Bootstrap tests of stochastic dominance with asymptotic similarity on the boundary

Oliver Linton
Kyungchul Song
Yoon-Jae Whang

Bootstrap Tests of Stochastic Dominance with Asymptotic Similarity on the Boundary

Oliver Linton[1], Kyungchul Song[2] and Yoon-Jae Whang[3]

London School of Economics, University of Pennsylvania and Seoul National University

February 19, 2008

Abstract

We propose a new method of testing stochastic dominance which improves on existing tests based on bootstrap or subsampling. Our test requires estimation of the contact sets between the marginal distributions. Our tests have asymptotic sizes that are exactly equal to the nominal level uniformly over the boundary points of the null hypothesis and are therefore valid over the whole null hypothesis. We also allow the prospects to be indexed by infinite as well as finite dimensional unknown parameters, so that the variables may be residuals from nonparametric and semiparametric models. Our simulation results show that our tests are indeed more powerful than the existing subsampling and recentered bootstrap.

Key words and Phrases: Set estimation; Size of test; Unbiasedness; Similarity; Bootstrap; Subsampling.

JEL Classifications: C12, C14, C52.

[1] Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. E-mail address: `o.linton@lse.ac.uk`. Thanks to the ESRC and Leverhulme foundations for financial support.

[2] Department of Economics, University of Pennsylvania, 528 McNeil Building, 3718 Locust Walk, Philadelphia, Pennsylvania 19104-6297. Email: `kysong@sas.upenn.edu`

[3] Department of Economics, Seoul National University, Seoul 151-742, Korea. E-mail address: `whang@snu.ac.kr`. Thanks to the Korea Research Foundation for financial support.

1   Introduction

There has been a growth of interest in testing stochastic dominance relations among variables such as investment strategies and income distributions. The ordering by stochastic dominance covers a large class of utility functions and hence is more useful in general than the notions of partial orderings that are specific to a certain class of utility functions. However, the main difficulty in testing stochastic dominance lies in the complexity that arises in computing approximate critical values of the test. A stochastic dominance relation is a weak inequality relation between distribution functions or integral transform of distribution functions. Unlike nonparametric or semiparametric tests that are based on the equality of functions, the convergence of test statistics of Kolmogorov-Smirnov type or Cramér-von Mises type is not uniform over the probabilities under the null hypothesis. Discontinuity of convergence arises precisely between the "interior points" of the null hypothesis and the "boundary points" of the null hypothesis, where boundary points indicate those probabilities under which all pairs of the distribution functions meet at least one point in the interior of the support and interior points indicate probabilities under which at least one pair of the distribution functions do not meet at any point in the interior of the support. In general, the boundary points do not coincide with the least favorable subset of null hypothesis, i.e., the set of probabilities under which all the pairs of competing distribution functions are equal. As noted by Linton, Maasoumi, and Whang (2005), henceforth LMW, the usual approach of recentered bootstrap or wild bootstrap does not provide tests with asymptotically valid sizes uniformly over the probabilities under the null hypothesis; the asymptotic size of the tests are exact only on the least favorable points and invalid for other points under the null hypothesis. LMW proposed a subsampling method to obtain tests with exact asymptotic sizes over the boundary points and asymptotically valid for each probability under the null hypothesis.

This paper proposes a bootstrap-based test that has asymptotically valid size uniformly under the null hypothesis and asymptotically exact size on the boundary points. The method is based on bootstrap test statistics that are constructed to mimic the phenomenon of discontinuous convergence by using the estimated boundary points. This requires the estimation of a certain set, we call the "contact set"; set estimation is a topic of considerable recent interest in microeconometrics, see for example Moon and Schorfheide (2006), Chernozhukov, Hong, and Tamer (2007), and Fan and Park (2007). Our method of bootstrap has the following remarkable properties. First, it is asymptotically similar on the boundary. Second, more importantly, the bootstrap method provides asymptotically exact sizes. Furthermore, we can make the rejection probability under the null hypothesis converge to the true rejection

probability only slightly less fast than the asymptotic approximation. This implies that the convergence of rejection probability can be made faster than that of subsampling.

The proposed test of stochastic dominance admits variables that contain unknown finite-dimensional and infinite-dimensional parameters as long as those parameters have consistent estimators. Hence the proposed test can be used to test stochastic dominance relations among variables in the form of regression errors from the semiparametric regression models such as single index restrictions or partially parametric regressions. For example, one can detect the influence of a certain factor upon the stochastic dominance relation by comparing test results with and without the factor of interest while controlling for other factors. In some situations, it is desirable to control for publicly observed variables. It is a potential concern for the researcher in such cases that the result of testing may rely sensitively on the regression specification chosen to control certain factors stochastically. Semiparametric specification of the regression can be useful in this situation.

We perform Monte Carlo simulations that compare three methods: recentered bootstrap, subsampling method, and the bootstrap on the boundary domain proposed in this paper. We verify the superior performance of our method.

Testing stochastic dominance has drawn attention in the literature over the last decade. McFadden (1986), Anderson (1996) and Davidson and Duclos (2000) are among the early works that considered testing stochastic dominance. Barrett and Donald (2003) proposed a consistent bootstrap test that has an asymptotically exact size on the least favorable points for the special case where the prospects are mutually independent. LMW suggested a subsampling method that has asymptotically valid size on the boundary points and applies to variables that contain unknown finite-dimensional parameters and allows the prospects to be mutually dependent.

In Section 2, we define the null hypothesis of stochastic dominance and introduce notations. In Section 3, we suggest test statistics and develop asymptotic theory both under the null hypothesis and local alternatives. Section 4 is devoted to the bootstrap procedure, explaining the method of obtaining bootstrap test statistics and establishing their asymptotic properties. In Section 5, we describe Monte Carlo simulation studies and discuss results from them. Section 6 concludes. All the technical proofs are relegated to the appendix.

2 Stochastic Dominance

2.1 The Null Hypothesis

Let $\{X_k\}_{k=1}^K$ be a set of continuous outcome variables which may, for example, represent different points in time or different regions. Let $F_k(x)$ be the distribution function of the k-th random variable X_k. Let $D_k^{(1)}(x) = F_k(x)$ and

$$D_k^{(s)}(x) = \int_{-\infty}^x D_k^{(s-1)}(t)dt.$$

Then we say that X_1 *stochastically dominates* X_2 *at order* s if $D_1^{(s)}(x) \leq D_2^{(s)}(x)$ for all x with strict inequality for some x. This is equivalent to an ordering of expected utility over a certain class of utility functions $\mathcal{U}_s$, see LMW.

Let $D_{kl}^{(s)}(x) = D_k^{(s)}(x) - D_l^{(s)}(x)$ and define

$$d_s^* = \min_{k \neq l} \sup_{x \in \mathcal{X}} D_{kl}^{(s)}(x) \text{ and } c_s^* = \min_{k \neq l} \int_{\mathcal{X}} \max\left\{D_{kl}^{(s)}(x), 0\right\}^2 dx,$$

where $\mathcal{X}$ denotes a set contained in the union of the supports of X_k, $k = 1, 2, \ldots, K$. We do not assume that the set $\mathcal{X}$ is bounded. The null hypothesis that this paper focuses on takes the following form:

$$H_{0,A} : d_s^* \leq 0 \text{ vs. } H_{1,A} : d_s^* > 0. \tag{1}$$

The null hypothesis represents the presence of a stochastic dominance relationship between a pair of variables in $\{X_k\}_{k=1}^K$. The alternative hypothesis corresponds to no such incidence. Alternatively, we might also consider the following form of the null and alternative hypotheses

$$H_{0,B} : c_s^* = 0 \text{ vs. } H_{1,B} : c_s^* > 0. \tag{2}$$

Obviously the null hypothesis $d_s^* \leq 0$ implies the null hypothesis $c_s^* = 0$. On the contrary, when $c_s^* = 0$, $D_{kl}^{(s)}(x) \leq 0$, μ-a.e., for some pair (k, l), where μ denotes the Lebesgue measure on $\mathcal{X}$. Therefore, if the $D_k^{(s)}$'s are continuous functions, both the null hypotheses in (1) and (2) are equivalent, but in general, the null hypothesis in (1) is stronger than that in (2).

Of crucial importance in the sequel are the "contact" sets, these are denoted by $B_{kl}^{(s)}$ the subset of $\mathcal{X}$ such that $D_k^{(s)}(x)$ and $D_l^{(s)}(x)$ are equal, i.e.,

$$B_{kl}^{(s)} = \{x \in \mathcal{X} : D_{kl}^{(s)}(x) = 0\}. \tag{3}$$

These sets can be empty, can contain a countable number of isolated points, or can be a

3

countable union of disjoint intervals. The limiting distributions of our test statistics depend only these subsets of $\mathcal{X}$.

2.2   Test Statistics and Asymptotic Theory

In many cases of practical interest the outcome variable maybe the residual from some model. This arises in particular when data is limited and one may want to use a model to adjust for systematic differences. Common practice is to cut the data into subsets, say of families with two children, and then make comparisons across homogenous populations. When data are limited this can be difficult and a modelling approach can overcome this difficulty. We suppose that the variables X_k depend on an unknown finite dimensional parameter $\theta_0 \in \mathbb{R}^{d_\Theta}$ and on an infinite dimensional parameter $\tau_0 \in \mathcal{T}$ (here, $\mathcal{T}$ is a totally bounded class of functions with respect to a certain metric), so that we write $X_k = X_k(\theta_0, \tau_0)$. For example, the variable X_k may be the residual from the partially parametric regression $X_k(\theta_0, \tau_0) = Y_k - Z_{1k}^\top \theta_0 - \tau_0(Z_{2k})$ or the single index framework $X_k(\theta_0, \tau_0) = Y_k - \tau_0(Z_{1k}^\top \theta_0)$. Generically we let $X_k(\theta, \tau)$ be specified as $X_k(\theta, \tau) = \varphi_k(W; \theta, \tau)$ where $\varphi_k(\cdot; \theta, \tau)$ is a real-valued function known up to the parameter $(\theta, \tau) \in \Theta \times \mathcal{T}$. In the example of $X_k(\theta_0, \tau_0) = Y_k - \tau_0(Z_{1k}^\top \theta_0)$, W denotes the vector of Y_k and Z_{1k}, $k = 1, \ldots, K$. This set up is more general than in LMW who allowed only linear regression.

   We next specify further the precise properties that we require of the data generating process. Let $B_{\Theta \times \mathcal{T}}(\delta) = \{(\theta, \tau) \in \Theta \times \mathcal{T} : ||\theta - \theta_0|| < \delta, \sup_{P \in \mathcal{P}} ||\tau - \tau_0||_{P,2} < \delta\}$, where the norm $|| \cdot ||_{P,2}$ denotes the usual $L_2(P)$-norm. We introduce a bounded weight function $q(x)$ (see Assumption 3(ii)) and define $h_x^{(s)}(\varphi) = (x - \varphi)^{s-1} 1\{\varphi \leq x\} q(x)$. Let $N_{[]}(\varepsilon, \mathcal{T}, || \cdot ||_{P,r})$ denote the bracketing number of $\mathcal{T}$ with respect to the $L_r(P)$-norm, i.e. the smallest number of ε- brackets that are needed to cover the space $\mathcal{T}$ (e.g. van der Vaart and Wellner (1996)). The conditions in Assumptions 1 and 2 below are concerned with the data generating process of W and the map φ_k. Let $\mathcal{P}$ be the collection of all the potential distributions of W that satisfy the conditions Assumptions 1-3 below.

Assumption 1 : (i) $\{W_i\}_{i=1}^N$ is a random sample.
(ii) $\sup_{P \in \mathcal{P}} \log N_{[]}(\varepsilon, \mathcal{T}, || \cdot ||_{P,2}) \leq C \varepsilon^{-d}$ for some $d \in (0, 1]$.
(iii) For some $\delta > 0$ and $s > 0$, $\sup_{P \in \mathcal{P}} \mathbf{E}_P[\sup_{(\theta, \tau) \in B_{\Theta \times \mathcal{T}}(\delta)} |\varphi_k(W; \theta, \tau)|^{2((s-1) \vee 1) + \delta}] < \infty$.
(iv) For some $\delta > 0$, there exists a functional $\Gamma_{k,P}(x)[\theta - \theta_0, \tau - \tau_0]$ of $(\theta - \theta_0, \tau - \tau_0)$, $(\theta, \tau) \in B_{\Theta \times \mathcal{T}}(\delta)$, such that

$$\left| \mathbf{E}_P\left[h_x^{(s)}(\varphi_k(W; \theta, \tau)) \right] - \mathbf{E}_P\left[h_x^{(s)}(\varphi_k(W; \theta_0, \tau_0)) \right] - \Gamma_{k,P}(x)[\theta - \theta_0, \tau - \tau_0] \right|$$
$$\leq C_1 ||\theta - \theta_0||^2 + C_2 ||\tau - \tau_0||_{P,2}^2$$

4

with constants C_1 and C_2 that do not depend on P, and for each $\varepsilon > 0$

$$\mathrm{limsup}_{N \geq 1} \mathrm{sup}_{P \in \mathcal{P}} P \left\{ \sup_{x \in \mathcal{X}} \left| \sqrt{N} \Gamma_{k,P}(x)[\hat{\theta} - \theta_0, \hat{\tau} - \tau_0] - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \psi_{x,k,P}(W_i) \right| > \varepsilon \right\} = 0,$$

(4)

where $\psi_{x,P}(\cdot)$ satisfies that there exist $\eta, C > 0$ and $s_1 \in (0,1]$ such that for all $x \in \mathcal{X}$, $\mathbf{E}_P \left[\psi_{x,k,P}(W_i) \right] = 0$, $\sup_{P \in \mathcal{P}} ||\sup_{x \in \mathcal{X}} ||\psi_{x,k,P}||| ||_{P,2+\eta} < \infty$, and

$$\sup_{P \in \mathcal{P}} \mathbf{E}_P \left[\sup_{x_2 \in \mathcal{X}: |x-x_2| \leq \varepsilon} \left| \psi_{x,k,P}(W_i) - \psi_{x_2,k,P}(W_i) \right|^2 \right] \leq C \varepsilon^{2s_1}, \text{ for all } \varepsilon > 0.$$

The bracketing entropy condition for $\mathcal{T}$ in (ii) is satisfied by many classes of functions. For example, when $\mathcal{T}$ is a Hölder class of smoothness α with the common domain of $\tau(\cdot) \in \mathcal{T}$ that is convex, and bounded in the $d_{\mathcal{T}}$-dim Euclidean space with $d_{\mathcal{T}}/\alpha \in (0,1]$ (e.g. Corollary 2.7.2 in van der Vaart and Wellner (1996)), the bracketing entropy condition holds. The condition is also satisfied when $\mathcal{T}$ is contained in a class of uniformly bounded functions of bounded variation. Condition (iii) is a moment condition with local uniform boundedness. The moment condition is widely used in the literature of semiparametric inferences. In the example of single-index restrictions where $Y_k = \tau_0(Z_{1k}^\top \theta_0) + \varepsilon_k$, we can write $\varphi(w; \theta, \tau) = \tau_0(z_{1k}^\top \theta_0) - \tau(z_{1k}^\top \theta) + \varepsilon_k$. If τ is uniformly bounded in the neighborhood of τ_0 in $L_2(P)$, the moment condition is immediately satisfied when $\mathbf{E}[|\varepsilon_i|^{2((s-1)\vee 1)+\delta}] < \infty$. Or if τ is Lipschitz continuous with a uniformly bounded Lipschitz coefficient and $|\tau_0(v) - \tau(v)| \leq C|v|^p$, $p \geq 1$, and $\mathbf{E}[||Z_{1k}||^{p \vee (2((s-1)\vee 1)+\delta)}] < \infty$, then the moment condition holds. We may check this condition for other semiparametric specifications in a similar manner. Condition (iv) is concerned with the pathwise differentiability of the functional $\int h_x^{(s)}(\varphi_k(W; \theta, \tau)) dP$ in $(\theta, \tau) \in B_{\Theta \times \mathcal{T}}(\delta)$. Let $\varphi_1 = \varphi(W; \theta, \tau)$ and $\varphi_2 = \varphi(W; \theta_0, \tau_0)$. Then we can write

$$\begin{aligned} h_x^{(s)}(\varphi_1) - h_x^{(s)}(\varphi_2) &= \{(x - \varphi_1)^{s-1} - (x - \varphi_2)^{s-1}\} 1\{\varphi_2 \leq x\} \\ &\quad + (x - \varphi_2)^{s-1} \{1\{\varphi_1 \leq x\} - 1\{\varphi_2 \leq x\}\}. \end{aligned}$$

Hence the pathwise differentiability is immediate when $\varphi(w; \theta, \tau)$ is continuously differentiable in (θ, τ) and W is continuous and has a bounded density. The condition in (4) indicates that the functional $\Gamma_{k,P}$ at the estimators has an asymptotic linear representation. This condition can be established using the standard method of expanding the functional in terms of the estimators, $\hat{\theta}$ and $\hat{\tau}$, and using the asymptotic linear representation of these estimators. This asymptotic linear representation for these estimators is available in many semiparametric models. Since our procedure does not make use of its particular characteristic beyond the condition in (4), we keep this condition at a high level for the sake of brevity.

Assumption 2 : (i) $\varphi(W;\theta_0,\tau_0)$ is absolutely continuous with bounded density.

(ii) Condition (A) below holds when $s = 1$, and Condition (B), when $s > 1$.

(A) There exist $\delta, C > 0$ and a subvector W_1 of W such that (a) the conditional density of $\varphi(W;\theta,\tau)$ given W_1 is bounded uniformly over $(\theta,\tau) \in B_{\Theta\times\mathcal{T}}(\delta)$, (b) for each (θ,τ) and $(\theta',\tau') \in B_{\Theta\times\mathcal{T}}(\delta)$, $\varphi_k(W;\theta,\tau) - \varphi_k(W;\theta',\tau')$ is measurable with respect to the σ-field of W_1, and (c) for each $(\theta_1,\tau_1) \in B_{\Theta\times\mathcal{T}}(\delta)$ and for each $\varepsilon > 0$,

$$\sup_{P\in\mathcal{P}}\sup_{w_1}\mathbf{E}_P\left[\sup_{(d_2,\tau_2)\in B_{\Theta\times\mathcal{T}}(\varepsilon)}|\varphi_k(W;\theta_1,\tau_1) - \varphi_k(W;\theta_2,\tau_2)|^2\,|W_1 = w_1\right] \leq C\varepsilon^{2s_2} \quad (5)$$

for some $s_2 \in (d,1]$ with d in Assumption 1(ii), where the supremum over w_1 runs in the support of W_1.

