Available online at www.sciencedirect.com
SCIENCE DIRECT
JOURNAL OF Econometrics
Journal of Econometrics 133 (2006) 191–205
www.elsevier.com/locate/jeconom

Testing for stochastic dominance using the weighted McFadden-type statistic

Lajos Horváth[a], Piotr Kokoszka[b], Ričardas Zitikis[c,*]

[a]*Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, UH 84121-6026, USA*
[b]*Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UH 84322-3900, USA*
[c]*Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ont., Canada N6A 5B7*

Accepted 21 March 2005
Available online 23 May 2005

Abstract

We demonstrate that when testing for stochastic dominance of order three and above, using a *weighted* version of the Kolmogorov–Smirnov-type statistic proposed by McFadden [1989. In: Fomby, T.B., Seo, T.K. (Eds.), Studies in the Economics of Uncertainty. Springer, New York, pp. 113–134] is necessary for obtaining a non-degenerate asymptotic distribution. Since the asymptotic distribution is complex, we discuss a bootstrap approximation for it in the context of a real application.

JEL classification: C12; D31; D63

Keywords: Stochastic dominance; Kolmogorov–Smirnov-type statistic; Asymptotic distribution; Brownian bridge; Bootstrap

*Corresponding author. Tel.: +1 519 661 2111x86983; fax: +1 519 661 3813.
E-mail address: zitikis@stats.uwo.ca (R. Zitikis).

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2005.03.013

1. Introduction

Stochastic dominance has been a prominent research topic in the econometric and actuarial literature for a number of decades. Many theoretical and applied contributions in these and related areas have been summarized by Mosler and Scarsini (1991), Levy (1992), Shaked and Shanthikumar (1994), Davidson and Duclos (2000), Müller and Stoyan (2002), Barrett and Donald (2003), Chakravarty and Muliere (2003), Linton et al. (2003).

McFadden (1989) proposed a Kolmogorov–Smirnov-type test for stochastic dominance. In the case of first-order stochastic dominance, McFadden (1989) derived the sampling distribution of the test statistic. In the case of second-order stochastic dominance, McFadden (1989) obtained a series of results that partially characterized the asymptotic distribution of the statistic. McFadden (1989) also proved bounds for the critical values of the test. Statistical inference for stochastic dominance at higher orders were subsequently considered by a number of authors including Anderson (1996), Davidson and Duclos (2000), Barrett and Donald (2003), Linton et al. (2003). In particular, Anderson (1996) based his test on making comparisons between distribution functions and their integrals at a finite number of fixed points. This approach reduced the original problem to an easier one involving finite dimensional distributions. Using this procedure, however, one inevitably loses the asymptotic consistency of the test statistic. Davidson and Duclos (2000) explicitly noted this drawback and made further progress in the area. Employing the theory of empirical processes, Barrett and Donald (2003) departed from the earlier used method of finite comparisons and obtained desired asymptotic results for the Kolmogorov–Smirnov-type statistic for stochastic dominance over any *compact* interval. Whether or not the compactness assumption could be removed remained an open problem. We note in this regard that most of the distributions that are used for modelling data in econometrics and actuarial science do not have compact supports. Hence, resolving the aforementioned problem is of interest and importance.

The main goal of this paper is to show that the compactness assumption cannot be removed without modifying the test statistic. Specifically, we show that, depending on the order of stochastic dominance, considering *weighted* versions of the aforementioned Kolmogorov–Smirnov-type statistic is necessary for obtaining non-degenerate limit distributions.

2. Stochastic dominance: definition and notation

Let X_A and X_B be two real-valued random variables with distribution functions F_A and F_B, respectively. For $C \in \{A, B\}$ and $x \in \mathbf{R}$, define the iterative process by first setting $D_C(1; x) := F_C(x)$ and then

$$D_C(s; x) = \int_{-\infty}^{x} D_C(s-1; y)\,\mathrm{d}y, \quad s = 2, 3, \dots .$$

The integral $D_C(s; x)$ is finite if $\mathbf{E}((X_C)_-^{s-1}) < \infty$, where $x_- := \max(-x, 0)$. We assume the latter moment assumption throughout the paper. Now, by definition, the distribution function F_A stochastically dominates F_B at the order $s \in \mathbf{N}$ if

$$D_A(s;x) \leqslant D_B(s;x) \quad \text{for all } x \in \mathbf{R}. \tag{1}$$

If we replace "for all $x \in \mathbf{R}$" by "for all $x \leqslant z$" in (1), then we have the notion of stochastic dominance at the order s and up to and including the "poverty line" z, which is a real number, usually a non-negative one.

The problem concerning testing for first-order stochastic dominance was thoroughly discussed by McFadden (1989), and Schmid and Trede (1996a). Therefore, from now on we restrict ourselves to the case $s \geqslant 2$ only. Under this assumption we check that statement (1) is equivalent to

$$\Delta_A(s;x) \leqslant \Delta_B(s;x) \quad \text{for all } x \in \mathbf{R}, \tag{2}$$

where

$$\Delta_C(s;x) := \int_{-\infty}^{x} (x-y)^{s-2} F_C(y)\,\mathrm{d}y, \quad C \in \{A,B\}.$$

We shall use this formulation of stochastic dominance in our considerations below.

