An improved bootstrap test of density ratio ordering

Brendan K. Beare[a,*], Xiaoxia Shi[b]

[a]*Department of Economics, University of California – San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0508, USA*
[b]*Department of Economics, University of Wisconsin – Madison, 1180 Observatory Drive, Madison, WI 53706-1393, USA*

Abstract

Two probability distributions with common support are said to exhibit density ratio ordering when they admit a nonincreasing density ratio. Existing statistical tests of the null hypothesis of density ratio ordering are known to be conservative, with null limiting rejection rates below the nominal significance level whenever the two distributions are unequal. It is shown that a bootstrap procedure can be used to raise the pointwise limiting rejection rate to the nominal significance level on the boundary of the null. This improves power against nearby alternatives. The proposed procedure is based on preliminary estimation of a contact set, the form of which is obtained from a novel representation of the Hadamard directional derivative of the least concave majorant operator. Numerical simulations indicate that improvements to power can be very large in moderately sized samples.

Keywords:
contact set, density ratio ordering, Hadamard directional derivative, least concave majorant, ordinal dominance curve

1. Introduction

Let F and G be cumulative distribution functions (cdfs) on the real line $\mathbb{R}$, with common support. When F and G admit a nonincreasing density ratio $\mathrm{d}F/\mathrm{d}G$, we say that there is *density ratio ordering* between F and G. Density ratio ordering implies, but is not implied by, first order stochastic dominance. While first order or higher order stochastic dominance provides a suitable ordering between distributions in many applications, there are times when economic or financial models indicate that density ratio ordering is the more appropriate property to consider. For instance, Beare (2011) shows that, in a simple one period pricing model, perverse behavior of contingent claims occurs if and only if there is a failure of density ratio ordering between the risk neutral and physical payoff distributions associated with the market portfolio. See also Beare and Schmidt (2016) and Beare and Dossani (2018) for related empirical analyses. Other contexts in which density ratio ordering plays a key role, including mechanism design and auction theory, are discussed by Roosen and Hennessy (2004).

Statistical methods for testing the null hypothesis of stochastic dominance between two cdfs are already well established; see e.g. Anderson (1996), Davidson and Duclos (2000), Barrett and Donald (2003), Linton et al. (2005, 2010), Donald and Hsu (2016) and the survey article by Maasoumi (2001). Less work has been done on testing the null hypothesis of density ratio ordering. Dykstra et al. (1995) and Roosen and Hennessy (2004) dealt with the case where F and G are discrete distributions. The more delicate case where F and G are continuous distributions was studied by Carolan and Tebbs (2005) and Beare and Moon (2015). These authors exploit the fact that, in the continuous case, density ratio ordering is equivalent to the concavity of the *ordinal dominance curve* (odc): the composition of F with G^{-1}, the quantile function for G. They consider a statistic constructed from the difference between an empirical estimate of the odc and its *least*

*Corresponding author.
Email addresses: `bbeare@ucsd.edu` (Brendan K. Beare), `xshi@ssc.wisc.edu` (Xiaoxia Shi)

concave majorant (lcm). It is compared to a fixed critical value that delivers a limiting rejection rate equal to nominal size when $F = G$, and below nominal size when $F \neq G$ but density ratio ordering is satisfied.

The contribution of this paper is a modification to the density ratio ordering test of Carolan and Tebbs (2005) and Beare and Moon (2015) that improves power. We retain the test statistic used by those authors, but compare it to a data dependent critical value computed using the bootstrap. This has the effect of raising the limiting rejection rate of the test to the nominal significance level on the boundary of the null; more precisely, at those points in the null where the limit distribution of the test statistic is nondegenerate. Consequently, power is improved at nearby points in the alternative. Our bootstrap procedure requires preliminary estimation of a *contact set*, and has a similar flavor to the bootstrap procedures used by Linton et al. (2010) and Donald and Hsu (2016) to improve the power of the test of stochastic dominance proposed by Barrett and Donald (2003).

The main technical hurdles we face when studying the asymptotic properties of our procedure relate to the differential properties of the lcm operator. Beare and Moon (2015) showed that this operator fails to be Hadamard differentiable at all points in the null, but instead satisfies a weaker smoothness condition dubbed *Hadamard directional differentiability* by Shapiro (1990). Hadamard directional differentiability suffices for the application of the functional delta method, which is the key device used by Beare and Moon (2015) to determine the asymptotic behavior of their test statistic. However, as shown by Dümbgen (1993) and discussed further in a recent working paper by Fang and Santos (2016), standard bootstrap inference can be problematic when working with operators that are Hadamard directionally differentiable but not Hadamard differentiable. We propose a modified bootstrap procedure with good asymptotic and finite sample properties. Our primary technical innovation is a new representation of the Hadamard directional derivative of the lcm operator that expresses the derivative at each point in the null in terms of an estimable subset of the unit cube: our contact set.

The remainder of our paper is structured as follows. In Section 2 we introduce our sampling framework and test statistic, including a discussion of the directional differentiability of the lcm operator, and an explanation of how this property can be used to derive relevant asymptotic results under the null. In Section 3 we present our main results, including our new representation of the directional derivative of the lcm operator. We explain how this representation can be used to develop a modified bootstrap procedure based on preliminary estimation of the contact set, and establish conditions under which this procedure raises the limiting rejection rate of our test to the nominal significance level on the boundary of the null. Section 4 provides a discussion of some practical issues that arise in the implementation of our procedure, including the numerical computation of suprema and integrals, and the selection of a tuning parameter used in the contact set estimation. Section 5 reports numerical evidence on the finite sample performance of our procedure, and final remarks are given in Section 6. Mathematical proofs of the results stated in Section 3 are collected in Appendix A. A variance estimator used in the construction of our estimated contact set is provided in Appendix B. Further numerical results to supplement those in Section 5 are provided in Appendix C.

2. Test statistic

Here we introduce the test of density ratio ordering studied by Carolan and Tebbs (2005) and Beare and Moon (2015), including details sufficient to provide a basis for our discussion of bootstrap critical values in Section 3. In Section 2.1 we define the null and alternative hypotheses, state the sampling framework, and explain the construction of the test statistic. In Section 2.2 we review results given by Beare and Moon (2015) on the differential properties of the lcm operator, including a discussion of the distinction between Hadamard differentiability and Hadamard directional differentiability. These results are used in Section 2.3 to give a brief derivation of the limit distribution of our test statistic under the null hypothesis, again following Beare and Moon (2015).

2.1. Statistical framework

Our data consist of two independent and identically distributed samples of real valued random variables $(X_1, \ldots, X_m)$ and $(Y_1, \ldots, Y_n)$, mutually independent of one another. (We briefly discuss the case of dependent samples in Section 3.4 below.) We let F denote the common cdf of the X_i's and G denote the common cdf of the Y_j's, and assume that F and G are continuous and strictly increasing on their common support. Our goal is to test the hypothesis that the odc $R = F \circ G^{-1}$ is concave, where $G^{-1}(u) = \inf\{y : G(y) \geq u\}$ is the quantile function corresponding to G. Let Θ denote the collection of strictly increasing, continuously differentiable maps $\theta : [0,1] \to [0,1]$ with $\theta(0) = 0$ and $\theta(1) = 1$, and let $\Theta_0 = \{\theta \in \Theta : \theta \text{ is concave}\}$. We maintain throughout that $R \in \Theta$, and write R' for its first derivative. We seek to test the null hypothesis $H_0 : R \in \Theta_0$ against the alternative hypothesis $H_1 : R \in \Theta \setminus \Theta_0$.

Let $\ell^\infty([a,b])$ denote the collection of uniformly bounded real valued functions on $[a,b]$ equipped with the uniform norm. The following definition is taken from Beare and Moon (2015, Def. 2.1).

Definition 2.1. Given a closed interval $[a,b] \subseteq [0,1]$, the lcm over $[a,b]$ is the operator $\mathcal{M}_{[a,b]} : \ell^\infty([0,1]) \to \ell^\infty([a,b])$ that maps each $f \in \ell^\infty([0,1])$ to the function

$$\mathcal{M}_{[a,b]} f(u) = \inf\{g(u) : g \in \ell^\infty([a,b]),\ g \text{ is concave, and } f \leq g \text{ on } [a,b]\},\ \ u \in [a,b].$$

We write $\mathcal{M}$ as shorthand for $\mathcal{M}_{[0,1]}$, and refer to $\mathcal{M}$ as the lcm operator.

Following Carolan and Tebbs (2005), we take as our estimator of R the empirical odc $R_{m,n} = F_m \circ G_n^{-1}$, where

$$F_m(\cdot) = \frac{1}{m}\sum_{i=1}^{m} 1(X_i \leq \cdot), \quad G_n(\cdot) = \frac{1}{n}\sum_{j=1}^{n} 1(Y_j \leq \cdot),$$

are the empirical cdfs of (X_i) and (Y_j) respectively. Our test statistic is

$$M_{m,n} = c_{m,n} \left\| \mathcal{M} R_{m,n} - R_{m,n} \right\|_p,$$

where $c_{m,n} = (mn/(m+n))^{1/2}$, $\|\cdot\|_p$ is the L^p-norm with respect to Lebesgue measure on $[0,1]$, and $p \in [1,\infty]$. This statistic was proposed by Carolan and Tebbs (2005) for $p = 1$ and $p = \infty$, while Beare and Moon (2015) considered the more general family of statistics indexed by $p \in [1,\infty]$

The empirical odc $R_{m,n}$ is unaffected with probability one if we replace our observations X_i and Y_j with $\psi(X_i)$ and $\psi(Y_j)$ for any real valued ψ strictly increasing on the common support of F and G. Taking $\psi = G$ normalizes the cdf of the $\psi(X_i)$'s to be R and the cdf of the $\psi(Y_j)$'s to be uniform on $[0,1]$, and so we see that the distribution of $M_{m,n}$ is uniquely determined by R. Consequently, it makes sense to talk about the distribution of $M_{m,n}$ at different points in Θ; different pairs of cdfs (F,G) give rise to the same distribution for $M_{m,n}$ whenever they correspond to the same odc $R \in \Theta$.

In the asymptotic theory to be developed shortly, we will let the two sample sizes m and n tend to infinity simultaneously, with $n/(m+n) \to \lambda \in (0,1)$. Formally, we can think of m as being implicitly a function of n, with $m(n) \to \infty$ and $n/(m(n)+n) \to \lambda \in (0,1)$ as $n \to \infty$. We might therefore consider indexing all sample statistics only by n, and never by m or m,n. However, for concreteness, we continue to index sample statistics by m and/or n where appropriate, consistent with Carolan and Tebbs (2005) and Beare and Moon (2015).

2.2. Differential properties of the lcm operator

The arguments used by Beare and Moon (2015) to determine the null limiting behavior of $M_{m,n}$ rely critically on an understanding of the differential properties of the operator $\mathcal{M}$. The following definition is adapted from Dümbgen (1993).

Definition 2.2. Let $\mathscr{X}$ and $\mathscr{Y}$ be real Banach spaces. A map $\phi : \mathscr{X}_\phi \subseteq \mathscr{X} \to \mathscr{Y}$ is said to be Hadamard directionally differentiable at $x \in \mathscr{X}_\phi$ tangentially to a linear space $\mathscr{X}_0 \subseteq \mathscr{X}$ if there exists a map $\phi'_x : \mathscr{X}_0 \to \mathscr{Y}$ such that

$$\phi'_x(z) = \lim_{n\to\infty} \frac{\phi(x + t_n z_n) - \phi(x)}{t_n}, \tag{2.1}$$

for any sequences $z_1, z_2, \ldots \in \mathscr{X}$ and $t_1, t_2, \ldots \in (0,1)$ with $z_n \to z \in \mathscr{X}_0$ and $t_n \downarrow 0$ as $n \to \infty$, and $x + t_n z_n \in \mathscr{X}_\phi$ for all n. We refer to $\phi'_x(z)$ as the Hadamard directional derivative of ϕ at x in direction z. If ϕ'_x is linear then we say that ϕ is Hadamard differentiable at x tangentially to $\mathscr{X}_0$, and we refer to $\phi'_x(z)$ as the Hadamard derivative of ϕ at x in direction z.

