nonempty because otherwise $P(A_{N,M}) = 1$ for all N, M. Consider the maximization problem,

$$\sup_{p \in P} (\hat{L}_G(p) - \hat{L}_F(z))$$

It is easy to show that $\bar{P}^*_\delta$ is a compact set. Consequently,

$$\sup_{p \in \bar{P}^*_\delta} (\hat{L}_G(p) - \hat{L}_F(p)) = -2\eta < 0 \tag{11}$$

because of Assumption 1. Pick an arbitrary $p^* \in P^*$ and note that the event $A_{N,M}$ is implied by the event,

$$\begin{aligned} \sup_{p \in \bar{P}^*_\delta} (\hat{L}_G(p) - \hat{L}_F(p)) &< -\eta \\ (\hat{L}_G(p^*) - \hat{L}_F(p^*)) &> -\eta \end{aligned}$$

so that,

$$\begin{aligned} P(A_{N,M}) &\geq P(\sup_{p \in \bar{P}^*_\delta} (\hat{L}_G(p) - \hat{L}_F(p)) < (\hat{L}_G(p^*) - \hat{L}_F(p^*))) \\ &\geq P(\{\sup_{p \in \bar{P}^*_\delta} (\hat{L}_G(p) - \hat{L}_F(p)) < -\eta\} \cap \{(\hat{L}_G(p^*) - \hat{L}_F(p^*)) > -\eta\}) \\ &\geq P((\hat{L}_G(p^*) - \hat{L}_F(p^*)) > -\eta) - P(\sup_{p \in \bar{P}^*_\delta} (\hat{L}_G(p) - \hat{L}_F(p)) > -\eta) \\ &\to 1 \end{aligned}$$

using Lemma 1, (3), (11) and CMT which implies that,

$$\begin{aligned} \sup_{p \in \bar{P}^*_\delta} (\hat{L}_G(p) - \hat{L}_F(p)) &= \sup_{p \in \bar{P}^*_\delta} \Big(\hat{L}_G(p) - \hat{L}_F(p) - (L_G(p) - L_F(p)) + (L_G(p) - L_F(p))\Big) \\ &\leq \sup_{p \in \bar{P}^*_\delta} |(\hat{L}_G(p) - \hat{L}_F(p)) - (L_G(p) - L_F(p))| \\ &\quad + \sup_{p \in \bar{P}^*_\delta} ((L_G(p) - L_F(p)) \\ &\quad \overset{a.s.}{\to} -2\eta \end{aligned}$$

Then,

$$\begin{aligned} P(\hat{S}_G \leq c) &= P(\{\hat{S}_G \leq c\} \cap A_{N,M}) + P(\{\hat{S}_G \leq c\} \cap \bar{A}_{N,M}) \qquad (12) \\ &\geq P(\{\sup_{p \in P^*} \hat{T}(p) + \sup_{p \in P^+} |\hat{T}(\hat{p}) - \hat{T}(p)| \leq c\} \cap A_{N,M}) + P(\{\hat{S}_G \leq c\} \cap \bar{A}_{N,M}) \end{aligned}$$

$$
\begin{aligned}
&\geq P(\{\sup_{p\in P^*}\hat{T}(p)+\sup_{|p_1-p_2|<\delta}|\hat{T}(\hat{p})-\hat{T}(p)|\leq c\}\cap A_{N,M})+P(\{\hat{S}_G\leq c\}\cap \bar{A}_{N,M})\\
&\geq P(\sup_{p\in P^*}\hat{T}(p)+\sup_{|p_1-p_2|<\delta}|\hat{T}(\hat{p})-\hat{T}(p)|\leq c)-P(\bar{A}_{N,M})\\
+\ &P(\{\hat{S}_G\leq c\}\cap \bar{A}_{N,M})
\end{aligned}
$$

where the second line follows from the fact that in the event $A_{N,M}$ the inequality in (8) holds and the third line follows from the fact that,

$$\sup_{p\in P^+}|\hat{T}(\hat{p})-\hat{T}(p)|\leq \sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)|$$

For the first term we use equation 25.12 of Billingsley (1968). In particular we have that,

$$
\begin{aligned}
P(\sup_{p\in P^*}\hat{T}(p) &\leq c')-P(\sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)|\geq\varepsilon_1)\\
&\leq P(\sup_{p\in P^*}\hat{T}(p)+\sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)|\leq c)
\end{aligned}
\tag{13}
$$

and,

$$
\begin{aligned}
P(\sup_{p\in P^*}\hat{T}(p)+\sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)| &\leq c)\\
\leq P(\sup_{p\in P^*}\hat{T}(p)\leq c'')-\quad & P(\sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)|\geq\varepsilon_1).
\end{aligned}
\tag{14}
$$