It is worth noting that the formulation of stochastic dominance presented in the previous paragraph naturally connects the econometric notion of stochastic dominance with the actuarial one. Namely, it is said that "risk" X_A (which is a non-negative random variable with c.d.f. F_A) is smaller than "risk" X_B (with c.d.f. F_B) in the sense of stochastic dominance at the order $s \geqslant 2$ if

$$\pi_A(s;x) \leqslant \pi_B(s;x) \quad \text{for all } x \in [0,\infty), \tag{3}$$

where

$$\pi_C(s;x) := \int_{x}^{\infty} (y-x)^{s-2} (1 - F_C(y))\,\mathrm{d}y, \quad C \in \{A,B\}.$$

The integral $\pi_C(s;x)$ is finite if $\mathbf{E}((X_C)_+^{s-1}) < \infty$, where $x_+ := \max(x,0)$. We note in passing that when $s=2$, then $\pi_C(s;x)$ is known in the actuarial literature as the stop-loss transform of the risk X_C (cf., e.g., Müller, 1996, and references therein). With G_C denoting the (left-continuous) distribution function of $-X_C$, we have that statement (3) holds if and only if

$$\Delta_{G_A}(s;x) \leqslant \Delta_{G_B}(s;x) \quad \text{for all } x \in (-\infty, 0). \tag{4}$$

We now see that statement (2) covers statement (4), just like it covers the usual in econometrics definition of stochastic dominance: $\Delta_{F_A}(s;x) \leqslant \Delta_{F_B}(s;x)$ for all $x \in [0,\infty)$.

3. Weighted Kolmogorov–Smirnov-type statistic

Let $X_{A,1}, \ldots, X_{A,n(A)}$ be independent and identically distributed random variables, each with distribution function F_A, and let $X_{B,1}, \ldots, X_{B,n(B)}$ be independent and identically distributed random variables, each with distribution function F_B. Let $F_{A,n(A)}$ and $F_{B,n(B)}$ be the empirical distribution functions corresponding to the two samples. Define

$$T_n(s;q) = \sup_{x\in\mathbf{R}} q(x)V_n(s;x),$$

where $q:\mathbf{R}\to[0,\infty)$ is a weight function and $\{V_n(s;x), x\in\mathbf{R}\}$ is a stochastic process defined by the formula

$$V_n(s;x) := \sqrt{\frac{n(A)n(B)}{n(A)+n(B)}}(\Delta_{A,n(A)}(s;x) - \Delta_{B,n(B)}(s;x))$$

with the notation

$$\Delta_{C,n(C)}(s;x) := \int_{-\infty}^{x} (x-y)^{s-2} F_{C,n(C)}(y)\,\mathrm{d}y, \quad C\in\{A,B\}.$$

In our following considerations we shall frequently use the indicator function $I_A(x)$ with various sets $A\subseteq\mathbf{R}$, meaning that $I_A(x)$ equals 1 when $x\in A$ and 0 otherwise.

When $s=2$ and $q(x)=I_{\mathbf{R}}(x)$, then $T_n(s;q)$ is the Kolmogorov–Smirnov-type statistic proposed by McFadden (1989) for testing second-order stochastic dominance. In this case, asymptotic properties of the statistic were investigated by Schmid and Trede (1996b, 1998) assuming that the underlying distributions have compact supports. Under the same assumption but for arbitrary $s\geqslant 2$, asymptotic properties of the statistic were thoroughly investigated by Barrett and Donald (2003).

The aforementioned results on compact intervals can also be interpreted in the following way. Namely, we assume no compactness of the supports but, instead, use the weight function $q(x)=I_{(-\infty,z]}(x)$. This leads to the situation investigated by Barrett and Donald (2003). The following natural question arises: can we use $q(x)=I_{\mathbf{R}}(x)$ instead of the aforementioned $q(x)=I_{(-\infty,z]}(x)$? The answer to the question is negative when $s>2$, because in this case the statistic $T_n(s;q)$ is infinite. This can be shown as follows. Let both $n(A)$ and $n(B)$ be equal to 1. Write the equalities:

$$\begin{aligned}\Delta_{A,1}(s;x) - \Delta_{B,1}(s;x) &= \int_{-\infty}^{x} (x-y)^{s-2}(I_{[X_{A,1},\infty)}(y) - I_{[X_{B,1},\infty)}(y))\,\mathrm{d}y \\ &= \frac{1}{s-1}((x-X_{A,1})_+^{s-1} - (x-X_{B,1})_+^{s-1}). \end{aligned} \tag{5}$$

The supremum over all $x\in\mathbf{R}$ of the right-hand side of Eq. (5) is infinite for any $s>2$, provided that, of course, $X_{A,1}\neq X_{B,1}$. This shows that when constructing a useful Kolmogorov–Smirnov-type statistic for testing stochastic dominance at the order $s>2$, we have to use a weight function q such that $q(x)\to 0$ when $x\to\infty$. In fact, we show below that any weight function q satisfying the condition

$$\sup_{x\in\mathbf{R}} q(x)(1+x_+)^{s-2} < \infty \tag{6}$$

makes $T_n(s;q)$ well defined. Moreover, we prove that assumption (6) is optimal in the context of the present paper.

We conclude this section with the note that other than Kolmogorov–Smirnov-type statistics have been proposed in the literature for testing stochastic dominance. For example, Deshpande and Singh (1985) proposed an integral-type test for second-order stochastic dominance. They assumed that one of the distribution functions F_A and F_B is known. The assumption was later removed by Eubank et al. (1993) by suggesting another integral-type statistic. Further developments followed by Kaur et al. (1994).

4. Asymptotic distribution of $T_n(s;q)$

Suppose that we want to test the following hypotheses (cf., e.g., Davidson and Duclos, 2000; Barrett and Donald, 2003; Linton et al., 2003):

$$\mathrm{H}_0 : \Delta_A(s;x) \leqslant \Delta_B(s;x) \quad \text{for all } x \in \mathbf{R}$$

vs.

$$\mathrm{H}_1 : \Delta_A(s;x) > \Delta_B(s;x) \quad \text{for some } x \in \mathbf{R}.$$

Under the alternative hypothesis H_1 the statistic $T_n(s;q)$ converges to ∞, unless the weight function $q(x)$ equals 0 for all those x for which $\Delta_A(s;x) > \Delta_B(s;x)$ is true. Indeed, since $\sqrt{n(A)}\,(\Delta_{A,n(A)}(s;x) - \Delta_A(s;x))$ and $\sqrt{n(B)}\,(\Delta_{B,n(B)}(s;x) - \Delta_B(s;x))$ have centered Gaussian limit distributions (cf., e.g., Barrett and Donald, 2003), we have that for any x such that $\Delta_A(s;x) > \Delta_B(s;x)$ the quantity $V_n(s;x)$ and thus $T_n(s;q)$ must converge in probability to ∞ when $n \to \infty$. Hence, under the null hypothesis we want to obtain the limiting distribution of $T_n(s;q)$ so that the asymptotic critical values z_α (not to be mixed up with the standard normal critical values) would be possible to calculate.

