Tails of Lorenz Curves*

Christian Schluter† University of Bristol

Mark Trede‡ Universität zu Köln

February 2001

Abstract

The Lorenz dominance criterion is the centre piece of inequality analysis. Yet, the appeal of this criterion, which requires considering Lorenz curves in their entirety, is undermined by the practical problem that many sample Lorenz curves intersect in the tails. The commonly used inferential methods, based on central limit theorem arguments, do not apply to the tails since these contain too few observations. By contrast, we propose a test procedure, based on a domain of attraction assumption, which fully takes into account the tail behaviour of Lorenz curves. Our experiments and empirical examples demonstrate the good performance of the proposed test: in many cases we are able to infer that despite sample tail crossings the population Lorenz curves do, in fact, exhibit Lorenz dominance.

Keywords: Lorenz curves; statistical inference; tail behaviour; regular variation

JEL classification: D31, D63, I32

*This paper forms part of the research programme of the TMR network **Living Standards, Inequality and Taxation** [Contract No. ERBFMRXCT 980248] of the European Communities whose financial support is gratefully acknowledged. Many thanks also to Kees Jan van Garderen, Christian Kleiber, and Richard Smith for valuable comments.

†Department of Economics, University of Bristol, 8 Woodland Road, Bristol, BS8 1TN, UK. Tel. +44 (0)117 928 8431, Fax. +44 (0)117 928 8577. Email: C.Schluter@bristol.ac.uk.

‡Corresponding author. Seminar für Wirtschafts- und Sozialstatistik, Universität zu Köln, 50923 Köln, Germany,Tel.: +49-221-470 2283, Fax: +49-221-470 5074, e-mail: trede@wiso.uni-koeln.de

1 Introduction

The main tool for analysing economic inequality is the Lorenz curve. In order to compare inequality between two distributions one draws their Lorenz curves and concludes that inequality is unanimously higher in one distribution if its Lorenz curve is everywhere below the curve of the other distribution: any inequality measure which satisfies the principles of transfer, of anonymity, and of mean independence will rank the two distributions in the same way as the Lorenz curves (Atkinson, 1970). The Lorenz curve provides, however, only a partial ordering of income distributions. If the curves cross, the ranking is indeterminate unless one is willing to make further assumptions about the social welfare function.

Population Lorenz curves are rarely known since one rarely has information about the entire population, and empirical Lorenz curves have to be estimated from sample data. The statistical theory for the main body of the Lorenz curve, which contains many observations, is well-developed (Beach and Davidson, 1983; Beach and Richmond, 1985; Davidson and Duclos, 1997). However, these methods do not apply to the tails of the Lorenz curve since the tails contain too few observations to permit invoking the usual central limit theorem arguments. However, the tail behaviour is of considerable interest, and it is precisely in the tails where crossings often occur in practise. Such crossings of sample Lorenz curves can occur although the population Lorenz curves do not cross. This possibility is illustrated by our experiments, reported below, with realistically calibrated parametric models: 45 percent of sample Lorenz curves intersect in the tails, although the population Lorenz curves do not.

To overcome this tail behaviour problem we develop a test which is based on extreme value theory and the theory of regular variation. For income distributions we have in mind, it is reasonable to assume that their tails lie in the domain of attraction of the Fréchet distribution, i.e. they decay like power functions. Examples of parametric models which exhibit this characteristic are the generalised beta distributions of the second kind (McDonald and Xu, 1995), and therefore the special cases of the Singh-Maddala distribution and the Dagum distribution, all of which fit real world income data reasonably well (Brachmann, Stich and Trede, 1996). We do not examine middle heavy and thin tailed distributions, which decay like exponential functions, such as log-normal distributions (whose Lorenz curves cannot cross). Moreover, the fit of parametric models based on power functions to the tails of real world income data is far superior to the fit of lognormal models.

The domain-of-attraction assumption permits us to estimate extreme quantiles outside the data range without imposing strong assumptions on the parametric form of the income distribution. The test procedure based on extreme value theory closes the vexing gap left by the conventional approach to statistical inference for Lorenz curves: using our test we are able to infer in many cases that despite sample tail crossings, the population Lorenz curves do, in fact, exhibit Lorenz dominance. Moreover, our experiments suggest that the empirical level of our test is close to its nominal value.

This paper is organised as follows. Before setting out our proposed test for Lorenz curve tails, we collect in Section 2 the relevant results from extreme value theory. The test itself is presented in section 2.4. Section 3 provides two illustrations of our testing procedure: a Monte-Carlo simulation and an empirical example using data on disposable personal income from the Luxembourg Income Study (LIS). Section 4 concludes.

2 Statistical Inference for Lorenz Curve Tails

2.1 Preliminaries

Let $X_1, \ldots, X_n$ be an i.i.d. sample from an absolutely continuous (income) distribution function F_X with $F_X(0) = 0$. As the Lorenz curve is scale invariant we assume without loss of generality that the mean of X is normalized to $E(X) = 1$. The upper tail of F_X is denoted by $\bar{F}_X(x) = 1 - F_X(x)$, and order statistics by $X_{(1)} \geq \ldots \geq X_{(n)}$.

The Lorenz curve of X is given by

$$\{(p, L_X(p)), 0 \leq p \leq 1\} \quad \text{with} \quad L_X(p) = \int_{x \geq 0} I\left(x \leq F_X^{-1}(p)\right) x dF_X(x),$$

where $I(.)$ is the indicator function. Let Y be a similarly defined random variable. X Lorenz dominates Y if $L_X(p) \geq L_Y(p)$ for all $p \in [0, 1]$ and $L_X(p_0) > L_Y(p_0)$ for at least one $p_0 \in [0, 1]$.

