Testing for Almost Stochastic Dominance

Wonwoo Bae* Yoon-Jae Whang†

November 30, 2022

Click here for the latest version.

Abstract

We propose a nonparametric test for the null hypothesis of almost stochastic dominance (ASD). The traditional stochastic dominance (SD) rule ranks distributions for *all* utility functions in a certain class, which can be restrictive in practice. To circumvent the limitation of the SD rule, Leshno and Levy (2002) developed the ASD rule that applies to *most* rather than *all* decision makers by eliminating economically pathological preferences. The ASD rule can be applied to many empirical economic problems including investment decisions and policy evaluations. Despite its usefulness, to the best of our knowledge, there has been no formal test of ASD available in the literature. In this paper, we propose an L_p-type test statistic based on empirical distribution functions and introduce bootstrap procedures to compute the critical values. We investigate the finite sample performance of the testing procedures by a set of Monte Carlo simulation experiments. We apply our test to compare the return distributions of stocks and bonds over different investment horizons. The ASD tests support the popular practice of adjusting the portfolios of stocks and bonds based on the investment horizons.

JEL Classifications: C12, C14, C15

Keywords: Almost Stochastic Dominance, Test Consistency, Bootstrap, Investment Horizon

*Department of Economics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea. E-mail: `bww1016@snu.ac.kr`.

†Department of Economics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea. E-mail: `whang@snu.ac.kr`.

1 Introduction

Many economic problems involve comparison between different prospects. Among other methods (e.g., mean-variance analysis in finance), there has been a considerable interest in ranking stochastic objects in terms of their distributions. The stochastic dominance (SD) rule of various orders provides a uniform weak ordering of distributions of many economic outcomes such as investment strategies and welfare outcomes. The ordering by SD covers a large class of utility functions. For example, the first order SD rule ranks decisions made by agents with strictly increasing utility functions and the second order SD rule corresponds to decisions of agents with strictly increasing and strictly concave utility functions.

The *uniform* ordering property over a large utility class is one of the main advantages of using the SD rule. However, this advantage cannot be fully exploited because the SD rule can be too strong to attain in practice. A small violation of the SD rule makes the ordering invalid, which may happen when there exists a crossing of distributions, or equivalently, when there exists an agent with "extreme" preference. For example, consider two prospects X_1 and X_2, where X_1 provides either \$2 or \$3 with equal probability and X_2 provides \$1 or $\$1,000,000$ with equal probability. It would be reasonable to think that "most" investors would prefer X_2 to X_1. However, neither X_1 nor X_2 stochastically dominates the other because their distributions cross each other.

To circumvent the limitation of the SD rule by eliminating such extreme preferences, Leshno and Levy (2002) introduced the almost stochastic dominance (ASD) rule which applies to *most* rather than *all* decision makers. Tzeng, Huang, and Shih (2013) showed that the original ASD does not satisfy the uniform ordering property and modified the definition so that it can be satisfied. Specifically, the class of utility functions over which the ASD rule applies now excludes "economically pathological" preferences, meaning mathematically valid but economically irrelevant.[1] The ASD rule allows for limited violations of the SD rule by shrinking the choice set to a set of utility functions with bounded derivatives. In other words, crossings between distributions, which are commonly found in empirical examples, are allowed. This allows expanding the set of rankable prospects, compared to the case under the traditional SD rule, with uniformity property maintained for a reasonably large class of utility functions.

There has been many applications of the ASD rule to financial decision making problems. For example, Bali, Demirtas, Levy, and Wolf (2009) use the concept of ASD to find evidence in favor of the popular practice of primarily allocating a greater proportion to stocks and then gradually relocating funds to bonds as the investment horizon shortens. The arugment for stocks for the long run is further investigated in Levy (2009).

[1] Leshno and Levy (2002) suggests one typical example of pathological preferences: a myopic utility function $u(x) = \frac{x^\alpha}{\alpha}$, with $0 < \alpha < 1$.

Levy (2012) develops algorithms for derivation of the ASD efficient investment sets and shows its efficacy. Do (2021) shows outperformance of socially responsible investing (SRI) portfolio over the market indexes using the ASD rule. See Levy (2016), for additional applications. However, empirical analyses using ASD have been mainly based on numerical computation of the violation ratio without considering sampling errors.