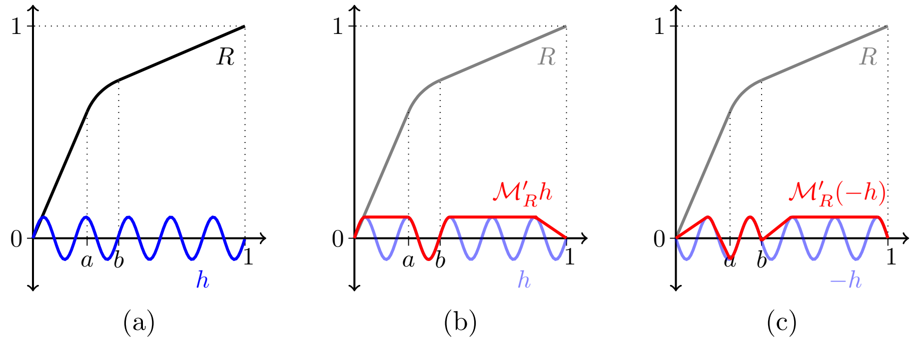


Figure 2.1: Panel (b) displays the Hadamard directional derivative $\mathcal{M}'_R h$ for the particular choice of R and h shown in panel (a). Panel (c) displays $\mathcal{M}'_R(-h)$.

A Hadamard directional derivative is automatically continuous and positive homogeneous of degree one, but may be nonlinear. Linearity turns out to be unimportant for applications of the functional delta method (Shapiro, 1991), but is vitally important for applying the functional delta method for the bootstrap (Dümbgen, 1993; Fang and Santos, 2016), commonly used to establish bootstrap consistency. We will say more about this later.

It turns out that, at points $R \in \Theta_0$, the lcm operator $\mathcal{M}$ is Hadamard directionally differentiable but not in general Hadamard differentiable. The following result, in which $C([0,1])$ denotes the space of continuous real valued functions on $[0,1]$ equipped with the uniform norm, was proved by Beare and Moon (2015, Lem. 3.2). See also Beare and Fang (2017, Prop. 2.1) for a closely related result that applies to the lcm of real valued functions on $[0,\infty)$.

Lemma 2.1. *If $R \in \Theta_0$ then $\mathcal{M}$ is Hadamard directionally differentiable at R tangentially to $C([0,1])$. Given $h \in C([0,1])$, if R is affine in a neighborhood of $u \in (0,1)$, then we have $\mathcal{M}'_R h(u) = \mathcal{M}_{[a_{R,u}, b_{R,u}]} h(u)$, where*

$$\begin{aligned} a_{R,u} &= \sup\{u' \in (0,u] : R \textit{ is not affine in a neighborhood of } u'\}, \\ b_{R,u} &= \inf\{u' \in [u,1) : R \textit{ is not affine in a neighborhood of } u'\}, \end{aligned}$$

and we define $\inf \varnothing = 1$ and $\sup \varnothing = 0$. If R is not affine in a neighborhood of $u \in (0,1)$, or if $u \in \{0,1\}$, then $\mathcal{M}'_R h(u) = h(u)$.

We illustrate the content of Lemma 2.1 with an example in Figure 2.1. In panel (a) we display the odc R at which we wish to differentiate $\mathcal{M}$. It is affine over the intervals $[0,a]$ and $[b,1]$, and strictly concave over the interval $[a,b]$. We also display the direction h in which we wish to differentiate, a sinusoid. In panel (b) we display $\mathcal{M}'_R h$, the Hadamard directional derivative of $\mathcal{M}$ at R in direction h. It has three distinct parts. Over the intervals $[0,a]$ and $[b,1]$, where R is affine, the directional derivative is given by the restricted lcms $\mathcal{M}_{[0,a]} h$ and $\mathcal{M}_{[b,1]} h$ respectively. Over the interval $[a,b]$, where R is strictly concave, the directional derivative is h. In panel (c) we display $\mathcal{M}'_R(-h)$, the Hadamard directional derivative of $\mathcal{M}$ at R in direction $-h$. Comparing $\mathcal{M}'_R h$ and $\mathcal{M}'_R(-h)$ in panels (b) and (c), we observe that $\mathcal{M}'_R h \neq -\mathcal{M}'_R(-h)$, implying that $\mathcal{M}'_R$ cannot be linear. Consequently, $\mathcal{M}$ is not Hadamard differentiable at R tangentially to $C([0,1])$ in the example depicted. In fact, as noted by Beare and Moon (2015), $\mathcal{M}$ is Hadamard differentiable at $R \in \Theta_0$ tangentially to $C([0,1])$ if and only if R is strictly concave.

An alternative generalization of Hadamard differentiability, called quasi-Hadamard differentiability, has been proposed by Beutner and Zähle (2010). The primary difference between Hadamard differentiability and quasi-Hadamard differentiability is that for the latter property we do not require the norm on $\mathscr{X}_0$ to extend to a norm on $\mathscr{X}$, which may be an arbitrary vector space. As discussed by Beutner and Zähle (2010, 2012), this allows the use of weighted norms on $\mathscr{X}_0$ that may not be finite-valued on $\mathscr{X}$, broadening the scope of application of the functional delta method due to the availability of weak convergence results for weighted empirical processes. If we were to tighten Definition 2.2 by requiring (2.1) to hold for all sequences $t_1, t_2, \ldots \in (-1,0) \cup (0,1)$ converging to zero, rather than merely for all sequences $t_1, t_2, \ldots \in (0,1)$ decreasing to zero, then Definition 2.2 would become a special case of quasi-Hadamard differentiability. However, such a tightening would force the directional derivative $\phi'_x(z)$ to satisfy $\phi'_x(-z) = -\phi'_x(z)$ for every $x \in \mathscr{X}_\phi$ and $z \in \mathscr{X}_0$ such that $x \pm z \in \mathscr{X}_\phi$, which in view of Figure 2.1 is clearly not the case when ϕ is the lcm operator. The lcm operator therefore fails to be quasi-Hadamard differentiable in the sense of Beutner and Zähle (2010, 2012). On the other hand, a slightly different definition of quasi-Hadamard differentiability is stated by Beutner et al. (2012) and Volgushev and Shao (2014), in which only sequences $t_1, t_2, \ldots$ converging to zero from above are considered. This weaker definition of quasi-Hadamard differentiability is weaker than Hadamard directional differentiability as defined here. Volgushev and Shao (2014) observe that it suffices for the application of the functional delta method but is insufficient for the application of the functional delta method for the bootstrap, whereas Beutner and Zähle (2016) observe that a version of the functional delta method for the bootstrap applies under the stronger definition of quasi-Hadamard differentiability.

2.3. Limit distribution under concavity

Let $\mathcal{A} : \ell^\infty([0,1]) \to \mathbb{R}$ be the operator

$$\mathcal{A}f = \|\mathcal{M}f - f\|_p, \quad f \in \ell^\infty([0,1]).$$

When R is concave our test statistic $M_{m,n}$ may be written as

$$M_{m,n} = c_{m,n}\left(\mathcal{A}R_{m,n} - \mathcal{A}R\right).$$

Two ingredients suffice for us to establish the limit distribution of $M_{m,n}$ at each $R \in \Theta_0$. First, we require weak convergence of the empirical odc process $c_{m,n}(R_{m,n} - R)$ to a suitable limit, and second, we require the operator $\mathcal{A}$ to satisfy a smoothness condition sufficient for the application of the functional delta method. The former ingredient has been available at least since Hsieh and Turnbull (1996, Thm. 2.2); the following statement is taken from Beare and Moon (2015, Lem. 3.1), with $\rightsquigarrow$ denoting weak convergence in a metric space—in this case $\ell^\infty([0,1])$, but elsewhere depending on context—in the sense of Hoffmann-Jørgensen.

Lemma 2.2. *Suppose $R \in \Theta$. Then as $m \wedge n \to \infty$ with $n/(m+n) \to \lambda \in (0,1)$, we have $c_{m,n}(R_{m,n} - R) \rightsquigarrow T$, where T has the form*

$$T(u) = \lambda^{1/2} B_1(R(u)) + (1-\lambda)^{1/2} R'(u) B_2(u), \quad u \in [0,1],$$

and B_1 and B_2 are independent standard Brownian bridges on $[0,1]$.

It remains to establish a smoothness condition on $\mathcal{A}$ sufficient for the application of the functional delta method. With Lemma 2.1 in hand, a routine application of the chain rule for Hadamard directionally differentiable operators (Shapiro, 1990, Prop. 3.6) establishes that $\mathcal{A}$ is Hadamard directionally differentiable at $R \in \Theta_0$ tangentially to $C([0,1])$, with directional derivative

$$\mathcal{A}'_R h = \|\mathcal{M}'_R h - h\|_p, \quad h \in C([0,1]).$$

Though textbook treatments of the functional delta method typically impose Hadamard differentiability upon the operator in question, it is sufficient to impose the weaker requirement of Hadamard directional differentiability. This was proved by Shapiro (1991, Thm. 2); for a more recent statement, see Fang and Santos (2016, Thm. 2.1). We thus arrive at the following result.

Theorem 2.1. *Suppose $R \in \Theta$. Then as $m \wedge n \to \infty$ with $n/(m+n) \to \lambda \in (0,1)$, we have $M_{m,n} \rightsquigarrow \mathcal{A}'_R T$.*

From Lemma 2.2 we see that the law of T is uniquely determined by R, and hence the law of $\mathcal{A}'_R T$ is also uniquely determined by R. Beare and Moon (2015, Thm. 4.1) proved that, for $p \in [1,2]$, $\mathcal{A}'_R T$ is stochastically dominated by $\mathcal{A}'_I T = \|\mathcal{M}B - B\|_p$, where I is the identity map on $[0,1]$ and B is a Brownian bridge. We therefore refer to $R = I$ as the *least favorable configuration* (lfc) and may construct a conservative test of concavity by using as a critical value the relevant quantile of the law of $\mathcal{A}'_I T$. If we reject the null hypothesis of concavity when M_n exceeds this critical value, then the limiting rejection rate of our test is α at the lfc $R = I$, and is no greater than α at all other $R \in \Theta_0$. The idea of using a fixed critical value to control size at the lfc is due to Carolan and Tebbs (2005), and requires us to choose $p \in [1, 2]$, as $R = I$ is no longer least favorable when $p \in (2, \infty]$ (Beare and Moon, 2015, Thm. 4.2).

The disadvantage of using a fixed critical value to set the limiting rejection rate equal to α at the lfc $R = I$ is that the limiting rejection rate may be well below α at other $R \in \Theta_0$. Indeed, since $\mathcal{A}'_R T = 0$ when R is strictly concave, the limiting rejection rate at all strictly concave $R \in \Theta_0$ is zero. Numerical results reported by Beare and Moon (2015) also indicate that, with $\alpha = 0.05$ and in sample sizes as large as 500, the rejection rate is effectively zero at some members of Θ_0 that are not strictly concave, and are in fact affine over wide portions of their domain. This is problematic because any concave member of Θ may be approximated arbitrarily well in the uniform metric by a nonconcave member of Θ, suggesting that power against relevant nonconcave alternatives may be close to zero.

3. Bootstrap critical values

Our main results are in this section. In Section 3.1 we give a novel representation of the Hadamard directional derivative of the lcm operator and explain how it can be used to express the null limit distribution of $M_{m,n}$ in terms of a contact set and the weak limit of the empirical odc process. In Section 3.2 we discuss the estimation of this contact set. In Section 3.3 we show how the estimated contact set can be used to bootstrap critical values in a way that yields a limiting rejection rate equal to the nominal significance level at all points in the null where R is not strictly concave. In Section 3.4 we show that our procedure leads to an improvement in local power. In Section 3.5 we briefly consider the case of dependent samples. Proofs of all results are collected in Appendix A.

3.1. An alternative representation of $\mathcal{M}'_R$

Begin by defining the set

$$A = \{(u, v, w) \in [0, 1]^3 : v \leq u \leq w\}.$$

Let $\mathcal{S} : \ell^\infty([0, 1]) \to \ell^\infty(A)$ be the operator

$$\mathcal{S}f(u, v, w) = \frac{(w - u)f(v) + (u - v)f(w)}{w - v}, \quad f \in \ell^\infty([0, 1]), \quad (u, v, w) \in A,$$

where for $v = w$ we define $\mathcal{S}f(u, v, w) = f(u)$. We may view $\mathcal{S}f(u, v, w)$ as the approximation to $f(u)$ obtained by linearly interpolating between the values taken by f at v and w. We note the following property of $\mathcal{S}$ for later use.