Lorenz dominance can equivalently be expressed as second order stochastic dominance of the normalised distributions, because, if $E(X) = E(Y) = 1$,

$$X \text{ Lorenz dominates } Y \Longleftrightarrow \int_x^{\infty} \overline{F}_X(t)\, dt \leq \int_x^{\infty} \overline{F}_Y(t)\, dt \text{ for all } x > 0. \tag{1}$$

Since this paper is concerned with inference for tail behaviour of Lorenz curves we assume the following:

Assumption A1: *X Lorenz dominates Y in the middle of the distribution.*

Tests of this assumption about the main body of the Lorenz curve are well-known (Beach and Davidson, 1983; Beach and Richmond, 1985; Davidson and Duclos, 1997). To test for overall Lorenz dominance, we therefore have to test multiple hypotheses: whether Lorenz dominance occurs (i) in the main body of the Lorenz curve – the conventional test – and (ii) in the tails. Depending on the formulation of the null and alternative hypotheses, this calls for either an intersection-union or union-intersection test (see e.g. Savin, 1993). We discuss this point in greater detail in Appendix B. In order to facilitate the exposition, we concentrate exclusively on (ii).

2.2 Results from Extreme Value Theory

A well known result concerning the distribution of the maximum is that if there exist norming constants $c_n > 0$ and $d_n \in \mathbb{R}$ such that

$$\frac{X_{(1)} - d_n}{c_n} \xrightarrow{D} Z,$$

then Z is distributed as either of the following three distributions: (i) the Gumbel distribution $\exp(-\exp(-x))$ for $x \in \mathbb{R}$, (ii) the Weibull distribution given by $\exp(-(-x)^{\alpha})$ for $x \leq 0$ and 1 otherwise, with $\alpha > 0$, and (iii) the Fréchet distribution

$$\Psi_{\alpha}(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases} \tag{2}$$

with $\alpha > 0$. We make the following

Assumption A2: *The distribution F_X lies in the domain of attraction of the Fréchet distribution Ψ_{α}.*

For income models we have in mind, the Fréchet distribution is the only relevant limiting distribution for reasons that will become clearer once we are able to translate the assumption about the maximum into a condition on the tail of the distribution. To this end, we use the concept of regular variation. Recall that a function g is called regularly varying at x_0 with index θ if

$$\lim_{x \to x_0} \frac{g(tx)}{g(x)} = t^{\theta}, \qquad t > 0.$$

The class of all distribution functions with regularly varying tails with parameter θ is denoted by R_{θ}. If $\theta = 0$, the function is said to be slowly varying.

If the distribution is in the domain of attraction of the Fréchet distribution Ψ_α with parameter α, the index of regular variation of the upper tail $\bar{F}_X(x)$ at infinity equals $\theta = -\alpha$, i.e.

$$\lim_{x\to\infty} \frac{\bar{F}_X(tx)}{\bar{F}_X(x)} = t^{-\alpha}, \qquad t > 0. \tag{3}$$

Hence, even though the approach is not parametric we can use the model $\bar{F}_X(x) = x^{-\alpha} L_0(x)$ with $L_0 \in R_0$ for the upper tail of the income distribution.

Assumption A2': *F_X satisfies for some $\alpha > 0$*

$$\bar{F}_X(x) = x^{-\alpha} L_0(x) \tag{4}$$

for some slowly varying function $L_0 \in R_0$.

Thus, the tails are heavy in that they decay like power functions. We do not examine distributions with middle heavy tails which decay exponentially fast, such as the lognormal distribution.

Similar arguments apply to the lower tail of $F_X(x)$ which we assume to be regularly varying at 0 with index β,

$$\lim_{x\to 0} \frac{F_X(tx)}{F_X(x)} = t^{\beta}, \qquad t > 0. \tag{5}$$

If F_X is regularly varying at zero with β then $\overline{F}_{X^{-1}}$ is regularly varying at infinity with $-\beta$:

$$\lim_{x\to\infty} \frac{\overline{F}_{X^{-1}}\left(tx^{-1}\right)}{\overline{F}_{X^{-1}}\left(x^{-1}\right)} = \lim_{x\to 0} \frac{F_X\left(t^{-1}x\right)}{F_X(x)} = \left(t^{-1}\right)^{\beta} = t^{-\beta}$$

since $\overline{F}_{X^{-1}}\left(x^{-1}\right) = F_X(x)$. This relationship allows us to deal with the statistical inference of upper tails only as the same results hold for the lower tails if we consider the reciprocals.

We will base our testing procedure on the following important result linking Lorenz dominance and the parameters of regular variation:

Theorem 1 *Under assumptions A1 and A2,*

$$X \text{ Lorenz dominates } Y \iff \alpha_X \geq \alpha_Y \text{ and } \beta_X \geq \beta_Y. \tag{6}$$

A simple proof of the sufficiency statement of the theorem is given in Kleiber (1999), which we reproduce here for the upper tail only. >From (1), Lorenz dominance of X over Y is equivalent to $g(x) = \int_x^\infty \overline{F}_Y(t)\,dt / \int_x^\infty \overline{F}_X(t)\,dt \geq$

1 for all $x > 0$. By assumption, the tail of $\overline{F}_Y$ is regularly varying with parameter $-\alpha_Y$, and hence its integral regularly varies with $-\alpha_Y+1$. Therefore g regularly varies with $\alpha_X - \alpha_Y$, but $\lim_{x\to\infty} g(x) \geq 1$ iff $\alpha_X \geq \alpha_Y$. The result for the lower tail is established similarly. The necessity statement is obvious considering assumption A1.

2.3 Estimating the Parameter of Regular Variation

The parameter α of regular variation at infinity in (4) can be estimated by Hill's estimator given by[1]

$$\hat{\alpha} = H_{k,n}^{-1} \tag{7}$$

$$H_{k,n} = \frac{1}{k}\sum_{i=1}^{k} \ln\left(X_{(i)}\right) - \ln\left(X_{(k)}\right), \tag{8}$$

where k is the number of extreme observations to be included. This estimator was originally proposed by Hill (1975) as the maximum likelihood estimator of the parameter α of the Pareto distribution model $\bar{F}(x; c, \alpha) = cx^{-\alpha}$ (i.e. the special case in which the slowly varying function in (4) is a constant). However, more general properties of Hill's estimator are well-known. For fixed k, the estimator $H_{k,n}$ converges in distribution to a gamma distribution as $n \to \infty$. It follows from a diagonalisation argument that for any $\bar{F}$ satisfying (4), $\sqrt{k}\left(H_{k,n} - \alpha^{-1}\right)$ converges to a normal distribution with variance α^{-2} provided k tends to infinity sufficiently slowly. Various theorems exist in the literature which make the last statement more precise. We present one of them below.[2]