The purpose of this paper is to develop a nonparametric test of the almost stochastic dominance hypothesis. To the best of our knowledge, no formal statistical inference method for ASD has been available in the literature.[2] This paper considers an L_p- or Supremum- type test statistic based on empirical distribution functions and suggests bootstrap methods under various sampling schemes to compute critical values and proves their asymptotic validity. Using the proposed test, this paper empirically evaluate the common practice of adjusting portfolios of stocks and bonds.

Our paper contributes to the large literature on testing the SD hypotheses. McFadden (1989), Klecan, McFadden, and McFadden (1991), Kaur, Rao, and Singh (1994), Anderson (1996), and Davidson and Duclos (2000) propose different approaches to testing SD. Barrett and Donald (2003) suggest a consistent bootstrap method to test SD of any order between two prospects under an independent sampling scheme. Linton, Maasoumi, and Whang (2005) develop a consistent subsampling test for SD under general sampling schemes allowing for time series dependence. Linton, Song, and Whang (2010) propose a bootstrap-based test with improved power performance by utilizing the information on the binding part of the inequality restrictions. Our paper is also related to tests of various weaker notions of the SD relation. Examples include Álvarez-Esteban, del Barrio, Cuesta-Albertos, and Matrán (2016) who propose testing the approximate SD relationship based on mixture (or contaminated) models and Knight and Satchell (2008) who proposed a test of the infinite order SD hypothesis. However, none of these papers formulate the inference method for the ASD hypothesis, although the ASD rule is the dominant notion in empirical research. See Whang (2019) for a comprehensive survey.

We investigate the finite sample performance of our testing procedure using Monte Carlo simulation experiments. We find that the size and power properties are generally acceptable even for the modest sample sizes. We also evaluate the finite sample property of the method proposed by Guo, Levy, and Wong (2015) in the case of first order ASD.

As an empirical example, we apply our test to evaluate the common practice of adjusting the portfolio with the stocks to bonds ratio dependent on the investment horizon. We first verify that the standard SD test does not support investors' preferences for one prospect over the other at most of the investment horizons considered. However, our test finds empirical support for investors' increasing preferences for stocks as the investment

[2] Guo, Levy, and Wong (2015) proposed an ASD test, but their hypotheses of interest include only one of the inequality restrictions that comprise the definition of higher order ASD and they do not provide a formal proof for the validity of their testing procedure. In the later sections, we provide theoretical and numerical comparisons of our ASD test with their test.

horizon extends. Our test is especially suitable for financial decision making problems because they often require rationality which is relevant both economically and mathematically.

The rest of this paper is organized as follows: In Section 2, we give definitions of ASD and introduce the hypotheses of interest. In Section 3, we define our test statistics and develop their asymptotic properties. In Section 4, we introduce our bootstrap inference method and prove its asymptotic validity. In Section 5, we investigate the local power property of our test. In Section 6, we conduct Monte Carlo experiments for evaluating the finite sample performance. In Section 7, we provide an empirical analysis of the popular investment practice using our test. In Section 8, we make concluding remarks.

2 Almost Stochastic Dominance and the Hypotheses of Interest

2.1 Almost Stochastic Dominance

We define almost stochastic dominance based on the modified version of ASD of Tzeng, Huang, and Shih (2013) which corrected an error in the original definition of Leshno and Levy (2002).

Let X_1 and X_2 be two prospects supported on $\mathcal{X} = [\underline{x}, \overline{x}]$, $-\infty < \underline{x} < \overline{x} < +\infty$ with distributions F_1 and F_2, respectively. For $k = 1, 2$, define the intergrated distribution functions $F_k^{(m)}(x) = \int_{\underline{x}}^{x} F_k^{(m-1)}(z)dz$ for $m \geq 2$ with the convention $F_k^{(1)}(x) = F_k(x)$. Let $[\cdot]_+ = \max\{\cdot, 0\}$ and $[\cdot]_- = \min\{\cdot, 0\}$.