Then we have that,

$$\lim\left(P(\sup_{p\in P^*}\hat{T}(p)\leq c')-P(\sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)|\geq\varepsilon_1)\right)>P(\sup_{p\in P^*}\bar{T}(p)\leq c)-2\varepsilon^*$$

and,

$$\lim\left(P(\sup_{p\in P^*}\hat{T}(p)\leq c'')-P(\sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)|\geq\varepsilon_1)\right)<P(\sup_{p\in P^*}\bar{T}(p)\leq c)+2\varepsilon^*$$

using (9), (10) and (5). Since ε^* is arbitrary (13) and (14) imply that

$$\lim P(\sup_{p\in P^*}\hat{T}(p)+\sup_{|p_1-p_2|<\delta}|\hat{T}(p_1)-\hat{T}(p_2)|\leq c)=P(\sup_{p\in P^*}\bar{T}(p)\leq c)\tag{15}$$

Note that for the third term in the last line of (12) we have that,

$$
\begin{aligned}
0 &\leq P(\{\hat{S}_G\leq c\}\cap\bar{A}_{N,M})\\
&\leq P(\bar{A}_{N,M})\to 0.
\end{aligned}
$$

using the fact that $P(\bar{A}_{N,M}) \to 0$. Then along with (12) and (15) we have that,

$$\liminf P(\hat{S}_G \leq c) \geq P(\sup_{p \in P^*} \bar{T}(p) \leq c)$$

Therefore using (7) we obtain the result that,

$$\lim P(\hat{S}_G \leq c) = P(\sup_{p \in P^*} \bar{T}(p) \leq c).$$

To show the result in (i) of Proposition 1 fix G and consider two situations that are consistent with the null hypothesis. The first is where the distribution generating the other Lorenz curve is denoted F and satisfies $L_G(p) \equiv L_F(p)$ for all $p \in P$. The second situation involves the distribution F' which is such that $L_G(p) \leq L_{F'}(p)$ for all p but $L_G(p) \equiv L_{F'}(p)$ for all $p \in P^* \subset P$. Assume for simplicity of argument (and without loss of generality – as will be demonstrated below) that the means of all three distributions are identical and equal to one – that is, $\mu_G = \mu_F = \mu_{F'} = 1$. It is clearly the case that,

$$\bar{T}^0 = \sqrt{\lambda}\mathcal{L}_G - \sqrt{1-\lambda}\mathcal{L}_F \overset{d}{\equiv} \mathcal{L}_G$$

since $\mathcal{L}_G \overset{d}{\equiv} \mathcal{L}_F$. Let,

$$\bar{T}' = \sqrt{\lambda}\mathcal{L}_G - \sqrt{1-\lambda}\mathcal{L}_{F'}$$

Note that the process $\mathcal{L}_G$ can be written (using a change of variables) as,

$$\mathcal{L}_G(p) \overset{d}{\equiv} -\int_{y_l}^{Q_G(p)} \mathcal{B}(G(y))dy + L_G(p)\int_{y_l}^{y_u} \mathcal{B}(G(y))dy$$

where we use the notation y_l and y_u for the smallest and largest income values in the G population. Similarly, write,

$$\mathcal{L}_{F'}(p) \overset{d}{\equiv} -\int_{x_l}^{Q_{F'}(p)} \mathcal{B}(F'(x))dx + L_{F'}(p)\int_{x_l}^{x_u} \mathcal{B}(F'(x))dx$$

where we use the notation x_l and x_u for the smallest and largest income values in the F' population. Consider two points p_1, p_2 in the set P^* (with $p_1 \leq p_2$) and let,

$$\begin{aligned} z_1 &= Q_G(p_1) \\ z_2 &= Q_G(p_2) \end{aligned}$$

Since

$$\int_0^p Q_G(t)dt = L_G(p) \leq L_{F'}(p) = \int_0^p Q_{F'}(t)dt \tag{16}$$

(and using continuity of the quantile functions) with equality at p_1 and p_2 then it must be the case

$$\begin{aligned} z_1 &= Q_{F'}(p_1) \\ z_2 &= Q_{F'}(p_2) \end{aligned}$$

Using the same arguments one can show that $y_l \leq x_l$ and $y_u \geq x_u$. Then the covariance kernel of the term $\mathcal{I}_G(p) = -\int_{y_l}^{Q_G(p)} \mathcal{B}(G(y))dy$ is given by,

$$\begin{aligned} E(\mathcal{I}_G(p_1)\mathcal{I}_G(p_2)) &= (z_2 - z_1)\int_{-\infty}^{z_1} G(t)dt \\ &\quad +2\int_{-\infty}^{z_1}\int_{-\infty}^{s} G(t)dtds - (\int_{-\infty}^{z_1} G(t)dt)(\int_{-\infty}^{z_2} G(t)dt) \end{aligned} \tag{17}$$