Theorem 2 *Under assumption A2':*

(a) (Weak consistency) If $k \to \infty$ *,* $k/n \to 0$*, for* $n \to \infty$

$$\hat{\alpha} \xrightarrow{p} \alpha$$

[1] Following the preceding remarks, Hill's estimator is readily adaptable for an estimation of β for the lower tails, being now based on the k smallest observations $X_{(n-k+1)}, \ldots, X_{(n)}$:

$$\hat{\beta} = \left(\frac{1}{k}\sum_{i=1}^{k} \ln\left(X_{(n-i+1)}^{-1}\right) - \ln\left(X_{(n-k+1)}^{-1}\right)\right)^{-1}.$$

[2] See, for instance Embrechts, Klüppelberg and Mikosch (1997, chap. 6.4). Theorem 2.c is due to de Haan and Peng (1999). Another version is given in Haeusler and Teugels (1985). See also Weissman (1978).

(b) (Strong consistency) If $k/n \to 0$, $k/\ln\ln n \to \infty$, as $n \to \infty$

$$\hat{\alpha} \xrightarrow{a.s.} \alpha$$

(c) (Asymptotic normality) Assume $\lim_{x\to\infty} \frac{\bar{F}(tx)/\bar{F}(x)-t^{-\alpha}}{\gamma(x)} = t^{-\alpha}\frac{t^{-\rho}-1}{-\rho}$, $t > 0$ exists where $\gamma(x)$ is a measurable function of constant sign. We refer to this as a "second order condition" with the second order parameter of regular variation $-\rho$. Let $U(t) = F^{-1}(1-t^{-1})$, and $\Gamma(x) = \alpha^{-2}\gamma(U(x))$ and $k \to \infty$ but $k/n \to 0$. If

$$\lim_{n\to\infty} \sqrt{k}\Gamma\left(\frac{n}{k}\right) = \lambda \in \mathbb{R}$$

then, as $n \to \infty$, the estimator $\hat{\alpha}$ is consistent and asymptotically normal with

$$\sqrt{k}\left(\hat{\alpha}-\alpha\right) \xrightarrow{D} N\left(\frac{\alpha^3}{-\rho-\alpha}\lambda, \alpha^2\right). \tag{9}$$

Asymptotic normality is also obtained by Hall (1982) using a different approach. He assumes that the true distribution satisfies

$$\bar{F}(x) = x^{-\alpha}c(1 + dx^{-\rho} + o(x^{-\rho})). \tag{10}$$

asymptotically, an assumption which more stringent than (4). The Fréchet distribution, for instance, can be expanded asymptotically into the above form, i.e. $\bar{F}(x) = cx^{-\alpha}(1 - 0.5cx^{-\alpha} + o(x^{-\alpha}))$. If the distribution can be expanded to $m+1$ terms, so that $\bar{F}(x) = cx^{-\alpha}(1 + d_1x^{-\alpha} + \ldots + d_mx^{-m\alpha} + o(x^{-m\alpha}))$, he shows that if $k \to \infty$ such that $k = o(n^{2m/(2m+1)})$ then $\sqrt{k}\left(\hat{\alpha}-\alpha\right) \to^D N\left(0, \alpha^2\sigma^2\right)$. In particular, if $\bar{F}(x) = cx^{-\alpha}(1+O(x^{-\rho}))$ as $x \to \infty$, if $k \to \infty$ and if $k = o\left(n^{2\rho/(2\rho+\alpha)}\right)$ as $n \to \infty$, then $\sqrt{k}\left(\hat{\alpha}-\alpha\right) \to N\left(0, \alpha^2\right)$.

Hall's result and theorem 2.c can be linked by observing that

$$L(x) = c\left(1 + dx^{-\rho} + O(x^{-2\rho})\right) \tag{11}$$

is a slowly varying function, $L \in R_0$. Moreover,

$$\begin{aligned} \frac{L(tx)}{L(x)} - 1 &= \frac{\left(1 + t^{-\rho}dx^{-\rho} + O(x^{-2\rho})\right)}{(1 + dx^{-\rho} + O(x^{-2\rho}))} - 1 \\ &= \left(t^{-\rho}-1\right)dx^{-\rho} + O\left(x^{-2\rho}\right), \end{aligned} \tag{12}$$

so $-\rho$ in (10) is in fact the second order variation parameter of theorem 2.c, and the required function is

$$\gamma(x) = (-\rho)\,dx^{-\rho}. \tag{13}$$

In order to implement the Hill estimator, it remains to choose k appropriately. For a sample with given size, there is no universal optimal choice, and different methods have been proposed. One method is a Hill's plot: plot the estimate $H_{k,n}^{-1}$ against k and select a value of k for which the plot is (roughly) constant. Embrechts et al. (1997, p. 194) observe that the Hill estimator can perform poorly if the slowly varying function in (4) is far from being a constant. This poor performance manifests itself in a volatile "Hill's horror plot". It is therefore informative in a parametric context to examine whether a given parametric model is close to the Pareto model asymptotically. We examine this point below. If the Hill's plot is too volatile, using a logarithmic scale for k may increase the display space taken up by $H_{k,n}^{-1}$ around the true value α. This alt(ernative)Hill's plot, proposed in Drees, de Haan and Resnick (2000) is thus given by $\{(\varepsilon, H_{[n^\varepsilon],n}^{-1}), 0 \leq \varepsilon \leq 1\}$.

In order to illustrate theorem 2 we discuss some parametric models in Appendix A. The results will also be of use in the simulation study below in which k is chosen by minimising the mean-squared error of the Hill's estimator $\hat{\alpha} = H_{k,n}^{-1}$.

2.4 The Test

The income distribution functions of X and Y are given by F_X and F_Y, respectively. Let $-\alpha_X$ be the index of regular variation of $\bar{F}_X$ (at infinity) and β_X the index of regular variation of F_X at zero, and define $-\alpha_Y$ and β_Y for F_Y similarly.