Define the nested classes of utility functions $\mathcal{U}_1 = \{u : u^{(1)} \geq 0\}$ and $\mathcal{U}_2 = \{u \in \mathcal{U}_1 : u^{(2)} \leq 0\}$, where $u^{(s)}$, $s \in \mathbb{Z}^+$, denote the s-th order derivative of u. The higher-order utility function classes are defined recursively as $\mathcal{U}_m = \{u \in \mathcal{U}_{m-1} : (-1)^m u^{(m)} \leq 0\}$ for $m \geq 2$. For $\epsilon \in (0, \frac{1}{2})$ and $m \geq 1$, let

$$\mathcal{U}_m(\epsilon) = \left\{ u \in \mathcal{U}_m : (-1)^{m+1} u^{(m)}(x) \leq \inf_{x \in \mathcal{X}} \left\{(-1)^{m+1} u^{(m)}(x)\right\} \left[\frac{1}{\epsilon} - 1\right], \forall x \in \mathcal{X} \right\}$$

be the set of utility functions with the additional restrictions on the ratio between the maximum and minimum values of $u^{(m)}(x)$ so that large changes in $u^{(m)}(x)$ with respect to x are excluded. Note that $\mathcal{U}_1(\epsilon)$ and $\mathcal{U}_2(\epsilon)$ exclude from $\mathcal{U}_1$ and $\mathcal{U}_2$ utility functions such as $u(x) = x \cdot 1(x \leq \frac{1}{2}) + \frac{1}{2} \cdot 1(x > \frac{1}{2})$ and $u(x) = \log(x)$ assigning relatively low marginal utility to large values of x and high marginal utility to very low values of x.

The first order ASD is defined as follows:

Definition 1. X_1 *ϵ-almost first order stochastic dominates* X_2, denoted as $X_1 \succeq_{A1S(\epsilon)} X_2$ for $0 < \epsilon < \frac{1}{2}$, if and only if,

(a) $\mathbf{E}_{F_1} u(X_1) \geq \mathbf{E}_{F_2} u(X_2)$, $\forall u \in \mathcal{U}_1(\epsilon)$, or

(b) $\int_{\mathcal{X}}[F_1(x) - F_2(x)]_+ dx \leq \epsilon \int_{\mathcal{X}} |F_1(x) - F_2(x)| dx$.

The definition can be extended to the higher-order ($m \geq 2$) ASD:

Definition 2. X_1 *ϵ-almost m-th order stochastic dominates* X_2, denoted as $X_1 \succeq_{AmS(\epsilon)} X_2$ for $0 < \epsilon < \frac{1}{2}$, if and only if,

(a) $\mathbf{E}_{F_1} u(X_1) \geq \mathbf{E}_{F_2} u(X_2)$, $\forall u \in \mathcal{U}_m(\epsilon)$, or

(b) $\int_{\mathcal{X}}[F_1^{(m)}(x) - F_2^{(m)}(x)]_+ dx \leq \epsilon \int_{\mathcal{X}} |F_1^{(m)}(x) - F_2^{(m)}(x)| dx$ and $F_1^{(j)}(\overline{x}) \leq F_2^{(j)}(\overline{x})$ for $j = 2, \ldots, m$.

For the proof of the equivalence of the definitions (a) and (b) in Definitions 1 and 2, see Leshno and Levy (2002, Theorem 1) Tzeng, Huang, and Shih (2013, Theorem 1 and 2). Definition 1 (b) and the first inequality of Definition 2 (b) controls the deviation from the SD relation by a prespecified constant ϵ. Let

$$\theta_m = \frac{\int_{\mathcal{X}}[F_1^{(m)}(x) - F_2^{(m)}(x)]_+ dx}{\int_{\mathcal{X}} |F_1^{(m)}(x) - F_2^{(m)}(x)| dx}.$$

This quantity takes values between 0 and 1 and can serve as a measure of deviation from the m-th order SD. That is, given $\epsilon \in (0, \frac{1}{2})$, the m-th order ASD requires $\theta_m \leq \epsilon$. Note that this condition is a necessary condition for m-th order ASD ($m \geq 2$).[3] Note that the quantity coincides with the Utopia index of Anderson, Post, and Whang (2020) when there are two prospects.