Lemma 3.1. *$\mathcal{S}$ is a linear isometry.*

With the operator $\mathcal{S}$ and odc R we define the set

$$B = \{(u, v, w) \in A : \mathcal{S}R(u, v, w) = R(u)\},$$

and the family of cross-sections

$$B(u) = \{(v, w) \in [0, 1]^2 : (u, v, w) \in B\}, \quad u \in [0, 1].$$

The set B always contains the main diagonal $u = v = w$ of the unit cube, and the cross-section $B(u)$ always includes the point (u, u).

Our alternative representation of the Hadamard directional derivative of the lcm operator—compare to Lemma 2.1 above—is as follows.

Lemma 3.2. *The Hadamard directional derivative of $\mathcal{M}$ at $R \in \Theta_0$ in direction $h \in C([0,1])$ satisfies*

$$\mathcal{M}'_R h(u) = \sup_{(v,w)\in B(u)} \mathcal{S}h(u,v,w), \quad u \in [0,1].$$

In view of Theorem 2.1 and Lemma 3.2, when $R \in \Theta_0$ the weak limit $\mathcal{A}'_R T$ of our test statistic $M_{m,n}$ satisfies

$$\mathcal{A}'_R T = \left\| \sup_{(v,w)\in B(\cdot)} \tilde{\mathcal{S}}T(\cdot, v, w) \right\|_p,$$

where $\tilde{\mathcal{S}} : \ell^\infty([0,1]) \to \ell^\infty(A)$ is the operator

$$\tilde{\mathcal{S}}f(u,v,w) = \mathcal{S}f(u,v,w) - f(u), \quad f \in \ell^\infty([0,1]), \quad (u,v,w) \in A.$$

The weak limit $\mathcal{A}'_R T$ is uniquely determined by the law of T and the set B. In this sense, B plays a similar role to the so-called contact set used by Linton et al. (2010) to characterize the null limit distribution of their statistic for testing stochastic dominance. We shall borrow their terminology and refer to B as our contact set. Contact sets also play a key role in the analyses of Anderson et al. (2012), Chang et al. (2015) and Lee et al. (2018), although in these papers there arise significant additional technical complications owing to the lack of a weak convergence result analogous to Lemma 2.2.

3.2. Contact set estimation

To implement our bootstrap procedure we require a preliminary estimate of the unknown contact set B. We now present two candidate estimators of B, denoted $B_{m,n}$ and $B'_{m,n}$. By construction, $B_{m,n} \subseteq B'_{m,n}$. Under the null hypothesis, the two estimators closely approximate B with probability approaching one; see Lemma 3.3 below. Under the alternative hypothesis, there can be large differences between the two that persist asymptotically. We will see later that a smaller estimated contact set delivers a smaller critical value, improving the probability of rejecting the null hypothesis. Our preferred contact set estimator is therefore $B_{m,n}$, but we also discuss $B'_{m,n}$ for expository purposes.

Our contact set estimators make use of a tuning parameter $\delta_{m,n} \in (0,\infty)$. The tuning parameter is required to converge to zero as the sample sizes m and n increase, but not too quickly.

Assumption 3.1. As $m \wedge n \to \infty$ with $n/(m+n) \to \lambda \in (0,1)$, we have $\delta_{m,n} \to 0$ and $c_{m,n}\delta_{m,n} \to \infty$.

The results given in this section are valid for any choice of $\delta_{m,n}$ that satisfies Assumption 3.1 with probability one. In Section 4.2 we suggest an approach to choosing $\delta_{m,n}$ in practice. Our first contact set estimator is

$$B'_{m,n} = \left\{ (u,v,w) \in A : |\tilde{\mathcal{S}}R_{m,n}(u,v,w)| \le \delta_{m,n} V^*_{m,n}(u,v,w)^{1/2} \right\}. \tag{3.1}$$

It contains those triples $(u,v,w) \in A$ for which $\mathcal{S}R_{m,n}(u,v,w)$ is close to $R_{m,n}(u)$, with closeness defined in terms of the threshold $\delta_{m,n}V^*_{m,n}(u,v,w)^{1/2}$. Here, $V^*_{m,n}(u,v,w)$ is a regularized estimate of the variance of $c_{m,n}\tilde{\mathcal{S}}R_{m,n}(u,v,w)$, which we set equal to

$$V^*_{m,n}(u,v,w) = \eta_u \wedge (\eta_\ell \vee (V_{m,n}(u,v,w))), \tag{3.2}$$

where η_u and η_ℓ are positive constants and $V_{m,n}(u,v,w)$ is a plug-in estimate of the asymptotic variance of $c_{m,n}\tilde{\mathcal{S}}R_{m,n}(u,v,w)$. A formula for $V_{m,n}(u,v,w)$ is given in Appendix B. The role of η_ℓ and η_u is technical: fixed positive lower and upper bounds on the threshold facilitate a simple proof of Lemma 3.3 below, which establishes good behavior of $B_{m,n}$ and $B'_{m,n}$. Lemma 3.3 in fact remains true for any choice of $V^*_{m,n}$ that is bounded from above and below by fixed positive values; indeed, we could simply set $V^*_{m,n} = 1$. However, we have found in numerical simulations that letting the threshold reflect the variance of the estimated curve yields a test with superior power, and works well even when $\eta_\ell = 0$ and $\eta_u = \infty$. Using a smaller η_u like $\eta_u = 1$ yields moderate power improvement over using $\eta_u = \infty$ in the simulations. Thus, we suggest using $\eta_\ell = 0$ and $\eta_u = 1$.

In large samples, $B'_{m,n}$ can be expected to provide a good approximation to B regardless of whether the null hypothesis is true. For our purposes, a better estimator of B is one that provides a good approximation to B when the null hypothesis is satisfied, but is as small as possible otherwise. Consider the possible contact sets B that may obtain when the null hypothesis is satisfied. When R is concave, if B contains some triple $(u, v, w) \in A$, then it must be the case that $R(t) = \mathcal{M}R(t)$ for $t \in \{u, v, w\}$. Our second contact set estimator is constructed to exclude members of $B'_{m,n}$ that appear very likely to violate this condition:

$$B_{m,n} = \left\{(u, v, w) \in B'_{m,n} : \mathcal{M}R_{m,n}(u) - R_{m,n}(u) \leq \delta_{m,n}\right\}.$$

Ideally, one would like to choose the threshold here to reflect the sample variance of $\mathcal{M}R_{m,n}(u) - R_{m,n}(u)$ as well but this variance is more difficult to characterize. Thus, for transparency, we simply use $\delta_{m,n}$.

Our next result states that, with high probability, $B_{m,n}$ provides a good outer-approximation to our contact set B when the null hypothesis is satisfied.

Lemma 3.3. *Suppose $R \in \Theta_0$ and Assumption 3.1 is satisfied. Then as $m \wedge n \to \infty$ with $n/(m+n) \to \lambda \in (0, 1)$, we have $P(B \subseteq B_{m,n} \subseteq B'_{m,n} \subseteq B^\epsilon) \to 1$ for any $\epsilon > 0$, where*

$$B^\epsilon = \left\{a \in A : \inf_{b \in B} \|b - a\| \leq \epsilon\right\},$$

the ϵ-enlargement of B.

3.3. Bootstrap procedure

In short, our bootstrap approximation to the weak limit $\mathcal{A}'_R T$ of $M_{m,n}$ works by simulating the distribution of $M^*_{m,n} = \hat{\mathcal{A}}'_{m,n} T^*_{m,n}$ conditional on our data, where $T^*_{m,n}$ is a bootstrap version of T, and $\hat{\mathcal{A}}'_{m,n} : \ell^\infty([0, 1]) \to \mathbb{R}$ is the data dependent operator

$$\hat{\mathcal{A}}'_{m,n} f = \left\| \sup_{(v,w) \in B_{m,n}(\cdot)} \tilde{\mathcal{S}} f(\cdot, v, w) \right\|_p, \quad f \in \ell^\infty([0, 1]).$$

The estimated operator $\hat{\mathcal{A}}'_{m,n}$ is determined by the estimated contact set $B_{m,n}$; note that $B_{m,n}(u)$ is a cross-section of $B_{m,n}$, defined in the same way as $B(u)$. Our approach places us in the general framework used by Fang and Santos (2016) to explore the use of bootstrap inference when standard differentiability conditions are violated.

To obtain $T^*_{m,n}$, we first construct bootstrap versions of F_m and G_n by setting

$$F^*_m(\cdot) = \frac{1}{m} \sum_{i=1}^m V^*_{i,m} 1(X_i \leq \cdot), \quad G^*_n(\cdot) = \frac{1}{n} \sum_{j=1}^n W^*_{j,n} 1(Y_j \leq \cdot), \tag{3.3}$$

where the weights $V^*_m = (V^*_{1,m}, \ldots, V^*_{m,m})$ and $W^*_n = (W^*_{1,n}, \ldots, W^*_{n,n})$ are drawn independently of the data and of one another from the multinomial distribution with probabilities spread evenly over the categories $1, \ldots, m$ and $1, \ldots, n$ respectively. This corresponds to letting F^*_m and G^*_n be the empirical cdfs of bootstrap samples of size m and n drawn from the discrete uniform distribution on $(X_1, \ldots, X_m)$ and $(Y_1, \ldots, Y_n)$ respectively. From F^*_m and G^*_n we construct $R^*_{m,n} = F^*_m \circ G^{*-1}_n$, our bootstrap version of $R_{m,n}$. We then set $T^*_{m,n} = c_{m,n}(R^*_{m,n} - R_{m,n})$.

The following result establishes that the law of $T^*_{m,n}$ conditional on the data provides an accurate approximation to the law of T with high probability. Weak convergence conditional on the data in probability is meant in the sense of Kosorok (2008, pp. 19-20).

Lemma 3.4. *Suppose $R \in \Theta$. Then as $m \wedge n \to \infty$ with $n/(m + n) \to \lambda \in (0, 1)$, we have $T^*_{m,n} \rightsquigarrow T$ conditional on the data in probability.*

The law of $T^*_{m,n}$ conditional on the data can be simulated: we simply compute large numbers of realizations of $T^*_{m,n}$ corresponding to repeated draws of the multinomial weights V^*_m and W^*_n. In order to obtain suitable critical values for our test statistic, we seek to approximate the law of its weak limit $\mathcal{A}'_R T$ when $R \in \Theta_0$. If $\mathcal{A}$ were Hadamard differentiable at $R \in \Theta_0$ tangentially to $C([0,1])$, we could deduce from the functional delta method for the bootstrap that $c_{m,n}(\mathcal{A}R^*_{m,n} - \mathcal{A}R_{m,n}) \rightsquigarrow \mathcal{A}'_R T$ conditional on the data in probability, which would justify the use of the law of $c_{m,n}(\mathcal{A}R^*_{m,n} - \mathcal{A}R_{m,n})$ conditional on the data as an approximation to the law of $\mathcal{A}'_R T$. Unfortunately we cannot apply the delta method for the bootstrap in this fashion unless R is strictly concave, because it is only at the strictly concave members of Θ_0 that $\mathcal{A}$ is Hadamard differentiable. Though $\mathcal{A}$ is Hadamard directionally differentiable at all $R \in \Theta_0$, it was shown by Dümbgen (1993) and Fang and Santos (2016) that directional differentiability does not suffice for the application of the functional delta method for the bootstrap, and that the naïve bootstrap typically fails when working with operators that are not fully Hadamard differentiable. The relevant results are Proposition 1 of Dümbgen (1993) and Theorems 3.1 and A.1 and Remark 3.2 of Fang and Santos (2016), which show that the law of $c_{m,n}(\mathcal{A}R^*_{m,n} - \mathcal{A}R_{m,n})$ conditional on the data provides a consistent approximation to the target weak limit $\mathcal{A}'_R T$ if and only if $\mathcal{A}$ is Hadamard differentiable at R tangentially to the support of T. When this condition fails, as it does here when R is not strictly concave, consistent approximation may instead be achieved using a modified bootstrap procedure based on an explicit or implicit estimate of the directional derivative $\mathcal{A}'_R$. In our case that estimate is $\hat{\mathcal{A}}'_{m,n}$ defined above, constructed from the estimated contact set $B_{m,n}$.

We will approximate the law of $\mathcal{A}'_R T$ using the law of $M^*_{m,n} = \hat{\mathcal{A}}'_{m,n} T^*_{m,n}$ conditional on the data. For a test with nominal size $\alpha \in (0, 1/2)$ we take as our critical value

$$\mu_{m,n}(\alpha) = \inf\{x : P(M^*_{m,n} \leq x \mid X_1, \ldots, X_m, Y_1, \ldots, Y_n) \geq 1 - \alpha\},$$

the $(1-\alpha)$-quantile of the distribution of $M^*_{m,n}$ conditional on the data.