Extreme upper and lower order statistics are asymptotically independent (David, 1981, p. 267). Let $\overline{k}$ and $\underline{k}$ denote the number of upper and lower extreme observations, respectively, to be included in the estimators. Under conditions of Theorem 2.c with $\overline{k}$ and $\underline{k}$ growing sufficiently slowly so that the bias term λ equals 0, the joint asymptotic distribution of $\hat{\alpha}_X$ and $\hat{\beta}_X$ is

$$\begin{bmatrix} \sqrt{\overline{k}}\,(\hat{\alpha}_X - \alpha_X) \\ \sqrt{\underline{k}}(\hat{\beta}_X - \beta_X) \end{bmatrix} \to N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \alpha_X^2 & 0 \\ 0 & \beta_X^2 \end{bmatrix} \right).$$

Further, if the samples of X and Y are independent, so are $\hat{\alpha}_X$ and $\hat{\alpha}_Y$, and $\hat{\beta}_X$ and $\hat{\beta}_Y$.

To test whether X Lorenz dominates Y throughout (under assumption A1), the appropriate null and alternative hypotheses concerning the tails are

$$\begin{aligned} H_0 &: \text{ the population Lorenz curves cross at the bottom or the top} \\ &= (\alpha_X < \alpha_Y \text{ or } \beta_X < \beta_Y) \\ H_1 &: \text{ not } H_0 \\ &= (\alpha_X \geq \alpha_Y \text{ and } \beta_X \geq \beta_Y) \end{aligned} \tag{14}$$

Rejecting the null hypothesis firmly establishes Lorenz dominance of X over Y because of theorem 1.

Let $\overline{k}_X, \underline{k}_X$ and $\overline{k}_Y, \underline{k}_Y$ denote the number of extreme observations for the estimation of $\hat{\alpha}_X, \hat{\beta}_X$ and $\hat{\alpha}_Y, \hat{\beta}_Y$, respectively. A suitable test is based on the two statistics

$$T_1 = \frac{\hat{\alpha}_X - \hat{\alpha}_Y}{\sqrt{\frac{\hat{\alpha}_X^2}{\overline{k}_X} + \frac{\hat{\alpha}_Y^2}{\overline{k}_Y}}}, \tag{15}$$

$$T_2 = \frac{\hat{\beta}_X - \hat{\beta}_Y}{\sqrt{\frac{\hat{\beta}_X^2}{\underline{k}_X} + \frac{\hat{\beta}_Y^2}{\underline{k}_Y}}}. \tag{16}$$

These test statistics are asymptotically normal, despite the dependence between numerator and denominator, as can be seen from an application of Slutsky's theorem[3].

As T_1 and T_2 are asymptotically independent and standard normally distributed, the null hypothesis is rejected when both T_1 and T_2 are too large. The critical level δ for significance level γ is chosen such that the probability $P(T_1 > \delta \text{ and } T_2 > \delta) \leq \gamma$ under the null hypothesis. Because of

$$\begin{aligned} P(T_1 > \delta \text{ and } T_2 > \delta) &= P(T_1 > \delta) \times P(T_2 > \delta) \\ &= (1 - \Phi(\delta))^2 \end{aligned}$$

the critical value is given by $\delta = \Phi^{-1}\left(1 - \sqrt{\gamma}\right)$ where Φ^{-1} is the quantile function of $N(0,1)$.

If the parameters are on the boundary of H_0 the (true) null hypothesis is rejected (asymptotically) with a probability of γ. If the parameters are

[3] Consider $\tau = (\hat{\alpha}_X - \hat{\alpha}_Y)\left(\alpha_X^2/\overline{k}_X + \alpha_Y^2/\overline{k}_Y\right)^{-0.5}$. It is easily seen that τ has a limiting Gaussian distribution, but it cannot be implemented since α_X and α_Y are, of course, unknown. Define the random sequence $c_{\overline{k}_X,\overline{k}_Y} = \left(\alpha_X^2/\overline{k}_X + \alpha_Y^2/\overline{k}_Y\right)^{0.5} / \left(\widehat{\alpha}_X^2/\overline{k}_X + \widehat{\alpha}_Y^2/\overline{k}_Y\right)^{0.5}$, which has the property that $c_{\overline{k}_X,\overline{k}_Y} \xrightarrow{p} 1$. By Slutsky's theorem, $T_1 = \tau c_{\overline{k}_X,\overline{k}_Y}$ converges to the same (Gaussian) distribution as τ. T_2 is dealt with similarly.

inside H_0 the error probability of the first kind is less than γ. The power of the test depends, of course, on the true parameter values, and we have approximately

$$\begin{aligned} & P\left(H_0 \text{ rejected} | \alpha_X \geq \alpha_Y \text{ and } \beta_X \geq \beta_Y\right) \\ = \; & P\left(T_1 > \delta \text{ and } T_2 > \delta\right) \qquad (17) \\ = \; & \left(1 - \Phi\left(\delta - \frac{\alpha_X - \alpha_Y}{\sqrt{\frac{\alpha_X^2}{k_X} + \frac{\alpha_Y^2}{k_X}}}\right)\right) \\ & \times \left(1 - \Phi\left(\delta - \frac{\beta_X - \beta_Y}{\sqrt{\frac{\beta_X^2}{k_X} + \frac{\beta_Y^2}{k_X}}}\right)\right). \qquad (18) \end{aligned}$$

3 Illustrations

We first present some evidence which reveals that sample Lorenz curves may intersect in the tails, although the population Lorenz curves do not cross: in our experiments 45% of sample Lorenz curves intersect in the tails. This is precisely the situation about which we would like to make statistical inference. Using our test we are able to infer in many cases that despite sample tail crossings, the population Lorenz curves exhibit statistically significant Lorenz dominance. Moreover, our experiments suggest that the empirical level of our test is close to its nominal value.

3.1 The Experiments

We let X and Y have Singh-Maddala distributions defined by

$$f\left(x; a, b, c\right) = \frac{bcx^{b-1}}{a^b \left[1 + \left(x/a\right)^b\right]^{c+1}}.$$

In Appendix A we show that for this function the parameters relevant for the experiment are given by Table 1.

tail:	upper	lower
parameter of regular variation (θ)	$-bc$	$-b$
second order parameter (ρ)	b	b
$\lambda_{k,n}$	$b^{-1}k^{3/2}n^{-1}$	$2^{1/3}n^{2/3}$

Table 1: Regular variation parameters for the Singh-Maddala distribution. For the definition of $\lambda_{k,n}$ see theorem 2.c.