2.2 The Hypotheses of Interest

The null hypothesis of m-th order almost stochastic dominance is given by

$$\begin{gathered} H_0^{(m)} : \int_{\mathcal{X}}[F_1^{(m)}(x) - F_2^{(m)}(x)]_+ dx \leq \epsilon \int_{\mathcal{X}} |F_1^{(m)}(x) - F_2^{(m)}(x)| dx \\ \text{and} \\ F_1^{(j)}(\overline{x}) \leq F_2^{(j)}(\overline{x}) \text{ for } 2 \leq j \leq m. \end{gathered} \tag{2.1}$$

which is equivalent to the uniform ordering of two prospects for individuals with $u \in \mathcal{U}_m(\epsilon)$. The alternative hypothesis $H_1^{(m)}$ is the negation of $H_0^{(m)}$, that is, there exists at least one person with $u \in \mathcal{U}_m(\epsilon)$ who ranks the prospects differently. For example, $H_0^{(2)}$ implies that most risk averse individuals whose utility function belongs to $\mathcal{U}_2(\epsilon)$ would prefer prospect X_1 to prospect X_2.

For $m \geq 2$, the null hypothesis consists of 1 inequality which concerns the deviation from the m-th order SD and $m - 1$ inequalities which act as boundary conditions. To

[3] The testing procedure proposed by Guo, Levy, and Wang (2015) is incomplete for $m \geq 2$ since their hypotheses of interest only concern this necesssary condition.

test these inequalities jointly, it is convenient to define the population quantity as a nonnegative and increasing function of each population quantity. Let $\Lambda_p : \mathbb{R}^m \to [0, \infty$ is a nonnegative and increasing function for $p \in \{1, 2\}$. We specifically focus on the following map:

$$\Lambda_p(d_{m,1}, \ldots, d_{m,m}) = \left(\max\left\{[d_{m,1}]_+, \ldots, [d_{m,m}]_+\right\}\right)^p, \tag{2.2}$$

or,

$$\Lambda_p(d_{m,1}, \ldots, d_{m,m}) = \sum_{j=1}^{m} [d_{m,j}]_+^p. \tag{2.3}$$

Then, we define the population quantity as a maximum or sum of each quantity for $p \in \{1, 2\}$:

$$d_m^* = \Lambda_p(d_{m,1}, \ldots, d_{m,m}), \tag{2.4}$$

where

$$\begin{aligned} d_{m,1} &= \int_{\mathcal{X}} \left\{[F_1^{(m)}(x) - F_2^{(m)}(x)]_+ - \epsilon|F_1^{(m)}(x) - F_2^{(m)}(x)|\right\} dx \\ d_{m,j} &= F_1^{(j)}(\overline{x}) - F_2^{(j)}(\overline{x}) \text{ for } 2 \leq j \leq m. \end{aligned}$$

Then, the hypotheses of interest can be equivalently stated as

$$H_0^{(m)} : d_m^* = 0 \;\text{ vs. } H_1^{(m)} : d_m^* > 0.$$

The test statistic defined below is based on the sample analogue of d_m^*.

3 Test Statistics and Large Sample Properties

3.1 Test Statistics

We define our test statistic based on data $\{X_{k,t} : k = 1, 2, t = 1, \ldots, T_k\}$. We first estimate F_k using the empirical cumulative distribution function (ECDF)

$$\bar{F}_k(x) := \frac{1}{T_k} \sum_{t=1}^{T_k} 1(X_{k,t} \leq x),\; k = 1, 2.$$

To test the null hypothesis $H_0^{(1)}$, we consider the following test statistic:

$$\begin{aligned} S_T &= \sqrt{T} \int_{\mathcal{X}} \left\{\left[\bar{F}_1(x) - \bar{F}_2(x)\right]_+ - \epsilon\left|\bar{F}_1(x) - \bar{F}_2(x)\right|\right\} dx \\ &= \sqrt{T} \int_{\mathcal{X}} \left\{(1 - \epsilon)\left[\bar{F}_1(x) - \bar{F}_2(x)\right]_+ + \epsilon\left[\bar{F}_1(x) - \bar{F}_2(x)\right]_-\right\} dx, \end{aligned} \tag{3.1}$$

where $T := T_1 T_2/(T_1 + T_2)$ when $T_1 \neq T_2$ and $T := T_1 = T_2$ otherwise, and the second equality holds since $|a| = [a]_+ - [a]_-$.