Theorem 3.1. *Suppose $R \in \Theta_0$ and Assumption 3.1 is satisfied. Then as $m \wedge n \to \infty$ with $n/(m+n) \to \lambda \in (0,1)$, we have $M^*_{m,n} \rightsquigarrow \mathcal{A}'_R T$ conditional on the data in probability. If in addition R is not strictly concave, we have $P(M_{m,n} > \mu_{m,n}(\alpha)) \to \alpha$.*

Theorem 3.1 establishes that our bootstrap procedure delivers a test with limiting rejection rate equal to nominal size whenever R is concave but not strictly concave. These R are precisely those points in Θ_0 at which the limit distribution of $M_{m,n}$ is nondegenerate, and form what Linton et al. (2010) refer to as the *boundary of the null*. Of course, this notion of boundary differs from the usual topological one; in the uniform topology, every member of Θ_0 is the limit of a sequence in Θ_1, and so Θ_0 is its own boundary.

A shortcoming of Theorem 3.1 is that it says nothing about the limiting rejection rate of our test when R is strictly concave. In this case, both $M_{m,n}$ and $\mu_{m,n}(\alpha)$ converge in probability to zero, and we cannot say much of substance about their relative magnitudes without investigating their higher order asymptotic behavior, which seems difficult. In a related context, Andrews and Shi (2013, p. 625) have proposed a technical remedy to this problem: instead of using $\mu_{m,n}(\alpha)$ as our critical value, we can use $\mu_{m,n}(\alpha) + \epsilon$ or $\mu_{m,n}(\alpha) \vee \epsilon$, where $\epsilon > 0$ is some small fixed value. The presence of ϵ prevents our critical value from converging in probability to zero alongside $M_{m,n}$ when R is strictly concave, ensuring a limiting rejection rate of zero. For further discussion, see Fang and Santos (2016, Rem. 3.12) and Donald and Hsu (2016, p. 13). We have found in numerical simulations with $p = 1$ and $p = 2$ that in practice it is unnecessary to modify the critical value in this fashion. Our test appears to be conservative at strictly concave choices of R, and also at many concave choices of R that are not strictly concave.

3.4. Power advantage of the bootstrap test

To demonstrate the asymptotic power advantage of our bootstrap test over the fixed critical value test of Beare and Moon (2015) we shall show that, when $p \in [1,2]$, the conditional weak limit of our bootstrapped statistic $M^*_{m,n}$ under a sequence of local alternatives is stochastically dominated by $\mathcal{A}'_I B$, where B is a standard Brownian bridge. Recall that the test of Beare and Moon (2015) involves comparing $M_{m,n}$ to the

$(1-\alpha)$-quantile of $\mathcal{A}_I' B$, whereas our bootstrap test involves comparing $M_{m,n}$ to the conditional $(1-\alpha)$-quantile of $M_{m,n}^*$. Let $\nu(\alpha)$ denote the former quantile and, as above, let $\mu_{m,n}(\alpha)$ denote the latter quantile.

Fix a point in the null $R \in \Theta_0$ and a sequence $R^{(1)}, R^{(2)}, \ldots$ in $\Theta \setminus \Theta_0$ such that, for all m, we have $R^{(m)} = R + m^{-1/2}\Delta$ for some $\Delta \in C([0,1])$. We can interpret $R^{(m)}$ as a local alternative to R obtained by perturbing R in direction Δ. Let T be as in Lemma 2.2. The following result characterizes the asymptotic behavior of $M_{m,n}$ and $M_{m,n}^*$, and the relative asymptotic power of tests based on the critical values $\nu(\alpha)$ and $\mu_{m,n}(\alpha)$, along our sequence of local alternatives.

Theorem 3.2. *Suppose that $p \in [1,2]$. Under the sequence $(R^{(m)} : m \geq 1)$, as $m \wedge n \to \infty$ with $n/(m+n) \to \lambda \in (0,1)$ we have*

(*a*) $M_{m,n} \rightsquigarrow \mathcal{A}_R'(T + \lambda^{1/2}\Delta)$,
(*b*) $M_{m,n}^* \rightsquigarrow \mathcal{A}_R' T$ *conditional on the data in probability, and*
(*c*) $\limsup P(M_{m,n} > \nu(\alpha)) \leq \liminf P(M_{m,n} > \mu_{m,n}(\alpha))$.

The limits superior and inferior in (*c*) *are proper limits. If R is not strictly concave and also not equal to I then the weak inequality in* (*c*) *may be replaced with a strict inequality.*

The conditional weak limit of $M_{m,n}^*$ given in part (b) of Theorem 3.2 can be shown (Beare and Moon, 2015, Theorem 4.1) to be stochastically dominated by $\mathcal{A}_I' B$. Thus $\mu_{m,n}(\alpha)$ converges to a value $\mu(\alpha) \leq \nu(\alpha)$, implying a weak increase in local power, as indicated by part (c) of Theorem 3.2. The increase in local power is strict when R belongs to the boundary of the null but does not satisfy the lfc $R = I$. It can be seen in panels (a) and (b) of Figure 3 of Beare and Moon (2015) that the difference between $\mu(\alpha)$ and $\nu(\alpha)$ can be very large for certain choices of $R \in \Theta_0$.

3.5. Extension to dependent samples

We have focused for clarity on the case where the data are independent between and within samples. However, most of our results generalize without much difficulty to the dependent case. It is more natural to consider dependence between samples when observations are paired as (X_i, Y_i), so we focus on the case $m = n$ and set $c_n = c_{m,n}$, $R_n = R_{m,n}$ and $R_n^* = R_{m,n}^*$. Suppose we drop the assumption of independence between and within samples and instead assume directly that

$$n^{1/2}((F_n, G_n) - (F, G)) \rightsquigarrow (B_1 \circ F, B_2 \circ G) \tag{3.4}$$

in $\ell^\infty([0,1]) \times \ell^\infty([0,1])$, where B_1 and B_2 are random elements of $C([0,1])$. Donsker's theorem implies that (3.4) is satisfied when the data are independent between and within samples, with B_1 and B_2 independent standard Brownian bridges. In the case where we have dependence between samples but the pairs (X_i, Y_i) are iid across i, a bivariate extension of Donsker's theorem ensures that (3.4) continues to be satisfied but the standard Brownian motions B_1 and B_2 are no longer independent. When there is dependence within samples, (3.4) may be satisfied under a suitable mixing condition or related property (Dehling and Philipp, 2002). In any case, if (3.4) is satisfied then Lemma 2.2 will continue to hold with B_1 and B_2 suitably redefined; this can be proved by applying the functional delta method using Lemma A.1 in Appendix A. Note that Lemma A.1 is applicable because we may assume without loss of generality that G is the uniform distribution on $[0,1]$, due to the fact that R_n is unaffected with probability one if we replace our pairs of observations (X_i, Y_i) with $(G(X_i), G(Y_i))$, as discussed in Section 2.1. Theorem 2.1 and Lemma 3.3 also remain valid and require essentially no change to the proofs.

The bootstrap procedure proposed in Section 3.3 cannot be expected to work under dependent sampling when F_n^* and G_n^* are defined as in (3.3). We instead require bootstrap versions of F_n and G_n that properly capture dependence between and within samples. In the case where we have dependence between samples but the pairs (X_i, Y_i) are iid across i, it is enough to replace the multinomial weights $W_{j,n}^*$ in (3.3) with $V_{j,n}^*$, so that the two bootstrapped cdfs use the same weights. When there is dependence within samples, a resampling procedure suitable for serially dependent data must be applied; see Radulović (2002) for a survey and discussion of such procedures. A suitable procedure must satisfy

$$n^{1/2}((F_n^*, G_n^*) - (F_n, G_n)) \rightsquigarrow (B_1 \circ F, B_2 \circ G) \tag{3.5}$$

in $\ell^\infty([0,1]) \times \ell^\infty([0,1])$ conditional on the data in probability, meaning that it consistently approximates the law of $n^{1/2}((F_n, G_n) - (F, G))$. We further require that R_n^* is unaffected with probability one if we replace the pairs (X_i, Y_i) with $(G(X_i), G(Y_i))$, as is the case when F_n^* and G_n^* are defined as in (3.3). With these additional conditions in place the proof of Lemma 3.4 goes through in the dependent case, but with (A.1) and (A.2) assumed rather than deduced. Theorem 3.1 then follows with no further changes to the proof required. Theorem 3.2 follows similarly if we assume (3.4) and (3.5) to hold under the sequence of local alternatives considered there.

4. Practical implementation

Here we provide some pragmatic guidelines for implementing our testing procedure. In Section 4.1 we provide a step-by-step guide to the computation of our test statistic and bootstrap critical value, avoiding abstract operations such as suprema over infinite sets and integration, and instead using only operations that are easily implementable using standard numerical software packages. A method for choosing the tuning parameter $\delta_{m,n}$ is suggested in Section 4.2.

4.1. Numerical computation

What follows is a step-by-step recipe for computing our test statistic and critical value. We allow the calculations to use a smaller number of grid points (N) than the sample sizes to make the computation easier in the case of very large n (e.g. $n > 1000$). When $N = n$, all steps below provide an exact calculation, with the exception of step 3(v), which uses a summation to numerically approximate an integral. When $N < n$, all steps use approximation.

1. Compute the test statistic.
   (i) Order the sample of X as $X_{(1)} \leq \cdots \leq X_{(m)}$.
   (ii) Set $R_{m,n}(0) = 0$ and for $i = 1, \ldots, N$ compute

$$R_{m,n}\left(\frac{i}{N}\right) = \frac{1}{m}\max\{j = 1, \ldots, m : X_{(j)} \leq G_n^{-1}(i/N)\},$$

   with the maximum over the empty set defined to be zero.
   (iii) For $j = 0, \ldots, N-1$ and $i = j+1, \ldots, N$ and $k = i, \ldots, N$ compute

$$\mathcal{S}R_{m,n}\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right) = \frac{(k-i)R_{m,n}(j/N) + (i-j)R_{m,n}(k/N)}{k-j},$$

   and for $i = 0, \ldots, N$ set $\mathcal{S}R_{m,n}(i/N, i/N, i/N) = R_{m,n}(i/N)$.
   (iv) Set $\mathcal{M}R_{m,n}(1) = R_{m,n}(1)$ and for $i = 1, \ldots, N$ compute

$$\mathcal{M}R_{m,n}\left(\frac{i-1}{N}\right) = \max_{j=1,\ldots,i}\max_{k=i,\ldots,N}\mathcal{S}R_{m,n}\left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N}\right).$$

   (v) Compute $M_{m,n}$. For $p = 1$ we have

$$M_{m,n} = \frac{c_{m,n}}{N}\sum_{i=1}^{N}\left[\frac{1}{2}\mathcal{M}R_{m,n}\left(\frac{i-1}{N}\right) + \frac{1}{2}\mathcal{M}R_{m,n}\left(\frac{i}{N}\right) - R_{m,n}\left(\frac{i}{N}\right)\right].$$

   For $p = 2$ we have

$$M_{m,n} = \frac{c_{m,n}}{N^{1/2}}\left(\sum_{i=1}^{N}\left\{\frac{1}{3}\left[\mathcal{M}R_{m,n}\left(\frac{i}{N}\right) - \mathcal{M}R_{m,n}\left(\frac{i-1}{N}\right)\right]^2 + \left[\mathcal{M}R_{m,n}\left(\frac{i-1}{N}\right) - R_{m,n}\left(\frac{i}{N}\right)\right]\left[\mathcal{M}R_{m,n}\left(\frac{i}{N}\right) - R_{m,n}\left(\frac{i}{N}\right)\right]\right\}\right)^{1/2}.$$

2. Determine which of the relevant points in the unit cube belong to the estimated contact set. For $i = 0, \ldots, N$ and $j = 0, \ldots, i$ and $k = i, \ldots, N$, set $b_{i,j,k} = 1$ if the inequalities

$$\left| \mathcal{S}R_{m,n} \left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N} \right) - R_{m,n} \left(\frac{i}{N} \right) \right| \leq \delta_{m,n} V^*_{m,n} \left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N} \right)$$

and

$$\mathcal{M}R_{m,n} \left(\frac{i}{N} \right) - R_{m,n} \left(\frac{i}{N} \right) \leq \delta_{m,n}$$

are satisfied, and set $b_{i,j,k} = 0$ otherwise. The variance estimator $V^*_{m,n}(i/N, j/N, k/N)$ is computed as described in Section 3.2 and Appendix B.