Also note that (using integration by parts),

$$\begin{aligned} E(\mathcal{I}_G(p_1)\mathcal{I}_G(1)) &= (y_u - z_1)\int_{-\infty}^{z_1} G(t)dt \\ &\quad +2\int_{-\infty}^{z_1}\int_{-\infty}^{s} G(t)dtds - (\int_{-\infty}^{z_1} G(t)dt)(\int_{-\infty}^{y_u} G(t)dt) \\ &= (\mu_G - z_1)\int_{-\infty}^{z_1} G(t)dt + 2\int_{-\infty}^{z_1}\int_{-\infty}^{s} G(t)dtds \\ &= (1 - z_1)\int_{-\infty}^{z_1} G(t)dt + 2\int_{-\infty}^{z_1}\int_{-\infty}^{s} G(t)dtds \end{aligned} \tag{18}$$

where the last line follows from the fact that we have assumed that $\mu_G = 1$. Finally note that,

$$E(\mathcal{I}_G(1)\mathcal{I}_G(1)) = 2\int_{-\infty}^{y_u}\int_{-\infty}^{s} G(t)dtds - (\int_{-\infty}^{y_u} G(t)dt)(\int_{-\infty}^{y_u} G(t)dt) \tag{19}$$

Since for the two points under consideration we have that,

$$\begin{aligned} \int_0^{p_1} Q_G(t)dt &= \int_0^{p_1} Q_{F'}(t)dt \\ \int_0^{p_2} Q_G(t)dt &= \int_0^{p_2} Q_{F'}(t)dt \end{aligned}$$

then it is also the case that,

$$\int_{-\infty}^{z_1} G(t)dt = \int_{-\infty}^{z_1} F'(t)dt \tag{20}$$

$$\int_{-\infty}^{z_2} G(t)dt = \int_{-\infty}^{z_2} F'(t)dt \tag{21}$$

by using integration by parts. Also note that since $x_u \leq y_u$ and $F'(t) = 1$ on the range $[x_u, y_u]$ then we have (by integration by parts),

$$\begin{aligned}
\int_{-\infty}^{y_u} F'(t)dt &= \int_{-\infty}^{x_u} F'(t)dt + \int_{x_u}^{y_u} 1dt \\
&= (x_u - \mu_{F'}) + (y_u - x_u) \\
&= y_u - \mu_{F'} \\
&= \int_{-\infty}^{y_u} G(t)dt
\end{aligned}$$

It is known that (16) implies second order stochastic dominance so that,

$$\int_{-\infty}^{z} G(t)dt \geq \int_{-\infty}^{z} F'(t)dt \tag{22}$$

for all z and that this implies third order stochastic dominance,

$$\int_{-\infty}^{z}\int_{-\infty}^{s} G(t)dtds \geq \int_{-\infty}^{z}\int_{-\infty}^{s} F'(t)dtds. \tag{23}$$

Define the following constants,

$$\begin{aligned}
c_1^* &= 2\int_{-\infty}^{z_1}\int_{-\infty}^{s} G(t)dtds - 2\int_{-\infty}^{z_1}\int_{-\infty}^{s} F'(t)dtds \\
c_2^* &= 2\int_{z_1}^{z_2}\int_{-\infty}^{s} G(t)dtds - 2\int_{z_1}^{z_2}\int_{-\infty}^{s} F'(t)dtds \\
c_3^* &= 2\int_{z_2}^{y_u}\int_{-\infty}^{s} G(t)dtds - 2\int_{z_2}^{y_u}\int_{-\infty}^{s} F'(t)dtds
\end{aligned}$$

and note that $c_j^* \geq 0$ using the fact noted in (22) and (23). Then we note that,

$$\begin{aligned}
2\int_{-\infty}^{z_2}\int_{-\infty}^{s} G(t)dtds - 2\int_{-\infty}^{z_2}\int_{-\infty}^{s} F(t)dtds &= c_1^* + c_2^* \\
2\int_{-\infty}^{y_u}\int_{-\infty}^{s} G(t)dtds - 2\int_{-\infty}^{y_u}\int_{-\infty}^{s} F(t)dtds &= c_1^* + c_2^* + c_3^*
\end{aligned}$$

Now,

$$\begin{aligned}
E((\mathcal{L}_G(p_1) - \mathcal{L}_G(p_2))^2) &= E(\mathcal{I}_G(p_1)^2) + L_G(p_1)^2 E(\mathcal{I}_G(1)^2) - 2L_G(p_1)E(\mathcal{I}_G(p_1)\mathcal{I}_G(1)) \\
&\quad + E(\mathcal{I}_G(p_2)^2) + L_G(p_2)^2 E(\mathcal{I}_G(1)^2) - 2L_G(p_2)E(\mathcal{I}_G(p_2)\mathcal{I}_G(1)) \\
&\quad - 2E(\mathcal{I}_G(p_1)\mathcal{I}_G(p_2)) + 2L_G(p_1)E(\mathcal{I}_G(p_2)\mathcal{I}_G(1)) \\
&\quad + 2L_G(p_2)E(\mathcal{I}_G(p_1)\mathcal{I}_G(1)) - 2L_G(p_1)L_G(p_2)E(\mathcal{I}_G(1)^2)
\end{aligned}$$

and use the basic expressions in (17), (18), (19) to see that,

$$\begin{aligned}
E(\mathcal{I}_G(p_1)^2) &= E(\mathcal{I}_{F'}(p_1)^2) + c_1^* \\
E(\mathcal{I}_G(p_1)\mathcal{I}_G(1)) &= E(\mathcal{I}_{F'}(p_1)\mathcal{I}_{F'}(1)) + c_1^* \\
E(\mathcal{I}_G(p_2)^2) &= E(\mathcal{I}_{F'}(p_1)^2) + c_1^* + c_2^* \\
E(\mathcal{I}_G(p_2)\mathcal{I}_G(1)) &= E(\mathcal{I}_{F'}(p_2)\mathcal{I}_{F'}(1)) + c_1^* + c_2^*
\end{aligned}$$