We parameterise the densities as $f_X(.; 100, 2.8, 1.7)$ and $f_Y(.; 100, 2.4, 1.8)$ so that X Lorenz dominates Y. This choice is further motivated by the facts that the Lorenz curves: (i) look similar to curves encountered in empirical applications and (ii) are far apart in the middle of the distribution, in order to satisfy assumption A1. Lorenz dominance follows immediately from the analytical form of the Lorenz curves, given for the Singh-Maddala distribution by

$$p \mapsto IB_{1-(1-p)^{1/c}}\left(\frac{1}{b}+1, c-\frac{1}{b}\right)$$

where $IB(\cdot,\cdot)$ is the incomplete Beta function (Schader and Schmid, 1988). The parameter choice is, of course, also consistent with Lorenz dominance in the tails as a comparison of the parameters of regular variation reveals.[4] The population Gini coefficients[5] are $Gini_X = 0.2887$ and $Gini_Y = 0.3275$.

X Lorenz dominates Y, but the corresponding empirical Lorenz curves $\hat{L}_X$ and $\hat{L}_Y$ may, of course, intersect. Our Monte Carlo simulation is based on $100,000$ replications of empirical Lorenz curves estimated from samples of size $5,000$. The result, reported in Table 2, is that the proportion of non-intersecting $\hat{L}_X$ and $\hat{L}_Y$ is as low as 0.55: In 45 percent of the cases the empirical curves cross even though the theoretical curves do not. The reason for this unsatisfactory state of affairs is the large number of intersections in the tails as shown in the table.

intersecting sample Lorenz curves	45 %
intersections in lower tail	22 %
intersections in upper tail	30 %

Table 2: Results of the Monte Carlo simulations: intersections of sample Lorenz curves when the population Lorenz curve of X dominates that of Y. The lower and upper tail regions are defined by the 0.05- and 0.95-quantiles.

We proceed to examine the performance of our test. For each replication we estimate the parameters α_X, α_Y, β_X and β_Y using the Hills's estimator. This requires choosing the number of extreme observations to be considered. Since it is impractical to evaluate a Hill's plot for each iteration of the simulation, we let $\overline{k}$ and $\underline{k}$ minimise the mean-squared error of the Hill's estimator in the parametric model on which the simulation is based.

[4] Using (6) since, as regards the upper tail, $\alpha_X = 2.8 \times 1.7 > \alpha_Y = 2.4 \times 1.8$ and, for the lower tail, $\beta_X = 2.8 > \beta_Y = 2.4$.

[5] Recall that the Gini coefficient is defined by $Gini = 1 - 2\int_0^1 L(p)\,dp$.

Clearly, this simplification is not possible in empirical applications where the population values are unknown. From theorem 2.c it is immediate that the mean-squared error of the upper tail parameter estimate is

$$MSE_{\hat{\alpha}} = \frac{1}{k}\left(\alpha^2 + \frac{\alpha^6 \lambda_{k,n}^2}{(-\rho-\alpha)^2}\right) \quad \text{and } \overline{k} = \arg\min_k MSE_{\hat{\alpha}}$$

where $-\rho$ is the second order regular variation parameter and $\lambda = \lambda_{k,n}$ is defined in the theorem. For the lower tail, let $\underline{k}$ minimise $MSE_{\widehat{\beta}}$. For the Singh-Maddala distributions we obtain, using Table 1, $\overline{k}$ and $\underline{k}$ reported in Table 3.

	$\overline{k}$	$\underline{k}$
$f_X(.;100,2.8,1.7)$	142	369
$f_Y(.;100,2.4,1.8)$	128	369

Table 3: Number of extreme observations included in the Hill's estimator. The sample size is $n = 5,000$.

From the estimates $\hat{\alpha}_X$, $\hat{\alpha}_Y$, $\hat{\beta}_X$, $\hat{\beta}_Y$ and from $\overline{k}_X$, $\underline{k}_X$, $\overline{k}_Y$ and $\underline{k}_Y$ we compute the statistics (15) and (16) for our test of the null hypothesis

H_0 : the population Lorenz curves cross at the bottom or the top.

The result is that in 61.5% of cases we reject H_0 at a significance level of $\gamma = 0.1$ and firmly conclude Lorenz dominance.[6]

We carried out further Monte-Carlo simulations to determine the empirical significance level of the test procedure, as power properties should not be assessed unless the nominal significance level is kept everywhere on H_0 (in particular at the border of H_0 where both sample are drawn from the same distribution). Drawing 100,000 samples of size 5,000 each from two Singh-Madalla distributions with equal densities $f_X(.;100,2.8,1.7) = f_Y(.;100,2.8,1.7)$ we find that the empirical significance level for $\gamma = 0.1$ is 0.09736.[7] Hence our test is slightly conservative and the results concerning the power are meaningful.

3.2 Empirical Examples

We now proceed to illustrate the merits of our test procedure with real world data taken from the Luxembourg Income Study (LIS)[8], which has

[6] Using (18) the (asymptotic) power equals 0.591.

[7] Similarly, using the second set of parameters for the Singh-Madalla distribution, i.e., $f_X(.;100,2.4,1.8) = f_Y(.;100,2.4,1.8)$, the empirical level is 0.09807.

[8] See http://www.lis.ceps.lu/summary.htm for a detailed description.

been used for many inequality analyses (see e.g. Atkinson, Rainwater and Smeeding, 1994). This database provides comprehensive and comparable information about household composition and income for many countries. The LIS definition of disposable income includes earnings, other factor income, means and non means tested social insurance transfers and public and private pension transfers; mandatory social insurance contributions and income tax are substracted.

We investigate Lorenz dominance relations of four major economies in 1994: the United States, Canada, Italy and, at the other end of the inequality spectrum, Germany. The left hand side of Table 4 reports summary statistics of the income distributions (number of observations, coefficient of variation and the Gini coefficient).