3. Generate the bootstrap critical value.
   (i) Generate bootstrap samples $X^*_1, \ldots, X^*_m$ and $Y^*_1, \ldots, Y^*_n$ by drawing with replacement from the original samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$.
   (ii) For $i = 0, \ldots, n$ and $j = 0, \ldots, i$ and $k = i, \ldots, N$ compute $R^*_{m,n}(i/N)$ and $\mathcal{S}R^*_{m,n}(i/N, j/N, k/N)$ by following the procedure in steps 1(i)-1(iii).
   (iii) For $i = 0, \ldots, N$ and $j = 0, \ldots, i$ and $k = i, \ldots, n$ compute

$$\tilde{S}T^*_{m,n} \left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N} \right) = c_{m,n} \left[\mathcal{S}R^*_{m,n} \left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N} \right) - R^*_{m,n} \left(\frac{i}{N} \right) \right. \\ \left. - \mathcal{S}R_{m,n} \left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N} \right) + R_{m,n} \left(\frac{i}{N} \right) \right].$$

   (iv) For $i = 0, \ldots, N$ compute

$$H^*_{m,n} \left(\frac{i}{N} \right) = \max_{j=0,\ldots,i} \max_{k=i,\ldots,N} b_{i,j,k} \tilde{S}T^*_{m,n} \left(\frac{i}{N}, \frac{j}{N}, \frac{k}{N} \right).$$

   (v) Exact computation of $M^*_{m,n}$ is complicated. We suggest using the numerical approximation

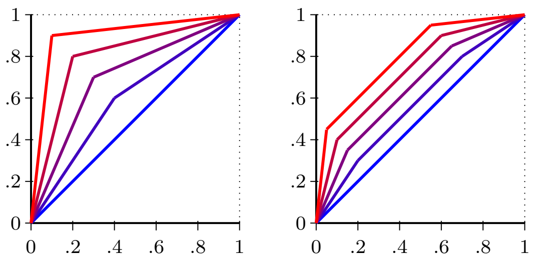
$$M^*_{m,n} \approx \left[\frac{1}{N} \sum_{i=1}^{n} H^*_{m,n} \left(\frac{i}{N} \right)^p \right]^{1/p}.$$

   (vi) Repeat steps 3(i)-3(v) S times, for some large S, to obtain a large number of realizations of $M^*_{m,n}$. Our bootstrap critical value $\mu_{m,n}(\alpha)$ is set equal to the $[\alpha S]$-th largest of these realizations. We reject the null if $M_{m,n} > \mu_{m,n}(\alpha)$. As a p-value we may take the smallest q such that $M_{m,n} > \mu_{m,n}(q)$.

4.2. Tuning parameter selection

Under Assumption 3.1 we are free to choose any tuning parameter $\delta_{m,n}$ that satisfies $\delta_{m,n} \to 0$ and $c_{m,n}\delta_{m,n} \to \infty$ as our sample sizes m and n increase. That is all well and good for the purposes of asymptotic thought experiments, but not a lot of help when it comes to choosing $\delta_{m,n}$ in practice. The following procedure for choosing $\delta_{m,n}$ has worked well for us in numerical simulations when $p = 1$ and $p = 2$, and thus is recommended for practical implementation.

For a grid of candidate tuning parameters, use Monte Carlo simulation to compute the rejection rate of the test when $R = I$, the least favorable case for $p = 1$ and $p = 2$. Then, choose the smallest tuning parameter that yields a rejection rate acceptably close to the nominal size α. Such a choice is possible because the rejection rate of our test is monotonically decreasing in $\delta_{m,n}$ and we have found in numerical simulations that the rejection rate is below α at $R = I$ when $\delta_{m,n}$ is chosen very large, and rises above α at $R = I$ when $\delta_{m,n}$ becomes sufficiently small. The selected tuning parameter will control the finite sample rejection rate at $R = I$ by construction, and we have found in numerical simulations that it delivers a finite sample rejection rate near or below nominal size at other points in the null.



(a) R_γ^{I} (b) R_γ^{II}

Figure 5.1: Concave ordinal dominance curves used to evaluate finite sample size. The parameter γ varies from 0 to 0.8 in even increments.

5. Finite sample performance

To investigate the finite sample performance of our proposed testing procedure we used Monte Carlo simulation to compute rejection rates at a range of odcs satisfying the null or alternative hypothesis. Here we report results obtained for equally sized samples with $m = n = 200$. Results for other sample sizes we investigated were qualitatively similar; we report results for $m = n = 400$ in Appendix C. For each odc considered, we used 10000 Monte Carlo replications to compute rejection rates. We used the method of Giacomini et al. (2013) to reduce computation time, so bootstrap critical values were based on 10000 bootstrap samples drawn over the full set of Monte Carlo replications. Rejection rates were computed using $p = 1$ and $p = 2$. A tuning parameter value of $\delta_{m,n} = .065$ was used for $p = 1$ and .06 was used for $p = 2$; at these values, preliminary simulations of the kind described in Section 4.2 indicated that the rejection rates at $R = I$ were close to .05. In view of the rate requirement on $\delta_{m,n}$, we expect the values that work well at larger sample sizes to be smaller, and at smaller sample sizes to be larger.

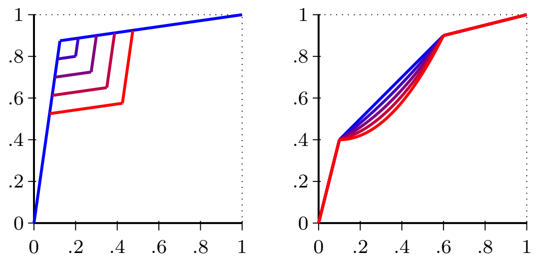
The odcs used in our simulations were drawn from four parametric families. To investigate the behavior of our test when R is concave, we considered two families of concave odcs parametrized by $\gamma \in [0, 1)$.

$$\text{(I)} \qquad R_\gamma^{\mathrm{I}}(u) = \begin{cases} \frac{1+\gamma}{1-\gamma} u & \text{if } 0 \le u \le \frac{1-\gamma}{2} \\ \frac{1-\gamma}{1+\gamma} u + \frac{2\gamma}{1+\gamma} & \text{if } \frac{1-\gamma}{2} \le u \le 1, \end{cases}$$

$$\text{(II)} \qquad R_\gamma^{\mathrm{II}}(u) = \begin{cases} \frac{1+\gamma}{1-\gamma} u & \text{if } 0 \le u \le \frac{1-\gamma}{4} \\ u + \gamma/2 & \text{if } \frac{1-\gamma}{4} \le u \le \frac{3-\gamma}{4} \\ \frac{1-\gamma}{1+\gamma} u + \frac{2\gamma}{1+\gamma} & \text{if } \frac{3-\gamma}{4} \le u \le 1. \end{cases}$$

In panel (a) of Figure 5.1 we graph R_γ^{I} for several values of γ. At $\gamma = 0$ the graph of R_γ^{I} is the 45° line, while for $\gamma > 0$ the graph is piecewise affine with a single kink located at a point that moves toward the upper-left corner of the unit square as $\gamma \to 1$. This is the same family of curves considered in numerical simulations reported by Beare and Moon (2015, Fig. 1), except that we have not bothered to smooth away the single kink appearing when $\gamma > 0$. This means that our kinked odcs violate the continuous differentiability condition imposed on members of Θ; however, we have found that applying a small degree of smoothing to R_γ^{I} to restore continuous differentiability makes essentially no difference to the rejection rates computed. In panel (b) of Figure 5.1, we graph the second family of odcs R_γ^{II} for several values of γ. At $\gamma = 0$, its graph is the 45° line, while for $\gamma > 0$, the graph is piecewise affine with two kinks located at points that move toward the upper-left corner of the unit square as γ increases.

To investigate the behavior of our test when R is not concave, we considered two additional families of



(a) R_γ^{III} (b) R_γ^{IV}

Figure 5.2: Nonconcave ordinal dominance curves used to evaluate finite sample power. The parameter γ varies from 0 to .4 in even increments in panel (a), and from 0 to 1 in even increments in panel (b).

odcs parametrized by $\gamma \in [0, 1]$.

$$\text{(III)} \qquad R_\gamma^{\mathrm{III}}(u) = \begin{cases} 7u & \text{if } 0 \le u \le \frac{1-\gamma}{8} \\ \frac{1}{7}u + \frac{6-6\gamma}{7} & \text{if } \frac{1-\gamma}{8} \le u \le \frac{1+6\gamma}{8} \\ 7u - 6\gamma & \text{if } \frac{1+6\gamma}{8} \le u \le \frac{1+7\gamma}{8} \\ \frac{1}{7}u + \frac{6}{7} & \text{if } \frac{1+7\gamma}{8} \le u \le 1, \end{cases}$$

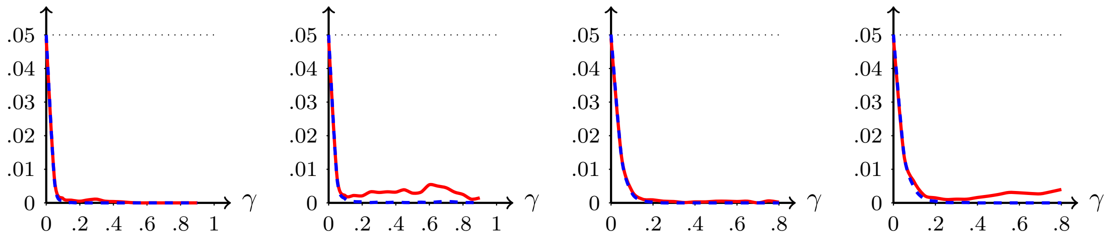
$$\text{(IV)} \qquad R_\gamma^{\mathrm{IV}}(u) = \begin{cases} 4u & \text{if } 0 \le u \le \frac{1}{10} \\ u + \frac{3}{10} + 2\gamma\left(u - \frac{1}{10}\right)\left(u - \frac{3}{5}\right) & \text{if } \frac{1}{10} \le u \le \frac{3}{5} \\ \frac{1}{4}u + \frac{3}{4} & \text{if } \frac{3}{5} \le u \le 1. \end{cases}$$

In panel (a) of Figure 5.2 we graph R_γ^{III} for several values of γ. When $\gamma = 0$ we see that R_γ^{III} is a piecewise affine concave function with a single kink, and in fact we have $R_0^{\mathrm{III}} = R_{.75}^{\mathrm{I}}$. When $\gamma > 0$, R_γ^{III} is a piecewise affine nonconcave function with three kinks. As γ increases, R_γ^{III} moves further away from the concave function R_0^{III}; intuitively, we can think of R_γ^{III} as moving deeper into the alternative region as γ increases. Strictly speaking R_γ^{III} does not belong to Θ due to the violation of continuous differentiability, but as with R_γ^{I} this is a purely technical issue that can be overcome by applying a negligible degree of smoothing at kink points. In panel (b) of Figure 5.2 we graph R_γ^{IV} for several values of γ. At $\gamma = 0$ we see that R_γ^{IV} is a piecewise affine concave function with two kinks, and in fact we have $R_0^{\mathrm{IV}} = R_{.6}^{\mathrm{II}}$. When $\gamma > 0$ the portion of the curve between 0.1 and 0.6 bows to the lower-right.

We report rejection rates using a fixed critical value as in Carolan and Tebbs (2005) and Beare and Moon (2015) as well as the bootstrap critical values proposed here. Test statistics were computed as described in Section 4.1, using $N = 100$ gridpoints. For fairness of comparison, we adjusted the fixed critical value downward so that the rejection rate of the fixed critical value test is exactly equal to the nominal size of .05 when $R = I$.