Since the null hypothesis H_0 is composite, we expect that the 'boundary' case $\Delta_A(s;x) = \Delta_B(s;x)$ plays a major role. Indeed, let x be such that $\Delta_A(s;x) < \Delta_B(s;x)$, which is a part of the null hypothesis H_0. In this case, $V_n(s;x)$ converges in probability to $-\infty$, and so the value of $V_n(s;x)$ with the x does not (asymptotically) contribute to the value of $T_n(s;q)$. To put the above note into the mathematical framework, we have the bound

$$T_n(s;q) \leqslant \sup_{x \in \mathbf{R}} q(x) \left(V_n(s;x) - \sqrt{\frac{n(A)n(B)}{n(A)+n(B)}} (\Delta_A(s;x) - \Delta_B(s;x)) \right) \tag{7}$$

under the null hypothesis H_0; with the equality holding in the above discussed 'boundary' case $\Delta_A(s;x) = \Delta_B(s;x)$ for all $x \in \mathbf{R}$, which is equivalent to $F_A(x) = F_B(x)$ for all $x \in \mathbf{R}$. The following theorem gives assumptions under which the right-hand side of (7) has a non-degenerate asymptotic distribution.

Theorem 1. *Let $s \geqslant 2$. Assume that the weight function $q : \mathbf{R} \to [0,\infty)$ is such that* (6) *holds, and let the distribution functions F_A and F_B satisfy the condition*

$$\int_{\mathbf{R}} (1 + y_-^{s-2}) \sqrt{F_C(y)(1 - F_C(y))}\, \mathrm{d}y < \infty, \quad C \in \{A, B\}. \tag{8}$$

Furthermore, assume that $n(B)/(n(A)+n(B)) \to \eta$ *for a constant* $\eta \in (0,1)$ *when both* $n(A)$ *and* $n(B)$ *converge to infinity. Then the right-hand side of* (7) *converges in distribution to the random variable*

$$T(s;q) := \sup_{x \in \mathbf{R}} q(x)\Gamma(s;x),$$

where $x \mapsto \Gamma(s;x)$ *is a Gaussian stochastic process defined by the formula*

$$\Gamma(s;x) := \int_{-\infty}^{x} (x-y)^{s-2}(\sqrt{\eta}\mathscr{B}_1(F_A(y)) + \sqrt{1-\eta}\mathscr{B}_2(F_B(y)))\,\mathrm{d}y \tag{9}$$

with two independent standard Brownian bridges $\mathscr{B}_1$ *and* $\mathscr{B}_2$ *on* $[0,1]$.

We shall now discuss the optimality of conditions in Theorem 1. For this purpose we restrict ourselves to the simpler case when both F_A and F_B are equal. Denote either of the two distribution functions by F for notational simplicity. We first note that condition (6) cannot be relaxed. Indeed, assuming that

$$\int_{0}^{\infty} \sqrt{1-F(y)}\,\mathrm{d}y < \infty, \tag{10}$$

we have that

$$x^{-(s-2)} \int_{0}^{x} (x-y)^{s-2}\mathscr{B}(F(y))\,\mathrm{d}y \to_{\text{a.s.}} \int_{0}^{\infty} \mathscr{B}(F(y))\,\mathrm{d}y, \quad x \to \infty. \tag{11}$$

The proof of this statement is given at the beginning of Section 6 below. Now a comment on assumption (10) follows. With X denoting a random variable with the distribution function F, we have that, for any $\delta > 0$, the moment condition $\mathbf{E}(X_+^{2+\delta}) < \infty$ implies assumption (10). In turn, assumption (10) implies $\mathbf{E}(X_+^2) < \infty$. Finally we note that condition (8) is satisfied under assumption (10) and

$$\int_{-\infty}^{0} y_-^{s-2}\sqrt{F(y)}\,\mathrm{d}x < \infty. \tag{12}$$

Assumption (12) follows if we have $\mathbf{E}(X_-^{s-1+\delta}) < \infty$ for some $\delta > 0$.

5. Bootstrap approximation and an illustrative application

Practical implementation of Theorem 1 requires an approximation to the distribution function $H(z) := \mathbf{P}\{T(s;q) \leqslant z\}$, or at least to its upper tail. We note in passing that the distribution function H is continuous except (cf. Tsirel'son, 1975) in such pathological cases as $q(x) = 0$ for all $x \in \mathbf{R}$, that is, when $T(s;q) = 0$. For further details and comments on this and other related issues we refer to Csörgő and Mason (1989), and references therein. The distribution of the (non-degenerate) random variable $T(s;q)$ can be approximated using bootstrap. For details on the latter topic we refer to the papers by Barrett and Donald (2003), Linton et al. (2003), and references therein.