and,

$$\begin{aligned}
E(\mathcal{I}_G(1)^2) &= E((-\int_{-\infty}^{y_u} \mathcal{B}(F'(y))dy)^2) + c_1^* + c_2^* + c_3^* \\
&= E(\mathcal{I}_{F'}(1)^2) + c_1^* + c_2^* + c_3^* + c_4^*
\end{aligned}$$

where,

$$\begin{aligned}
c_4^* &= E((-\int_{-\infty}^{y_u} \mathcal{B}(F'(y))dy)^2) - E(\mathcal{I}_{F'}(1)^2) \\
&= E((-\int_{-\infty}^{y_u} \mathcal{B}(F'(y))dy)^2) - E((-\int_{-\infty}^{x_u} \mathcal{B}(F'(y))dy)^2) \\
&\geq 0
\end{aligned}$$

with the last line following from the fact that $y_u \geq x_u$ and the fact that the variance of $-\int_{-\infty}^{z} \mathcal{B}(F'(y))dy$ is an increasing function of z. Then we have that,

$$\begin{aligned}
E((\mathcal{L}_G(p_1) - \mathcal{L}_G(p_2))^2) &= E(\mathcal{I}_{F'}(p_1)^2) + c_1^* + L_{F'}(p_1)^2(E(\mathcal{I}_{F'}(1)^2) + c_1^* + c_2^* + c_3^* + c_4^*) \\
&\quad -2L_{F'}(p_1)(E(\mathcal{I}_{F'}(p_1)\mathcal{I}_{F'}(1)) + c_1^*) \\
&\quad E(\mathcal{I}_{F'}(p_1)^2) + c_1^* + c_2^* + L_{F'}(p_2)^2(E(\mathcal{I}_{F'}(1)^2) + c_1^* + c_2^* + c_3^* + c_4^*) \\
&\quad -2L_{F'}(p_2)(E(\mathcal{I}_{F'}(p_2)\mathcal{I}_{F'}(1)) + c_1^* + c_2^*) \\
&\quad -2(E(\mathcal{I}_{F'}(p_1)\mathcal{I}_{F'}(p_2)) + c_1^*) + 2L_{F'}(p_1)(E(\mathcal{I}_{F'}(p_2)\mathcal{I}_{F'}(1)) + c_1^* + c_2^*) \\
&\quad +2L_{F'}(p_2)(E(\mathcal{I}_G(p_1)\mathcal{I}_G(1)) + c_1^*) \\
&\quad -2L_{F'}(p_1)L_{F'}(p_2)(E(\mathcal{I}_{F'}(1)^2) + c_1^* + c_2^* + c_3^* + c_4^*) \\
&= E((\mathcal{L}_{F'}(p_1) - \mathcal{L}_{F'}(p_2))^2) + c_1^*(L_{F'}(p_2) - L_{F'}(p_1))^2 \\
&\quad +c_2^*(1 - (L_{F'}(p_2) - L_{F'}(p_1)))^2 + (c_3^* + c_4^*)(L_{F'}(p_2) - L_{F'}(p_1))^2 \\
&\geq E((\mathcal{L}_{F'}(p_1) - \mathcal{L}_{F'}(p_2))^2)
\end{aligned}$$

where the last inequality follows from the fact that all of the c_j^* are non-negative. Then consider,

$$\begin{aligned} E((\bar{T}'(p_2) - \bar{T}'(p_1))^2) &= \lambda E((\mathcal{L}_G(p_2) - \mathcal{L}_G(p_1))^2) + (1-\lambda)E((\mathcal{L}_{F'}(p_2) - \mathcal{L}_{F'}(p_1))^2) \\ &\leq \lambda E((\mathcal{L}_G(p_2) - \mathcal{L}_G(p_1))^2) + (1-\lambda)E((\mathcal{L}_G(p_2) - \mathcal{L}_G(p_1))^2) \\ &= E((\mathcal{L}_G(p_2) - \mathcal{L}_G(p_1))^2) \\ &= E((\bar{T}^0(p_2) - \bar{T}^0(p_1))^2) \end{aligned}$$

Since the stochastic processes are separable, mean zero and Gaussian then Proposition A.2.6 of Van der Vaart and Wellner (1996) (the Slepian, Fernique, Marcus and Shepp inequality) implies that,

$$P(\sup_{p \in P^*} \bar{T}'(p) > c) \leq P(\sup_{p \in P^*} \bar{T}^0(p) > c) \leq P(\sup_{p \in P} \bar{T}_2^0(p) > c)$$

for any $c > 0$ where the second inequality follows from the fact that $P^* \subset P$. But $P(\sup_{p \in P} \bar{T}^0(p) > c)$ is the asymptotic probability of rejection in the case where $L_F(p) \equiv L_G(p)$ for all $p \in P$ and so the result follows in the case where we have assumed that $\mu_{F'} = \mu_G = 1$. Since the empirical and population Lorenz curves are invariant to scalar multiplication of the variables one can consider the comparison of Lorenz curves for the transformed variables,

$$\begin{aligned} Y^* &= \frac{Y}{\mu_G} \\ X^* &= \frac{X}{\mu_{F'}} \end{aligned}$$

and the preceding arguments hold since the transformed variables have a mean of one. **Q.E.D.**