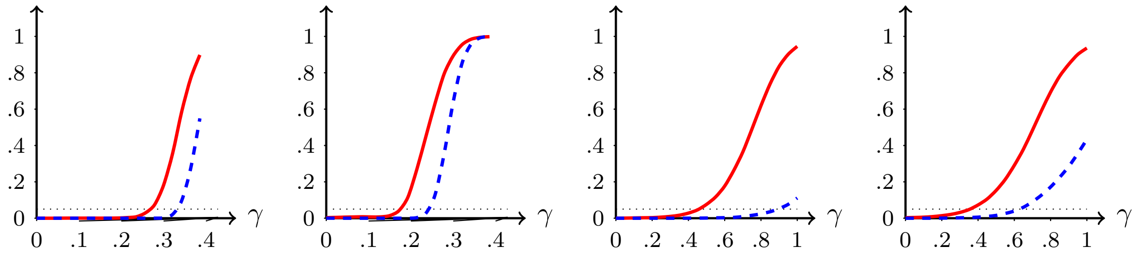
Figure 5.3 displays the rejection rates we computed for the families of concave odcs R_γ^{I} and R_γ^{II}. For each family of concave odcs, we plot the rejection rate with the L^1 ($p = 1$) and the L^2 ($p = 2$) tests against the parameter γ. We see that in all cases the rejection rates using the fixed critical values drop very rapidly to zero as γ increases, becoming indistinguishable from zero by at most $\gamma = .2$, and staying at that level as γ rises to one. The rejection rates using the bootstrap critical values are not flat at .05 either. When $p = 1$ the bootstrap rejection rates look similar to the fixed critical value rejection rates. When $p = 2$ the situation is more interesting: the bootstrap rejection rates drop rapidly to zero as γ rises above zero, but bounce back somewhat as γ continues to rise; they still remain well below nominal size, however.

Figure 5.4 displays power curves for the families of odcs R_γ^{III} and R_γ^{IV}. The results for $p = 1$ and $p = 2$ are qualitatively similar: power curves for both tests rise as γ increases, with the test using bootstrap critical



(a) R^{I}_{γ}, $p = 1$ (b) R^{I}_{γ}, $p = 2$ (c) R^{II}_{γ}, $p = 1$ (d) R^{II}_{γ}, $p = 2$

Figure 5.3: Null rejection rates with fixed critical values (dashed) and bootstrap critical values (solid), with sample sizes $m = n = 200$.



(a) $R^{\mathrm{III}}_{\gamma}$, $p = 1$ (b) $R^{\mathrm{III}}_{\gamma}$, $p = 2$ (c) R^{IV}_{γ}, $p = 1$ (d) R^{IV}_{γ}, $p = 2$

Figure 5.4: Power curves with fixed critical values (dashed) and bootstrap critical values (solid), with sample sizes $m = n = 200$.

values easily outperforming the test using fixed critical values. For example, with $R^{\mathrm{III}}_{\gamma}$ and $p = 2$ at $\gamma = 0.25$, the improvement in power brought about by our bootstrap procedure is about 0.8. Comparing the power curves for $p = 1$ and $p = 2$, we see better performance with $p = 2$.

Why are the null rejection rates for the bootstrap test plotted in Figure 5.3 not approximately flat at .05, as suggested by Theorem 3.1? One plausible source of the under-rejection is the fact that the contact set corresponding to the odcs R^{I}_{γ} and R^{II}_{γ} changes discontinuously as γ rises above zero. This causes the limit distribution of $M_{m,n}$ to discontinuously shrink toward zero when γ rises above zero (see e.g. Beare and Moon, 2015, Fig. 2). On the other hand, by construction, our contact set estimator should be asymptotically unable to detect deviations of γ from zero that are of smaller order than the tuning parameter $\delta_{m,n}$. This may lead our bootstrap procedure to deliver critical values that more closely approximate the (much larger) upper quantiles of the limit distribution of $M_{m,n}$ at $\gamma = 0$ when in fact γ is positive, leading to rejection rates well below nominal size. While the above reasoning is informal and heuristic, it seems broadly consistent with findings of Andrews and Guggenberger (2010), who showed that subsampling procedures can fail to achieve uniform similarity (while still being pointwise similar on the boundary, like our procedure) in situations where the limit distribution to be approximated varies discontinuously with a nuisance parameter.

6. Final remarks

We have been concerned in this paper with the problem of testing whether a ratio of pdfs is nonincreasing. We proposed a bootstrap procedure based on preliminary estimation of a contact set that can deliver substantially greater power than existing tests based on fixed critical values. Numerical simulations indicate that our procedure remains conservative when $p = 1$ or $p = 2$.

It may be possible to adapt the methods developed here to more general hypothesis testing problems that can be formulated in terms of the concavity of some estimable function R, not necessarily an odc.

If we have an estimator R_n of R such that $n^{1/2}(R_n - R)$ converges weakly to a continuous limit then, following the approach taken in this paper, it should be possible to use the functional delta method to determine the limit distribution of a test statistic $M_n := n^{1/2}\|\mathcal{M}R_n - R_n\|_p$, and to use Lemma 3.2 to motivate a bootstrap procedure based on preliminary estimation of a suitable contact set. A recent paper by Seo (2018) takes this approach to construct a more powerful bootstrap version of a test of stochastic monotonicity proposed by Delgado and Escanciano (2012). There is an additional level of dimensionality to her problem, so the relevant contact set turns out to be a subset of the four dimensional unit hypercube. Similar improvements can presumably be made to a test of conditional stochastic dominance also proposed by Delgado and Escanciano (2013). More broadly, our results may be relevant in any situation where the lcm operator is used to construct a statistical test of concavity.

Acknowledgements

We thank Zhonglin Li, Juwon Seo, and Wooyoung Kim for research assistance, and Andres Santos and seminar participants at the University of Texas at Austin, Hong Kong University of Science and Technology, University of Tokyo, University of Sydney, Pennsylvania State University, University College London, University of Copenhagen, Aarhus University, Hitotsubashi University, National University of Singapore, Center for Monetary and Fiscal Studies, New York University, and Northwestern University for helpful comments.

Appendix A. Proofs

Here we provide proofs of all results stated in Section 3.

Proof of Lemma 3.1. Linearity is obvious, so we have $\sup|\mathcal{S}f_1 - \mathcal{S}f_2| = \sup|\mathcal{S}(f_1 - f_2)|$ for $f_1, f_2 \in \ell^\infty([0,1])$. Let $g = f_1 - f_2$. Since $\mathcal{S}g(u,v,w)$ is a convex combination of $g(v)$ and $g(w)$, it is bounded in absolute value by $\max\{|g(v)|, |g(w)|\} \leq \sup|g|$. And since $\mathcal{S}g(u,u,u) = g(u)$, we have $g(u) \leq \sup|\mathcal{S}g|$. Consequently, $\sup|\mathcal{S}g| = \sup|g|$, and our claim is proved. □

Proof of Lemma 3.2. Suppose first that R is affine in a neighborhood of u. In this case Lemma 2.1 implies that $\mathcal{M}'_R h(u) = \mathcal{M}_{[a_{R,u}, b_{R,u}]}h(u)$. Applying a result of Carolan (2002, Lemma 1) expressing the lcm as a supremum of secant segments, we may write

$$\mathcal{M}_{[a_{R,u}, b_{R,u}]}h(u) = \sup_{a_{R,u} \leq v \leq u} \sup_{u \leq w \leq b_{R,u}} \mathcal{S}h(u,v,w).$$

Since R is concave, the rectangle $[a_{R,u}, u] \times [u, b_{R,u}]$ is precisely the cross-section $B(u)$, and our claim is proved. Next suppose that R is not affine in a neighborhood of u. Since R is concave, for all $(v,w) \in B(u)$ we must have either $v = u$ or $w = u$, or both, and so $\sup_{(v,w)\in B(u)} \mathcal{S}h(u,v,w) = h(u)$. But Lemma 2.1 implies that $\mathcal{M}'_R h(u) = h(u)$, and so our claim is proved in this case also. □

Proof of Lemma 3.3. Since $B_{m,n} \subseteq B'_{m,n}$ by construction, it suffices to show that $P(B'_{m,n} \subseteq B^\epsilon) \to 1$ and that $P(B \subseteq B_{m,n}) \to 1$. We first show that $P(B'_{m,n} \subseteq B^\epsilon) \to 1$. Since $\tilde{\mathcal{S}}R$ is continuous and is equal to zero precisely on the contact set B, we have $\inf_{a \in A \setminus B^\epsilon} |\tilde{\mathcal{S}}R(a)| > 0$. The operator $\tilde{S}$ is continuous, so Lemma 2.2 and the continuous mapping theorem imply that $\tilde{\mathcal{S}}R_{m,n}$ converges uniformly in probability to $\tilde{\mathcal{S}}R$, yielding

$$\sup_{a \in B'_{m,n}} |\tilde{\mathcal{S}}R(a)| = \sup_{a \in B'_{m,n}} |\tilde{\mathcal{S}}R_{m,n}(a)| + o_p(1) \leq \delta_{m,n}\eta_u^{1/2} + o_p(1) = o_p(1),$$

the last equality following from Assumption 3.1. It follows that

$$P\left\{\sup_{a \in B'_{m,n}} |\tilde{\mathcal{S}}R(a)| < \inf_{a \in A \setminus B^\epsilon} |\tilde{\mathcal{S}}R(a)|\right\} \to 1.$$

Consequently, $P(B'_{m,n} \cap (A \setminus B^\epsilon) = \varnothing) \to 1$, and so $P(B'_{m,n} \subseteq B^\epsilon) \to 1$.

We next show that $P(B \subseteq B_{m,n}) \to 1$. Using the linearity of $\tilde{\mathcal{S}}$ and the fact that $\tilde{\mathcal{S}}R(a) = 0$ for all $a \in B$, we find that

$$\sup_{a\in B} |\tilde{\mathcal{S}}R_{m,n}(a)| = c_{m,n}^{-1} \sup_{a\in B} |\tilde{\mathcal{S}}\left(c_{m,n}(R_{m,n} - R)\right)(a)|.$$

Therefore, since $\tilde{\mathcal{S}}(c_{m,n}(R_{m,n} - R)) \rightsquigarrow \tilde{\mathcal{S}}T$ by Lemma 2.2 and the continuous mapping theorem, we conclude in view of Assumption 3.1 that $\sup_{a\in B} |\tilde{\mathcal{S}}R_{m,n}(a)| = o_p(\delta_{m,n})$. This shows that $P(B \subseteq B'_{m,n}) \to 1$. Further, since R is concave, we may use the triangle inequality to write

$$\sup_{u\in[0,1]} |\mathcal{M}R_{m,n}(u) - R_{m,n}(u)| \leq \sup_{u\in[0,1]} |\mathcal{M}R_{m,n}(u) - \mathcal{M}R(u)| + \sup_{u\in[0,1]} |R_{m,n}(u) - R(u)|.$$

Both terms on the right-hand side of this inequality are $o_p(\delta_{m,n})$ under Assumption 3.1, and so $P(\mathcal{M}R_{m,n}(u) \leq R_{m,n}(u) + \delta_{m,n}) \to 1$ for every $u \in [0,1]$. Combined with the fact that $P(B \subseteq B'_{m,n}) \to 1$, this shows that $P(B \subseteq B_{m,n}) \to 1$. □

The proof of Lemma 3.4 makes use of the following lemma establishing Hadamard differentiability of the quantile-quantile transformation. We provide a short proof applying lemmas of van der Vaart and Wellner (1996) and van der Vaart (1998) for convenience, but the result is not really new, and can be traced back to a doctoral thesis of Reeds (1976); see Dudley (1992, p. 405).