In what follows we shall discuss a bootstrap approximation in the context of a real example that concerns pre-tax household income in the US, as discussed in the paper by Bandourian et al. (2002). We note at the outset that this problem suggests testing the following hypotheses:

$$\mathrm{H}_0^{\mathrm{eq}} : F_A(x) = F_B(x) \quad \text{for all } x \in \mathbf{R}$$

vs.

$$\mathrm{H}_1 : \Delta_A(s; x) > \Delta_B(s; x) \quad \text{for some } x \in \mathbf{R}.$$

That is, we want to test if the distribution of income during year "B" changed if compared to year "A" so that $\Delta_A(s; x) > \Delta_B(s; x)$ holds for at least one $x \in \mathbf{R}$. We note that even though the null hypothesis $\mathrm{H}_0^{\mathrm{eq}}$ is only a part of the earlier considered H_0, we can use the result of Theorem 1. For this, we first note that under the null hypothesis $\mathrm{H}_0^{\mathrm{eq}}$ the distribution of the process $\sqrt{\eta}\mathscr{B}_1(F_A(\cdot)) + \sqrt{1-\eta}\mathscr{B}_2(F_B(\cdot))$ coincides with that of $\mathscr{B}(F(\cdot))$, where $\mathscr{B}$ is a standard Brownian bridge on $[0, 1]$. Hence, we have the formula

$$\Gamma^{\mathrm{eq}}(s; x) := \int_{-\infty}^{x} (x - y)^{s-2} \mathscr{B}(F(y))\,\mathrm{d}y$$

for the earlier introduced process $x \mapsto \Gamma(s; x)$. In turn, under the null hypothesis $\mathrm{H}_0^{\mathrm{eq}}$ the asymptotic distribution of the statistic $T_n(s; q)$ is equal to that of the random variable

$$T^{\mathrm{eq}}(s; q) := \sup_{x \in \mathbf{R}} q(x)\Gamma^{\mathrm{eq}}(s; x).$$

To estimate the critical values of the distribution function, say H^{eq}, of the limiting random variable $T^{\mathrm{eq}}(s; q)$, we use the following bootstrap approximation. Let

$$F_{\mathrm{pool}} := \frac{n(A)}{n(A) + n(B)} F_{A,n(A)} + \frac{n(B)}{n(A) + n(B)} F_{B,n(B)}.$$

Generate two sets of independent and identically distributed random variables, $X^*_{A,1}, \ldots, X^*_{A,n(A)}$ and $X^*_{B,1}, \ldots, X^*_{B,n(B)}$, each having the distribution function F_{pool}. Let $F^*_{A,n(A)}$ and $F^*_{B,n(B)}$ be the empirical distribution functions corresponding to these two samples. We now define the bootstrapped weighted Kolmogorov–Smirnov-type statistic:

$$T^*_n(s; q) := \sqrt{\frac{n(A)n(B)}{n(A) + n(B)}} \sup_{x \in \mathbf{R}} q(x)(\Delta^*_{A,n(A)}(s; x) - \Delta^*_{B,n(B)}(s; x)),$$

where for $C \in \{A, B\}$, the quantity $\Delta^*_{C,n(C)}(s; x)$ is defined just like $\Delta_{C,n(C)}(s; x)$ but now with $F^*_{C,n(C)}$ instead of $F_{C,n(C)}$. The empirical distribution of the $T^*_n(s; q)$ serves as an approximation to the distribution function H^{eq} of $T^{\mathrm{eq}}(s; q)$. In particular, the upper αth quantile $z_\alpha := \inf\{z \geqslant 0 : H^{\mathrm{eq}}(z) \geqslant 1 - \alpha\}$ can be approximated by

$$z^*_\alpha := \inf\{z \geqslant 0 : H^*(z) \geqslant 1 - \alpha\},$$

where $H^*(z) := \mathbf{P}^*\{T^*_n(s; q) \leqslant z\}$ with $\mathbf{P}^*$ denoting the probability $\mathbf{P}$ conditioned on the random variables $X_{A,1}, \ldots, X_{A,n(A)}$ and $X_{B,1}, \ldots, X_{B,n(B)}$. In view of Theorem 1 with

$F_A = F_B = F$ we have that $\mathbf{P}\{T_n(s;q) > z^*_\alpha\} \to \alpha$. On the other hand, under the alternative hypothesis H_1 we have that $\mathbf{P}\{T_n(s;q) > z^*_\alpha\} \to 1$. The asymptotic P-value (empirical size) $1 - H^{\mathrm{eq}}(T_n(s;q))$ is approximated by

$$\mathbf{P}^*\{T^*_n(s;q) > T_n(s;q)\}. \tag{13}$$

Before applying the above bootstrap procedure in practice, we need to choose a weight function q. To make the presentation more transparent, suppose we are interested in testing for stochastic dominance at the order $s = 3$. In view of condition (6), which is needed for the validity of Theorem 1, we can choose from the class of weight functions $q(x) > 0$ such that, when $x \to \infty$,

$$q(x) = \mathcal{O}(1/x). \tag{14}$$

When looking at the class of weight functions q described by condition (14), one naturally thinks about two cases:

(1) For some fixed constant z, the weight function $q(x)$ equals 1 for all $x \leqslant z$ and 0 for all $x > z$.
(2) For some fixed constants z and a, the weight function $q(x)$ equals 1 for all $x \leqslant z$ and is of the form $a/(a + x - z)$ for all $x > z$. (The parameter $a > 0$ regulates the decay of the weight function.)

The test that involves the first weight function q will be called 'truncated', and the test with the second choice of q will be called 'weighted'. We certainly acknowledge that there are many other possible choices of weight functions, but a full scale analysis of such functions would be beyond the scope of the present paper.

We shall now present our findings demonstrating the potential advantage of using the weighted test rather than the truncated one in the above bootstrap procedure. We work with the Weibull distribution

$$F(x) = 1 - \exp\left\{-\left(\frac{x}{\mu}\right)^\gamma\right\}, \quad x \geqslant 0$$

with the shape parameter γ and the scale parameter μ taking the following values (cf. explanation below):

Year	γ	μ
1991	1.326	40,167
1994	1.304	45,246
1997	1.300	50,287

These three sets of parameters are indexed by a year because they are maximum likelihood estimates for the distribution of the pre-tax household income in the US in the given year, as presented in the paper by Bandourian et al. (2002). For the values of γ considered here, the parameter μ is roughly equal to the mean of the Weibull distribution. Numerical integration shows that, at the order $s = 3$, the 1997 distribution dominates the 1994 distribution which in turn dominates the 1991 distribution. We wish to compare the ability of the above introduced 'truncated' and 'weighted' tests to detect these dominances.