Proof of Lemma 2: Clearly,

$$\tilde{Z}_G = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} (Y_i - \mu_G) \Rightarrow \mathcal{G}_G(1).$$

The function $\tilde{B}_G$ at the value $p \in [0,1]$ is given by,

$$\begin{aligned}\tilde{B}_G(p) &= \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(1(Y_i \leq Q_G(p)) - p)\\ &= \sqrt{M}(\hat{G}(Q_G(p)) - G(Q_G(p)))\end{aligned}$$

using the fact that $G(Q_G(p)) = p$. For the third component we note that,

$$\begin{aligned}\frac{1}{M}\sum_{i=1}^{M} Y_i 1(Y_i &\leq Q_G(p)) = Q_G(p)\hat{G}(Q_G(p)) - \int_{y_l}^{Q_G(p)} \hat{G}(y)dy\\ E(Y_i 1(Y_i &\leq Q_G(p))) = Q_G(p)G(Q_G(p)) - \int_{y_l}^{Q_G(p)} G(y)dy\end{aligned}$$

Then we can write,

$$\tilde{C}_G(p) = Q_G(p)\sqrt{M}(\hat{G}(Q_G(p)) - G(Q_G(p))) - \int_{y_l}^{Q_G(p)} \sqrt{M}(\hat{G}(y) - G(y))dy$$

Consequently we have that,

$$\begin{aligned}Q_G(p)\tilde{B}_G(p) - \tilde{C}_G(p) &= \int_{y_l}^{Q_G(p)} \sqrt{M}(\hat{G}(y) - G(y))dy\\ &\Rightarrow \int_{y_l}^{Q_G(p)} \mathcal{B}_G(G(y))dy\end{aligned}$$

by CMT. Consequently we have that,

$$\begin{aligned}\tilde{L}_{G,M}(p) &\Rightarrow -\frac{1}{\mu_G}(\int_{y_l}^{Q_G(p)} \mathcal{B}_G(G(y))dy) - \frac{L_G(p)}{\mu_G}\mathcal{G}_G(1)\\ &\equiv \frac{1}{\mu_G}\mathcal{G}_G(p) - \frac{L_G(p)}{\mu_G}\mathcal{G}_G(1)\\ &= \mathcal{L}_G(p)\end{aligned}$$

where the second line follows using a change of variable and the last line follows from the definition of $\mathcal{L}_G(p)$ **Q.E.D.**

Proof of Lemma 3: First note that,

$$\tilde{Z}_G^* = \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(Y_i - \hat{\mu}_G)U_i = \frac{1}{\sqrt{M}}\sum_{i=1}^{M}(Y_i - \mu_G)U_i + (\hat{\mu}_G - \mu_G)\frac{1}{\sqrt{M}}\sum_{i=1}^{M} U_i$$

The first term satisfies,

$$\frac{1}{\sqrt{M}}\sum_{i=1}^{M}(Y_i - \mu_G)U_i \Rightarrow \mathcal{G}_G(1)$$

by the Lindeberg Levy Central Limit Theorem using the facts that

$$\begin{aligned} E((Y_i - \mu_G)U_i) &= 0 \\ E(((Y_i - \mu_G)^2 U_i^2) &= E((Y_i - \mu_G)^2) \end{aligned}$$

using the independence of the U_i. For the second term we have by the Strong Law of Large Numbers that, $\hat{\mu}_G - \mu_G \overset{a.s.}{\rightarrow} 0$. Therefore $\hat{\mu}_G - \mu_G \rightarrow 0$ for almost all samples. Condition on the sample and note that,

$$\begin{aligned} E_U((\hat{\mu}_G - \mu_G)U_i) &= 0 \\ E_U((\hat{\mu}_G - \mu_G)^2 U_i^2) &= (\hat{\mu}_G - \mu_G)^2 \\ &\rightarrow 0 \end{aligned}$$

Therefore by the Markov inequality,

$$\lim_{M\rightarrow\infty} P_U(|(\hat{\mu}_G - \mu_G)\frac{1}{\sqrt{M}}\sum_{i=1}^{M} U_i| > \varepsilon) = 0.$$