(B) There exist $\delta, C > 0$ such that Condition (c) above is satisfied.

Assumption 2(ii) contains two different conditions that are suited to each case of $s = 1$ or $s > 1$. This different treatment is due to the nature of the function $h_x(\varphi) = (x - \varphi)^{s-1}1\{\varphi \leq x\}q(x)$ that is discontinuous in φ when $s = 1$ and smooth in φ when $s > 1$. Condition (A) can be viewed as a generalization of the set up of LMW. Condition (A)(a) is analogous to Assumption 1(iii) of LMW. Condition (ii)(A)(b) is satisfied by many semiparametric models. For example, in the case of a partially parametric specification: $X_k(\theta_0,\tau_0) = Y_k - Z_{1k}^\top\theta_0 - \tau_0(Z_{2k})$, we take $W = (Y, Z_1, Z_2)$ and $W_1 = (Z_1, Z_2)$. In the case of single index restrictions: $X_k(\theta_0,\tau_0) = Y_k - \tau_0(Z_k^\top\theta_0)$, $W = (Y, Z)$ and $W_1 = Z$. The condition in (5) requires the function $\varphi(W;\theta,\tau)$ to be (conditionally) locally uniformly $L_2(P)$-continuous in $(\theta,\tau) \in B_{\Theta\times\mathcal{T}}(\delta)$ uniformly over $P \in \mathcal{P}$. Sufficient conditions and discussions can be found in Chen, Linton, and van Keilegom (2003). We can weaken this condition to the unconditional version when we consider only the case $s > 1$.

We now turn to the definition of the test statistics. Let $\bar{F}_{kN}(x,\theta,\tau) = \frac{1}{N}\sum_{i=1}^N 1\{X_{ki}(\theta,\tau) \leq x\}$ and

$$\bar{D}_{kl}^{(s)}(x,\theta,\tau) = \bar{D}_k^{(s)}(x,\theta,\tau) - \bar{D}_l^{(s)}(x,\theta,\tau) \text{ for } s \geq 1,$$

where $\bar{D}_k^{(1)}(x,\theta,\tau) = \bar{F}_{kN}(x,\theta,\tau)$ and $\bar{D}_k^{(s)}(x,\theta,\tau)$ is defined through the following recursive relation:

$$\bar{D}_{kl}^{(s)}(x,\theta,\tau) = \int_{-\infty}^x \bar{D}_{kl}^{(s-1)}(t,\theta,\tau)dt \text{ for } s \geq 2. \quad (6)$$

The test statistics we consider are based on the weighted empirical analogues of d_s^* and c_s^*,

namely,

$$D_N^{(s)} = \min_{k \neq l} \sup_{x \in \mathcal{X}} q(x)\sqrt{N}\bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) \text{ and}$$

$$C_N^{(s)} = \min_{k \neq l} \int_{\mathcal{X}} \max\left\{q(x)\sqrt{N}\bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}), 0\right\}^2 dx.$$

Horváth, Kokoszka, and Zitikis (2006) showed that in the case of the Kolmogorov-Smirnov functional with $s \geq 3$, it is necessary to employ a weight function $q(x)$ such that $q(x) \to 0$ as $x \to \infty$ in order to obtain nondegenerate asymptotics. The numerical integration in (6) can be cumbersome in practice. Integrating by parts, we have an alternative form

$$\bar{D}_k^{(s)}(x, \theta, \tau) = \frac{1}{N(s-1)!} \sum_{i=1}^{N} (x - X_{ki}(\theta, \tau))^{s-1} \, 1\{X_{ki}(\theta, \tau) \leq x\}.$$

Since $\bar{D}_k^{(s)}$ and $D_k^{(s)}$ are obtained by applying a linear operator to $\bar{F}_{kN}$ and F_k, the estimated function $\bar{D}_k^{(s)}$ is an unbiased estimator for $D_k^{(s)}$. Regarding the estimators $\hat{\theta}$ and $\hat{\tau}$ and the choice of the weight function q, we assume the following.

Assumption 3 : (i) $||\hat{\theta} - \theta_0|| = o_P(N^{-1/4})$, $||\hat{\tau} - \tau_0||_{P,2} = o_P(N^{-1/4})$ uniformly in $P \in \mathcal{P}$, and $\sup_{P \in \mathcal{P}} P\{\hat{\tau} \in \mathcal{T}\} \to 1$ as $N \to \infty$.
(ii) $\sup_{x \in \mathcal{X}} \left(1 + |x|^{(s-1) \vee (1+\delta)}\right) q(x) < \infty$, for some $\delta > 0$ and for q, nonnegative, first order continuously differentiable function on $\mathcal{X}$ with a bounded derivative.

When $\hat{\theta}$ is a solution from the M- estimation problem, its rate of convergence can be obtained by following the procedure of Theorem 3.2.5 of van der Vaart and Wellner (1996). The uniformity in P in this case can be ensured by combining the uniform (in P) oscillation behavior of the population objective function and the associated empirical processes. For instance, one may employ a uniform oscillation exponential bound for empirical processes established by Giné (1997), Theorem 6.1. The rate of convergence for the nonparametric component $\hat{\tau}$ can also be established by controlling the bias and variance part uniformly in P. In order to control the variance part, one may employ the framework of Giné and Zinn (1991) and establish the central limit theorem that is uniform in P. The sufficient conditions for the last condition in (i) can be checked, for example, from the results of Andrews (1994). Condition (ii) is very convenient and is fulfilled by an appropriate choice of the weight function. The use of the weight function is convenient as $\mathcal{X}$ is allowed to be unbounded. Condition (ii) is stronger than that of Horváth, Kokoszka, and Zikitis (2006) who under a set-up simpler than this paper, imposed that $\sup_{x \in \mathcal{X}} \left(1 + (\max(x, 0)^{(s-2) \vee 1}\right) q(x) < \infty$. Note that the condition in (ii) implies that when $\mathcal{X}$ is a bounded set, we may simply take $q(x) = 1$.

When $s = 1$ so that our focus is on the first order stochastic dominance relation, we may transform the variable $X_{ki}(\theta, \tau)$ into one that has a bounded support by taking a smooth strictly monotone transform. After this transformation, we can simply take $q(x) = 1$.

The first result below is concerned with the convergence in distribution of $D_N^{(s)}$ and $C_N^{(s)}$ under the null hypothesis uniformly in a subset of $\mathcal{P}$. As for $h_x^{(s)}$ defined prior to Assumption 1, we define $h_{kl,i}^{\Delta}(x) = h_x^{(s)}(X_{ki}) - h_x^{(s)}(X_{li})$ and $\psi_{kl,i}^{\Delta}(x) = \psi_{x,k,P}(W_i) - \psi_{x,l,P}(W_i)$. Let $\nu_{kl}^{(s)}(\cdot)$ be a Gaussian process on $\mathcal{X}$ with covariance kernel given by

$$C_P(x_1, x_2) = Cov_P(h_{kl,i}^{\Delta}(x_1) + \psi_{kl,i}^{\Delta}(x_1), h_{kl,i}^{\Delta}(x_2) + \psi_{kl,i}^{\Delta}(x_2)),$$

where Cov_P denotes the covariance under P. The asymptotic critical values are based on this Gaussian process. By convention, we assume that the supremum over an empty set is equal to $-\infty$.

Let $\mathcal{P}_0$ be the set of probabilities under the null hypothesis and let $\mathcal{P}_{00}^{KS}$ be the set of probabilities in $\mathcal{P}_0$ such that the contact set $B_{kl}^{(s)}$ is nonempty for all pairs $k \neq l$.[4] The set of probabilities $\mathcal{P}_{00}^{KS}$ represents the "boundary" of the null hypothesis. In other words, the boundary points are the probabilities under which the contact set $B_{kl}^{(s)}$ is nonempty for all $k \neq l$. We define the set of boundary points differently for Cramér-von Mises tests, $C_N^{(s)}$. Let $\mathcal{P}_{00}^{CM}$ be the set of probabilities in $\mathcal{P}_0$ such that for each $k \neq l$, the contact set $B_{kl}^{(s)}$ has a positive Lebesgue measure. Hence, in general $\mathcal{P}_{00}^{CM} \subset \mathcal{P}_{00}^{KS}$.

Theorem 1 : Suppose that Assumptions 1-3 hold. Then under the null hypothesis, as $N \to \infty$,

$$D_N^{(s)} \underset{D}{\to} \begin{cases} \min_{k \neq l} \sup_{x \in B_{kl}^{(s)}} (\nu_{kl}^{(s)}(x)), & \text{if } P \in \mathcal{P}_{00}^{KS} \\ -\infty, & \text{if } P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS} \end{cases} \quad \text{and}$$

$$C_N^{(s)} \underset{D}{\to} \begin{cases} \min_{k \neq l} \int_{B_{kl}^{(s)}} \max\{\nu_{kl}^{(s)}(x), 0\}^2 dx, & \text{if } P \in \mathcal{P}_{00}^{CM} \\ 0, & \text{if } P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{CM} \end{cases}.$$

Under a fixed probability $P \in \mathcal{P}_{00}^{KS}$ on the boundary, the test statistic is asymptotically tight uniformly over $P \in \mathcal{P}_{00}^{KS}$. But under a fixed probability $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS}$, we have $\sup_{x \in \mathcal{X}} D_{kl}^{(s)}(x) < 0$ for some k, l, and hence the test statistics diverge to $-\infty$. The limiting distribution of the test statistic $D_N^{(s)}$ exhibits discontinuity; they are not asymptotically tight at each $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS}$, while they are so at each $P \in \mathcal{P}_0^{KS}$. This phenomenon of discontinuity arises often in moment inequality models. (e.g. Moon and Schorfheide (2006),

---

[4]Hence, that $P \in \mathcal{P}_{00}^{KS}$ implies $d_s^* = 0$ because $\mathcal{P}_{00}^{KS} \subset \mathcal{P}_0$. However, the converse is not true, in particular when $\mathcal{X}$ is unbounded.

Chernozhukov, Hong, and Tamer (2007), and Andrews and Guggenberger (2006)).

As for $C_N^{(s)}$, similar discontinuity arises in the limiting distribution. When $P \in \mathcal{P}_{00}^{CM}$, $C_N^{(s)}$ has a nondegenerate limiting distribution. However, when $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{CM}$, the limiting distribution becomes degenerate at zero. We introduce the following definition of a test having an asymptotically exact size.

Definition 1 : (i) A test φ_α with a nominal level α is said to have an *asymptotically exact size* if there exists a nonempty subset $\mathcal{P}_0' \subset \mathcal{P}_0$ such that:

$$\mathrm{limsup}_{N \to \infty} \mathrm{sup}_{P \in \mathcal{P}_0} \mathbf{E}_P \varphi_\alpha \leq \alpha.$$

$$\mathrm{limsup}_{N \to \infty} \mathrm{sup}_{P \in \mathcal{P}_0'} \mathbf{E}_P \varphi_\alpha = \alpha. \tag{7}$$

(ii) When a test φ_α satisfies (7), we say that the test is *asymptotically similar on* $\mathcal{P}_0'$.

The Gaussian process $\nu_{kl}^{(s)}$ in Theorem 1 depends on the unknown elements of the null hypothesis, precluding the use of first order asymptotic critical values in practice. Barrett and Donald (2003) suggested a bootstrap procedure and LMW, a subsampling approach. Both studies have not paid attention to the issue of uniformity in the convergence of tests. In order to deal with uniformity, we separate $\mathcal{P}_0$ into the boundary points and the interior points and deal with asymptotic theory of the test statistic separately. The proof of Theorem 1 establishes the weak convergence of $q(\cdot)\sqrt{N}\{\bar{D}_{kl}^{(s)}(\cdot, \hat{\theta}, \hat{\tau}) - D_{kl}(\cdot, \theta_0, \tau_0)\}$ to $\nu_{kl}^{(s)}(\cdot)$ uniformly over $P \in \mathcal{P}$. This result is obtained by showing that the class of functions indexing the empirical process is a uniform Donsker class (Giné and Zinn (1991)). The remaining step is to deal with the limiting behavior of the test statistic uniformly over the interior points.

One might consider a typical bootstrap procedure for this situation. The difficulty for the bootstrap method tests of stochastic dominance lies mainly in the fact that it is hard to impose the null hypothesis upon the test. There have been approaches that consider only *least favorable* subset of the models of the null hypothesis as in the following.

$$F_1(x) = \cdots = F_K(x) \text{ for all } x \in \mathcal{X}. \tag{8}$$

This leads to the problem of asymptotic nonsimilarity in the sense that when the true data generating process lies away from the least favorable subset of the models of the null hypothesis, and yet still lie at the boundary points, the bootstrap sizes become misleading even in large samples. LMW calls this phenomenon *asymptotic nonsimilarity on the boundary*. Bootstrap procedures that employ the usual recentering implicitly impose restrictions that do not hold outside the least favorable set and hence asymptotically biased against such

9

probabilities outside the set.

To illustrate heuristically why a test that uses critical values from the least favorable case of a composite null hypothesis can be asymptotically biased, let us consider a simple example in the finite dimensional case. Suppose that the observations $\{X_i = (X_{1i}, X_{2i}) : i = 1, \ldots, N\}$ are mutually independently and identically distributed with unknown mean $\mu = (\mu_1, \mu_2)$ and known variance $\Sigma = diag(1, 1)$. Let the hypotheses of interest be given by:

$$H_0 : \mu_1 \leq 0 \text{ and } \mu_2 \leq 0 \quad \text{vs.} \quad H_1 : \mu_1 > 0 \text{ or } \mu_2 > 0.$$

These hypotheses are equivalent to:

$$H_0 : d^* \leq 0 \quad \text{vs.} \quad H_1 : d^* > 0, \tag{9}$$

where $d^* = \max\{\mu_1, \mu_2\}$. The "boundary" of the null hypothesis is given by $\mathcal{B}_{BD} = \{(\mu_1, \mu_2) : \mu_1 \leq 0 \text{ and } \mu_2 \leq 0\} \cap \{(\mu_1, \mu_2) : \mu_1 = 0 \text{ or } \mu_2 = 0\}$, while the "least favorable case (LFC)" is given by $\mathcal{B}_{LF} = \{(\mu_1, \mu_2) : \mu_1 = 0 \text{ and } \mu_2 = 0\} \subset \mathcal{B}_{BD}$. To test (9), one may consider the following the t-statistic:

$$T_N = \max\{N^{1/2}\overline{X}_1, N^{1/2}\overline{X}_2\},$$

where $\overline{X}_k = \sum_{i=1}^{N} X_{ki}/N$. Then, the asymptotic null distribution of T_N is non-degenerate provided the true μ lies on the boundary $\mathcal{B}_{BD}$, but the distribution depends on the location of μ. That is, we have

$$T_N \xrightarrow{d} \begin{cases} \max\{Z_1, Z_2\} & \text{if} & \mu = (0, 0) \\ Z_1 & \text{if} & \mu_1 = 0, \mu_2 < 0 \\ Z_2 & \text{if} & \mu_1 < 0, \mu_2 = 0 \end{cases},$$

where Z_1 and Z_2 are mutually independent standard normal random variables. On the other hand, T_N diverges to $-\infty$ under the "interior" of the null hypothesis. Suppose that z_α^* satisfies $P(\max\{Z_1, Z_2\} > z_\alpha^*) = \alpha$ for $\alpha \in (0, 1)$. Then, the test based on the least favorable case is asymptotically non-similar on the boundary because, for example, $\lim_{N \to \infty} P(T_N > z_\alpha^* | \mu = (0, 0)) \neq \lim_{N \to \infty} P(T_N > z_\alpha^* | \mu = (0, -1))$. Now consider the following sequence of local alternatives: for $\delta > 0$,

$$H_N : \mu_1 = \frac{\delta}{\sqrt{N}} \text{ and } \mu_2 < 0.$$

Then, under H_N, it is easy to see that $T_N \xrightarrow{d} N(\delta, 1)$. However, the test based on the LFC critical value may be biased against this local alternatives, because $\lim_{N \to \infty} P(T_N > z_\alpha^*) =$

$P(N(\delta, 1) > z_\alpha^*) < \alpha$ for some values of δ. To see this, in Figure 1, we draw a c.d.f.'s of $\max\{Z_1, Z_2\}$ and $N(\delta, 1)$ for $\delta = 0.0$, 0.2, and 1.5. Clearly, the distribution of $\max\{Z_1, Z_2\}$ first-order stochastic dominates that of $N(0.2, 1)$, i.e., T_N is asymptotically biased against H_N for $\delta = 0.2$.
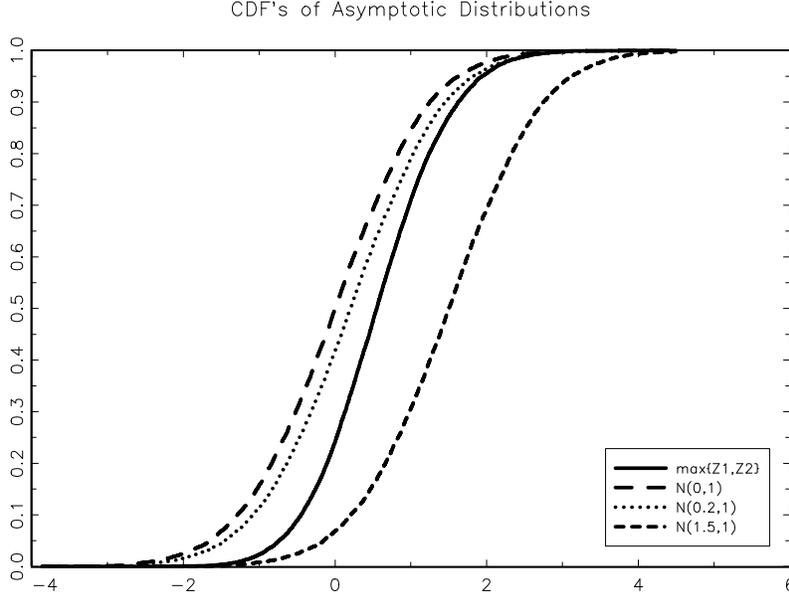

Figure 1. Shows that T_N is asymptocally biased against H_N for $\delta = 0.2$.