We used $z = 70,000$, which corresponds to a large historical household income. The value of a was chosen to be 100,000. Note that in the weighted test, incomes greater than $z + a$ receive less than half the weight of incomes smaller that z. The value of a thus controls the rate of decay of the weight function: the larger the value of a, the slower the decay. Furthermore, in our numerical experiments we used $n(A) = n(B) = m$ for $m = 50, 100, 250$. The bootstrap P-values (13) were approximated using $B = 1000$ bootstrap replications. In each experimental setting, we generated $R = 100$ Monte Carlo replications. Because of the way we present our findings, larger values of R are not needed.

Fig. 1 compares the empirical distribution of the R P-values for the two tests with $m = 100$. The test statistics were computed with the 1991 distribution as distribution "A" and the distribution of the later year as distribution "B". Since the alternative is known to be true, a superior test should have smaller P-values leading to more rejections. The weighted test is seen to have superior power. The boxes in the plots contain the central 50% of the P-values, the central bar indicates the median. The whiskers indicate the range containing "most" P-values (the precise definition is somewhat technical); the values not in this range are indicated by individual bars. The plots for other values of m are similar.

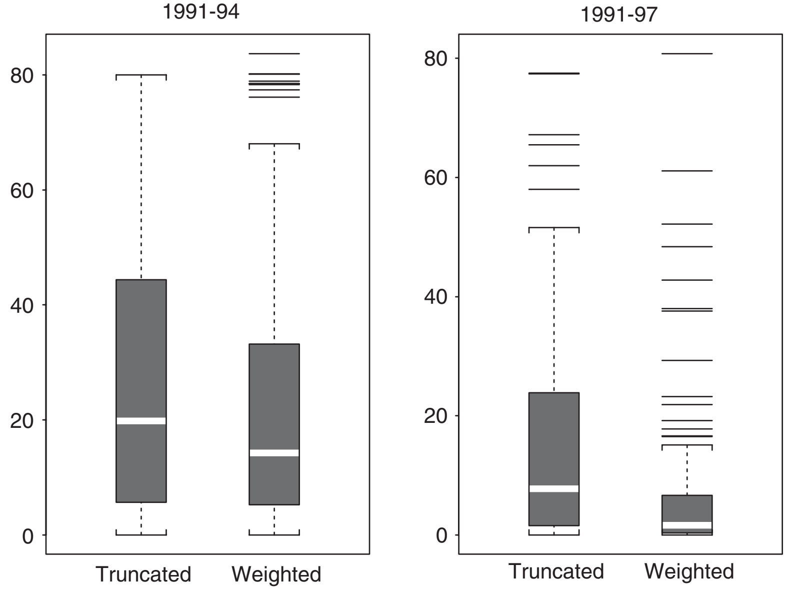

Fig. 1. Boxplots of the P-values for the truncated and weighted tests.

The inferior power of the truncated test is due to the fact that it does not use observations over the truncation level z. The weighted test uses all observations, even though large observations are used with decreasing weight. Its power does not increase noticeably for a greater than the selected value of $a = 100,000$ as there are very few incomes greater than this level. As a decreases, the power decreases. For small a the weight function decays so fast that the weighted test becomes very similar to the truncated test. For z in the range from 50,000 to 100,000 the power of the truncated test increases with z.

The above argument, however, does not clearly indicate which test should have better size. Rather than comparing empirical sizes for any given nominal size, we again use a graphical comparison arrived at as follows. If a null hypothesis is true, the P-values of a test whose empirical size is equal to the nominal size at every significance level have uniform distribution on the interval $[0, 1]$. Thus the empirical distribution function of a finite number R of these P-values should be close to the 45° line.

Fig. 2 shows the empirical distributions of the P-values for the two tests under the null hypothesis in which both distributions "A" and "B" are equal to the 1991 Weibull distribution. It is seen that both tests have similar size properties; the upper tail of a test statistic corresponds to the left bottom corner of each panel in Fig. 2, so the large discrepancies in the right upper corner are not relevant in practice. The size of the weighted test is not affected by the value of a.

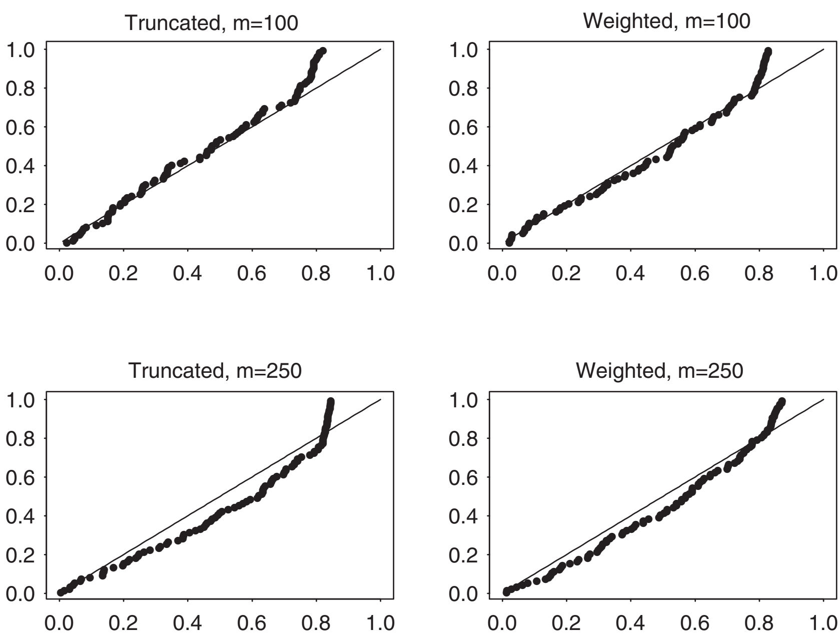

Fig. 2. Empirical distribution functions of the P-values under the null hypothesis $F_A = F_B$.