Therefore for this particular sample we have that (conditionally) $\tilde{Z}_G^* \Rightarrow \mathcal{G}_G'(1)$. But this holds for almost every sample so that unconditionally $\tilde{Z}_G^* \overset{a.s.}{\Rightarrow} \mathcal{G}_G'(1)$. Next consider the term,

$$\begin{aligned} \hat{Q}_G(p)\tilde{B}_G^*(p) - \tilde{C}_G^*(p) &= \int_{y_l}^{\hat{Q}_G(p)} \frac{1}{\sqrt{M}}\sum_{i=1}^{M}\left(1(Y_i \leq \hat{Q}_G(p)) - \hat{G}(\hat{Q}_G(p))\right) U_i \\ &= W_0(p) + W_1(p) + W_2(p) + W_3(p) \end{aligned}$$

where we can write,

$$W_j(p) = \frac{1}{\sqrt{M}}\sum_{i=1}^{M} U_i \hat{w}_{ji}(p)$$

$$\hat{w}_{0i}(p) = \int_{y_l}^{Q_G(p)} \left(1(Y_i \leq Q_G(p)) - G(Q_G(p))\right) dp$$

$$\hat{w}_{1i}(p) = \int_{Q_G(p)}^{\hat{Q}_G(p)} \left(1(Y_i \leq Q_G(p)) - G(Q_G(p))\right) dp$$

$$\hat{w}_{2i}(p) = \int_{y_l}^{\hat{Q}_G(p)} (1(Y_i \leq \hat{Q}_G(p)) - 1(Y_i \leq Q_G(p)))dp$$

$$\hat{w}_{3i}(p) = \int_{y_l}^{\hat{Q}_G(p)} (\hat{G}(\hat{Q}_G(p)) - G(Q_G(p)))dp$$

noting that U_i is independent of $\hat{w}_{ji}$. The following inequalities can be shown for each of the last three terms,

$$\begin{aligned}
\sup_p |\hat{w}_{1i}(p)| &\leq \sup_p |\hat{Q}_G(p) - Q_G(p)| \\
\sup_p |\hat{w}_{2i}(p)| &\leq \sup_p |\hat{Q}_G(p) - Q_G(p)| \\
\sup_p |\hat{w}_{1i}(p)| &\leq \Delta \left(\sup_p |\hat{G}(y) - G(y)| + \sup_p |\hat{Q}_G(p) - Q_G(p)| \right)
\end{aligned}$$

for some positive finite constant. By Lemma 1 all of the terms on the right hand side converge to zero almost surely. Consequently we have that for $j = 1, 2, 3$,

$$\sup_p W_j(p) = o_p(1)$$

for almost every sample using Markov inequality. Applying Corollary 2.9.3 of Van der Vaart and Wellner (1996) to the term $W_0(p)$, we have the result,

$$\hat{Q}_G(p)\tilde{B}^*_G(p) - \tilde{C}^*_G(p) \overset{a.s.}{\Rightarrow} \mathcal{G}'_G(1)$$

which is an independent copy of $\mathcal{G}_G(p)$. Combining results and using the CMT we have that, $\tilde{L}^*_G \overset{a.s.}{\Rightarrow} \mathcal{L}'_G$. **Q.E.D.**

Proof of Proposition 2: Let, $\hat{P}_M(t)$ be the c.d.f. of the process (conditional on the original sample of Y_i) generated by $\sup_p \tilde{L}^*_G(p)$. By Lemma 2 and the CMT we have that,

$$\sup_p \tilde{L}^*_G(p) \overset{a.s.}{\Rightarrow} \sup_p \mathcal{L}'_G(p) \tag{24}$$

where the process $\sup_p \mathcal{L}'_G(p)$ is identical to (but independent of) the process $\sup_p \mathcal{L}_G(p)$. Note that the median of the distribution of $\sup_p \mathcal{L}'_G(p)$ (denoted $\bar{P}^0_G(t)$) is strictly positive and finite. By Tsirel'son (1975) $\bar{P}^0_G(t)$ is absolutely continuous on $(0, \infty)$ and, moreover, the $(1 - \alpha)$ quantile, denoted $c_G(\alpha)$ is finite and positive for any fixed $\alpha < 1/2$ using (for instance) Proposition A.2.7 of Van der Vaart and Wellner (1996). Note that event $\{\hat{p}_G < \alpha\}$ is equivalent to the event that $\{\hat{S}_G > \hat{c}(\alpha)\}$ where,

$$\hat{c}_G(\alpha) = \inf\{t : \hat{P}_M(t) > \alpha\}$$

Also note that $\hat{c}_G(\alpha) \overset{a.s.}{\to} c_G(\alpha)$ by (24) and the properties of $\bar{P}^0(t)$ noted above. Then,

$$\begin{aligned} P(\text{reject } H_0^G | H_0^G) &= P(\hat{S}_G > \hat{c}_G(\alpha)) \\ &= P(\hat{S}_G > c_G(\alpha)) + (P(\hat{S}_G > \hat{c}_G(\alpha)) - P(\hat{S}_G > c_G(\alpha))) \end{aligned}$$

Consider the second term in brackets and note that,

$$P(\hat{S}_G > \hat{c}_G(\alpha)) - P(\hat{S}_G > c_G(\alpha)) = P(\hat{c}_G(\alpha) \leq \hat{S}_G < c_G(\alpha)) - P(c_G(\alpha) \leq \hat{S}_G < \hat{c}_G(\alpha))$$