3   Bootstrap Procedure

In this section we discuss our method for obtaining asymptotically valid critical values. We propose the wild bootstrap procedure as follows. Let

$$X_{ki}(\theta, \tau) = Y_{ki} - g(Z_{ki}; \theta, \tau),$$

for some function g. Among the necessary identification conditions for θ_0 and τ_0 is the condition that $\mathbf{E}\left[Y_{ki} - g(Z_{ki}; \theta, \tau)|Z_{ki}\right] = 0$. Let $X_i(\theta, \tau)$ be a K- dimensional vector whose k-th entry is given by $X_{ki}(\theta, \tau)$. Let $(\hat{X}_i)_{i=1}^N = (X_i(\hat{\theta}, \hat{\tau}))_{i=1}^N$. Following the wild bootstrap procedure (Härdle and Mammen (1993)), we draw $\{(\varepsilon_{i,b}^*)_{i=1}^N\}_{b=1}^B$ such that $\mathbf{E}^*[\varepsilon_{ki,b}^*|Z_{ki}] = 0$ and $\mathbf{E}^*[\varepsilon_{ki,b}^{*2}|Z_{ki}] = \hat{X}_{ki}^2$, where $\mathbf{E}^*[\cdot|Z_{ki}]$ denotes the bootstrap conditional distribution given Z_{ki} and $\varepsilon_{ki,b}^*$ is the k-th element of $\varepsilon_{i,b}^*$. Let $Y_{ki,b}^* = g(Z_{ki}; \hat{\theta}, \hat{\tau}) + \varepsilon_{ki,b}^*$, $k = 1, \ldots, K$ and take $W_{i,b}^* = (Y_{ki,b}^*, Z_{ki})_{k=1}^K$. Using this bootstrap sample, construct estimators $\hat{\theta}_b^*$ and $\hat{\tau}_b^*$, for each $b = 1, \ldots, B$. Then, we define $X_{ki,b}^* = Y_{ki,b}^* - g(Z_{ki}; \hat{\theta}_b^*, \hat{\tau}_b^*)$.

11

Now, we introduce the following bootstrap empirical process

$$\bar{D}_{kl,b}^{(s)*}(x) = \frac{1}{N(s-1)!} \sum_{i=1}^{N} \left\{ h_x(X_{ki,b}^*) - h_x(X_{li,b}^*) - \mathbf{E}_N \left[h_x(\hat{X}_{ki}) - h_x(\hat{X}_{li}) \right] \right\}, \ b = 1, 2, \ldots, B,$$

where $\mathbf{E}_N$ denotes the expectation with respect to the empirical measure of $\{\hat{X}_i\}_{i=1}^{N}$. The quantity $\bar{D}_{kl,b}^{(s)*}(x)$ denotes the bootstrap counterpart of $\bar{D}_{kl}^{(s)}(x)$. Given a sequence $c_N \to 0$ and $c_N \sqrt{N} \to \infty$, we construct the estimated contact set:

$$\hat{B}_{kl}^{(s)} = \left\{ x \in \mathcal{X} : q(x)|\bar{D}_{kl}^{(s)}(x)| < c_N \right\}. \tag{10}$$

As for the weight function $q(x)$, we may consider the following type of function. For $z_1 < z_2$ and for a constant $a > 0$, we set

$$q(x) = \begin{cases} 1 \text{ if } x \in [z_1, z_2] \\ a/(a + |x - z_2|^{(s-1)\vee 1}) \text{ if } x > z_2 \\ a/(a + |x - z_1|^{(s-1)\vee 1}) \text{ if } x < z_1. \end{cases}$$

This is a modification of the weighting function considered by Horváth, Kokoszka, and Zikitis (2006).

We propose the Kolmogorov-Smirnov type bootstrap test statistic as

$$D_{N,b}^{(s)*} = \begin{cases} \min_{k \neq l} \sup_{x \in \hat{B}_{kl}^{(s)}} \sqrt{N} q(x) \bar{D}_{kl,b}^{(s)*}(x), \text{ if } \hat{B}_{kl}^{(s)} \neq \varnothing \text{ for all } k \neq l \\ \\ \pi_N \text{ if } \hat{B}_{kl}^{(s)} = \varnothing \text{ for some } k \neq l, \end{cases} \tag{11}$$

where π_N is a positive sequence such that $\pi_N \to \infty$. From this, we obtain the bootstrap critical values, $c_{\alpha,N,B}^{*KS} = \inf\{t : B^{-1}\Sigma_{b=1}^{B}1\{D_{N,b}^{(s)*} \leq t\} \geq 1 - \alpha\}$, yielding an α-level bootstrap test: $\varphi_\alpha^{KS} \triangleq 1\{D_N^{(s)} > c_{\alpha,N,B}^*\}$. The sequence π_N governs the asymptotic behavior of the size of the test. When π_N increases faster, the size of the test converges to zero faster in the interior points $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS}$.

Likewise, we can define the Cramér-von Mises type bootstrap test statistic as

$$C_{N,b}^{(s)*} = \begin{cases} N \min_{k \neq l} \int_{\hat{B}_{kl}^{(s)}} \max \left\{ q(x) \bar{D}_{kl,b}^{(s)*}(x), 0 \right\}^2 dx, \text{ if } \hat{B}_{kl}^{(s)} \neq \varnothing \text{ for all } k \neq l \\ \\ \pi_N \text{ if } \hat{B}_{kl}^{(s)} = \varnothing \text{ for some } k \neq l, \end{cases} \tag{12}$$

where π_N is a positive sequence such that $\pi_N \to \infty$. The bootstrap critical values are

obtained by $c_{\alpha,N,B}^{*CM} = \inf\{t : B^{-1}\Sigma_{b=1}^{B}1\{C_{N,b}^{(s)*} \le t\} \ge 1-\alpha\}$, yielding an α-level bootstrap test: $\varphi_{\alpha}^{CM} \triangleq 1\{C_N^{(s)} > c_{\alpha,N,B}^{*CM}\}$.

The main step we need to justify the bootstrap procedure uniformly in $P \in \mathcal{P}$ under this general semiparametric environment is to establish the bootstrap uniform central limit theorem for the empirical processes in the test statistic, where uniformity holds over probability measures in $\mathcal{P}$. The uniform central limit theorem for empirical processes has been developed by Giné and Zinn (1991) and Sheehy and Wellner (1992) through the characterization of uniform Donsker classes.

The bootstrap central limit theorem for empirical processes was established by Giné and Zinn (1990) who considered a nonparametric bootstrap procedure. See also a nice lecture note on the bootstrap by Giné (1997). We cannot directly use these results because the bootstrap distribution for each sample $\{W_{i,b}^{*}\}_{i=1}^{N}$ is non-identically distributed both in the case of wild-bootstrap or residual-based bootstrap. In the Appendix, we show how we can make the bootstrap CLT result of Giné (1997) accommodate this case by slightly modifying his result.[5]

This paper's bootstrap procedure does not depend on a specific estimation method for θ_0 and τ_0. Due to this generality of our framework, we introduce the following additional assumptions about the bootstrap sample W_{ib}^{*}, the bootstrap estimators $(\hat{\theta}^{*}, \hat{\tau}^{*})$, and the function $g(z; \theta, \tau)$ introduced at the beginning of this section. Let $\mathcal{G}_N$ be the σ-field generated by $\{W_i\}_{i=1}^{N}$.

Assumption 4: (i) For arbitrary $\varepsilon > 0$,

$$P\left\{||\hat{\theta}^{*} - \hat{\theta}|| + ||\hat{\tau}^{*} - \hat{\tau}||_{P,2} > N^{-1/4}\varepsilon|\mathcal{G}_N\right\} \to_P 0, \text{ uniformly in } P \in \mathcal{P}.$$

(ii) For $\psi_{x,k,P}(\cdot)$ in Assumption 1(iv), $\mathbf{E}[\psi_{x,k,P}(W_i)|Z_{ki}] = 0$, and for some $r_N \to 0$ with $\limsup_{N \to \infty} r_N/c_N < \infty$,

$$P\left\{\sup_{x \in \mathcal{X}}\left|\sqrt{N}\hat{\Gamma}_{k,P}(x) - \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\psi_{x,k,P}(W_{i,b}^{*})\right| > r_N\varepsilon|\mathcal{G}_N\right\} \to_P 0$$

uniformly in $P \in \mathcal{P}$ for any $\varepsilon > 0$, where $\hat{\Gamma}_{k,P}(x) = \mathbf{E}_N[h_x^{(s)}(\varphi_k(W_i^{*}; \hat{\theta}^{*}, \hat{\tau}^{*}))] - \mathbf{E}_N[h_x^{(s)}(\varphi_k(W_i^{*}; \hat{\theta}, \hat{\tau}))]$ and $\mathbf{E}^{*}[\psi_{x,k,P}(W_{i,b}^{*})|Z_{ki}] = 0$.

---

[5]For a weak convergence of residual-based bootstrap process in a regression model, see Koul and Lahiri (1994). For a parametric bootstrap procedure for goodness-of-fit tests, see, among others, Andrews (1997). See also Härdle and Mammen (1993) and Whang (2000) for a wild-bootstrap procedure that contains finite-dimensional estimators.

(iii) There exists C such that for each $(\theta_1, \tau_1) \in B_{\Theta \times \mathcal{T}}(\delta)$ and for each $\varepsilon > 0$,

$$\sup_{z \in \mathcal{Z}_k} \sup_{(d_2, \tau_2) \in B_{\Theta \times \mathcal{T}}(\varepsilon)} |g(z; \theta_1, \tau_1) - g(z; \theta_2, \tau_2)| \leq C \varepsilon^{2s_2}, \tag{13}$$

for some $s_2 \in (d, 1]$, where $\mathcal{Z}_k$ is the support of Z_{ki}.

(iv) For each $k \neq l$, there exists $\delta > 0$ and constants $c_{kl} > 0$ and $p_{kl} \geq 1$ such that for each $x, x' \in \mathcal{X}$,

$$\left| q(x)|D_{kl}^{(s)}(x)| - q(x')|D_{kl}^{(s)}(x')| \right| \geq c_{kl} \{ \mathbf{E}[h_x^{(s)}(X_{ki}) - h_x^{(s)}(X_{li}) - \{h_{x'}^{(s)}(X_{ki}) - h_{x'}^{(s)}(X_{li})\}]^2 \wedge \delta \}^{p_{kl}/2}.$$

Condition (i) requires that the bootstrap estimators should be consistent at a rate faster than $N^{-1/4}$ uniformly over $P \in \mathcal{P}$. The bootstrap validity of M-estimators is established by Arcones and Giné (1992). Infinite dimensional Z-estimators are dealt with by Wellner and Zhan (1996). See also Abrevaya and Huang (2005) for a bootstrap inconsistency result for a case where the estimators are not asymptotically linear. Condition (iii) implies the condition in (5) by choosing $W_{1i} = Z_{ki}$ and can be weakened to the locally uniform L_2-continuity (e.g. Chen, Linton, and van Keilegom (2003)) when we confine our attention to the case $s > 1$. Condition (iv) is similar to Condition C.2 in Chernozhukov, Hong, and Tamer (2007) and is used to control the convergence rate of the estimated set with respect an appropriate norm. This condition is needed only for the convergence rate of the rejection probability.

The wild bootstrap procedure is introduced to ensure Condition (ii) which assumes the bootstrap analogue of the asymptotic linearity of $\sqrt{N}\Gamma_{k,P}(x)[\hat{\theta}^* - \hat{\theta}, \hat{\tau}^* - \hat{\tau}]$ with the same influence function $\psi_{x,P}$. In the case of parametric regression specification: $X_{ki}(\theta_0) = Y_{ki} - g(Z_{ki}; \theta_0)$ without the nonparametric function τ_0, the parameter θ_0 is typically identified by the condition $\mathbf{E}[X_{ki}(\theta_0)] = 0$ and in this case, we may use the residual-based bootstrap procedure by replacing $\varepsilon_{ki,b}^*$ by those resampled from $\{\hat{X}_i - \bar{X}_N\}_{i=1}^N$ where $\bar{X}_N = \frac{1}{N}\sum_{i=1}^N X_i(\hat{\theta}, \hat{\tau})$. Then we need Condition (ii) with the centered influence function $\frac{1}{\sqrt{N}}\sum_{i=1}^N \{\psi_{x,k,P}(W_{i,b}^*) - \mathbf{E}_N\psi_{x,k,P}(\hat{W}_i)\}$ where $\hat{W}_i = (\hat{Y}_i, Z_i)$ and $\hat{Y}_{ki} = g(Z_{ki}; \hat{\theta}) + X_{ki}(\theta_0)$. (See. e.g. Koul and Lahiri (1994)).

One of the important conditions that are needed in this case is the uniform oscillation behavior of the bootstrap empirical process, as established in the proof of Theorem 2.2 of Giné (1997). We slightly modify his result so that empirical processes constructed from the residual-based bootstrap procedure are accommodated. This latter result of Giné is also crucial for our establishing the bootstrap central limit theorem that is uniform in P.

Theorem 2 : Suppose that the conditions of Theorem 1 and Assumption 4 hold and that $c_N \to 0$ and $c_N\sqrt{N} \to \infty$. Then the following holds.

14

(i) Uniformly over $P \in \mathcal{P}_0$,

$$P\left\{D_N^{(s)} > c_{\alpha,\infty}^{*KS}\right\} \leq \alpha + O_P(c_N\sqrt{-\log c_N}), \text{ and}$$

$$P\left\{C_N^{(s)} > c_{\alpha,\infty}^{*CM}\right\} \leq \alpha + O_P(c_N\sqrt{-\log c_N}).$$

(ii) Furthermore, suppose that for all $P \in \mathcal{P}_{00}^{KS}$, the distribution of $\min_{k \neq l} \sup_{x \in B_{kl}^{(s)}}(\nu_{kl}^{(s)}(x))$ is continuous, and for all $P \in \mathcal{P}_{00}^{CM}$, the distribution of $\min_{k \neq l} \int_{B_{kl}^{(s)}} \max\{\nu_{kl}^{(s)}(x), 0\}^2 dx$ is continuous. Then

$$P\left\{D_N^{(s)} > c_{\alpha,\infty}^{*KS}\right\} = \alpha + O_P(c_N\sqrt{-\log c_N}) \text{ uniformly over } P \in \mathcal{P}_{00}^{KS} \text{ and}$$

$$P\left\{C_N^{(s)} > c_{\alpha,\infty}^{*CM}\right\} = \alpha + O_P(c_N\sqrt{-\log c_N}) \text{ uniformly over } P \in \mathcal{P}_{00}^{CM},$$

The first result says that the bootstrap tests have asymptotically correct sizes. The second result tells us that the bootstrap test is asymptotically similar on the boundary. The second result combined with the first result establishes that the bootstrap test has exact asymptotic size equal to α. The result is uniform over $P \in \mathcal{P}_0$. The caveats of the absolute continuity of the limiting distributions are necessary because in certain cases the distribution can be discrete or even degenerate at one point. For example, consider the case where we have $K = 2$, $s = 1$, and $\mathcal{X} = [0,1]$ which is the common support for X_1 and X_2 whose distribution functions meet only at 0 and 1 within $\mathcal{X}$. Then, the Gaussian process $\nu_{kl}^{(s)}(\cdot)$ is a variant of a Brownian bridge process taking values 0 at end points of 0 and 1 and hence $\min_{k \neq l} \sup_{x \in B_{kl}^{(s)}}(\nu_{kl}^{(s)}(x))$ is a discrete random variable degenerate at zero. In this case, the asymptotic size is zero, violating the equality in (ii).

The limit behavior of the bootstrap test statistic mimics the discontinuity in Theorem 1. Both the original and bootstrap test statistics diverge when the data generating process is away from the boundary points. However, the direction of divergence is opposite. When the bootstrap distribution is away from "the boundary points," i.e., if $\hat{B}_{kl}^{(s)} = \varnothing$ for some $k \neq l$, the bootstrap test statistic diverges to ∞. However, outside the boundary points, the original test statistic diverges to $-\infty$. This opposite direction is imposed to keep the asymptotic size of the bootstrap test below α uniformly over all the probabilities $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS}$.

Our requirement for the sequence π_N to increase to infinity is necessary for our results that are uniform in P. To see this, consider the following:

$$\sqrt{N}q(x)\bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) = \sqrt{N}q(x)\left\{\bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) - D_{kl}^{(s)}(x)\right\} + \sqrt{N}q(x)D_{kl}^{(s)}(x). \quad (14)$$

The first component on the right-hand side is uniformly asymptotically tight and its tail

probability goes to zero only when we increase the cut-off value to infinity. The behavior of the last term depends on the sequence of probabilities $P_N \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS}$. The worst case arises when P_N converges to a boundary point so fast that for some $k \neq l$, $\sup_{x \in \mathcal{X}} \sqrt{N} q(x) D_{kl}^{(s)}(x)$ also converges to zero. Considering uniformity in P, we should take this case into account. Therefore, π_N should increase to infinity in order to control the tail probability of the supremum of the combined quantities in (14) appropriately. On the other hand, for only pointwise results for each $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS}$, it suffices to set π_N to be any nonnegative number.

It is worth noting that the convergence of the rejection probability on the boundary points is slightly slower than c_N. Since we can take c_N converge to zero faster than $N^{-1/3}$, the convergence of the rejection probability can be made to be faster than the subsampling-based test.[6]

Suppose that the null hypotheses in (1) and (2) are equivalent. Let us compare the asymptotic similarity property of two tests $D_N^{(s)}$ and $C_N^{(s)}$ that use bootstrap critical values as proposed above. In general $\mathcal{P}_{00}^{CM} \subset \mathcal{P}_{00}^{KS}$. Hence the Kolmogorov-Smirnov tests are asymptotically similar over a wider set of probabilities than the Cramér-von Mises type tests. Such a comparison is relevant when for all $k \neq l$, the contact sets $B_{kl}^{(s)}$ are not empty and for some $k \neq l$, $B_{kl}^{(s)}$ is a finite set, as would happen, for example, with $D_k^{(s)}$ and $D_l^{(s)}$ contact on a finite set of points and $D_k^{(s)}(x) \leq D_l^{(s)}(x)$. In that case, the Kolmogorov-Smirnov test has an exact asymptotic size of α while the asymptotic size of Cramér-von Mises test is zero.