In conclusion, the weighted test has better power than the truncated test while its size does not deteriorate.

6. Proofs

Proof of statement (11). We start with the note that when (10) holds, then the integral $\int_0^\infty |\mathscr{B}(F(y))|\,\mathrm{d}y$ converges almost surely. Indeed, this is true since the expectation $\mathbf{E}(\int_0^\infty |\mathscr{B}(F(y))|\,\mathrm{d}y)$ does not exceed the integral $\int_0^\infty \sqrt{1-F(y)}\,\mathrm{d}y$, which is finite. To proceed with the proof of (11), we first fix $L>0$, but later we shall let it go to infinity. Since we are interested in the case when $x \to \infty$, we assume without loss of generality that $L<x$. We have

$$\left|\frac{1}{x^{s-2}}\int_0^x (x-y)^{s-2}\mathscr{B}(F(y))\,\mathrm{d}y - \int_0^L \left(1-\frac{y}{x}\right)^{s-2}\mathscr{B}(F(y))\,\mathrm{d}y\right|$$
$$\leqslant \left(1-\frac{L}{x}\right)^{s-2}\int_L^\infty |\mathscr{B}(F(y))|\,\mathrm{d}y. \tag{15}$$

Using the Lebesgue's dominated convergence theorem we have that, when $x \to \infty$, the integral $\int_0^L (1-y/x)^{s-2}\mathscr{B}(F(y))\,\mathrm{d}y$ converges to $\int_0^L \mathscr{B}(F(y))\,\mathrm{d}y$. Furthermore, when $x \to \infty$, then the right-hand side of bound (15) converges to $\int_L^\infty |\mathscr{B}(F(y))|\,\mathrm{d}y$, and the latter integral converges to 0 when $L \to \infty$ almost surely, since the integral $\int_0^\infty |\mathscr{B}(F(y))|\,\mathrm{d}y$ is finite almost surely. Collecting the above statements, we complete the proof of statement (11). □

Proof of Theorem 1. We start with the notation

$$\widetilde{e}_n := e_{A,n(A)} - e_{B,n(B)},$$

where

$$e_{C,n(C)}(y) := \sqrt{\frac{n(A)n(B)}{n(A)+n(B)}}\,(F_{C,n(C)}(y) - F_C(y))$$

for $C \in \{A, B\}$. Hence, for the quantity inside the supremum on the right-hand side of bound (7), we have the following representation:

$$V_n(s;x) - \sqrt{\frac{n(A)n(B)}{n(A)+n(B)}}\,(\varDelta_A(s;x) - \varDelta_B(s;x))$$
$$= V_n^{\mathrm{c}}(s;x) := \int_{-\infty}^x (x-y)^{s-2}\widetilde{e}_n(y)\,\mathrm{d}y, \tag{16}$$

where the superscript 'c' in $V_n^{\mathrm{c}}(s;x)$ refers to 'centered'. Next, we fix any $L \geqslant 1$ (later we shall let it converge to infinity) and write $q(x)V_n^{\mathrm{c}}(s;x)$ as the sum of the following six quantities:

$$Q_n^{[1]}(L,x) := I_{(-\infty,-L]}(x)\int_{-\infty}^{x} q(x)(x-y)^{s-2}\widetilde{e}_n(y)\,\mathrm{d}y,$$

$$Q_n^{[2]}(L,x) := I_{(-L,L]}(x)\int_{-\infty}^{-L} q(x)(x-y)^{s-2}\widetilde{e}_n(y)\,\mathrm{d}y,$$

$$Q_n^{[3]}(L,x) := I_{(-L,L]}(x)\int_{-L}^{x} q(x)(x-y)^{s-2}\widetilde{e}_n(y)\,\mathrm{d}y,$$

$$Q_n^{[4]}(L,x) := I_{(L,\infty)}(x)\int_{-\infty}^{-L} q(x)(x-y)^{s-2}\widetilde{e}_n(y)\,\mathrm{d}y,$$

$$Q_n^{[5]}(L,x) := I_{(L,\infty)}(x)\int_{-L}^{L} q(x)(x-y)^{s-2}\widetilde{e}_n(y)\,\mathrm{d}y,$$

$$Q_n^{[6]}(L,x) := I_{(L,\infty)}(x)\int_{L}^{x} q(x)(x-y)^{s-2}\widetilde{e}_n(y)\,\mathrm{d}y.$$

Hence, we have that, for some $|\theta|\leqslant 1$,

$$\begin{aligned}\sup_{x\in\mathbf{R}} q(x)V_n^{\mathrm{c}}(s;x) = {} & \sup_{x\in\mathbf{R}}(Q_n^{[3]}(L,x)+Q_n^{[5]}(L,x)) + \theta\Bigg(\sup_{x\in\mathbf{R}}|Q_n^{[1]}(L,x)| \\ & + \sup_{x\in\mathbf{R}}|Q_n^{[2]}(L,x)| + \sup_{x\in\mathbf{R}}|Q_n^{[4]}(L,x)| + \sup_{x\in\mathbf{R}}|Q_n^{[6]}(L,x)|\Bigg). \end{aligned} \tag{17}$$

A similar equality (with possibly different $|\theta|\leqslant 1$) holds for the limit $T(s;q)$:

$$\begin{aligned}T(s;q) = {} & \sup_{x\in\mathbf{R}}(Q^{[3]}(L,x)+Q^{[5]}(L,x)) + \theta\Bigg(\sup_{x\in\mathbf{R}}|Q^{[1]}(L,x)| \\ & + \sup_{x\in\mathbf{R}}|Q^{[2]}(L,x)| + \sup_{x\in\mathbf{R}}|Q^{[4]}(L,x)| + \sup_{x\in\mathbf{R}}|Q^{[6]}(L,x)|\Bigg), \end{aligned} \tag{18}$$

where for example, we used $Q^{[1]}(L,x)$ to denote the quantity $Q_n^{[1]}(L,x)$ with $\widetilde{e}_n$ replaced by $\sqrt{\eta}\mathscr{B}_1(F_A(\cdot))+\sqrt{1-\eta}\mathscr{B}_2(F_B(\cdot))$. In what follows we shall prove the statement

$$\sup_{x\in\mathbf{R}} q(x)V_n^{\mathrm{c}}(s;x)\to_d T(s;q) \tag{19}$$

by obtaining appropriate convergence statements concerning the quantities on the right-hand sides of Eqs. (17) and (18).