Likewise, define the empirical analogue of the general integrated CDF as

$$\bar{F}_k^{(m)}(x) := \frac{1}{T_k} \sum_{t=1}^{T_k} \frac{(x - X_{k,t})^{m-1} 1(X_{k,t} \leq x)}{(m-1)!}, \; k = 1, 2.$$

To test the null hypothesis $H_0^{(m)}$, we consider the following max-type or sum-type test statistic based on the sample analogue of d_m^* defined in (2.3):

$$S_T = \Lambda_p \left(S_{T,1}, \ldots, S_{T,j} \right), \tag{3.2}$$

where

$$\begin{aligned} S_{T,1} &= \sqrt{T} \int_{\mathcal{X}} \left\{ [\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)]_+ - \epsilon |\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)| \right\} dx \Big/ \widehat{\sigma}_1 \\ S_{T,j} &= \sqrt{T} \left[\bar{F}_1^{(j)}(\overline{x}) - \bar{F}_2^{(j)}(\overline{x}) \right] \Big/ \widehat{\sigma}_j \text{ for } 2 \leq j \leq m. \end{aligned}$$

Here, $\widehat{\sigma}_j$'s are normalizing factors which are proportional to the standard deviation of $S_{T,j}$'s to prevent scaling issues. In the following sections, we consider $\widehat{\sigma}_j = 1$ for $2 \leq j \leq m$ for the sake of simplicity.

Remark. For testing the first order ASD, the test statistic coincides with that of Guo, Levy, and Wang (2015, hereafter GLW). Thus, the difference between our test and GLW's test lies in the testing procedure to calculate the critical values for the first order ASD test. On the other hand, the test statistic differs from that of GLW for higher order ASD since our test is a joint test for the entire set of restrictions of ASD, while GLW's test is a test for the first inequality restriction of ASD.

3.2 Large Sample Properties

We present the regularity conditions to derive the asymptotic properties of S_T. In this subsection, we mainly focus on $S_{T,1}$, which corresponds to the first inequality restriction concerning the violation from the m-th order SD. Once we obtain the large sample property of this term, we can test the main hypothesis of interest since the asymptotic normality of each $S_{T,j}$ for $2 \leq j \leq m$ is straightforward.[4] To derive the asymptotic property of our test statistics, we specify conditions for the data generating process of $\{X_{k,t}\}_{t=1}^{T_k}$, $k = 1, 2$. We assume that the observed data are generated under either of the following sampling schemes.

[4]By Theorem 1.4.8 of van der Vaart and Wellner (1996), we have joint convergence since the asymptotic processes of $S_{T,j}$'s are separable. Thus, we can obtain the limit result of our test statistic by applying the continuous mapping theorem.

Assumption 1 (Type I sampling).

(a) $\{X_{k,t}\}_{t=1}^{T_k}$ is an i.i.d. sequence for $k = 1, 2$.

(b) The union of supports of $X_{k,t}$, $k = 1, 2$, is $\mathcal{X} = [\underline{x}, \overline{x}]$, $-\infty < \underline{x} < \overline{x} < \infty$, and the distribution of $X_{k,t}$ is absolutely continuous with respect to the Lebesgue measure and has bounded density, for $k = 1, 2$.

(c) As $T_1, T_2 \to \infty$, $T_1/(T_1 + T_2) \to \lambda \in (0, 1)$.

Assumption 2 (Type II sampling).

(a) $\{(X_{1,t}, X_{2,t})^T : t = 1, \ldots, T\}$ is a strictly stationary and α-mixing sequence with $\alpha(m) = O(m^{-A})$ for some $A > (q-1)(1 + q/2)$, where q is an even integer that satisfies $q > 4$.

(b) Assumption 1 (b) holds.