The result of Theorem 2 allows for a wide range of choice for the sequence π_N in (11) and (12) in the bootstrap test statistic. In practice, we may consider two extreme cases $\pi_N = \log(N)$ and $\pi_N = \infty$, and see the robustness of the results. The latter case with $\pi_N = \infty$ is tantamount to the rule of setting the test $\varphi = 0$ (i.e. do not reject the null) when the contact set $\hat{B}$ is empty.

4   Asymptotic Power Properties

In this section, we establish asymptotic power properties of the bootstrap test. First, we consider consistency of the test.

Theorem 3 : Suppose that the conditions of Theorem 2 hold and that we are under a fixed alternative $P \in \mathcal{P} \backslash \mathcal{P}_0$. Then,

$$\lim_{N \to \infty} P \left\{ D_N^{(s)} > c_{\alpha, \infty}^{*KS} \right\} \to 1 \text{ and } \lim_{N \to \infty} P \left\{ C_N^{(s)} > c_{\alpha, \infty}^{*CM} \right\} \to 1.$$

---

[6]The slower convergence rate for the bootstrap test than the asymptotic approximation is due to the presence of the estimation error in the contact set $\hat{B}_{kl}^{(s)}$.

Therefore, the bootstrap test is consistent against all types of alternatives. This property is shared by other tests of LMW and Barrett and Donald (2003) for example.

Let us turn to asymptotic local power properties. We consider a sequence of probabilities $P_N \in \mathcal{P} \backslash \mathcal{P}_0$ and denote $D_{k,N}^{(s)}(x)$ to be $D_k^{(s)}(x)$ under P_N. That is, $D_{k,N}^{(s)}(x) = \mathbf{E}_{P_N} h_x^{(s)}(\varphi_k(W; \theta_0, \tau_0))$ using the notation of $h_x^{(s)}$ and the specification of X_k in a previous section, where $\mathbf{E}_{P_N}$ denotes the expectation under P_N. We confine our attention to $\{P_N\}$ such that for each $k \in \{1, \ldots, K\}$, there exists functions $H_k^{(s)}(\cdot)$ and $\delta_k^{KS(s)}(\cdot)$ such that

$$D_{k,N}^{KS(s)}(x) = H_k^{(s)}(x) + \delta_k^{KS(s)}(x)/\sqrt{N} \tag{15}$$

as $N \to \infty$.

In the case of Cramér-von Mises type tests, we consider the following local alternatives: for the functions $H_k^{(s)}(\cdot)$ in (15), we assume that

$$D_{k,N}^{CM(s)}(x) = H_k^{(s)}(x) + \delta_k^{CM(s)}(x)/\sqrt{N}. \tag{16}$$

We assume the following for the functions $H_k^{(s)}(\cdot)$, $\delta_k^{KS(s)}(\cdot)$ and $\delta_k^{CM(s)}(\cdot)$.

Assumption 5: (i) $C_{kl} = \{x \in \mathcal{X} : H_k^{(s)}(x) - H_l^{(s)}(x) = 0\}$ is nonempty for all $k \neq l$.
(ii) $\min_{k \neq l} \sup_{x \in \mathcal{X}} (H_k^{(s)}(x) - H_l^{(s)}(x)) \leq 0$.
(iii) $\sup_{x \in C_{kl}} (\delta_k^{KS(s)}(x) - \delta_l^{KS(s)}(x)) > 0$ and $\int_{C_{kl}} (\max\{(\delta_k^{CM(s)}(x) - \delta_l^{CM(s)}(x)), 0\})^2 dx > 0$, for all $k \neq l$.

Assumption 5 comprises conditions for the sequences in (15) and (16) to constitute non-void local alternatives. Note that when $H_k^{(s)}(x) = \mathbf{E}_P h_x^{(s)}(\varphi_k(W; \theta_0, \tau_0))$ for some $P \in \mathcal{P}$, these conditions for $H_k^{(s)}$ imply that $P \in \mathcal{P}_0$, that is, the probability associated with $H_k^{(s)}$ belongs to the null hypothesis, in particular, the boundary points.[7] The conditions for $\delta_k^{KS(s)}(x)$ and $\delta_k^{CM(s)}(x)$ indicate that for each N, the probability P_N belongs to the alternative hypothesis $\mathcal{P} \backslash \mathcal{P}_0$. Therefore, the sequence P_N represents local alternatives that converge to the null hypothesis (in particular, to the boundary points $\mathcal{P}_{00}^{KS}$) maintaining the convergence of $D_{k,N}^{KS(s)}(x)$ to $H_k^{(s)}(x)$ at the rate of $\sqrt{N}$ in the direction of $\delta_k^{KS(s)}(x)$. The following theorem shows that the tests are asymptotically unbiased against such sequences of local alternatives.

Theorem 4 : Suppose that the conditions of Theorem 2 and Assumption 5 hold.

---

[7] Since the distributions are continuous, the contact sets are nonempty under the alternatives. Therefore, any local alternatives that converge to the null hypothesis will have a limit in the set of boundary points.

(i) Under the local alternatives $P_N \in \mathcal{P} \backslash \mathcal{P}_0$ satisfying the condition in (15),

$$\lim_{N \to \infty} P_N \left\{ D_N^{(s)} > c_{\alpha,\infty}^{*KS} \right\} \geq \alpha.$$

(ii) Under the local alternatives $P_N \in \mathcal{P} \backslash \mathcal{P}_0$ satisfying the condition in (16),

$$\lim_{N \to \infty} P_N \left\{ C_N^{(s)} > c_{\alpha,\infty}^{*CM} \right\} \geq \alpha.$$

Similarly as in LMW, the asymptotic unbiasedness of the test follows from Anderson's Lemma (Anderson (1952)).[8] It is interesting to compare the asymptotic power of the bootstrap procedure that is based on the least favorable set of the null hypothesis. Using the same bootstrap sample $\{X_{ki,b}^* : k = 1, \cdots, K\}_{i=1}^n$, $b = 1, \ldots, B$, this bootstrap procedure alternatively consider the following bootstrap test statistics

$$
\begin{aligned}
D_{N,b}^{(s)*LF} &= \min_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} q(x) \bar{D}_{kl,b}^{(s)*}(x) \text{ and} \\
C_{N,b}^{(s)*LF} &= N \min_{k \neq l} \int_{\mathcal{X}} \max \left\{ q(x) \bar{D}_{kl,b}^{(s)*}(x), 0 \right\}^2 dx.
\end{aligned}
\tag{17}
$$

Let the bootstrap critical values with $B \to \infty$ denoted by $c_{\alpha,\infty}^{*KS-LF}$ and $c_{\alpha,\infty}^{*CM-LF}$ respectively. The results of this paper easily imply the following fact that this bootstrap procedure is certainly inferior to the procedure that this paper proposes. For brevity, we state the result for $D_{N,b}^{(s)*LF}$. We can obtain a similar result for $C_{N,b}^{(s)*LF}$.

Corollary 5 : Suppose that the conditions of Theorem 4 hold. Under the local alternatives $P_N \in \mathcal{P} \backslash \mathcal{P}_0$ satisfying the condition in (15),

$$\lim_{N \to \infty} P_N \left\{ D_N^{(s)} > c_{\alpha,\infty}^{*KS} \right\} \geq \lim_{N \to \infty} P_N \left\{ D_N^{(s)} > c_{\alpha,\infty}^{*KS-LF} \right\}.$$

Furthermore, assume that for all $k \neq l$, the union of the closures of the contact sets C_{kl} is a proper subset of the interior of $\mathcal{X}$ and the Gaussian processes $\nu_{kl}^{(s)}(x)$ in Theorem 1 satisfy that $\mathbf{E}[\{\sup_{x \in C_{kl}} \nu_{kl}^{(s)}(x)\}^2] > 0$. Then the inequality above is strict.

The result of Corollary 5 is remarkable that the bootstrap test of this paper weakly dominates the bootstrap in (17) regardless of the Pitman local alternative directions. Furthermore, when the union of the closures of the contact sets is a proper subset of the interior

---

[8] Anderson's Lemma does not apply to one-sided Kolmogorov-Smirnov type tests. Song (2008) demonstrates that there exist a class of $\sqrt{n}$-converging Pitman local alternatives against which one-sided Kolmogorov-Smirnov type tests are asymptotically biased. Also see the discussion at the end of Section 2.

of $\mathcal{X}$, our test strictly dominates the bootstrap in (17) uniformly over the Pitman local alternatives. However, this latter condition for contact sets can be strong when the number K of prospects are many. In this case, a full characterization of the case for strict dominance appears to be complicated. The result of Corollary 5 is based on the nonsimilarity of the bootstrap tests in (17) on the boundary. In fact, Corollary 5 implies that the test based on the bootstrap procedure using (17) is inadmissible. This result is related to Hansen (2003)'s finding in an environment of finite-dimensional composite hypothesis testing that a test that is not asymptotically similar on the boundary is asymptotically inadmissible.

5   Monte Carlo Experiments

5.1   Simulation Designs

In this section, we examine the finite sample performance of our tests using Monte Carlo simulations. In particular, we compare our procedure with the subsampling method and recentered bootstrap method which were suggested by LMW.

For a fair comparison of the simulation results, the simulation designs we consider are exactly same as those considered by LMW: the Burr distributions also examined by Tse and Zhang (2004), the lognormal distributions also considered by Barrett and Donald (2003) and the exchangeable normal processes of Klecan et. al. (1991).

We first describe the details of the simulation designs. We consider the five different designs from the Burr Type XII distribution, $B(\alpha, \beta)$ which is a two parameter family with cdf defined by:

$$F(x) = 1 - (1 + x^\alpha)^{-\beta}, \quad x \geq 0.$$

The parameters and the population values of d_1^*, d_2^* of the Burr designs are given below.

Design	X_1	X_2	d_1^*	d_2^*
1a	$B(4.7, 0.55)$	$B(4.7, 0.55)$	$0.000(FSD)$	$0.0000(SSD)$
1b	$B(2.0, 0.65)$	$B(2.0, 0.65)$	$0.0000(FSD)$	$0.0000(SSD)$
1c	$B(4.7, 0.55)$	$B(2.0, 0.65)$	0.1395	0.0784
1d	$B(4.6, 0.55)$	$B(2.0, 0.65)$	0.1368	0.0773
1e	$B(4.5, 0.55)$	$B(2.0, 0.65)$	0.1340	0.0761

On the other hand, we consider four different designs from the lognormal distribution $LN(\mu_k, \sigma_k^2)$. That is, we consider the distributions of $X_k = \exp(\mu_k + \sigma_k Z_k)$, where Z_k are i.i.d. $N(0, 1)$. The parameters and the population values of d_1^*, d_2^* of the lognormal designs

are given below.

Design	X_1	X_2	d_1^*	d_2^*
2a	$LN(0.85, 0.6^2)$	$LN(0.85, 0.6^2)$	$0.0000(FSD)$	$0.0000(SSD)$
2b	$LN(0.85, 0.6^2)$	$LN(0.7, 0.5^2)$	$0.0000(FSD)$	$0.0000(SSD)$
2c	$LN(0.85, 0.6^2)$	$LN(1.2, 0.2^2)$	0.0834	$0.0000(SSD)$
2d	$LN(0.85, 0.6^2)$	$LN(0.2, 0.1^2)$	0.0609	0.0122

To define the multivariate normal designs, we let

$$X_{ki} = (1 - \lambda) \left[\alpha_k + \beta_k \left(\sqrt{\rho} Z_{0i} + \sqrt{1 - \rho} Z_{ki} \right) \right] + \lambda X_{k,i-1},$$

where (Z_{0i}, Z_{1i}, Z_{2i}) are i.i.d. standard normal random variables, mutually independent. The parameters $\lambda = \rho = 0.1$ determine the mutual correlation of X_{1i} and X_{2i} and their autocorrelation. The parameters α_k, β_k are actually the mean and standard deviation of the marginal distributions of X_{1i} and X_{2i}. The marginals and the true values of the multivariate normal designs are:

Design	X_1	X_2	d_1^*	d_2^*
3a	$N(0, 1)$	$N(-1, 16)$	0.1981	$0.0000(SSD)$
3b	$N(0, 16)$	$N(1, 16)$	$0.0000(FSD)$	$0.0000(SSD)$
3c	$N(0, 1)$	$N(1, 16)$	0.1981	0.5967

In computing the suprema in $D_N^{(s)}$ we took a maximum over an equally spaced grid of size n on the range of the pooled empirical distribution. To compute the integral for $C_N^{(2)}$, we took the sum over the same grid points. We chose a total of 5 different subsamples for each sample size $N \in \{50, 500, 1000\}$. We took an equally spaced grid of subsample sizes: for $N = 50$, the subsample sizes are $\{20, 25, \ldots, 45\}$; for $N = 500$ the subsample sizes are $\{50, 100, \ldots, 250\}$; for $N = 1000$ the subsample sizes are $\{100, 200, \ldots, 500\}$.[9] This grid of subsamples are then used to compute the mean of the critical values from the grid, which had the best overall performance among the automatic methods considered by LMW. We used a total of 200 bootstrap repetitions in each case. In computing the suprema and integral in each subsample and bootstrap test statistic, we took the same grid of points as was used in the original test statistic.

---

[9]LMW considered a finer grid, i.e., 20, of different subsample sizes but their simulation results are very close to ours.

To estimate the contact set

$$\hat{B}_{kl}^{(s)} = \left\{ x \in \mathcal{X} : |\bar{D}_{kl}^{(s)}(x)| < c_N \right\},$$

we took the tuning parameter to be $c_N = cN^{-1/3}$, where $c \in \{0.25, 0.50, 0.75\}$ for $s = 1$ and $c \in \{2.0, 3.0, 4.0\}$ for $s = 2$. To compute the test statistics and the contact set, we took the weight function $q(x) = 1$ for all $x \in \mathcal{X}$, i.e., no weighting. In each experiment, the number of replications was $1,000$.

5.2   Simulation Results

Tables 1F - 3S present the rejection probabilities for the tests with nominal size 5%. The simulation standard error is approximately 0.007.

The overall impression is that all the methods considered work reasonably well in samples above 500. The full- sample methods work better than the subsample method under the null hypothesis in small samples . Our procedure is always more powerful than the recentered bootstrap method. In cases (1c,1d,1e,2d below) where the subsample method is better than the recentered bootstrap, our procedure is strictly more powerful than the subsample method. The performance of our procedure depend on the choice of the tuning parameter c_N. If it takes larger values, then the rejection probabilities get closer to those of the recentered bootstrap, but the rejection probabilities increase as c_N decreases. This is consistent with our theory because the rejection probability is a monotonically non-increasing function of c_N. In a reasonable range of the tuning parameter, our procedure strictly dominates the the subsample and recentered bootstrap methods in terms of power without sacrificing size in all of the designs we considered.

The first two designs in Tables 1F and 1S are to evaluate the size performance, especially in the least favorable case of the equality of two distributions, while the other three are to evaluate the power performance. The subsample method tends to over-reject under the null when $N = 50$ but the size distortions are negligible when $N \geq 500$. Our procedure has a good size performance over all tuning parameters we considered in Table 1F, but the case 1b in Table 1S shows that it tends to over-reject when the tuning parameter is chosen to be too small. The last three designs (case 1c, 1d, and 1e in Tables 1F and 1S) are quite conclusive. For moderate and large samples ($N \geq 500$), the subsample method is more powerful than the recentered bootstrap method. However, in these cases, our procedure strictly dominates the subsample method uniformly over all values of c_N we considered. This is so even at small samples ($N = 50$). This remarkable results forcefully demonstrate the power of our procedure.

Tables 2F and 2S give the rejection probabilities under the lognormal designs. The results under the least favorable case 2a are similar to those of 1a and 1b. The design 2b (in Table 2F) is quite instructive because it corresponds to the boundary case under which the two distributions "kiss" at points in the interior of the support. The recentered bootstrap tends to under-reject the null hypothesis and perform very differently from the least favorable case, while our procedure tends to have correct size for a suitable value of c_N. The design 2c in Table 2F shows that, in small sample $N = 50$, the full sample method is more powerful than the subsample method. In design 2d (Tables 2F and 2S), the subsample method is more powerful than the recentered bootstrap method for all sample sizes, but again our procedure dominates the subsample method.

Tables 3F and 3S consider the multivariate normal designs. The designs 3a (Table 2F) and 3b (Tables 3F and 3S) are to evaluate the size characteristics of the tests. The latter designs correspond to the "interior" of the null hypothesis and the rejection probabilities tend to zero as the sample size increases. This is consistent with our theory. The other designs are to evaluate the power performance. They show that the subsample method is less powerful than the recentered bootstrap and the latter is again dominated by our procedure.

Among the Kolmogorov-Smirnov and Cramer-von Mises type tests, the power performance depends on the alternatives. In designs 1c,1d,1e, and 2d, the Kolmogorov-Smirnov type tests are more powerful, while in designs 2c, 3a, and 3c, the Cramer-von Mises type tests are more powerful.

6   Conclusion

This paper proposes a new method for testing stochastic dominance that improves on the existing methods. Specifically, our tests have asymptotic sizes that are exactly correct uniformly over the entire null hypothesis. In addition, we have extended the domain of applicability of our tests to a more general class of situations where the outcome variable is the residual from some semiparametric model. Our simulation study demonstrates that our method works better than existing methods for quite modest sample sizes.

Our setting throughout has been i.i.d. data, but many time series applications call for the treatment of dependent data. In that case, one may have to use a block bootstrap algorithm in place of our wild bootstrap method. We expect, however, that similar results will obtain in that case.

7 Appendix: Mathematical Proofs

Throughout the proof, C denotes a constant that can assume different values in different places.