Quantities $Q_n^{[1]}(L,x)$ *and* $Q^{[1]}(L,x)$: Since both x and y are negative, the integrand $(x-y)^{s-2}$ does not exceed y_-^{s-2}. The weight function q is bounded on the negative half-line $(-\infty,0]$. Thus, the quantity in $Q_n^{[1]}(L,x)$ does not exceed $c\int_{-\infty}^{-L} y_-^{s-2}|\widetilde{e}_n(F(y))|\,\mathrm{d}y$ with a constant $c=c(q)\geqslant 0$ that depends only on q. For any fixed $\varepsilon>0$, we have that

$$\begin{aligned}\sup_n \mathbf{P}\left\{\int_{-\infty}^{-L} y_-^{s-2}|\widetilde{e}_n(y)|\,\mathrm{d}y\geqslant\varepsilon\right\} & \leqslant \frac{1}{\varepsilon}\int_{-\infty}^{-L} y_-^{s-2}\sqrt{\mathbf{E}(\widetilde{e}_n^2(y))}\,\mathrm{d}y \\ & \leqslant \frac{c}{\varepsilon}\sum_{C\in\{A,B\}}\int_{-\infty}^{-L} y_-^{s-2}\sqrt{F_C(y)(1-F_c(y))}\,\mathrm{d}y \end{aligned} \tag{20}$$

with a universal (i.e., not depending on any parameter) constant $c<\infty$. Because of assumption (8) [cf., also (12)], the integral on the right-hand side of bound (20) converges to 0 when $L\to\infty$. Hence, for any fixed $\varepsilon>0$, we have that

$$\sup_n \mathbf{P}\left\{\sup_{x\in\mathbf{R}} |Q_n^{[1]}(L,x)| \geqslant \varepsilon\right\} \to 0, \quad L\to\infty. \tag{21}$$

Statement (21) also holds with $Q_n^{[1]}(L,x)$ replaced by $Q^{[1]}(L,x)$.

Quantities $Q_n^{[2]}(L,x)$ *and* $Q^{[2]}(L,x)$: We first estimate the integrand $q(x)(x-y)^{s-2}$ by considering the two cases $-L\leqslant x\leqslant 1$ and $1<x\leqslant L$ separately. In the first case we have that $(x-y)^{s-2}\leqslant(1+y_-)^{s-2}$. In the second case we have that $q(x)(x-y)^{s-2}$ equals $q(x)x^{s-2}(1-yx^{-1})^{s-2}$, and the latter quantity does not exceed $c(1+y_-)^{s-2}$ with a constant $c=c(q,s)<\infty$ that depends only on q and s. Thus, in both cases we have that the integrand $q(x)(x-y)^{s-2}$ does not exceed $c(1+y_-)^{s-2}$ with a constant $c=c(q,s)<\infty$. This implies that $Q_n^{[2]}(L,x)$ does not exceed $c\int_{-\infty}^{-L} y_-^{s-2}|\widetilde{e}_n(y)|\,\mathrm{d}y$. In view of bound (20) we therefore have that, for any fixed $\varepsilon>0$,

$$\sup_n \mathbf{P}\left\{\sup_{x\in\mathbf{R}} |Q_n^{[2]}(L,x)| \geqslant \varepsilon\right\} \to 0, \quad L\to\infty. \tag{22}$$

Statement (22) also holds with $Q_n^{[2]}(L,x)$ replaced by $Q^{[2]}(L,x)$.

Quantities $Q_n^{[4]}(L,x)$ *and* $Q^{[4]}(L,x)$: Since x is positive, we estimate the integrand $(x-y)^{s-2}$ by $c(x^{s-2}+y_-^{s-2})$, where $c=c(s)<\infty$ depends only on s. In view of condition (6), we have that $q(x)x^{s-2}$ is bounded over the half-line $[0,\infty)$. Thus, for all x and y as they are in quantity $Q_n^{[4]}(L,x)$, we have that $q(x)(x-y)^{s-2}$ does not exceed $c(1+y_-^{s-2})$ for some $c=c(q,s)<\infty$. Hence, $Q_n^{[4]}(L,x)$ does not exceed $c\int_{-\infty}^{-L} y_-^{s-2}|\widetilde{e}_n(y)|\,\mathrm{d}y$. Therefore, for any fixed $\varepsilon>0$,

$$\sup_n \mathbf{P}\left\{\sup_{x\in\mathbf{R}} |Q_n^{[4]}(L,x)| \geqslant \varepsilon\right\} \to 0, \quad L\to\infty. \tag{23}$$

Statement (23) also holds with $Q_n^{[4]}(L,x)$ replaced by $Q^{[4]}(L,x)$.