Country	No. obs.	CV	Gini	$\overline{k}$	$\underline{k}$	$\hat{\alpha}$	$\hat{\beta}$
Canada (CN)	100180	0.5558	0.2833	317	563	4.27	2.13
Germany (GE)	15076	0.5373	0.2456	520	199	4.68	2.16
Italy (IT)	23681	0.7788	0.3419	698	154	2.86	1.64
USA (US)	162122	0.7270	0.3619	122	1337	4.00	1.23

Table 4: Summary statistics of income distributions, number of extreme values, and Hill estimates.

We proceed to investigate the inequality orderings of these countries based on their Lorenz curves. Before applying our test, we first verify whether our test is applicable, i.e. we verify whether Lorenz dominance occurs in the main body of the distributions (assumption A1). To this end, we consider the sample Lorenz curves at deciles. Table 5 reveals that in all cases but one Lorenz dominance appears to prevail. Our test is then appropriate for all pairs except Italy and the USA. Table 5 also makes clear that the tails cannot be ignored: apart from the pair Canada-Italy all other pairs have intersecting sample Lorenz curves, in all cases but one the intersection occurs in the tails. For instance, Canada does not appear to Lorenz dominate the USA, even despite the large difference between the Gini coefficients.

In order to test whether the tail crossings are statistically significant we apply our test procedure. The number of extreme observations to be included into the estimators are determined by investigating the alternative Hill's plots (see Appendix C). The right part of Table 4 gives the numbers of upper and lower extremes ($\overline{k}$ and $\underline{k}$) as well as the resulting Hill estimates of the index of regular variation for the upper tail ($\hat{\alpha}$) and for the lower tail ($\hat{\beta}$).

Table 6 states the null hypotheses that the Lorenz curves cross, given

	at deciles			entire curve		
	Canada	Germany	Italy	Canada	Germany	Italy
Germany	$>$			x		
Italy	$<$	$<$		$<$	x	
USA	$<$	$<$	x	x	x	x

Table 5: Lorenz dominance. Note: “$<$” means that the row country is dominated by the column country, “$>$” the reverse, “x” indicates crossing.

that there is Lorenz dominance in the main body of the distribution (hence Italy-US is disregarded). Further, we provide the values of the test statistics T_1 and T_2, the test results at 10% significance level are reported in the right-most column.

Pair	H_0 (Lorenz curves cross)	T_1	T_2	Test result
CN-GE	$\alpha_{CN} > \alpha_{GE}$ or $\beta_{CN} > \beta_{GE}$	-1.2789	-0.1478	do not reject
CN-IT	$\alpha_{CN} < \alpha_{IT}$ or $\beta_{CN} < \beta_{IT}$	5.3842	3.0752	reject
GE-IT	$\alpha_{GE} < \alpha_{IT}$ or $\beta_{GE} < \beta_{IT}$	7.8545	2.5608	reject
CN-US	$\alpha_{CN} < \alpha_{US}$ or $\beta_{CN} < \beta_{US}$	0.6280	9.3956	reject
GE-US	$\alpha_{GE} < \alpha_{US}$ or $\beta_{GE} < \beta_{US}$	1.6256	5.9223	reject

Table 6: Null hypotheses, test statistics, and test results

We conclude that there is strong statistical evidence that Canada Lorenz dominates Italy, Germany Lorenz dominates Italy, and that Canada Lorenz dominates the USA, even if the tails are taken into account. This demonstrates the good performance of the proposed test: in many cases we are able to infer that despite sample tail crossings the population Lorenz curves do, in fact, exhibit Lorenz dominance.

4 Conclusion

The appeal of the Lorenz dominance criterion is undermined by the practical problem that many sample Lorenz curves intersect in the tails. Our experiments also suggest that sample tail intersections may easily occur for population Lorenz curves which are “far apart”. The usual inferential methods, based on central limit theorem arguments, do not apply to these tails since they contain too few observations. By contrast, we have proposed a test procedure, based on a domain of attraction assumption, which fully takes into account the tail behaviour of Lorenz curves. Our experiments and empirical examples demonstrate the good performance of the proposed test: in many cases are we able to infer that despite sample tail crossings the population Lorenz curves do, in fact, exhibit Lorenz dominance. Moreover, our experiments suggest that the empirical level of our test is close to its nominal value.

References

Atkinson, A. (1970). On the measurement of inequality, *Journal of Economic Theory* **2**: 244–263.

Atkinson, A., Rainwater, L. and Smeeding, T. (1994). *Income Distribution in OECD Countries: The Evidence From Luxembourg Income Study*, OECD, Paris.

Beach, C. and Davidson, R. (1983). Distribution-free statistical inference with Lorenz curves and income shares, *Review of Economic Studies* **50**: 723–735.

Beach, C. and Richmond, J. (1985). Joint confidence intervals for income shares and Lorenz curves, *International Economic Review* **26**: 439–450.

Brachmann, K., Stich, A. and Trede, M. (1996). Evaluating parametric income distribution models, *Allgemeines Statistisches Archiv* **80**: 285–298.

David, H. (1981). *Order Statistics*, 2 edn, Wiley, New York.

Davidson, R. and Duclos, J.-Y. (1997). Statistical inference for the measurement of the incidence of taxes and transfers., *Econometrica* **65**: 1453–1465.

de Haan, L. and Peng, L. (1999). Exact rates of convergence to a stable law, *Journal of the London Mathematical Society* **59**: 1134–1152.

Drees, H., de Haan, L. and Resnick, S. (2000). How to make a Hill plot, *Annals of Statistics* **28**: 254–274.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events*, Springer, Berlin.

Haeusler, E. and Teugels, J. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation, *Annals of Statistics* **13**: 743–756.

Hall, P. (1982). On some simple estimates of an exponent of regular variation, *Journal of the Royal Statistical Society B* **44**: 37–42.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics* **3**: 1163–1174.

Kleiber, C. (1999). Tail dominance of income distributions, working paper.

McDonald, J. (1984). Some generalized functions for the size distribution of income, *Econometrica* **52**: 647–663.

McDonald, J. and Xu, Y. (1995). A generalization of the beta distribution with applications, *Journal of Econometrics* **66**: 133–152.