Lemma A.1. *Let $D \subset \ell^\infty([0,1])$ be the collection of all restrictions to $[0,1]$ of cdfs assigning probability one to $(0,1]$. Consider the map $\phi : D \times D \to \ell^\infty([0,1])$ defined by*

$$\phi(F,G) = F \circ G^{-1}, \quad (F,G) \in D \times D,$$

where $G^{-1}(u) = \inf\{x \in [0,1] : G(x) \geq u\}$. This map is Hadamard differentiable tangentially to $C([0,1]) \times C([0,1])$ at any $(F,G) \in D \times D$ such that F has continuous derivative f and G is the uniform distribution on $[0,1]$. Its derivative is

$$\phi'_{(F,G)}(h_1,h_2) = h_1 - f \cdot h_2, \quad (h_1,h_2) \in C([0,1]) \times C([0,1]).$$

Proof of Lemma A.1. The map $D \times D \ni (F,G) \mapsto (F, G^{-1}) \in D \times \ell^\infty([0,1])$ is Hadamard differentiable tangentially to $\ell^\infty([0,1]) \times C([0,1])$ at any $(F,G) \in D \times D$ such that G is the uniform distribution on $[0,1]$ by Lemma 21.4(ii) of van der Vaart (1998), with derivative $(h_1,h_2) \mapsto (h_1, -h_2)$. And the map $D \times \ell^\infty([0,1]) \ni (F,G) \mapsto F \circ G \in \ell^\infty([0,1])$ is Hadamard differentiable tangentially to $C([0,1]) \times \ell^\infty([0,1])$ at any $(F,G) \in D \times \ell^\infty([0,1])$ such that F has continuous derivative f by Lemma 3.9.27 of van der Vaart and Wellner (1996), with derivative $(h_1,h_2) \mapsto h_1 \circ G + (f \circ G) \cdot h_2$. Since ϕ is the composition of these two maps, our desired result follows from Lemma 3.9.3 of van der Vaart and Wellner (1996), a chain rule for Hadamard derivatives. □

Proof of Lemma 3.4. The odcs $R_{m,n}$ and $R^*_{m,n}$ are unaffected with probability one if we replace our observations X_i and Y_j with $G(X_i)$ and $G(Y_j)$, due to the fact that G is continuous and strictly increasing on the common support of F and G. We may therefore assume without loss of generality that G is the uniform distribution on $[0,1]$. Let B_1 and B_2 denote independent standard Brownian bridges on $[0,1]$. Due to the independence of the X_i's and Y_j's and our assumption that $n/(m+n) \to \lambda$, we have

$$c_{m,n}\left((F_m, G_n) - (F,G)\right) \rightsquigarrow \left(\lambda^{1/2} B_1 \circ F, (1-\lambda)^{1/2} B_2 \circ G\right) \tag{A.1}$$

in $\ell^\infty([0,1]) \times \ell^\infty([0,1])$ by Donsker's theorem, and

$$c_{m,n}\left((F^*_m, G^*_n) - (F_m,G_n)\right) \rightsquigarrow \left(\lambda^{1/2} B_1 \circ F, (1-\lambda)^{1/2} B_2 \circ G\right) \tag{A.2}$$

in $\ell^\infty([0,1]) \times \ell^\infty([0,1])$ conditional on the data in probability by Theorem 2.6 of Kosorok (2008). Let ϕ be the map defined in Lemma A.1, established there to be Hadamard differentiable at (F,G) tangentially to

$C([0,1]) \times C([0,1])$. The functional delta method for the bootstrap (Kosorok, 2008, Thm. 2.9) allows us to deduce from this Hadamard differentiability and the convergences (A.1) and (A.2) that

$$c_{m,n}\left(\phi(F_m^*, G_n^*) - \phi(F_m, G_n)\right) \rightsquigarrow \phi'_{(F,G)}\left(\lambda^{1/2} B_1 \circ F, (1-\lambda)^{1/2} B_2 \circ G\right)$$

in $\ell^\infty([0,1])$ conditional on the data in probability. Our desired result now follows by observing that

$$T^*_{m,n} = c_{m,n}\left(\phi(F_m^*, G_n^*) - \phi(F_m, G_n)\right),$$

and that

$$T = \phi'_{(F,G)}\left(\lambda^{1/2} B_1 \circ F, (1-\lambda)^{1/2} B_2 \circ G\right)$$

in law when G is the uniform distribution on $[0,1]$. □

Proof of Theorem 3.1. By Lemma 3.3 there exists a sequence $\epsilon_n \downarrow 0$ such that

$$P(B \subseteq B_{m,n} \subseteq B^{\epsilon_n}) \to 1.$$

Let $g_n : \ell^\infty([0,1]) \to \mathbb{R}$ be the map $g_n(f) = \| \sup_{(v,w)\in B^{\epsilon_n}(\cdot)} \tilde{\mathcal{S}} f(\cdot, v, w)\|_p$, and let $g = \mathcal{A}'_R$, so that in view of Lemma 3.2 we have

$$P(g(T^*_{m,n}) \leq M^*_{m,n} \leq g_n(T^*_{m,n})) \to 1. \tag{A.3}$$

We will show that, for any sequence f_n in $\ell^\infty([0,1])$ with $f_n \to f_\infty \in C([0,1])$, we have

$$g_n(f_n) \to g(f_\infty). \tag{A.4}$$

The convergence (A.4) is the result of the following argument:

$$\begin{aligned}
|g_n(f_n) - g(f_\infty)| &\leq |g_n(f_n) - g_n(f_\infty)| + |g_n(f_\infty) - g(f_\infty)| \\
&\leq \sup_{a\in B^{\epsilon_n}} |\tilde{S} f_n(a) - \tilde{S} f_\infty(a)| + |g_n(f_\infty) - g(f_\infty)| \\
&\leq 2\|f_n - f_\infty\|_\infty + |g_n(f_\infty) - g(f_\infty)| \\
&\leq 2\|f_n - f_\infty\|_\infty + \sup_{(a_n,a'_n)\in B\times B^{\epsilon_n}:\|a_n - a'_n\|\leq \epsilon_n} |\tilde{\mathcal{S}} f_\infty(a_n) - \tilde{\mathcal{S}} f_\infty(a'_n)| \to 0.
\end{aligned}$$

Here, the first and second inequalities follow from the triangle inequality, the third inequality holds by Lemma 3.1, the fourth inequality holds by the definition of g_n and g, and the convergence to zero holds because $f_n \to f_\infty$, $\epsilon_n \downarrow 0$ and $\tilde{S} f_\infty$ is uniformly continuous. Lemma 3.4 together with (A.4) allows us to apply the extended continuous mapping theorem (see e.g. Dümbgen, 1993, p. 136) to obtain $g_n(T^*_{m,n}) \rightsquigarrow g(T)$ and $g(T^*_{m,n}) \rightsquigarrow g(T)$ conditional on the data in probability. In view of (A.3) and the definition of g, we conclude that $M^*_{m,n} \rightsquigarrow \mathcal{A}'_R T$ conditional on the data in probability.

It is clear from Theorem 3.1 of Beare and Moon (2015) that when R is not strictly concave the distribution function of $\mathcal{A}'_R T$ is continuous everywhere and strictly increasing on $[0,\infty)$. Continuity everywhere combined with the weak convergence $M^*_{m,n} \rightsquigarrow \mathcal{A}'_R T$ conditional on the data in probability implies (Kosorok, 2008, Lemma 10.11(i)) that

$$\sup_{x\in\mathbb{R}} \left| P(M^*_{m,n} \leq x \mid X_1,\dots,X_m,Y_1,\dots,Y_n) - P(\mathcal{A}'_R T \leq x) \right| = o_p(1). \tag{A.5}$$

Let $\mu(\alpha) = \inf\{x : P(\mathcal{A}'_R T \leq x) \geq 1-\alpha\}$, the $(1-\alpha)$-quantile of $\mathcal{A}'_R T$. Since the distribution function of $\mathcal{A}'_R T$ is strictly increasing at $\mu(\alpha)$, the continuous mapping theorem applied to (A.5) yields $\mu_{m,n}(\alpha) = \mu(\alpha) + o_p(1)$. It now follows from the weak convergence $M_{m,n} \rightsquigarrow \mathcal{A}'_R T$ ensured by Theorem 2.1, and the continuity of the distribution function of $\mathcal{A}'_R T$ at $\mu(\alpha)$, that $P(M_{m,n} > \mu_{m,n}(\alpha)) \to \alpha$ as claimed. □

The proof of Theorem 3.2 makes use of Lemmas A.2, A.3, and A.4 below, which are local alternative versions of Lemmas 2.2, 3.3 and 3.4 respectively. The setting is that of Section 3.4.

Lemma A.2. *Under the sequence* $(R^{(m)} : m \geq 1)$*, as* $m \wedge n \to \infty$ *with* $n/(m+n) \to \lambda \in (0,1)$*, we have* $c_{m,n}(R_{m,n} - R) \rightsquigarrow T + \lambda^{1/2}\Delta$*, where* T *is the process given in Lemma 2.2.*

Proof. Without loss of generality we again normalize G to be the uniform distribution on $[0,1]$ and set $F = R$ and $F^{(m)} = R^{(m)}$ for each m. A version of Donsker's theorem applicable under drifting sequences of distributions (van der Vaart and Wellner, 1996, Lem. 2.8.7) ensures that

$$\sqrt{m}(F_m - F^{(m)}) \rightsquigarrow B_1 \circ F \quad \text{and} \quad \sqrt{n}(G_n - I) \rightsquigarrow B_2, \tag{A.6}$$

where $I(u) = u$ for all $u \in [0,1]$. The two convergences hold jointly and B_1 and B_2 are independent from each other because the empirical cdfs are constructed from independent samples. Since $\sqrt{m}(F^{(m)} - F) = \Delta$, we thus obtain

$$\sqrt{m}(F_m - F) \rightsquigarrow B_1 \circ F + \Delta \quad \text{and} \quad \sqrt{n}(G_n - I) \rightsquigarrow B_2 \tag{A.7}$$

jointly. In view of the Hadamard differentiability of the quantile-quantile transform established in Lemma A.1, we may now apply the functional delta method to obtain

$$c_{m,n}(R_{m,n} - R) \rightsquigarrow \lambda^{1/2}(B_1 \circ R + \Delta) - (1-\lambda)^{1/2} R' \cdot B_2. \tag{A.8}$$

We conclude by observing that the weak limit in (A.8) is equal in law to $T + \lambda^{1/2}\Delta$. □

Lemma A.3. *Suppose Assumption 3.1 is satisfied. Under the sequence* $(R^{(m)} : m \geq 1)$*, as* $m \wedge n \to \infty$ *with* $n/(m+n) \to \lambda \in (0,1)$*, we have* $P(B \subseteq B_{m,n} \subseteq B^{\epsilon}) \to 1$ *for any* $\epsilon > 0$*.*

Proof. The proof is the same as that of Lemma 3.3, but with Lemma A.2 invoked in place of Lemma 2.2, and $T + \lambda^{1/2}\Delta$ used in place of T. □

Lemma A.4. *Under the sequence* $(R^{(m)} : m \geq 1)$*, as* $m \wedge n \to \infty$ *with* $n/(m+n) \to \lambda \in (0,1)$*, we have* $T^*_{m,n} \rightsquigarrow T$ *conditional on the data in probability.*

Proof. Without loss of generality we again normalize G to be the uniform distribution on $[0,1]$ and set $F = R$ and $F^{(m)} = R^{(m)}$ for each m. The joint weak convergence

$$\sqrt{m}(F^*_m - F_m) \rightsquigarrow B_1 \circ F \quad \text{and} \quad \sqrt{n}(G^*_n - G_n) \rightsquigarrow B_2 \tag{A.9}$$

conditional on the data in probability that obtained when $\Delta = 0$ continues to hold under our drifting sequence of distributions due to the uniform Donsker property of the collection of indicators of half-intervals (Giné and Zinn, 1991). Therefore, in view of the Hadamard differentiability of the quantile-quantile transform established in Lemma A.1, we may deduce the desired conditional weak convergence $T^*_{m,n} \rightsquigarrow T$ by applying the functional delta method for the bootstrap (for a version that applies under drifting sequences of distributions, see e.g. the second part of Proposition 1 of Dümbgen (1993) applied to Hadamard differentiable maps). □

Proof of Theorem 3.2. Since $M_{m,n} = c_{m,n}(\mathcal{A}R_{m,n} - \mathcal{A}R)$, part (a) follows from Lemma A.2 by applying the functional delta method. Part (b) can be proved in the same way as Theorem 3.1, but with Lemmas A.3 and A.4 invoked in place of Lemmas 3.3 and 3.4 respectively. Part (c) is trivial when R is strictly concave because $M_{m,n}$ converges in probability to zero, implying that $\limsup P(M_{m,n} > \mu(\alpha)) = 0$. When R is not strictly concave, by appealing to part (b) for the conditional weak convergence $M^*_{m,n} \rightsquigarrow \mathcal{A}'_R T$ and arguing as in the final paragraph of the proof of Theorem 3.1, we deduce that $\mu_{m,n}(\alpha) = \mu(\alpha) + o_p(1)$. It then follows from the weak convergence $M_{m,n} \rightsquigarrow \mathcal{A}'_R(T + \lambda^{1/2}\Delta)$ established in part (a) that

$$P(M_{m,n} > \nu(\alpha)) \to P(\mathcal{A}'_R(T + \lambda^{1/2}\Delta) > \nu(\alpha)) \tag{A.10}$$

and

$$P(M_{m,n} > \mu_{m,n}(\alpha)) \to P(\mathcal{A}'_R(T + \lambda^{1/2}\Delta) > \mu(\alpha)). \tag{A.11}$$

Theorem 4.1 of Beare and Moon (2015) implies that $\mu(\alpha) \leq \nu(\alpha)$, so we deduce that part (c) is satisfied, with the limits superior and inferior being proper limits.