Quantities $Q_n^{[6]}(L,x)$ *and* $Q^{[6]}(L,x)$: Since $(x-y)^{s-2}$ does not exceed x^{s-2}, we have that $q(x)(x-y)^{s-2}$ does not exceed $c=c(q,s)<\infty$ that depends only on q and s. This implies that $Q_n^{[6]}(L,x)$ does not exceed $c\int_L^\infty |\widetilde{e}_n(y)|\,\mathrm{d}y$. Analogously to bound (20), we prove that

$$\sup_n \mathbf{P}\left\{\int_L^\infty |\widetilde{e}_n(y)|\,\mathrm{d}y \geqslant \varepsilon\right\} \leqslant \frac{c}{\varepsilon}\sum_{C\in\{A,B\}}\int_L^\infty \sqrt{F_C(y)(1-F_C(y))}\,\mathrm{d}y \tag{24}$$

with a universal constant $c<\infty$. Because of assumption (8) [cf., also (10)], the integral on the right-hand side of bound (24) converges to 0 when $L\to\infty$. Hence, for any fixed $\varepsilon>0$, we have that

$$\sup_n \mathbf{P}\left\{\sup_{x\in\mathbf{R}} |Q_n^{[6]}(L,x)| \geqslant \varepsilon\right\} \to 0, \quad L\to\infty. \tag{25}$$

Statement (25) also holds $Q_n^{[6]}(L,x)$ replaced by $Q^{[6]}(L,x)$.

The remaining quantities Q_n and Q_F: We want to show that

$$\sup_{x\in\mathbf{R}}(Q_n^{[3]}(L,x)+Q_n^{[5]}(L,x))\to_d \sup_{x\in\mathbf{R}}(Q^{[3]}(L,x)+Q^{[5]}(L,x)) \tag{26}$$

for every fixed $L\geqslant 1$. Since uniform empirical processes convergence weakly to Brownian bridges, the continuous mapping theorem implies statement (26) provided that we have

$$\sup_{x\in\mathbf{R}} I_{(-L,L]}(x)\int_{-L}^{x} q(x)(x-y)^{s-2}\,\mathrm{d}y<\infty, \tag{27}$$

$$\sup_{x\in\mathbf{R}} I_{(L,\infty)}(x)\int_{-L}^{L} q(x)(x-y)^{s-2}\,\mathrm{d}y<\infty. \tag{28}$$

Statement (27) holds since both x and y are in the finite interval $[-L,L]$. To verify statement (28), we note that $q(x)(x-y)^{s-2}$ does not exceed $q(x)x^{s-2}(1+|y|x^{-1})^{s-2}$ and the latter does not exceed $q(x)x^{s-2}2^{s-2}$. Hence, assumption (6) completes the proof of statement (28). We have proved statement (26).

In view of the above analyzes of the quantities $Q_n^{[1]}(L,x),\dots,Q_n^{[6]}(L,x)$ and their "Gaussian" counterparts $Q^{[1]}(L,x),\dots,Q^{[6]}(L,x)$, Eqs. (17) and (18) imply statement (19). This completes the proof of Theorem 1. □

Acknowledgements

The authors sincerely thank the co-editor, an associate editor, and two referees for constructive criticism and suggestions that helped to revise the paper.

Lajos Horváth and Piotr Kokoszka were partially supported by the NATO Grant PST.EAP.CLG 980599 and by the NSF Grant INT-0223262. Ričardas Zitikis was partially supported by a Discovery Research Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

Anderson, G., 1996. Nonparametric tests of stochastic dominance in income distributions. Econometrica 64, 1183–1193.

Bandourian, R., McDonald, J.B., Turley, R.S., 2002. A comparison of parametric models of income distribution across countries and over time. Technical Report, Department of Economics, Bringham Young University.

Barrett, G.F., Donald, S.G., 2003. Consistent tests for stochastic dominance. Econometrica 71, 71–104.

Chakravarty, S.R., Muliere, R., 2003. Welfare indicators: a review and new perspectives. 1. Measurement of inequality. Metron 61, 457–497.

Csörgő, S., Mason, D.M., 1989. Bootstrapping empirical functions. Annals of Statistics 17, 1447–1471.

Davidson, R., Duclos, J.-Y., 2000. Statistical inference for stochastic dominance and for the measurement of poverty and inequality. Econometrica 68, 1435–1464.

Deshpande, J.V., Singh, H., 1985. Testing for second-order stochastic dominance. Communications in Statistics—Theory and Methods 14, 887–893.

Eubank, R., Schechtman, E., Yitzhaki, S., 1993. A test for second-order stochastic dominance. Communications in Statistics—Theory and Methods 22, 1893–1905.

Kaur, A., Prakasa Rao, B.L.S., Singh, H., 1994. Testing for second-order stochastic dominance of two distributions. Econometric Theory 10, 849–866.

Levy, H., 1992. Stochastic dominance and expected utility: survey and analysis. Management Science 38, 555–593.

Linton, O., Maasoumi, E., Whang, Y., 2003. Consistent testing for stochastic dominance under general sampling schemes. Manuscript, London School of Economics, London, UK.

McFadden, D., 1989. Testing for stochastic dominance. In: Fomby, T.B., Seo, T.K. (Eds.), Studies in the Economics of Uncertainty. Springer, New York, pp. 113–134.

Mosler, K., Scarsini, M., 1991. Stochastic orders and decision under risk. Institute of Mathematical Statistics, Hayward, CA.

Müller, A., 1996. Orderings of risks: a comparative study via stop-loss transforms. Insurance: Mathematics & Economics 17, 215–222.

Müller, A., Stoyan, D., 2002. Comparison Methods for Stochastic Models and Risks. Wiley, Chichester.

Schmid, F., Trede, M., 1996a. Testing for first order stochastic dominance in either direction. Computational Statistics 11, 165–173.

Schmid, F., Trede, M., 1996b. Nonparametric inference for second-order stochastic dominance Discussion Paper no. 2/96. Seminar of Economics and Social Statistics, University of Cologne, Cologne.

Schmid, F., Trede, M., 1998. A Kolmogorov-type test for second-order stochastic dominance. Statistics & Probability Letters 37, 183–193.

Shaked, M., Shanthikumar, J.G., 1994. Stochastic Orders and their Applications. Academic Press, Boston, MA.

Tsirel'son, B.S., 1975. The density of the distribution of the maximum of a Gaussian process. Theory of Probability and its Applications 20, 847–856.