Savin, N. (1993). Multiple hypothesis testing, *in* Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, North-Holland, Amsterdam, chapter 14.

Schader, M. and Schmid, F. (1988). Zur Messung der relativen Konzentration aus gruppierten Daten, *Jahrbücher für Nationalökonomie und Statistik* **204**: 437–455.

Singh, S. and Maddala, G. (1976). A function for size distribution of incomes, *Econometrica* **44**: 963–970.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association* **73**: 812–815.

A Tail Behaviour of some Parametric Models

In this appendix we discuss the tail behaviour of the generalised beta distribution (which nests many important cases such as the Singh-Maddala distribution and the Dagum distribution) in order to illustrate the application of Theorem 2. These distributions have regularly varying tails and fit real world income data reasonably well (Brachmann et al., 1996).

The generalised beta distribution, proposed in McDonald (1984), has density

$$f(x;a,b,c,d) = \frac{bx^{bd-1}}{a^{bd}B(d,c)\left[1+(x/a)^b\right]^{d+c}} \tag{19}$$

where $B(\cdot,\cdot)$ denotes the Beta function, and nests various distributions as special cases. For instance, if $d=1$ then (19) reduces to the Singh-Maddala distribution, which captures many actual income distributions, both as regards both tails and the main body (Singh and Maddala, 1976). Its tail is given explicitly by $\bar{F}(x;a,b,c) = (1+(x/a)^b)^{-c}$.[9] Another example is the Dagum distribution (for $c=1$).

To obtain an approximation to the upper tail of the distribution function, divide (19) by $\left[(x/a)^b\right]^{d+c}$, expand $\left[(x/a)^b\right]^{d+c} / \left[1+(x/a)^b\right]^{d+c}$ to second order and integrate:

$$\bar{F}(x;a,b,c,d) = g_1 x^{-bc}\left(1+g_2x^{-b}+O\left(x^{-2b}\right)\right) \tag{20}$$

for some constants g_i. Thus, the upper tail is regularly varying with parameter $-\alpha=-bc$, and using (11) and (12), the second order parameter is $-\rho=-b$.[10] Equation (20) is also in a form which permits direct application of Hall's result, so that k of the Hill's estimator $\hat{\alpha}=H_{k,n}^{-1}$ must satisfy $o\left(n^{2/(2+c)}\right)$ to ensure unbiasedness. To apply theorem 2.c directly, it follows from (13) that $\gamma(x) = (-b)\,g_2x^{-b}$. In order to derive $U(.)$, just consider the first order term in (20) and invert to get $U(x) \propto x^{1/bc}$. Hence, $k^{0.5}\Gamma\left(\frac{n}{k}\right) \propto k^{0.5+1/c}n^{-1/c}$, so to obtain no bias we require $k=o\left(n^{2/(2+c)}\right)$.

[9] Note that this distribution is of the Pareto type for large x since $\bar{F}(x) = a^{bc}x^{-bc} + O(x^{-b(1+c)})$. Thus x needs to be large to avoid Hill's horror plots for the upper tail estimation. As regards the lower tail, we observe that $F(x) = x^b\left(ca^{-b}+O\left(x^b\right)\right)$. Hence a good result for the lower tail estimation is to be expected.

[10] Note that the first order result could also have been obtained directly from (19) using the lemmas in Embrechts et al. (1997, p. 564) by observing that its numerator is regularly varying at infinity with parameter $bd-1$, the denominator with $bd+bc$, so the ratio regularly varies with $-bc-1$, and the tail of the distribution function with $-bc$.

As regards the lower tail, the usual expansion yields

$$F(x,a,b,c,d) = g_3 x^{bd}\left(1 + g_4 x^b + O\left(x^{2b}\right)\right)$$

for some constants g_i. Hence the lower tail varies with parameter $\beta = bd$, and the second order parameter is b. Direct application of Hall's result shows that k of the Hill's estimator $\widehat{\beta} = H_{k,n}^{-1}$ must satisfy $o\left(n^{2/(2+d)}\right)$ to ensure unbiasedness. The "second order condition" can be verified in a similar fashion.

B The Multiple Hypotheses Test

In this appendix we consider how a test of assumption A1 can be taken into account for a test of whether X Lorenz dominates Y. This then constitutes a multiple hypotheses testing problem (see e.g. Savin, 1993) since Lorenz dominance must occur in both (i) the main body of the distribution and (ii) the tails. The implementation of the test – either as a union-intersection or an intersection-union test – depends on the various ways in which the individual hypotheses are formulated.

If the overall null hypothesis is in line with (14) the two individual null hypotheses are H_0^1 : "Y dominates X or a crossing occurs in the main body of the Lorenz curve" and H_0^2 : "Y dominates X or a crossing occurs in the tails of the Lorenz curve" with the alternatives H_1^1 : "X dominates Y in the main body" and H_1^2 : "X dominates Y in the tails". The overall null hypothesis is rejected if both H_0^1 and H_0^2 are rejected. Determining the overall significance level γ of this union-intersection test is facilitated by the fact that the individual test statistics are asymptotically independent. Hence, setting both significance individual levels to $\sqrt{\gamma}$ results in an overall level of γ.

If the overall null hypothesis is reversed, the individual hypotheses are H_0^1 : "X dominates Y in the main body" and H_0^2 : "X dominates Y in the tails" with the alternatives H_1^1 : "Y dominates X or a crossing occurs in the main body" and H_1^2 : "Y dominates X or a crossing occurs in the tails". This constitutes an intersection-union test, and the overall null hypothesis is rejected if H_0^1 or H_0^2 is rejected. The individual significance levels need to be set to $1 - \sqrt{1-\gamma}$ to keep an overall level of γ.

Since our aim is to firmly establish Lorenz dominance of X over Y the appropriate test is the union-intersection test.

C Alternative Hill's Plots

The alternative Hill's plots for Canada (CN), Germany (GE), Italy (IT), and the USA (US) are shown in Figures 1 for the upper tail and 2 for the lower tail.