Using Tsirel'son (1975) the distribution of $\sup_p \mathcal{L}'_G(p)$, say $\bar{P}_G^0(t)$, is absolutely continuous on $(0, \infty)$, so that because $c_G(\alpha) > 0$ it is a continuity point of the distribution. Then for a fixed (arbitrary) $\varepsilon > 0$, pick $\delta > 0$ such that $|c_G(\alpha) - c'| < \delta$ implies $|\bar{P}_G^0(c_G(\alpha)) - \bar{P}_G^0(c')| < \varepsilon$. Denote the events,

$$\begin{aligned} A_1 &= \{\hat{c}_G(\alpha) \leq \hat{S}_G < c_G(\alpha)\} \\ A_{1,\delta} &= \{c_G(\alpha) - \delta \leq \hat{S}_G < c_G(\alpha)\} \\ A_2 &= \{c_G(\alpha) \leq \hat{S}_G < \hat{c}_G(\alpha)\} \\ A_{2,\delta} &= \{c_G(\alpha) \leq \hat{S}_G < c_G(\alpha) + \delta\} \\ A_3 &= \{\hat{c}_G(\alpha) \in B(c_G(\alpha), \delta))\} \end{aligned}$$

where $B(c_G(\alpha), \delta)$ is the open ball around the point $c_G(\alpha)$ with radius δ. Then we have,

$$\begin{aligned} P(A_1) &= P(A_1 \cap A_3) + P(A_1 \cap \bar{A}_3) \\ &\leq P(A_{1,\delta}) + P(\bar{A}_3) \\ &\to P(c_G(\alpha) - \delta \leq \sup_{p \in P^*} \bar{T}(p) < c_G(\alpha)) \\ &\leq \varepsilon \end{aligned}$$

and,

$$\begin{aligned} P(A_2) &= P(A_2 \cap A_3) + P(A_2 \cap \bar{A}_3) \\ &\leq P(A_{2,\delta}) + P(\bar{A}_3) \\ &\to P(c_2(\alpha) \leq \sup_p \sup_{p \in P^*} \bar{T}(p) < c_G(\alpha) + \delta) \\ &\leq \varepsilon \end{aligned}$$

using the consistency of $\hat{c}_G(\alpha)$ and the convergence result for $\hat{S}_G$ shown in Proposition 1. Consequently,

$$\limsup |P(\hat{S}_G > \hat{c}_G(\alpha)) - P(\hat{S}_G > c_G(\alpha))| \leq 2\varepsilon$$

But ε was arbitrary so using the definition of $c_G(\alpha)$ and the results shown in Proposition 1 we have that,

$$\begin{aligned} P(\text{reject } H_0^G | H_0^G) &\rightarrow P(\sup_{p \in P^*} (\lambda^{1/2} \mathcal{L}_G(p) - (1-\lambda)^{1/2} \mathcal{L}_F(p)) > c_G(\alpha)) \\ &\leq P(\sup_p \mathcal{L}_G(p) > c_G(\alpha)) \\ &= \alpha \end{aligned}$$

On the other hand Proposition 1(ii) and finiteness of $c_G(\alpha)$ imply that $\lim P(\text{reject } H_0^G | H_1^G) = 1$. The result for the SD3 test follows similarly. **Q.E.D.**

Proof of Proposition 3: By Theorem 3.6.3 of Van der Vaart and Wellner (1996) we have that,

$$\sqrt{k}(\hat{G}_k^* - \hat{G}_k) \overset{p}{\Rightarrow} \mathcal{B}_{G,s}^* \circ G$$

provided that $k \rightarrow \infty$ where the process $\mathcal{B}_{G,s}^* \circ G$ has the same distribution as $\mathcal{B}_G \circ G$. This convergence is in the sense that,

$$\sup_{h \in BL_1} |E_C(h(\sqrt{k}(\hat{G}_k^* - \hat{G}_k))) - E(h(\mathcal{B}_G \circ G))| \xrightarrow{p} 0$$

where BL_1 is the space of bounded Lipschitz functions mapping $C[0,1]$ into $[0,1]$, and where E_C is the expectation given the sample $\mathcal{Y}$. Define the functional $\psi : D[y_l, y_u] \rightarrow C(0,1)$, by

$$\psi(G)(p) = \int_0^p Q_G(t) dt$$

and note that $\psi(G)(1) = \mu_G$. Using the definition in the text we can similarly define,

$$\psi(\hat{G}_k)(p) = \int_0^p \hat{Q}_{G,k}(t) dt$$

Using Assumptions 1 and 2, we have by Lemma 3.9.23 and Lemma 3.9.3 of Van der Vaart and Wellner (1996) and the linearity of the functional ψ in Q_G that considered as a functional of G, ψ is Hadamard differentiable at G with Hadamard derivative given by, $\psi'(\theta)$ where,

$$\psi'(\theta)(p) = -\int_0^p \frac{\theta(Q_G(t))}{g(Q_G(t))}dt$$

Using this fact Theorem 3.9.11 of Van der Vaart and Wellner (1996) implies that,

$$\sqrt{k}(\psi(\hat{G}_k^*) - \psi(\hat{G}_k) \overset{p}{\Rightarrow} \psi'(\mathcal{B}_{G,s}^* \circ G). \tag{25}$$