7.1 Proofs of the Main Results

Lemma A1 : *Let $V(x) = xq(x)$, and $\tilde{V}(x) = |D_{kl}(x)|q(x)$ $x \in \mathcal{X}$, and introduce pseudo metrics $d_V(x, x') = |V(x) - V(x')|$ and $d_{\tilde{V}}(x, x') = |\tilde{V}(x) - \tilde{V}(x')|$. Then,*

$$\log N_{[]}(\varepsilon, \mathcal{X}, d_V) \leq C\left\{\log h^{-1}(C/\varepsilon) - \log(\varepsilon)\right\} \text{ and}$$
$$\log N_{[]}(\varepsilon, \mathcal{X}, d_{\tilde{V}}) \leq C\left\{\log h^{-1}(C/\varepsilon) - \log(\varepsilon)\right\},$$

where $h(\cdot)$ is an arbitrary strictly increasing function.

Proof of Lemma A1 : Fix $\varepsilon > 0$. Note that $|V|$ is uniformly bounded by Assumption 3(ii) and hence for some $C > 0$, $V(x) < C/h(x)$ for all $x > 1$ where $h(x)$ is an increasing function. Put $x = h^{-1}(2C/\varepsilon)$ where $h^{-1}(y) = \inf\{x : h(x) \geq y : x \in \mathcal{X}\}$, so that we obtain

$$V(h^{-1}(2C/\varepsilon)) < \varepsilon/2.$$

Then, we can choose $L(\varepsilon) \leq Ch^{-1}(2C/\varepsilon)$ such that $L(\varepsilon) \geq 1$ and $\sup_{|x| \geq L(\varepsilon)} V(x) < \varepsilon/2$. Partition $[-L(\varepsilon), L(\varepsilon)]$ into intervals $\{\mathcal{X}_m\}_{m=1}^{N_L}$ of length ε with the number of intervals $N_{L(\varepsilon)}$ not greater than : $2L(\varepsilon)/\varepsilon$. Let $\mathcal{X}_0 = \mathcal{X} \setminus [-L(\varepsilon), L(\varepsilon)]$. Then $(\cup_{m \geq 1} \mathcal{X}_m) \cup \mathcal{X}_0$ constitutes the partition of $\mathcal{X}$. In terms of d_V, the diameter of the set $\mathcal{X}_0$ is bounded by ε because $\sup_{x, x' \in \mathcal{X}_0} |V(x) - V(x')| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$, and the diameter of $\mathcal{X}_m$, $m \geq 1$, is also bounded by $C\varepsilon$ because $d_V(x, x') = |V(x) - V(x')| \leq C|x - x'| \leq C\varepsilon$ for all $x, x' \in I_m$. The second inequality uses the fact that $q(x)$ is Lipschitz continuous. Therefore, we have $N(\varepsilon, \mathcal{X}, d_V) \leq 2L(\varepsilon)/\varepsilon + 1 \leq Ch^{-1}(2C/\varepsilon)/\varepsilon + 1$. Hence $\log N(\varepsilon, \mathcal{X}, d_V) \leq C \log h^{-1}(C/\varepsilon) - \log(\varepsilon)$.

We can obtain the same result for $d_{\tilde{V}}$ precisely in the same manner. Note that the properties of $V(x)$ that we used are that it is uniformly bounded and that $|V(x) - V(x')| \leq C|x - x'|$. By Assumption 2(i), $\tilde{V}(x)$ also satisfy these properties. ∎

Lemma A2 : *Let $\mathcal{F} = \left\{h_x^{(s)}(\varphi(\cdot; \theta, \tau)) : (x, \theta, \tau) \in \mathcal{X} \times B_{\Theta \times \mathcal{T}}(\delta)\right\}$. Then,*

$$\sup_{P \in \mathcal{P}} \log N_{[]}(\varepsilon, \mathcal{F}, ||\cdot||_{P,2}) \leq C\varepsilon^{-2d/s_2} + C \log \varepsilon, \text{ if } s = 1, \text{ and}$$
$$\sup_{P \in \mathcal{P}} \log N_{[]}(\varepsilon, \mathcal{F}, ||\cdot||_{P,2}) \leq C\varepsilon^{-d/s_2} + C/\varepsilon, \text{ if } s > 1.$$

Proof of Lemma A2 : Let $\mathcal{H} = \{h_x^{(s)} : x \in \mathbb{R}\}$ and $\gamma_x(y) = 1\{y \leq x\}$. Observe that the function $(x - \varphi)^{s-1}\gamma_x(\varphi)$ is monotone decreasing in φ for all x, z. Define $\Phi = \{\varphi(\cdot; \theta, \tau) : (\theta, \tau) \in \Theta \times \mathcal{T}\}$. Then by the local uniform L_p-continuity condition in Assumption 2(i)(c), we have

$$\log N_{[]}(\varepsilon, \Phi, ||\cdot||_{P,2}) \leq \log N_{[]}(C\varepsilon^{1/s_2}, \Theta \times \mathcal{T}, ||\cdot||_{P,2}) \leq C\varepsilon^{-d/s_2}. \tag{18}$$

Choose ε-brackets $(\varphi_j, \Delta_{1,j})_{j=1}^{N_1}$ of Φ such that $\int \Delta_{1,j}^2 dP \leq \varepsilon^2$. For each j, let $Q_{j,P}$ be the distribution of $\varphi_j(W)$ where W_j is distributed as P.

Suppose $s = 1$. In Lemma A1, we take $h(x) = (\log x)^{s_2/d}$ and obtain $\log N(\varepsilon, \mathcal{X}, d_V) \leq C\varepsilon^{-d/s_2} - C\log(\varepsilon)$. Fix $\varepsilon > 0$ and choose the partition $\mathcal{X} = \cup_{m=0}^{N_L} \mathcal{X}_m$ as above. Let x_m be the center of $\mathcal{X}_m, m \geq 1$,

and let x_0 be any point in $\mathcal{X}_0$. Write

$$\sup_{\varphi \in \Phi_j} \sup_{x \in \mathcal{X}_m} \left| \gamma_x(\varphi(w))q(x) - \gamma_{x_m}(\varphi_j(w))q(x_m) \right| \tag{19}$$
$$\leq \sup_{\varphi \in \Phi_j} \sup_{x \in \mathcal{X}_m} \left| \gamma_x(\varphi(w))q(x) - \gamma_{x_m}(\varphi_j(w))q(x) \right|$$
$$+ \sup_{\varphi \in \Phi_j} \sup_{x \in \mathcal{X}_m} \left| \gamma_{x_m}(\varphi_j(w))q(x) - \gamma_{x_m}(\varphi_j(w))q(x_m) \right| = \Delta^*_{m,j}(w), \text{ say.}$$

Then,

$$\mathbf{E}\left[\sup_{\varphi \in \Phi_j} \sup_{x \in \mathcal{X}_m} \left| \gamma_x(\varphi(W)) - \gamma_{x_m}(\varphi_j(W)) \right|^2 q(x)^2 | W_1 \right]$$
$$\leq \mathbf{E}\left[\sup_{x \in \mathcal{X}_m} 1\{x - \Delta_{1,j}(W_1) - |x - x_m| \leq \varphi_j(W) \leq x + \Delta_{1,j}(W_1) + |x - x_m|\} q(x)^2 | W_1 \right]$$
$$\leq \sup_{x \in \mathcal{X}_0} q(x)^2 1\{m = 0\} + C\{\varepsilon + \Delta_{1,j}(W_1)\} 1\{m \geq 1\}$$
$$\leq C\varepsilon^2 1\{m = 0\} + C\{\varepsilon + \Delta_{1,j}(W_1)\} 1\{m \geq 1\} \leq C\varepsilon + C\Delta_{1,j}(W_1).$$

The second inequality is obtained by splitting the supremum into the case $m = 0$ and the case $m \geq 1$. The third and fourth inequalities follow because $\sup_{x \in \mathcal{X}_0} |q(x)|^2 \leq \sup_{|x| > L(\varepsilon)} |x|^2 |q(x)|^2 \leq \varepsilon^2/4$. The last inequality follows by subsuming $C\varepsilon^2$ into $C\varepsilon$. Similarly we can show that

$$\mathbf{E}\left[\sup_{\varphi \in \Phi_j} \sup_{x \in \mathcal{X}_m} \left| \gamma_{x_m}(\varphi_j(W))q(x) - \gamma_{x_m}(\varphi_j(W))q(x_k) \right|^2 | W_1 \right] \leq C|q(x) - q(x_k)|^2 \leq C\varepsilon^2.$$

We conclude that $||\Delta^*_{k,j}||_{P,2} \leq C\varepsilon^{1/2}$. Hence

$$\sup_{P \in \mathcal{P}} \log N_{[]}(C\varepsilon^{1/2}, \mathcal{F}, L_p(P)) \leq \log N_{[]}(\varepsilon, \Phi, L_p(P)) + \log N_{[]}(\varepsilon, \mathcal{X}, d_V)$$
$$\leq C\varepsilon^{-d/s_2} - C\log\varepsilon.$$

Suppose $s > 1$. Then $(x - \varphi)^{s-1}\gamma_x(\varphi)q(x)$ is infinite times continuously differentiable in φ with bounded derivatives. Choose ε-brackets $(\varphi_j, \Delta_{1,j})_{j=1}^{N_1}$ of Φ as before. Then for each j, we choose also ε-brackets $(h_{k,j}, \Delta_{2,k,j})_{k=1}^{N_2(j)}$ such that for all $h \in \mathcal{H}$, $|h - h_{k,j}| \leq \Delta_{2,k,j}$ and $\int \Delta^p_{2,k,j} dQ_{j,P} \leq \varepsilon^p$. Since $\mathcal{H}$ is monotone increasing and uniformly bounded, we can take $N_{2(j)} \leq N_{[]}(\varepsilon, \mathcal{H}, L_p(Q_{j,P})) \leq C/\varepsilon$ by Birman and Solomjak (1967). Now, observe that

$$\left| h(\varphi(w)) - h_{k,j}(\varphi_j(w)) \right| \leq \left| h(\varphi(w)) - h(\varphi_j(w)) \right| + \left| h(\varphi_j(w)) - h_{k,j}(\varphi_j(w)) \right|$$
$$\leq \left| h(\varphi_j(w) - \Delta_{1,j}(w)) - h(\varphi_j(w) + \Delta_{1,j}(w)) \right| + \Delta_{2,k,j}(\varphi_j(w))$$
$$\leq C\Delta_{1,j}(w) + \Delta_{2,k,j}(\varphi_j(w)).$$

The last inequality is due to the fact that h is Lipschitz with a uniformly bounded coefficient. Take

$$\Delta^*_{k,j}(w) = C\Delta_{1,j}(w) + \Delta_{2,k,j}(\varphi_j(w)).$$

and note that $\int (\Delta^*_{k,j})^2 dP \leq C\varepsilon$. Therefore, $\log N_{[]}(\varepsilon, \mathcal{F}, ||\cdot||_{P,2}) \leq \log N_{[]}(C\varepsilon, \Phi, ||\cdot||_{P,2}) + C/\varepsilon$. Observe that the constants above do not depend on the choice of the measure P. By (18), we obtain the wanted results. ∎

Proof of Theorem 1 : Consider the following empirical process

$$v_{kN}^{(s)}(f) = \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \{f(W_i) - \mathbf{E}\left[f(W_i)\right]\}, \ f \in \mathcal{F} \tag{20}$$

where $\mathcal{F}$ is as defined in Lemma A2 above. Recall that $\mathcal{F}$ is uniformly bounded. Therefore, by Lemma A2 and $d/s_2 < 1$, $\mathcal{F}$ is asymptotically equicontinuous uniformly in $P \in \mathcal{P}$ and totally bounded for all $P \in \mathcal{P}$. By Proposition 3.1 and Theorem 2.3 of Giné and Zinn (1991) (See also Theorem 2.8.4 of van der Vaart and Wellner (1996)), the weak convergence of $v_{kN}^{(s)}(\cdot)$ in $l_\infty(\mathcal{F})$ uniform over $P \in \mathcal{P}$ immediately follows. Here $l_\infty(\mathcal{F})$ denotes the space of uniformly bounded real functions on $\mathcal{F}$.

Let $\mathcal{F}_N = \{f(\cdot; \theta, \tau) : (\theta, \tau) \in \Theta_N \times \mathcal{T}_N\}$. Let us consider the covariance kernel of the process $v_{kN}^{(s)}(f)$ in $f \in \mathcal{F}_N$. Choose $f_{1,N}(w) = f(w; \theta_{1,N}, \tau_{1,N})$ and $f_{2,N}(w) = f(w; \theta_{2,N}, \tau_{2,N})$, and let $f_{1,0}(w) = f(w; \theta_0, \tau_0)$ and $f_{2,0}(w) = f(w; \theta_0, \tau_0)$. Then,

$$\begin{aligned}
&|Pf_{1,N}f_{2,N} - Pf_{1,N}Pf_{2,N} - \{Pf_{1,0}f_{2,0} - Pf_{1,0}Pf_{2,0}\}| \\
\leq \ & C\left|P(f_{1,N} - f_{1,0})\right| + C\left|P(f_{2,N} - f_{2,0})\right| \\
\leq \ & C\left\{||\theta_{1,N} - \theta_0|| + ||\tau_{1,N} - \tau_{1,0}||_{P,2}\right\} + C\left\{||\theta_{2,N} - \theta_0|| + ||\tau_{2,N} - \tau_{2,0}||_{P,2}\right\} \to 0.
\end{aligned}$$

The convergence is uniform over all $P \in \mathcal{P}$. Hence, the uniform weak convergence of $v_{kN}^{(s)}(\cdot)$ in $l_\infty(\mathcal{F}_N)$ to a Gaussian process follows by Theorem 2.11.23 of van der Vaart and Wellner (1996). The covariance kernel of the Gaussian process is given by $Pf_{1,0}f_{2,0} - Pf_{1,0}Pf_{2,0}$.

Suppose that we are under the null hypothesis. Note that

$$\begin{aligned}
&\sqrt{N}\bar{D}_k^{(s)}(x, \hat{\theta}, \hat{\tau}) - \frac{\sqrt{N}}{(s-1)!}\mathbf{E}h_x^{(s)}(\varphi_k(W; \theta_0, \tau_0)) \\
= \ & \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h_x^{(s)}(\varphi_k(W_i; \hat{\theta}, \hat{\tau})) - \mathbf{E}\left[h_x^{(s)}(\varphi_k(W_i; \hat{\theta}, \hat{\tau}))\right] \right\} \\
& + \frac{\sqrt{N}}{(s-1)!} \left\{ \mathbf{E}\left[h_x^{(s)}(\varphi_k(W_i; \hat{\theta}, \hat{\tau})) - h_x^{(s)}(\varphi_k(W_i; \theta_0, \tau_0))\right] \right\} \\
= \ & \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h_x^{(s)}(\varphi_k(W_i; \theta_0, \tau_0)) - \mathbf{E}\left[h_x^{(s)}(\varphi_k(W; \theta_0, \tau_0))\right] \right\} \\
& + \frac{\sqrt{N}}{(s-1)!}\Gamma_{k,P}(x)[\hat{\theta} - \theta_0, \hat{\tau} - \tau_0] + o_P(1)
\end{aligned}$$

uniformly over $P \in \mathcal{P}$. Hence we write

$$\begin{aligned}
&\sqrt{N}\bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) - \frac{\sqrt{N}}{(s-1)!}\mathbf{E}h_{x,kl}^{\Delta}(W) \\
= \ & \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h_{x,kl}^{\Delta}(W_i) + \psi_{x,kl}^{\Delta}(W_i) - \mathbf{E}\left[h_{x,kl}^{\Delta}(W) + \psi_{x,kl}^{\Delta}(W)\right] \right\} + o_P(1).
\end{aligned}$$

25

Under $P \in \mathcal{P}_0$, let $\delta_{kl}(x) = \mathbf{E}\left[h^{\Delta}_{x,kl}(W)\right]$. Therefore,

$$\sqrt{N} \bar{D}^{(s)}_{kl}(x, \hat{\theta}, \hat{\tau}) - \frac{\sqrt{N} \delta_{kl}(x)}{(s-1)!}$$

$$= \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h^{\Delta}_{x,kl}(W_i) + \psi^{\Delta}_{x,kl}(W_i) - \mathbf{E}\left[h^{\Delta}_{x,kl}(W) + \psi^{\Delta}_{x,kl}(W)\right] \right\} + o_P(1).$$

Let $\Psi_{kl} = \{\psi^{\Delta}_{x,kl} : x \in \mathcal{X}\}$ and $\mathcal{H}_{kl} = \{h^{\Delta}_{x,kl} : x \in \mathcal{X}\}$. Consider the class of functions $\mathcal{F}_{kl} = \{h + \psi : (h, \psi) \in \mathcal{H}_{kl} \times \Psi_{kl}\}$. By Lemma A2 above and Theorem 6 of Andrews (1994), we have $\sup_{P \in \mathcal{P}} \log N_{[]}(\varepsilon, \mathcal{H}_{kl}, ||\cdot||_{P,2}) < C\varepsilon^{-((2d/s_2) \vee 1)}$ and by Assumption 1(iv), $\sup_{P \in \mathcal{P}} \log N_{[]}(\varepsilon, \Psi_{kl}, ||\cdot||_{P,2}) < C\varepsilon^{-d}$. Note that $\mathcal{H}_{kl}$ is uniformly bounded and Ψ_{kl} has an envelope $2\bar{\psi}$ such that $||2\bar{\psi}||_{P,2+\delta} < C$. Therefore, $\mathcal{F}_{kl}$ is a P-Donsker class. The computation of the covariance kernel of the limiting Gaussian process is straightforward. Therefore,

$$\sqrt{N} \bar{D}^{(s)}_{kl}(\cdot, \hat{\theta}, \hat{\tau}) - \frac{\sqrt{N} \delta_{kl}(\cdot)}{(s-1)!} \Longrightarrow \nu^{(s)}_{kl},$$

uniformly in $P \in \mathcal{P}$. This yields the limit results both for the Kolmogorov-Smirnov type test and the Cramér-von Mises type test. ∎

Let $\mathcal{F}$ be as defined in Lemma A2 above. Let the bootstrap process

$$\nu^{*(s)}_{k,b}(x) = \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h^{(s)}_x(\varphi_k(W^*_{i,b}; \hat{\theta}^*_b, \hat{\tau}^*_b)) - \mathbf{E}_N\left[h^{(s)}_x(\varphi_k(W_i; \hat{\theta}, \hat{\tau}))\right] \right\}, \quad x \in \mathcal{X},$$

where $W^*_i = (Y^*_i, Z_i)$. The following lemma is a key step for the bootstrap consistency of the test (Theorem 2). In the following, we use the usual stochastic convergence notation o_{P^*} and O_{P^*} that is with respect to the conditional distribution given $\mathcal{G}_N$.