It remains to verify that the limit in (A.10) is strictly less than the limit in (A.11) when we have neither R strictly concave nor $R = I$. This will be the case if $\mu(\alpha) < \nu(\alpha)$. The proof of Theorem 4.1 of Beare and Moon (2015) constructs a random variable Z equal in law to $\mathcal{A}'_R T$ such that $Z \leq \mathcal{A}'_I B$. One step in that proof applies a collection of inequalities between geometric and arithmetic means. All inequalities hold with equality if and only if all affine segments of R have slope one. But inequality (A.10) in the proof of Theorem 4.1 of Beare and Moon (2015) holds with equality if and only if R is comprised purely of affine segments. Thus if $R \neq I$ then at least one of the aforementioned inequalities must be strict, implying that $Z < \mathcal{A}'_I B$. Consequently, all interior quantiles of $\mathcal{A}'_R T$ lie strictly below the corresponding quantiles of $\mathcal{A}'_I B$, and hence $\mu(\alpha) < \nu(\alpha)$. □

Appendix B. Variance formula used in threshold construction

As discussed in Section 3.2, to construct the threshold in the estimated contact set it is advisable to use an estimator of the variance of $c_{m,n}\tilde{\mathcal{S}}R_{m,n}(u,v,w)$. From the covariance kernel of the process T appearing in Lemma 2.2, we derive the following expression for the limiting variance of $c_{m,n}\tilde{\mathcal{S}}R_{m,n}(u,v,w)$ when $v < w$:

$$\begin{aligned} V(u,v,w) = {} & \frac{\lambda(w-u)^2}{(w-v)^2}(R(u)-R(v)) + \frac{\lambda(u-v)^2}{(w-v)^2}(R(w)-R(u)) \\ & + \frac{(1-\lambda)(w-u)^2}{(w-v)^2}(R'(v)^2 v + R'(u)^2 u - 2R'(u)R'(v)v) \\ & + \frac{(1-\lambda)(u-v)^2}{(w-v)^2}(R'(w)^2 w + R'(u)^2 u - 2R'(u)R'(w)u) \\ & + \frac{2(1-\lambda)(w-u)(u-v)}{(w-v)^2}(R'(w)-R'(u))(R'(v)v - R'(u)u). \end{aligned}$$

When $v = w$, we instead have $V(u,v,w) = 0$. We estimate $V(u,v,w)$ with $V_{m,n}(u,v,w)$, which has the same formula as $V(u,v,w)$ except that λ is replaced by $n/(m+n)$, R by $R_{m,n}$, and R' by

$$R'_{m,n}(u) = \frac{R_{m,n}(\min\{1, u+b_n\}) - R_{m,n}(\max\{0, u-b_n\})}{\min\{1, u+b_n\} - \max\{0, u-b_n\}},$$

where $b_n = [n^{1/3}]/n$ with $[\cdot]$ rounding up to the nearest integer.

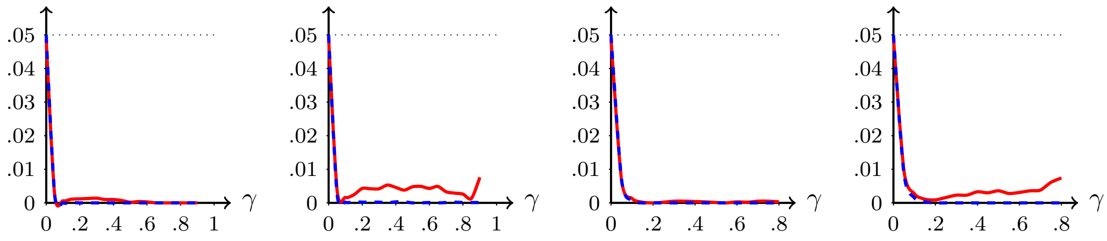
Appendix C. Additional numerical simulations

We repeated the numerical simulations reported in Section 5 with larger sample sizes. Specifically, we increased m and n from 200 to 400. Figures C.1 and C.2 are the counterparts to Figures 5.3 and 5.4 with $m = n = 400$ instead of $m = n = 200$. The null rejection rates with $m = n = 400$ reported in Figure C.1 are similar to those with $m = n = 200$ reported in Figure 5.3. In particular, we again see that the rejection rates drop rapidly to zero as γ rises above zero. With $p = 2$, the rejection rates tend to be a little higher at γ values further away from zero when the sample sizes are larger. Comparing Figures 5.4 and C.2, we see that the power of all tests is larger when the sample sizes are larger, as expected. Our bootstrap tests continue to be much more powerful than the fixed critical value tests at the larger samples sizes.

References

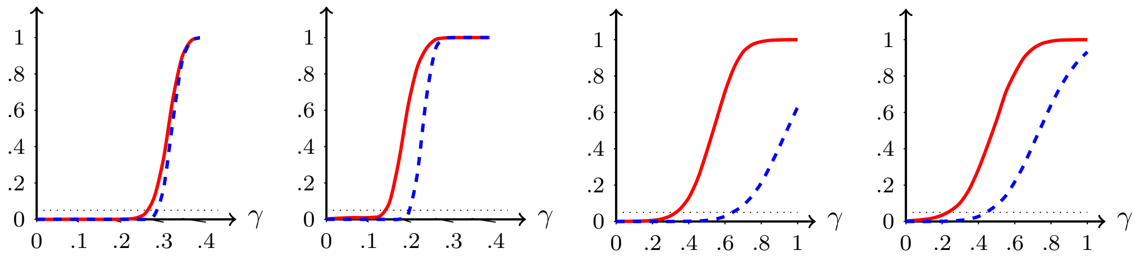
Anderson, G., 1996. Nonparametric tests of stochastic dominance in income distributions. Econometrica 64, 1183–1193.

Anderson, G., Linton, O., Whang, Y. -J., 2012. Nonparametric estimation and inference about the overlap of two distributions. J. Econometrics 171, 1–23.



(a) R^{I}_{γ}, $p = 1$ (b) R^{I}_{γ}, $p = 2$ (c) R^{II}_{γ}, $p = 1$ (d) R^{II}_{γ}, $p = 2$

Figure C.1: Null rejection rates with fixed critical values (dashed) and bootstrap critical values (solid), with sample sizes $m = n = 400$ and $N = 100$ gridpoints.



(a) $R^{\mathrm{III}}_{\gamma}$, $p = 1$ (b) $R^{\mathrm{III}}_{\gamma}$, $p = 2$ (a) R^{IV}, $p = 1$ (b) R^{IV}, $p = 2$

Figure C.2: Power curves with fixed critical values (dashed) and bootstrap critical values (solid), with sample sizes $m = n = 400$ and $N = 100$ gridpoints.

Andrews, D. W. K., Guggenberger, P., 2010. Asymptotic size and a problem with subsampling and with the m out of n bootstrap. Econometric Theory 26, 426–468.

Andrews, D. W. K., Shi, X., 2013. Inference based on conditional moment inequalities. Econometrica 81, 609–666.

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

Beare, B. K., 2011. Measure preserving derivatives and the pricing kernel puzzle. J. Math. Econom. 47, 689–697.

Beare, B. K., Moon, J. -M., 2015. Nonparametric tests of density ratio ordering. Econometric Theory 31, 471–492.

Beare, B. K., Schmidt, L. D. W., 2016. An empirical test of pricing kernel monotonicity. J. Appl. Econometrics 31, 338–356.

Beare, B. K., Dossani, A., 2018. Option augmented density forecasts of market returns with monotone pricing kernel. Quant. Finance 18, 623–635.

Beare, B. K., Fang, Z., 2017. Weak convergence of the least concave majorant of estimators for a concave distribution function. Electron. J. Stat. 11, 3841–3870.

Beutner, E., Wu, W. B., Zähle, H., 2012. Asymptotics for statistical functionals of long-memory sequences. Stochastic Process. Appl. 122, 910–929.

Beutner, E., Zähle, H., 2010. A modified functional delta method and its application to the estimation of risk functionals. J. Multivariate Anal. 101, 2452–2463.

Beutner, E., Zähle, H., 2012. Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. Bernoulli 18, 803–822.

Beutner, E., Zähle, H., 2016. Functional delta-method for the bootstrap of quasi-Hadamard differentiable functions. Electron. J. Stat. 10, 1181–1222.

Carolan, C. A., 2002. The least concave majorant of the empirical distribution function. Canad. J. Statist. 30, 317–328.

Carolan, C. A., Tebbs, J. M., 2005. Nonparametric tests for and against likelihood ratio ordering in the two sample problem. Biometrika 92, 159–171.

Chang, M., Lee, S., Whang, Y. -J., 2015. Nonparametric tests of conditional treatment effects with an application to single-sex schooling on academic achievements. Econom. J. 18, 307–346.

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

Dehling, H., Philipp, W., 2002. Empirical process techniques for dependent data. In: Dehling, H., Mikosch, T., Sørensen, M. (Eds.), Empirical Process Techniques for Dependent Data, Birkhäuser, pp. 3–113.

Delgado, M. A., Escanciano, J. C., 2012. Distribution-free tests of stochastic monotonicity. J. Econometrics 170, 68–75.

Delgado, M. A., Escanciano, J. C., 2013. Conditional stochastic dominance testing. J. Bus. Econom. Statist. 31, 16–28.
Donald, S. G., Hsu, Y. -C., 2016. Improving the power of tests of stochastic dominance. Econometric Rev. 35, 553–585.
Dudley, R. M., 1992. Nonlinear functions of empirical measures. In: Dudley, R. M., Hahn, M. G., Kuelbs, J. (Eds.), Probability in Banach Spaces, 8: Proceedings of the Eighth International Conference, Springer, pp. 403–410.
Dümbgen, L., 1993. On nondifferentiable functions and the bootstrap. Probab. Theory Related Fields 95, 125–140.
Dykstra, R., Kochar, S., Robertson, T., 1995. Inference for likelihood ratio ordering in the two-sample problem. J. Amer. Statist. Assoc. 90, 1034–1040.
Fang, Z., Santos, A., 2016. Inference on directionally differentiable functions. arXiv preprint arXiv:1404.3763v2 [math.ST].
Giacomini, R., Politis, D. N., White, H., 2013. A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. Econometric Theory 29, 567–589.
Giné, E., Zinn, J., 1991. Gaussian characterization of uniform Donsker classes of functions. Ann. Probab. 19, 758–782.
Hsieh, F., Turnbull, B. W., 1996. Nonparametric and semiparametric estimation of the receiver operating characteristic curve. Ann. Statist. 24, 25–40.
Kosorok, M. R., 2008. Introduction to Empirical Processes and Semiparametric Inference. Springer.
Lee, S., Song, K., Whang, Y. -J., 2018. Testing for a general class of functional inequalities. Econometric Theory, in press.
Linton, O., Maasoumi, E., Whang, Y. -J., 2005. Consistent testing for stochastic dominance under general sampling schemes. Rev. Econom. Stud. 72, 735–765.
Linton, O., Song, K., Whang, Y. -J., 2010. An improved bootstrap test of stochastic dominance. J. Econometrics 154, 186–202.
Maasoumi, E., 2001. Parametric and nonparametric tests of limited domain and ordered hypotheses in economics. In: Baltagi, B. (Ed.), A Companion to Theoretical Econometrics, Blackwell, pp. 538–556.
Radulović, D., 2002. On the bootstrap and empirical processes for dependent sequences. In: Dehling, H., Mikosch, T., Sørensen, M. (Eds.), Empirical Process Techniques for Dependent Data, Birkhäuser, pp. 345–364.
Reeds, J. A., 1976. On the definition of von Mises functionals. Ph.D. thesis, Harvard University.
Roosen, J., Hennessy, D. A., 2004. Testing for the monotone likelihood ratio assumption. J. Bus. Econom. Statist. 28, 358–366.
Seo, J., 2018. Tests of stochastic monotonicity with improved size and power properties. J. Econometrics, in press.
Shapiro, A., 1990. On concepts of directional differentiability. J. Optim. Theory Appl. 66, 477–487.
Shapiro, A., 1991. Asymptotic analysis of stochastic programs. Ann. Oper. Res. 30, 169–186.
van der Vaart, A., 1998. Asymptotic Statistics. Cambridge University Press.
van der Vaart, A. W., Wellner, J. A., 1996. Weak Convergence and Empirical Processes. Springer.
Volgushev, S., Shao, X., 2014. A general approach to the joint asymptotic analysis of statistics from subsamples. Electron. J. Stat. 8, 390–431.