The mapping that define the Lorenz curve has the form,

$$L_G = \frac{\psi(G)}{\psi(G)(1)}$$

so that,

$$L_G(p) = \frac{\psi(G)(p)}{\psi(G)(1)} = \frac{\int_0^p Q_G(t)dt}{\int_0^1 Q_G(t)dt} = \frac{\int_0^p Q_G(t)dt}{\mu_G}$$

We will treat this mapping as being a functional such that, $L_G = \lambda(\Theta_1, \Theta_2)$ where $\Theta_1 = \psi(G)$ and $\Theta_2 = \psi(G)(1)$. Considered as a functional $\lambda(\Theta_1, \Theta_2)$ is Hadamard differentiable with derivative,

$$\lambda'(\theta_1, \theta_2) = \frac{1}{\Theta_2}\theta_1 - \frac{\Theta_1}{\Theta_2^2}\theta_2$$

provided that $\Theta_2 > \varepsilon > 0$ which holds under Assumption 1. To show this consider the sequences $\Theta_{1t} = \Theta_1 + t\theta_{1t}$ and $\Theta_{2t} = \Theta_2 + t\theta_{2t}$ with $t \to 0$ (with $t \in R$) where θ_{1t} is such that $\theta_{1t} \to \theta_1$ and $\Theta_{1t} \in C(0,1)$ and $\theta_{2t} \to \theta_2$ and $\Theta_{2t} > \varepsilon$. Then,

$$\frac{\frac{\Theta_{1t}}{\Theta_{2t}} - \frac{\Theta_1}{\Theta_2}}{t} - \lambda'(\theta_1, \theta_2) = \frac{\theta_{1t} - \theta_1}{\Theta_{2t}} - \frac{t\theta_{2t}\theta_1}{\Theta_2\Theta_{2t}} - \frac{\Theta_1(\theta_{2t} - \theta_2)}{\Theta_2\Theta_{2t}} + \frac{t\Theta_1\theta_2\theta_{2t}}{\Theta_2^2\Theta_{2t}}$$

Each term is easily seen to converge to zero uniformly so that Hadamard differentiability holds. Using this fact, the result (25) and a further application of Theorem 3.9.11 of Van der Vaart and Wellner (1996) gives the result that,

$$\sqrt{k}(\hat{L}_{G,k}^* - \hat{L}_{G,k}) \overset{p}{\Rightarrow} \frac{1}{\mu_G}\psi'(\mathcal{B}_{G,s}^* \circ G) - \frac{L_G}{\mu_G}\psi'(\mathcal{B}_{G,s}^* \circ G)(1) \equiv \mathcal{L}_{G,s}$$

where, $\mathcal{L}_{G,s} \overset{d}{\equiv} \mathcal{L}_G$. The remainder of the proof follows the proof of Proposition 2 (using $\overset{p}{\Rightarrow}$ instead of $\overset{a.s}{\Rightarrow}$ so that $\hat{c}_G(\alpha) \xrightarrow{p} c_G(\alpha)$. **Q.E.D.**

References

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

[2] Barrett, G.F. and S.G. Donald (1999a) "Statistical Inference with Generalized Gini Indices of Inequality and Poverty", mimeo.

[3] Barrett, G.F. and S.G. Donald (1999b) "Consistent Tests for Stochastic Dominance", mimeo.

[4] Basmann, R.L., K.J. Hayes and D.J. Slottje (1993) *Some New Methods for Measuring and Describing Economic Inequality,* JAI Press.

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

[6] Billingsley, P. (1968) *Convergence of Probability Measures*, John Wiley and Sons: New York.

[7] 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.

[8] 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.

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

[10] Dasgupta, P., A.Sen and D.Starrett, (1973), "Notes on the Measurement of Inequality" *Journal of Economic Theory*, **6**, 180-187.

[11] Gail, M. H. and J. L. Gastwirth (1978) "A Scale-Free Goodness-of- Fit Test for the Exponential Distribution Based on the Lorenz Curve" *Journal of the American Statistical Association*, **73,** 787-793.

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

[13] Hansen, B. E. (1996) "Inference When a Nuisance Parameter Is Not Identified under the Null Hypothesis" *Econometrica*, **64**, 413-430.

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

[15] 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.

[16] Organization for Economic Co-operation and Development (1993) *Purchasing Power Parities and Real Expenditures for Canada and the United States*, Organization for Economic Co-operation and Development:Paris.

[17] Ryu, H. and D.J. Slottje (1996) "Two Flexible Functional Form Approaches for Approximating the Lorenz Curve" *Journal of Econometrics,* **72**, 251-74.

[18] Shorrocks,A.F. and J.E.Foster (1987),"Transfer sensitive inequality measures" *Review of Economic Studies*, **54**, 485-97.

[19] Tsirel'son, (1975) "The Density of the Distribution of the Maximum of a Gaussian Process", *Theory of Probability and its Applications,* **16**, 847-856.

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