Lemma A3 : *Under the assumptions of* Theorem 2,

$$\nu^{*(s)}_{k,b}(x) = \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h^{(s)}_x(\varepsilon^*_{ki,b}(\theta_0, \tau_0)) - \mathbf{E}_N[h^{(s)}_x(\varepsilon^*_{ki,b}(\theta_0, \tau_0))] \right\}$$

$$+ \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \psi_{x,P}(\varepsilon^*_{ib}(\theta_0, \tau_0)) + o_{P^*}(1)$$

in P uniformly in $P \in \mathcal{P}$. Furthermore, let $\mathcal{F}_N$ be the class defined in the proof of Theorem 1. *Then $\nu^{*(s)}_{kl,b} \to \nu^{(s)}_{kl}$ weakly in $l_\infty(\mathcal{F})$ in P uniformly in $P \in \mathcal{P}$.*[10]

Proof of Lemma A3 : Observe that

$$X^*_{ki,b}(\theta, \tau, \tilde{\theta}, \tilde{\tau}) = g(Z_{ki}, \tilde{\theta}, \tilde{\tau}) - g(Z_{ki}, \theta, \tau) + \varepsilon^*_{ki,b}(\tilde{\theta}, \tilde{\tau}).$$

---

[10]For the definition of the bootstrap uniform weak convergence, see Section 7.2 below.

Then, we are interested in the weak convergence of the following bootstrap empirical process

$$
\frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}_b^*, \hat{\tau}_b^*)) - \mathbf{E}_N \left[h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}, \hat{\tau})) \right] \right\} \tag{21}
$$

$$
= \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}_b^*, \hat{\tau}_b^*)) - h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}, \hat{\tau})) \right\}
$$

$$
+ \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}, \hat{\tau})) - \mathbf{E}_N[h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}, \hat{\tau}))] \right\}.
$$

We consider the bootstrap empirical process

$$
\nu_1^*(x, \theta, \tau) = \frac{1}{\sqrt{N}(s-1)!} \sum_{i=1}^{N} \left\{ h_x^{(s)}(\varepsilon_{ki,b}^*(\theta, \tau)) - \mathbf{E}_N[h_x^{(s)}(\varepsilon_{ki,b}^*(\theta, \tau))] \right\}.
$$

We can view the process as indexed by $h_x^{(s)} \circ \varphi(\cdot; \theta, \tau) \in \mathcal{F}$ and the process as a $l_\infty(\mathcal{F})$-valued random (potentially nonmeasurable) element, where $\mathcal{F}$ is the class defined in Lemma A2. By Proposition 3.1 of Giné and Zinn (1991), Lemma A2 implies that $\mathcal{F}$ is finitely uniformly pregaussian. Theorem 2.3 of Giné and Zinn (1991) combined with Proposition B1 below gives the central limit theorem for the bootstrap empirical process ν_1^* and so, the law of the process converges weakly to the law of a centered Gaussian process ν_1 on $\mathcal{X}$ conditionally on $\mathcal{G}_N$ in P uniformly in $P \in \mathcal{P}$. The Gaussian process ν_1 has a covariance kernel

$$
Cov_P(h_{x_1,kl}^\Delta(W; \theta_1, \tau_1), h_{x_2,kl}^\Delta(W; \theta_2, \tau_2)),
$$

where $h_{x,kl}^\Delta(W; \theta, \tau) = h_x^{(s)}(\varphi_k(W; \theta, \tau)) - h_x^{(s)}(\varphi_l(W; \theta, \tau))$. The asymptotic equicontinuity associated with this bootstrap uniform CLT implies that

$$
\sup_x \left| \nu_1^*(x, \hat{\theta}, \hat{\tau}) - \nu_1^*(x, \theta_0, \tau_0) \right| = o_{P^*}(1) \text{ in } P,
$$

uniformly in P. We turn to the term in the second line of (21). Then write it as

$$
\frac{\sqrt{N}}{(s-1)!} \left\{ \mathbf{E}_N h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}_b^*, \hat{\tau}_b^*)) - \mathbf{E}_N h_x^{(s)}(X_{ki,b}^*(\hat{\theta}, \hat{\tau}, \hat{\theta}, \hat{\tau})) \right\}. \tag{22}
$$

By Assumptions 1(iv) and 4(ii), we deduce that the process in (22) is asymptotically equivalent to

$$
\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \psi_{x,k,P}(W_{i,b}^*) + o_{P^*}(c_N) \text{ in } P,
$$

uniformly in $P \in \mathcal{P}$. Recall that $W_{i,b}^* = ([Y_{ki} - g(Z_{ki}, \hat{\theta}, \hat{\tau})]^* + g(Z_{ki}, \hat{\theta}, \hat{\tau}), Z_{ki})_{k=1}^K$ where $[\cdot]^*$ is the part obtained through the wild-bootstrap procedure. The above process is a bootstrap empirical process already centered. The wanted result follows from the asymptotic equicontinuity condition, Assumption 3(i), and the bootstrap CLT in Proposition B1. ∎

Proof of Theorem 2 : (i) Let $B_{kl}^{(s)} = \{x : D_{kl}^{(s)}(x) = 0\}$. Denote $[B_{kl}^{(s)}]^\varepsilon = \{x : q(x)|D_{kl}^{(s)}(x)| < \varepsilon\}$. We will

show the following at the end of the proof:

$$P\left\{B_{kl}^{(s)} \subset \hat{B}_{kl}^{(s)} \subset [B_{kl}^{(s)}]^{2c_N}\right\} \to 1, \text{ uniformly in } P \in \mathcal{P}. \tag{23}$$

Note that $[B_{kl}]^{2c_N} \subset \mathcal{X}$ from some sufficiently large N on by Assumption 1(i)(b). Let $\bar{\nu}_{kl,b}^*(x) = \sqrt{N}\bar{D}_{kl,b}^{(s)*}(x)$ and define $1_K = \{B_{kl} \neq \varnothing \text{ for all } k \neq l\}$ and $1_K^{2c_N} = 1\{[B_{kl}]^{2c_N} \neq \varnothing \text{ for all } k \neq l\}$. Also define $\hat{1}_K = \{\hat{B}_{kl} \neq \varnothing \text{ for all } k \neq l\}$ and $\hat{1}_K^{2c_N} = 1\{[\hat{B}_{kl}]^{2c_N} \neq \varnothing \text{ for all } k \neq l\}$. Since $P\left\{B_{kl}^{(s)} \subset \hat{B}_{kl}^{(s)} \subset [B_{kl}^{(s)}]^{2c_N}\right\} \to 1$ by (23) and $c_N \to 0$, it follows that with probability (uniformly over $P \in \mathcal{P}$) approaching one,

$$\sup_{x \in B_{kl}} \bar{\nu}_{kl,b}^*(x)1_K + \pi_N(1 - 1_K^{2c_N}) \tag{24}$$
$$\leq \sup_{x \in \hat{B}_{kl}} \bar{\nu}_{kl,b}^*(x)\hat{1}_K + \pi_N(1 - \hat{1}_K)$$
$$\leq \sup_{x \in [B_{kl}]^{2c_N}} \bar{\nu}_{kl,b}^*(x)1_K^{2c_N} + \pi_N(1 - 1_K).$$

We can apply the weak convergence of the process $\bar{\nu}_{kl,b}^*$ in Lemma A3 so that for each $c > 0$,

$$P\left\{\sup_h |\mathbf{E}[h(\bar{\nu}_{kl,b}^*)|\mathcal{G}_N] - \mathbf{E}h(\nu_{kl}^{(s)})| > c, \text{i.o.}\right\} \to 0,$$

where the supremum is over the space of bounded linear functionals h on $l_\infty(\mathcal{F})$ and the convergence is uniform over $P \in \mathcal{P}$. (Corollary 2.7 of Giné and Zinn (1991). See also Giné (1997)). Furthermore, since the class of functions indexing the bootstrap empirical process $\bar{\nu}_{kl,b}^*$ is uniformly pregaussian (as implied by the bracketing entropy conditions in Lemma A2 and Proposition 3.1 of Giné and Zinn (1991)), the process ν satisfies that

$$\lim_{\varepsilon \to 0} \sup_{P \in \mathcal{P}} \mathbf{E}_P \sup_{f_1, f_2 \in \mathcal{F}: ||f_1 - f_2||_{P,2} < \varepsilon} |\nu(f_1) - \nu(f_2)| = 0. \tag{25}$$

(See Definition 2.4 of Giné and Zinn (1991).)

Suppose that for all k, l, $B_{kl} \neq \varnothing$ so that $P \in \mathcal{P}_{00}^{KS}$. From (24), with probability approaching one,

$$\sup_{x \in B_{kl}} \bar{\nu}_{kl,b}^*(x) \leq \sup_{x \in \hat{B}_{kl}} \bar{\nu}_{kl,b}^*(x)\hat{1}_K + \pi_N(1 - \hat{1}_K) \leq \sup_{x \in [B_{kl}]^{2c_N}} \bar{\nu}_{kl,b}^*(x).$$

Therefore, we deduce that with probability approaching one,

$$\left|\sup_{x \in \hat{B}_{kl}} \bar{\nu}_{kl,b}^*(x)\hat{1}_K + \pi_N(1 - \hat{1}_K) - \sup_{x \in B_{kl}} \bar{\nu}_{kl,b}^*(x)\right| \tag{26}$$
$$\leq \sup_{x \in [B_{kl}]^{2c_N}} \bar{\nu}_{kl,b}^*(x) - \sup_{x \in B_{kl}} \bar{\nu}_{kl,b}^*(x)$$
$$\leq \sup_{x \in \mathcal{X}: d_{\bar{V}}(x,x') \leq 2c_N} \left|\bar{\nu}_{kl,b}^*(x) - \bar{\nu}_{kl,b}^*(x')\right|.$$

Given samples $\{W_i\}_{i=1}^N$, we define a random metric:

$$\hat{d}_{kl}(x, x') = \left\{\frac{1}{N}\sum_{i=1}^N \left|\{h_x^{(s)}(\hat{X}_{ki}) - h_x^{(s)}(\hat{X}_{li})\} - \{h_{x'}^{(s)}(\hat{X}_{ki}) - h_{x'}^{(s)}(\hat{X}_{li})\}\right|^2\right\}^{1/2}.$$

From the proof of Theorem 1, the class $\mathcal{F}$ is a uniform Donsker, and hence it follows that

$$\sup_{d_{\bar{V}}(x,x') < \delta} \hat{d}_{kl}(x, x') = \sup_{d_{\bar{V}}(x,x') < \delta} d_{kl}(x, x') + O_P(N^{-1/2})$$

uniformly in $P \in \mathcal{P}$, where $d_{kl}(x, x') = \{\mathbf{E}|\{h_x^{(s)}(X_{ki}) - h_x^{(s)}(X_{li})\} - \{h_{x'}^{(s)}(X_{ki}) - h_{x'}^{(s)}(X_{li})\}|^2\}^{1/2}$. Let us

define the following process:

$$\nu_{k,b}^{*(s)}(x;\hat{\theta},\hat{\tau}) = \frac{1}{\sqrt{N}(s-1)!}\sum_{i=1}^{N}\left\{h_x^{(s)}(\varepsilon_{ki,b}^*(\hat{\theta},\hat{\tau})) - \mathbf{E}_N\left[h_x^{(s)}(\varepsilon_{ki,b}^*(\hat{\theta},\hat{\tau}))\right]\right\}$$
$$+\frac{1}{\sqrt{N}(s-1)!}\sum_{i=1}^{N}\psi_{x,k,P}(W_{i,b}^*),$$

where $W_{i,b}^* = [g(Z_{ki};\hat{\theta},\hat{\tau}) + \varepsilon_{ki}^*(\hat{\theta},\hat{\tau}), Z_{ki}]_{k=1}^K$. By Assumption 4(iv), we have

$$c_{kl}\left\{\hat{d}_{kl}(x,x')\right\}^{p_{kl}} \le Cd_{\tilde{V}}(x,x') + O_P(N^{-p_{kl}/2}) = Cd_{\tilde{V}}(x,x') + O_P(N^{-1/2}). \tag{27}$$

Therefore,

$$\sup_{x\in\mathcal{X}:\ d_{\tilde{V}}(x,x')\le 2c_N}\left|\bar{\nu}_{kl,b}^*(x) - \bar{\nu}_{kl,b}^*(x')\right|$$
$$\le \sup\left\{\left|\bar{\nu}_{kl,b}^*(x) - \bar{\nu}_{kl,b}^*(x')\right| : x\in\mathcal{X}:\ c_{kl}\left\{\hat{d}_{kl}(x,x')\right\}^{p_{kl}}\le 3c_N\right\},$$

with probability approaching one. The sequence $3c_N$ is introduced to take into account the term $O_P(N^{-1/2})$ in (27). Following the proof of Lemma A3, we obtain that for some $\delta > 0$,

$$\mathbf{E}_P\left[\sup_{x\in\mathcal{X}:\ c_{kl}\{\hat{d}_{kl}(x,x')\}^{p_{kl}}\le 3c_N}\left|\bar{\nu}_{kl,b}^*(x) - \bar{\nu}_{kl,b}^*(x')\right|\Big|\mathcal{G}_N\right]$$
$$= \mathbf{E}_P\left[\sup_{x\in\mathcal{X}:\ c_{kl}\{\hat{d}_{kl}(x,x')\}^{p_{kl}}\le 3c_N}\left|\nu_{k,b}^{*(s)}(x;\hat{\theta},\hat{\tau}) - \nu_{k,b}^{*(s)}(x';\hat{\theta},\hat{\tau})\right|\Big|\mathcal{G}_N\right] + o_{P^*}(c_N) \text{ in } P$$

uniformly in P. Conditional on $\mathcal{G}_N$, the expected supnorm in the second line is bounded by that of a symmetrized one by the symmetrization lemma. This symmetrized process is sub-Gaussian with respect to $\hat{d}_{kl}$ by Hoeffding's inequality (e.g. van der Vaart and Wellner (1996), p.101), and hence by Corollary 2.2.8 of van der Vaart and Wellner (1996), we have

$$\mathbf{E}_P\left[\sup_{x\in\mathcal{X}:\ \hat{d}_{kl}(x,x')\le 3c_N^{1/p_{kl}}/c_{kl}}\left|\bar{\nu}_{kl,b}^*(x) - \bar{\nu}_{kl,b}^*(x')\right|\Big|\mathcal{G}_N\right] \le C\int_0^{Cc_N^{1/p_{kl}}}\sqrt{\log D(\varepsilon,\hat{d}_{kl})}d\varepsilon$$

where $D(\varepsilon,\hat{d}_{kl})$ is the packing number of $\mathcal{X}$ with respect to $\hat{d}_{kl}$. Note that $d_{\tilde{V}}^{1/p_{kl}}(x,x')/c_{kl} < \max\{\varepsilon/2, c_N/2\}$ implies $\hat{d}_{kl}(x,x') < \varepsilon$ with probability approaching one by (31). Hence we observe that

$$\log D(\varepsilon,\hat{d}_{kl}) \le \log D(\max\{\varepsilon/2, c_N/2\}, Cd_{\tilde{V}}^{1/p_{kl}}) \le \log D(\max\{\varepsilon^{p_{kl}}/2, c_N^{p_{kl}}/2\}, Cd_{\tilde{V}})$$
$$\le \log D(\varepsilon^{p_{kl}}/2, Cd_{\tilde{V}}) \le -C\log(\varepsilon),$$

with probability approaching one. The last inequality is obtained by choosing $h(x) = C_1 x^{1-C_2}$ with appropriate constants C_1 and C_2, and applying Lemma A1. Therefore,

$$\mathbf{E}_P\left[\sup_{x\in\mathcal{X}:\ \hat{d}_{kl}(x,x')\le Cc_N}\left|\bar{\nu}_{kl,b}^*(x) - \bar{\nu}_{kl,b}^*(y)\right|\Big|\mathcal{G}_N\right] = O_P(c_N\sqrt{-\log c_N})$$

29

uniformly in $P \in \mathcal{P}$. From (26), we conclude that

$$\sup_{x \in \hat{B}_{kl}} \bar{\nu}_{kl,b}^*(x) \hat{1}_K + \pi_N (1 - \hat{1}_K)$$
$$= \sup_{x \in B_{kl}} \bar{\nu}_{kl,b}^*(x) + O_{P^*}(c_N \sqrt{-\log c_N}) \text{ in } P \text{ uniform over } P \in \mathcal{P}_{00}^{KS}.$$

Since $\sup_{x \in B_{kl}^{(s)}} \bar{\nu}_{kl,b}^*(x) \xrightarrow[D^*]{} \sup_{x \in B_{kl}^{(s)}} \nu_{kl}^{(s)}(x)$ uniformly in P, we deduce that

$$c_{\alpha,\infty}^{*KS} = c_\alpha^{KS} + O_P(1/\sqrt{N}) + O_P(c_N \sqrt{-\log c_N}) = c_\alpha^{KS} + O_P(c_N \sqrt{-\log c_N}),$$

uniformly in $P \in \mathcal{P}_{00}^{KS}$, where c_α^{KS} is such that

$$c_\alpha^{KS} = \inf_{c \in \mathbb{R}} \left\{ P \left\{ \min_{k \neq l} \sup_{x \in B_{kl}} \nu_{kl}^{(s)}(x) \leq c_\alpha^{KS} \right\} \geq 1 - \alpha \right\}.$$

Hence $P\{D_N^{(s)} > c_{\alpha,\infty}^{*KS}\} = \alpha + O_P(c_N \sqrt{-\log c_N})$. This result yields the first result of asymptotic similarity.