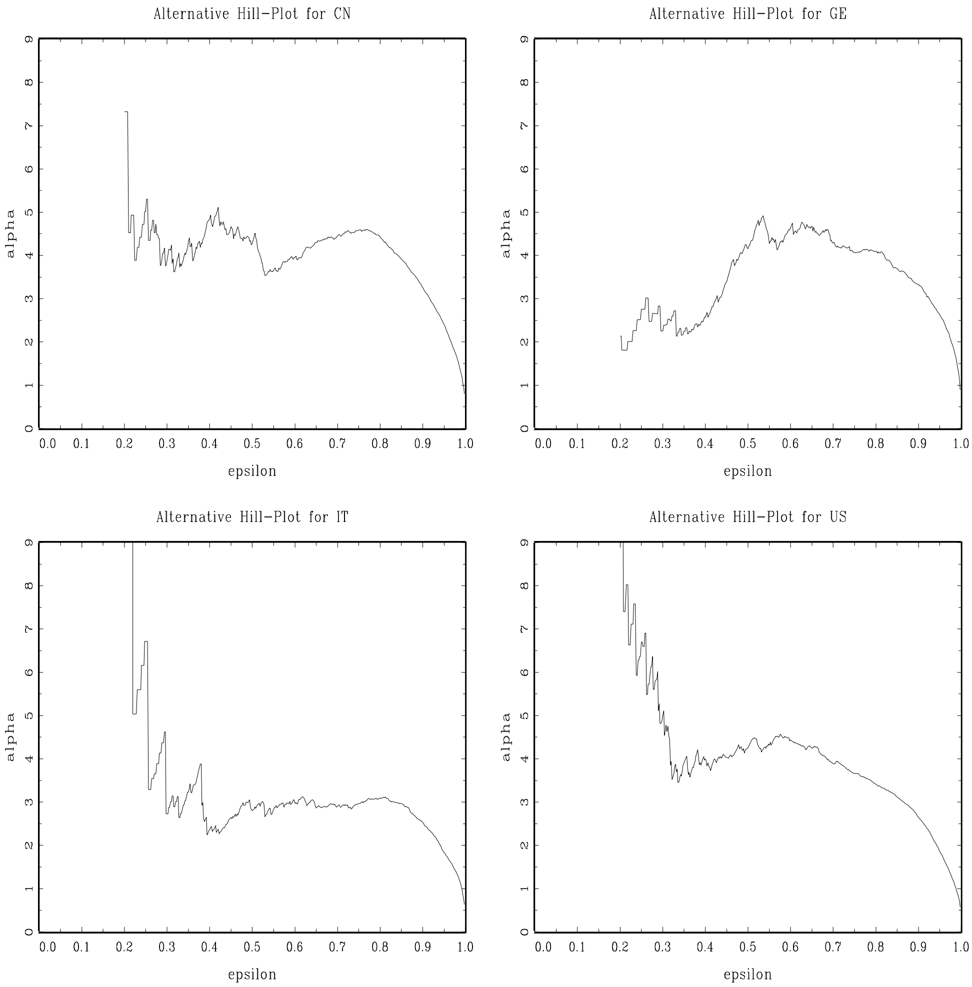

Figure 1: Alternative Hill plots for the upper tail parameter (α)

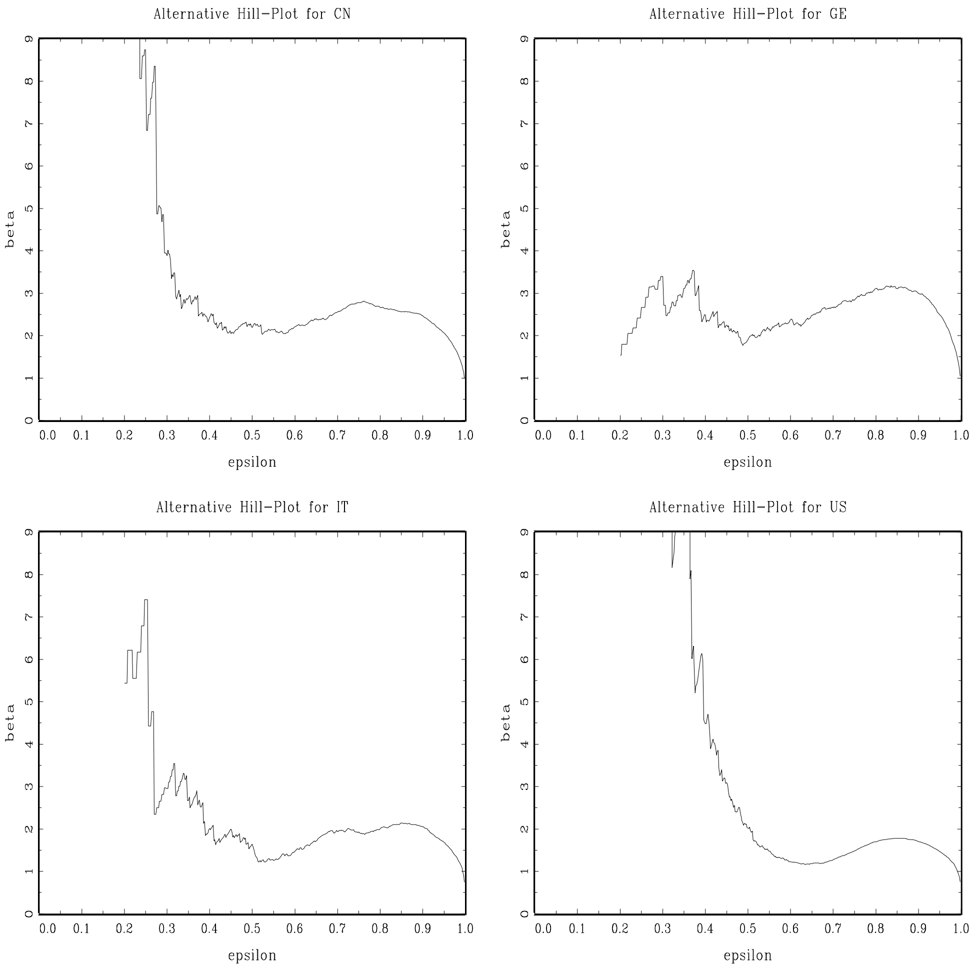

Figure 2: Alternative Hill plots for the lower tail parameter (β)