Now, suppose that $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{KS}$. In this case, for some k, l, we have $D_{kl}^{(s)}(x) < 0$ for all $x \in \mathcal{X}$ so that for some k, l, $B_{kl} = \varnothing$ and $1_K = 0$. From (24),

$$\pi_N(1 - 1_K^{2c_N}) \leq \min_{k \neq l} \sup_{x \in \hat{B}_{kl}} \bar{\nu}_{kl,b}^*(x) \hat{1}_K + \pi_N(1 - \hat{1}_K), \tag{28}$$

with probability approaching one. Recall that $1_K^{2c_N} = 0$ when $[B_{kl}]^{2c_N} = \varnothing$ for some $k \neq l$. For k, l such that $B_{kl} = \varnothing$ or $\sup_{x \in \mathcal{X}} |D_{kl}^{(s)}(x)| > 0$, there exists N_0 such that for all $N \geq N_0$, $\sup_{x \in \mathcal{X}} |D_{kl}^{(s)}(x)| > 2c_N > 0$ so that $[B_{kl}]^{2c_N} = \varnothing$. Hence from some large N on, we have $[B_{kl}]^{2c_N} = \varnothing$ or $1_K^{2c_N} = 0$. The bootstrap test statistic on the right-hand side of (28) is bounded from below by π_N from some large N on. Therefore,

$$c_{\alpha,\infty}^{*KS} \geq \pi_N, \tag{29}$$

from some large N on. Now, by Theorem 1,

$$P\left[\min_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} q(x) \bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) \geq c_{\alpha,\infty}^{*KS} \right]$$
$$\leq P\left[\min_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} q(x) \left(\bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) - D_{kl}^{(s)}(x) + D_{kl}^{(s)}(x) \right) \geq \pi_N \right] \text{ (from some large } N \text{ on)}$$
$$\leq P\left[\max_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} q(x) \{ \bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) - D_{kl}^{(s)}(x) \} \geq \pi_N/2 \right]$$
$$+ P\left[\min_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} q(x) D_{kl}^{(s)}(x) \geq \pi_N/2 \right].$$

Since the term $\max_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} q(x) \{ \bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) - D_{kl}^{(s)}(x) \}$ is uniformly asymptotically tight uniform over $P \in \mathcal{P}$ and $\pi_N \to \infty$, the first probability vanishes. Under H_0, the term $\min_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} q(x) D_{kl}^{(s)}(x)$ is less than or equal to zero uniformly over $P \in \mathcal{P}_0$, and hence the second probability also vanishes uniformly over $P \in \mathcal{P}_0$. Hence we obtain the wanted result.

Now, let us turn to the result for Cramér-von Mises type tests. Noting that $\mathcal{P}_{00}^{CM} \subset \mathcal{P}_{00}^{KS}$ and using the previous arguments we can show that uniformly in $P \in \mathcal{P}_{00}^{CM}$,

$$c_{\alpha,N}^{*CM} = c_\alpha^{CM} + O_P(c_N \sqrt{-\log c_N}),$$

where c_α^{CM} is such that

$$c_\alpha^{CM} = \inf_{c \in \mathbb{R}} \left\{ P \left\{ \min_{k \neq l} \int_{B_{kl}^{(s)}} \left(\max\{\nu_{kl}^{(s)}(x), 0\} \right)^2 dx \leq c_\alpha^{CM} \right\} \geq 1 - \alpha \right\}.$$

For $P \in \mathcal{P}_0 \backslash \mathcal{P}_{00}^{CM}$, observe that we have $c_{\alpha,N}^{*CM} \geq \pi_N$ as before,

$$P \left[\min_{k \neq l} N \int_{B_{kl}^{(s)}} (\max\{q(x) \bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}), 0\})^2 dx \geq c_{\alpha,N}^{CM} \right].$$

Observe that $\max(a + b, 0) \leq \max(a, 0) + \max(b, 0)$. Since we can bound the probability above by

$$P \left[\max_{k \neq l} 2N \int_{B_{kl}^{(s)}} \left(\max\{q(x)(\bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) - D_{kl}^{(s)}(x)), 0\} \right)^2 dx \geq \pi_N/2 \right]$$

$$+ P \left[\min_{k \neq l} 2N \int_{B_{kl}^{(s)}} \left(\max\{q(x) D_{kl}^{(s)}(x), 0\} \right)^2 dx \geq \pi_N/2 \right],$$

for some sufficiently large N on. By the asymptotic tightness of the empirical process, the first probability converges to zero. The second probability also converges to zero because under the null hypothesis, $D_{kl}^{(s)}(x) \leq 0$ for some $k \neq l$.

Let us turn to (23). Since the empirical process $\sqrt{N} q(x) \{D_k^{(s)}(x) - \bar{D}_k^{(s)}(x)\}$ is asymptotically tight for each $k = 1, \ldots, K$, $P\{\sup_x(q(x)|D_l^{(s)}(x) - \bar{D}_l^{(s)}(x)| + q(x)|D_k^{(s)}(x) - \bar{D}_k^{(s)}(x)|) > c_N\} \to 0$ by the choice of $c_N \to 0$ and $c_N \sqrt{N} \to \infty$. For any $x \in B_{kl}$ so that $D_{kl}^{(s)}(x) = 0$, by the triangular inequality,

$$q(x) \left| \bar{D}_{kl}^{(s)}(x) \right| \leq q(x) \left| D_l^{(s)}(x) - \bar{D}_l^{(s)}(x) \right| + q(x) \left| D_k^{(s)}(x) - \bar{D}_k^{(s)}(x) \right| \leq c_N$$

with probability approaching one. Thus we deduce that $P\{B_{kl}^{(s)} \subset \hat{B}_{kl}^{(s)}\} \to 1$. Now, for any $x \in \hat{B}_{kl}$, by the triangular inequality,

$$q(x) \left| D_{kl}^{(s)}(x) \right| \leq c_N + q(x) \left| D_l^{(s)}(x) - \bar{D}_l^{(s)}(x) \right| + q(x) \left| D_k^{(s)}(x) - \bar{D}_k^{(s)}(x) \right| \leq 2c_N,$$

with probability approaching one. Therefore, $P\{\hat{B}_{kl}^{(s)} \subset [B_{kl}^{(s)}]^{2c_N}\} \to 1$. The convergence uniform over $P \in \mathcal{P}$ immediately falls from the fact that $\sqrt{N}\{D_k^{(s)}(x) - \bar{D}_k^{(s)}(x)\}$ is asymptotically tight uniformly over $P \in \mathcal{P}$. ∎

Proof of Theorem 3: (i) The contact sets $B_{kl}^{(s)}$ are not empty for all $k \neq l$ under each alternative hypothesis because the marginal distributions of X_k and X_l are continuous. Following the proof of Theorem 2 in the case of $d_*^{(s)} = 0$, we have for the bootstrap critical values $c_{\alpha,\infty}^{*KS}$,

$$c_{\alpha,\infty}^{*KS} = c_\alpha^{KS} + O_P(c_N \sqrt{-\log c_N}).$$

Now, following the proof of Theorem 1, we can show that when $d_*^{(s)} > 0$,

$$\min_{k \neq l} \sup_{x \in \mathcal{X}} \sqrt{N} \bar{D}_{kl}^{(s)}(x, \hat{\theta}, \hat{\tau}) \to \infty.$$

Hence the result follows.

(ii) The proof is very similar to (i) and hence it is omitted. ∎

31

Proof of Theorem 4: (i) For any sequence $\{P_N\}_{N=1}^{\infty} \in \mathcal{P}$, the weak convergence result in the proof of Theorem 1 gives

$$\sqrt{N}\bar{D}_k^{(s)}(x,\hat{\theta},\hat{\tau}) - \frac{\sqrt{N}}{(s-1)!}\mathbf{E}_N h_x^{(s)}(\varphi_k(W;\theta_0,\tau_0))$$

$$= \frac{1}{\sqrt{N}(s-1)!}\sum_{i=1}^{N}\left\{h_x^{(s)}(\varphi_k(W_i;\hat{\theta},\hat{\tau})) - \mathbf{E}_N\left[h_x^{(s)}(\varphi_k(W_i;\hat{\theta},\hat{\tau}))\right]\right\}$$

$$+ \frac{\sqrt{N}}{(s-1)!}\left\{\mathbf{E}_N\left[h_x^{(s)}(\varphi_k(W_i;\hat{\theta},\hat{\tau})) - h_x^{(s)}(\varphi_k(W_i;\theta_0,\tau_0))\right]\right\}$$

$$\implies \nu_k^{(s)}(x),$$

where $\mathbf{E}_N$ denotes the expectation with respect to P_N. This result follows because the functions constituting the process is a uniform Donsker class. Under any sequence P_N such that

$$\mathbf{E}_N h_x^{(s)}(\varphi_k(W;\theta_0,\tau_0))/(s-1)! = H_k^{(s)}(x) + \delta_k(x)/\sqrt{N}$$

as $N \to \infty$, we obtain that $\sqrt{N}\bar{D}_k^{(s)}(x,\hat{\theta},\hat{\tau}) - \sqrt{N}H_k^{(s)}(x) \implies \delta_k(x) + \nu_k^{(s)}(x)$ or that under the local sequence $\{P_N\}$,

$$\sqrt{N}\bar{D}_{kl}^{(s)}(x,\hat{\theta},\hat{\tau}) - \sqrt{N}H_{kl}^{(s)}(x) \Rightarrow \nu_{kl}^{(s)}(x) + \delta_{kl}(x),$$

where $H_{kl}^{(s)}(x) = H_k^{(s)}(x) - H_l^{(s)}(x)$. Hence

$$\min_{k\neq l}\sup_{x\in\mathcal{X}}\left\{\sqrt{N}\bar{D}_{kl}^{(s)}(x,\hat{\theta},\hat{\tau})\right\} = \min_{k\neq l}\sup_{x\in\mathcal{X}}\left\{\sqrt{N}H_{kl}^{(s)}(x) + \nu_{kl}^{(s)}(x) + \delta_{kl}(x)\right\} + o_P(1).$$

$$= \min_{k\neq l}\sup_{x\in\mathcal{X}:H_{kl}^{(s)}(x)=0}\left\{\sqrt{N}H_{kl}^{(s)}(x) + \nu_{kl}^{(s)}(x) + \delta_{kl}(x)\right\} + o_P(1)$$

$$= \min_{k\neq l}\sup_{x\in\mathcal{X}:H_{kl}^{(s)}(x)=0}\left\{\nu_{kl}^{(s)}(x) + \delta_{kl}(x)\right\} + o_P(1).$$

Therefore, the asymptotic unbiasedness result follows from Theorem 2 by applying Anderson's Lemma (e.g. Bickel, Klaassen, Ritov, and Wellner (1993), p.466.)

(ii) The asymptotic unbiasedness follows similarly by Anderson's Lemma. Observe that the set of functions $\{(f)_{k=1}^{K} : \min_{k\neq j}\int(\max\{f_k - f_j, 0\})^2 dx < c\}$ is convex and symmetric around zero. ∎

Proof of Corollary 5 : Since the test statistics are the same, it suffices to compare the bootstrap critical values as $B \to \infty$. By the construction of the local alternatives, we are under the probability on the boundary. Since this bootstrap test statistic is recentered, it converges in distribution (conditional on $\mathcal{G}_N$) to the distribution of $\min_{k\neq l}\sup_{x\in\mathcal{X}}\nu_{kl}^{(s)}(x)$, while the distribution of the bootstrap test statistic $D_{N,b}^{(s)*}$ converges to that of $\min_{k\neq l}\sup_{x\in C_{kl}}\nu_{kl}^{(s)}(x)$. Note that

$$\min_{k\neq l}\sup_{x\in\mathcal{X}}\nu_{kl}^{(s)}(x) \geq \min_{k\neq l}\sup_{x\in C_{kl}}\nu_{kl}^{(s)}(x)$$

because $C_{kl} \subset \mathcal{X}$. Hence $c_{\alpha,\infty}^{*KS-LF} \geq c_{\alpha,\infty}^{*KS}$. This implies that

$$\lim_{N\to\infty}P_N\left\{D_N^{(s)} > c_{\alpha,\infty}^{*KS}\right\} = P\left\{\min_{k\neq l}\sup_{x\in C_{kl}}\left\{\nu_{kl}^{(s)}(x) + \delta_{kl}(x)\right\} > c_{\alpha,\infty}^{*KS}\right\} \tag{30}$$

$$\geq P\left\{\min_{k\neq l}\sup_{x\in C_{kl}}\left\{\nu_{kl}^{(s)}(x) + \delta_{kl}(x)\right\} > c_{\alpha,\infty}^{*KS-LF}\right\}.$$

Now, assume that the set $\mathcal{X}\backslash\cup_{k\neq l}\mathrm{cl}(C_{kl})$ contains a nonempty interior, so that

$$\sup_{x\in\mathcal{X}\backslash\cup_{k\neq l}\mathrm{cl}(C_{kl})}\nu_{kl}^{(s)}(x) > \sup_{x\in\cup_{k\neq l}\mathrm{cl}(C_{kl})}\nu_{kl}^{(s)}(x) \text{ for all } k\neq l$$

with positive probability because $\nu_{kl}^{(s)}(x)$ is a Gaussian process. Therefore, $c_{\alpha,\infty}^{*KS-LF} > c_{\alpha,\infty}^{*KS}$. In this case, by Gaussianity of the process $\nu_{kl}^{(s)}(x) + \delta_{kl}(x)$, the inequality in (30) is strict. ∎

7.2   Bootstrap Uniform Central Limit Theorem

In this section, we show how we can extend the bootstrap CLT result of Giné (1997) to accommodate the situation where bootstrap samples are non-identically distributed. The essential nature of the residual based bootstrap is to resample only from a marginal distribution of a random vector while leaving the remaining part intact. For example, consider the nonlinear regression model $Y_i = g(X_i,\theta)+\varepsilon_i$. Then, bootstrap sample is obtained as $\{Y_i^*, X_i\}_{i=1}^n$ or as $\{g(X_i,\hat{\theta})+[Y_i - g(X_i,\hat{\theta})]^*, X_i\}_{i=1}^n$ where $[\cdot]^*$ indicates the variable obtained from resampling. The bootstrap central limit theorem in this section is appropriate for this situation.

We adapt the notations and definitions of Giné and Zinn (1991) and Giné (1997) to our case of residual based bootstrap processes. Let $(S,\mathcal{S},P)$ be a probability space on a Banach space S, and let (X_i, Z_i) : $(S^{\mathbb{N}},\mathcal{S}^{\mathbb{N}},P^{\mathbb{N}}) \to (S,\mathcal{S},P)$ be the i-th coordinate function. Hence $\{(X_i,Z_i)\}_{i\in\mathbb{N}}$ is a sequence of pairs of B-valued (i.e. taking values from a Banach space) random elements taking values from $\mathcal{X}\times\mathcal{Z}$. Let $\mathcal{F}$ be a class of uniformly bounded measurable functions on $(S,\mathcal{S})$ with an envelope F, and let $\mathcal{F}' = \{f(\cdot,z) : (f,z) \in \mathcal{F}\times\mathcal{Z}\}$. Then, we define

$$P_{n,X}(\omega) \triangleq \frac{1}{n}\sum_{i=1}^n \delta_{X_i}(\omega) \text{ and } P_n(\omega) \triangleq \frac{1}{n}\sum_{i=1}^n \left\{\delta_{(\hat{X}_i,Z_i)}(\omega) - \frac{1}{n}\sum_{j=1}^n \delta_{(X_j,Z_i)}(\omega)\right\}$$

where $\hat{X}_1,\ldots,\hat{X}_n$ i.i.d. $P_{n,X}(\omega)$. The focus of this paper is the distribution of $(\hat{X}_i, Z_i)_{i=1}^n$. The main difference from the usual nonparametric bootstrap is that the conditional distribution of $(\hat{X}_i, Z_i)$ given $\{X_i, Z_i\}_{i\in\mathbb{N}}$ is not identical across different i's.

The measure $P_n(\omega)$ is a signed measure, and we write $\mathbf{E}_{P_n}f(X,Z) = \frac{1}{n}\sum_{i=1}^n\{f(\hat{X}_i,Z_i)-\frac{1}{n}\sum_{j=1}^n f(X_j,Z_i)\}$. Let G_P be a P-Brownian bridge, i.e. the centered Gaussian process on $L_2(P)$ with the covariance kernel given by $Cov_P(f,g)$. Let $\nu_n^* \triangleq \sqrt{n}P_n$ and define

$$d_{P_n}(\nu_n^*, G_P; \mathcal{F}) \triangleq \sup_{H\in BL(\mathcal{F})}|\mathbf{E}_{P_n}H(\nu_n^*) - \mathbf{E}_{P_n}H(G_P)|,$$

where $BL(\mathcal{F}) = \{H : l_\infty(\mathcal{F}) \to \mathbb{R}: ||H(v_1) - H(v_2)|| \leq \sup_{f\in\mathcal{F}}||v_1(f) - v_2(f)||$ and $\sup_{v\in l_\infty(\mathcal{F})}||H(\nu)|| < \infty\}$. Conditional on $\{X_i, Z_i\}_{i\in\mathbb{N}}$, the process ν_n^* is an empirical process of independent but nonidentically distributed random elements. Similarly we define $\nu_n \triangleq \sqrt{n}(P_{n,X} - P_X)$ and $d_P(\nu_n, G_P; \mathcal{F}')$. Then, we say $\mathcal{F} \in$ BUCLT($\mathcal{P}$) if $\mathcal{F}$ is P-pregaussian and $\sup_{P\in\mathcal{P}} d_{BL,P_n}(\nu_n^*, G_P; \mathcal{F}) \to 0$ in P uniformly in $P \in \mathcal{P}$. (For details, see Giné and Zinn (1991)). The primary focus is to identify sufficient conditions for $\mathcal{F} \in$ BUCLT($\mathcal{P}$). We follow the proof for the direct implication part of Theorem 2.2. of Giné (1997). For this we need the following lemma which is a slight variant of Proposition 2.5 of Giné (1997).

Lemma B1: *Let B be a Banach space, and for any $n \in \mathbb{N}$, let $w_{ji} = (x_j, z_i)$, where $x_j, z_i,\ i,j \in \mathbb{N}$, are points in B and let $w_i \triangleq \frac{1}{n}\Sigma_{j=1}^n(x_j, z_i)$. Let $W_1,\ldots,W_n$ be independent, non-identically distributed B-valued random variables with each W_i having law $P\{W_i = w_{ji}\} = 1/n,\ j = 1,\ldots,n$. Let $N_{ji},\ i,j \in \{1,\ldots,n\}$,*

be i.i.d. Poisson random variables with parameter 1. Assume these random vectors and variables are all independent. Then,

$$\mathbf{E} \left\| \sum_{i=1}^{n} (W_i - w_i) \right\| \leq \frac{e}{n(e-1)} \mathbf{E} \left\| \sum_{j=1}^{n} \sum_{i=1}^{n} (N_{ji} - 1)(w_{ji} - w_i) \right\|$$

Proof of Lemma B1: For any probability measure $\mathcal{L}$, define $\mathrm{Pois}(\mathcal{L}) = \exp(\mathcal{L} - 1)$. Let $\mathcal{L}(X)$ denote the distribution of X. By the Poissonization lemma (e.g. (2.5') of Giné (1997)), we obtain that

$$\mathbf{E} \left\| \sum_{i=1}^{n} (W_i - w_i) \right\| \leq \frac{e}{e-1} \int ||x|| d\mathrm{Pois} \left(\sum_{i=1}^{n} \mathcal{L}(W_i^b - w_i) \right). \tag{31}$$

Note that $\sum_{i=1}^{n} \mathcal{L}(W_i - w_i) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{w_{ji} - w_i}$ and that

$$
\begin{aligned}
\mathrm{Pois} \left(\sum_{j=1}^{n} \delta_{w_{ji} - w_i} \right) &= \mathrm{Pois}(N_{1i}(w_{1i} - w_i)) * \cdots * \mathrm{Pois}(N_{ni}(w_{ni} - w_i)) \\
&= \mathcal{L} \left(\sum_{j=1}^{n} N_{ji}(w_{ji} - w_i) \right) = \mathcal{L} \left(\sum_{j=1}^{n} (N_{ji} - 1)(w_{ji} - w_i) \right).
\end{aligned}
$$

Therefore, $\mathrm{Pois} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{w_{ji} - w_i} \right) = \mathcal{L} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} (N_{ji} - 1)(w_{ji} - w_i) \right)$. By (2.7) of Giné (1997) and by (31), we obtain the wanted result. ∎

Proposition B1: *Suppose that* $\sup_{P \in \mathcal{P}} \int_0^\infty \sqrt{\log N_{[]}(\varepsilon, \mathcal{F}, || \cdot ||_{P,2})} d\varepsilon < \infty$ *and* $\sup_{P \in \mathcal{P}} \int F^2 dP < \infty$. *Then* $\mathcal{F} \in \mathrm{BUCLT}(\mathcal{P})$.

Proof of Proposition B1: We follow the proof Theorem 2.2 of Giné (1997). A close examination of the proof shows that the only modification needed for this extension is bounding the oscillation of the bootstrap empirical process. Let $\mathbf{E}^*$ denote the expectation with respect to the bootstrap distribution. Similarly as in the proof there, we proceed as follows:

$$
\begin{aligned}
\mathbf{EE}^* ||\nu_n^*||_{\mathcal{F}_{\delta'}} &= \mathbf{EE}^* \left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\delta_{(\hat{X}_i, Z_i)} - P_{n,i} \right) \right\|_{\mathcal{F}_{\delta'}} \tag{32} \\
&\leq \frac{e}{e-1} \mathbf{E} \left\| \frac{1}{n\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{n} (N_{ji} - 1) \left\{ \delta_{(X_j, Z_i)} - \frac{1}{n} \sum_{k=1}^{n} \delta_{(X_k, Z_i)} \right\} \right\|_{\mathcal{F}_{\delta'}} \\
&\leq \frac{e}{e-1} \left\{ \mathbf{E} \left\| \frac{1}{n\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{n} (N_{ji} - 1)\delta_{(X_j, Z_i)} \right\|_{\mathcal{F}_{\delta'}} + \mathbf{E} \left\| \frac{1}{n^2\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{n} (N_{ji} - 1) \sum_{k=1}^{n} \delta_{(X_k, Z_i)} \right\|_{\mathcal{F}_{\delta'}} \right\}.
\end{aligned}
$$

We can show that the $\lim_{\delta \to 0} \limsup_{n \to \infty}$ of both the terms are equal to zero for each $P \in \mathcal{P}$ by proceeding exactly in the same manner as in the proof of Theorem 4.10 of Arcones and Giné (1993). Uniformity over $P \in \mathcal{P}$ can be ensured by applying to the chaining argument a uniform version of Bernstein's inequality (Proposition 2.3 of Arcones and Giné (1993)) that involves $\sup_{P \in \mathcal{P}} \int F_k^2 dP$, where F_k is the class of functions that are bracketed by the k-th bracket. The number of the brackets can be chosen not to depend on P due

to the condition that $\sup_{P \in \mathcal{P}} \int_0^\infty \sqrt{\log N_{[]}(\varepsilon, \mathcal{F}, ||\cdot||_{P,2})} d\varepsilon < \infty$. Details are omitted. ∎

References

[1] Abadie, A. (2002), "Bootstrap tests of distributional treatment effects in instrumental variables," *The Journal of the American Statistical Association*, 97, 284-292.

[2] Anderson, G. J. (1996), "Nonparametric tests of stochastic dominance in income distributions," *Econometrica*, 1183-1193.

[3] Andrews, D. W. K. (1994), "Empirical process methods in econometrics," in *The Handbook of Econometrics*, Vol. IV ed. by R. F. Engle and D. L. McFadden, Amsterdam: North-Holland.

[4] Andrews, D. W. K. (1997), "A conditional Kolmogorov test," *Econometrica,* 65, 1097-1128.

[5] Andrews, D. W. K. and P. Guggenberger (2007), "The limit of exact size and a problem with subsampling and with the m out of n bootstrap," Working paper.

[6] Arcones, M. A. and E. Giné (1992), "On the bootstrap of M-estimators and other statistical functionals," in *Exploring the Limits of Bootstrap*, ed. by R. LePage and L. Billard, Wiley, 13-47.

[7] Arcones, M. A. and E. Giné (1993), "Limit theorems for U-processes," *Annals of Statistics*, 21, 1494-1542.

[8] Barrett G. and S. Donald (2003), "Consistent tests for stochastic dominance," *Econometrica*, 71,71-104.

[9] Bickel, P. J. , A.J. Klaassen, Y. Ritov, and J. A. Wellner (1993): *Efficient and Adaptive Estimation for Semiparametric Models*, Springer Verlag, New York.

[10] Birman, M. S. and M. Z. Solomjak (1967), "Piecewise-polynomial approximation of functions of the classes W_p^α," *Mathematics of the USSR-Sbornik* 2, 295-317.

[11] Chen, X., O. Linton, and I. van Keilegom (2003), "Estimation of semiparametric models when the criterion function is not smooth," *Econometrica* 71, 1591-1608.

[12] Chernozhukov, V., H. Hong, and E. Tamer (2007) Estimation and confidence regions for parameter sets in econometric models. *Econometrica*

[13] Davidson, R. and J. Y. Duclos (2000), "Statistical inference for stochastic dominance and measurement for the poverty and inequality," *Econometrica*, 68, 1435-1464.

[14] Giné, E. (1997), *Lecture Notes on Some Aspects of the Bootstrap, Ecole de Éte de Calcul de Probabilités de Saint-Flour. Lecture Notes in Mathematics*, 1665, Springer, Berlin.

[15] Giné, E. and J. Zinn (1991), "Gaussian characterization of uniform Donsker classes of functions," *Annals of Probability*, 19, 758-782.

[16] Hansen, P. (2003), "Asymptotic tests of composite hypotheses," Working Paper.

[17] Härdle, W. and E. Mammen (1993), "Comparing nonparametric versus parametric regression fits," *Annals of Statistics*, 21, 1926-1947.

[18] Horváth, L., P. Kokoszka, and R. Zikitis (2006), "Testing for stochastic dominance using the weighted McFadden-type statistic," *Journal of Econometrics*, 133, 191-205.

[19] Kim, J. and D. Pollard (1990), "Cube-root asymptotics," *Annals of Statistics*, 18, 191-219.

[20] Koul, H. L. and S. N. Lahiri (1994), "On bootstrapping M-estimated residual processes in multiple linear regression models," *Journal of Multivariate Analysis*, 49, 255-265.

[21] Linton, O., E. Maasoumi, and Y-J. Whang (2005), "Consistent testing for stochastic dominance under general sampling schemes," *Review of Economic Studies* 72, 735-765.

[22] McFadden, D. (1989), "Testing for stochastic dominance," in *Studies in the Economics of Uncertainty: In Honor of Josef Hadar*, ed. by T. B. Fomby and T. K. Seo. New York, Berlin, London and Tokyo: Springer.

[23] Sheehy, A. and J. A. Wellner (1992), "Uniform Donsker classes of functions," *Annals of Probability*, 20, 1983-2030.

[24] Song, K. (2008), "Testing distributional inequalities and asymptotic bias," Working Paper.

[25] Tse, Y. K. and X. B. Zhang (2004), "A Monte Carlo investigation of some tests for stochastic dominance," *Journal of Statistical Computation and Simulation*, 74, 361-378.

[26] van der Vaart, A. W. and J. A. Wellner (1996), *Weak Convergence and Empirical Processes*, Springer-Verlag.

[27] Wellner, J. A. and Y. Zhan (1996), "Bootstrapping Z-estimators," Working Paper.

[28] Whang, Y. J. (2000), "Consistent bootstrap tests of parametric regression functions," *Journal of Econometrics*, 98, 27-46.

Design	N	KS					CM				
					LSW					LSW	
		SUB	BTS	I	II	III	SUB	BTS	I	II	III
1a, $d_1^* = 0$	50	.137	.063	.070	.063	.063	.105	.064	.084	.064	.064
	500	.046	.056	.057	.056	.056	.050	.057	.058	.057	.057
	1000	.052	.049	.049	.049	.049	.046	.060	.061	.060	.060
1b, $d_1^* = 0$	50	.122	.055	.061	.055	.055	.082	.071	.080	.072	.071
	500	.052	.051	.051	.051	.051	.058	.059	.061	.059	.059
	1000	.048	.059	.060	.051	.051	.056	.046	.048	.046	.046
1c, $d_1^* > 0$	50	.399	.685	.762	.707	.689	.279	.411	.581	.457	.424
	500	.981	.983	.983	.983	.983	.975	.981	.983	.982	.982
	1000	.994	.995	.995	.995	.995	.993	.994	.994	.994	.994
1d, $d_1^* > 0$	50	.381	.684	.781	.706	.689	.262	.426	.578	.464	.437
	500	.986	.984	.987	.985	.985	.985	.982	.983	.982	.982
	1000	.995	.994	.994	.994	.994	.995	.993	.993	.993	.993
1e, $d_1^* > 0$	50	.397	.646	.738	.670	.653	.265	.371	.528	.408	.387
	500	.986	.992	.992	.992	.992	.975	.991	.991	.991	.991
	1000	.987	.992	.992	.992	.992	.987	.989	.991	.991	.991

Table 1F. Rejection frequencies for the test of First Order Stochastic Dominance for Design 1. SUB refers to the subsampling method with critical values computed by the automatic method "Mean" described by LMW(2005) for the 5% null rejection probabilities. BT refers to the recentered bootstrap. LSW refers to the recentered bootstrap with set estimation, where the tuning parameter is given by $c_N = c \cdot N^{-1/3}$ with $c = 0.25, 0.50$, and 0.75 for cases I, II, and III, respectively.

		KS						CM				
					LSW						LSW		
Design	N	SUB	BTS	I	II	III	SUB	BTS	I	II	III
1a, $d_2^* = 0$	50	.110	.066	.068	.066	.066	.090	.062	.063	.062	.062
	500	.055	.055	.055	.055	.055	.049	.057	.057	.057	.057
	1000	.052	.050	.050	.050	.050	.048	.050	.050	.050	.050
1b, $d_2^* = 0$	50	.084	.061	.132	.075	.066	.064	.066	.154	.080	.071
	500	.068	.060	.178	.091	.071	.063	.060	.210	.096	.070
	1000	.073	.050	.177	.084	.058	.060	.051	.219	.086	.063
1c, $d_2^* > 0$	50	.261	.336	.846	.525	.368	.159	.200	.850	.380	.212
	500	.942	.451	.995	.985	.983	.738	.003	.996	.986	.983
	1000	.993	.545	.997	.996	.995	.969	.003	.997	.996	.995
1d, $d_2^* > 0$	50	.233	.329	.841	.544	.368	.149	.197	.851	.370	.209
	500	.958	.423	.994	.987	.985	.710	.004	.994	.987	.984
	1000	.995	.524	.999	.996	.994	.955	.003	.999	.998	.994
1e, $d_2^* > 0$	50	.233	.299	.798	.478	.326	.138	.179	.810	.346	.189
	500	.933	.424	.995	.993	.992	.653	.009	.996	.993	.991
	1000	.984	.484	.997	.995	.992	.934	.001	.998	.997	.992

Table 1S. Rejection frequencies for the test of Second Order Stochastic Dominance for Design 1. SUB refers to the subsampling method with critical values computed by the automatic method "Mean" described by LMW(2005) for the 5% null rejection probabilities. BT refers to the recentered bootstrap. LSW refers to the recentered bootstrap with set estimation, where the tuning parameter is given by $c_N = c \cdot N^{-1/3}$ with $c = 2.0, 3.0,$ and 4.0 for cases I, II, and III, respectively.

Design	N	KS					CM				
					LSW					LSW	
		SUB	BTS	I	II	III	SUB	BTS	I	II	III
	50	.130	.054	.058	.054	.054	.082	.058	.071	.059	.058
2a, $d_1^* = 0$	500	.049	.055	.055	.055	.055	.054	.057	.057	.057	.057
	1000	.070	.044	.044	.044	.044	.050	.061	.062	.061	.061
	50	.078	.072	.088	.075	.073	.031	.056	.068	.056	.056
2b, $d_1^* = 0$	500	.010	.026	.104	.046	.028	.002	.004	.028	.005	.004
	1000	.020	.018	.170	.076	.034	.007	.001	.031	.009	.002
	50	.300	.453	.787	.626	.574	.272	.718	.918	.872	.831
2c, $d_1^* > 0$	500	.979	1.00	1.00	1.00	1.00	.988	1.00	1.00	1.00	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	50	.257	.173	.555	.427	.377	.184	.016	.396	.223	.159
2d, $d_1^* > 0$	500	.969	.988	1.00	1.00	1.00	.954	.644	1.00	1.00	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	.995	1.00	1.00	1.00

Table 2F. Rejection frequencies for the test of First Order Stochastic Dominance for Design 2. SUB refers to the subsampling method with critical values computed by the automatic method "Mean" described by LMW(2005) for the 5% null rejection probabilities. BT refers to the recentered bootstrap. LSW refers to the recentered bootstrap with set estimation, where the tuning parameter is given by $c_N = c \cdot N^{-1/3}$ with $c = 0.25, 0.50,$ and 0.75 for cases I, II, and III, respectively.

Design	N	KS		LSW			CM		LSW		
		SUB	BTS	I	II	III	SUB	BTS	I	II	III
2a, $d_2^* = 0$	50	.077	.056	.063	.056	.056	.068	.061	.074	.061	.061
	500	.061	.046	.050	.046	.046	.054	.058	.066	058.	.058
	1000	.063	.065	.065	.065	.065	.057	.064	.064	.064	.064
2b, $d_2^* = 0$	50	.063	.078	.119	.078	.078	.037	.077	.135	.077	.077
	500	.002	.006	.154	.018	.006	.001	.002	.171	.006	.002
	1000	.004	.001	.215	.023	.001	.001	.000	.239	.014	.000
2c, $d_2^* = 0$	50	.032	.006	.007	.006	.006	.001	.006	.006	.006	.006
	500	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
2d, $d_2^* > 0$	50	.177	.027	.823	.499	.069	.145	.001	.851	.407	.006
	500	.936	.332	1.00	1.00	1.00	.857	.000	1.00	1.00	1.00
	1000	1.00	.860	1.00	1.00	1.00	.991	.000	1.00	1.00	1.00

Table 2S. Rejection frequencies for the test of Second Order Stochastic Dominance for Design 2. SUB refers to the subsampling method with critical values computed by the automatic method "Mean" described by LMW(2005) for the 5% null rejection probabilities. BT refers to the recentered bootstrap. LSW refers to the recentered bootstrap with set estimation, where the tuning parameter is given by $c_N = c \cdot N^{-1/3}$ with $c = 2.0, 3.0,$ and 4.0 for cases I, II, and III, respectively.

Design	N	KS					CM				
					LSW					LSW	
		SUB	BTS	I	II	III	SUB	BTS	I	II	III
	50	.637	.959	.993	.986	.977	.612	.972	.997	.992	.988
3a, $d_1^* > 0$	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	50	.060	.025	.032	.026	.025	.037	.026	.040	.028	.026
3b, $d_1^* = 0$	500	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	50	.611	.949	.993	.989	.985	.575	.970	.996	.992	.990
3c, $d_1^* > 0$	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3F. Rejection frequencies for the test of First Order Stochastic Dominance for Design 3. SUB refers to the subsampling method with critical values computed by the automatic method "Mean" described by LMW(2005) for the 5% null rejection probabilities. BT refers to the recentered bootstrap. LSW refers to the recentered bootstrap with set estimation, where the tuning parameter is given by $c_N = c \cdot N^{-1/3}$ with $c = 0.25, 0.50$, and 0.75 for cases I, II, and III, respectively.

| Design | N | KS | | | | | CM | | | | |
| | | SUB | BTS | I | II | III | SUB | BTS | I | II | III |
					LSW					LSW	
3a, $d_2^* = 0$	50	.031	.021	.025	.022	.021	.000	.016	.022	.016	.016
	500	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
3b, $d_2^* = 0$	50	.048	.044	.123	.053	.046	.039	.050	.149	.065	.051
	500	.000	.000	.020	.005	.002	.000	.000	.030	.006	.002
	1000	.000	.000	.011	.001	.000	.000	.000	.018	.002	.000
3c, $d_2^* > 0$	50	.546	.934	.942	.940	.938	.458	.920	.939	.926	.924
	500	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	1000	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3S. Table 1F. Rejection frequencies for the test of Second Order Stochastic Dominance for Design 3. SUB refers to the subsampling method with critical values computed by the automatic method "Mean" described by LMW(2005) for the 5% null rejection probabilities. BT refers to the recentered bootstrap. LSW refers to the recentered bootstrap with set estimation, where the tuning parameter is given by $c_N = c \cdot N^{-1/3}$ with $c = 2.0, 3.0$, and 4.0 for cases I, II, and III, respectively.