Under either assumption, we can show that the asymptotic null distribution of the test statistic is a functional of Gaussian processes. Note that $T := T_1T_2/(T_1 + T_2)$ under the Type I sampling and $T := T_1 = T_2$ under the Type II sampling. Define the empirical process in $x \in \mathcal{X}$ as $\nu_T^{(m)}(x) := \sqrt{T}[(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)) - (F_1^{(m)}(x) - F_2^{(m)}(x))]$. Let $\nu_{1,2}^{(m)}$ be a mean zero Gaussian process with covariance function given by

$$
\begin{aligned}
C(x_1, x_2) :&= \lim_{T\to\infty} E[\nu_T^{(m)}(x_1)\nu_T^{(m)}(x_2)] \\
&= \lim_{T\to\infty} T \cdot \mathrm{Cov}\left[\left(\bar{F}_1^{(m)}(x_1) - \bar{F}_2^{(m)}(x_1)\right), \left(\bar{F}_1^{(m)}(x_2) - \bar{F}_2^{(m)}(x_2)\right)\right], \qquad (3.3)
\end{aligned}
$$

for $x_1, x_2 \in \mathbb{R}$. Using Lemma B.1 in Appendix, it can be shown that $\nu_T^{(m)}(\cdot)$ weakly converges to $\nu_{1,2}^{(m)}(\cdot)$. Thus, under the least favorable case (LFC) of the null hypothesis $H_0^{(m)}$ (i.e., $d_{m,j} = 0$, $\forall j = 1, \ldots, m$.), we have

$$S_T \Rightarrow S_0 := \Lambda_p(S_{0,1}, \ldots, S_{0,m}), \qquad (3.4)$$

where

$$S_{0,1} = \int_{\mathcal{C}_0} \left\{(1-\epsilon)\left[\nu_{1,2}^{(m)}(x)\right]_+ + \epsilon\left[\nu_{1,2}^{(m)}(x)\right]_-\right\} dx + (1-\epsilon)\int_{\mathcal{C}_+} \nu_{1,2}^{(m)}(x)dx + \epsilon\int_{\mathcal{C}_-} \nu_{1,2}^{(m)}(x)dx$$

$$S_{0,j} = \nu_{1,2}^{(j)}(\overline{x}) \text{ for } 2 \le j \le m,$$

and

$$\mathcal{C}_0 = \{x \in \mathcal{X} : F_1^{(m)}(x) = F_2^{(m)}(x)\} \qquad (3.5)$$

$$\mathcal{C}_+ = \{x \in \mathcal{X} : F_1^{(m)}(x) > F_2^{(m)}(x)\} \tag{3.6}$$

$$\mathcal{C}_- = \{x \in \mathcal{X} : F_1^{(m)}(x) < F_2^{(m)}(x)\}. \tag{3.7}$$

This suggests that the limiting null distribution is generally non-pivotal and so the method to conduct inference should be settled. We suggest a bootstrap procedure to compute the critical values in the next section.

4 Bootstrap Procedure

The asymptotic distribution of the test statistic depends on the true data generating process. Here, we take an approach for obtaining critical values by mimicking the asymptotic null distribution of an approximation to the test statistic, which exploits information of each inequality restriction of the main hypothesis.

Since $S_{0,1}$, the weak limit of $S_{T,1}$, depends on the binding part of the inequality restrictions (i.e., the "*contact set*") of the support of $\mathcal{X}$, we need to estimate these contact sets. Before introducing the estimators for the contact sets, we introduce related notations. Specifically, we define the r-enlargement of the contact sets for $r > 0$ as follows:

$$\begin{aligned}
\mathcal{C}_0(r) &:= \{x \in \mathcal{X} : \sqrt{T}\left|F_1^{(m)}(x) - F_2^{(m)}(x)\right| \leq r\} \\
\mathcal{C}_+(r) &:= \{x \in \mathcal{X} : \sqrt{T}(F_1^{(m)}(x) - F_2^{(m)}(x)) > r\} \\
\mathcal{C}_-(r) &:= \{x \in \mathcal{X} : \sqrt{T}(F_1^{(m)}(x) - F_2^{(m)}(x)) < -r\}.
\end{aligned}$$

To describe our joint testing procedure, we first introduce selection functions following the moment selection idea of Andrews and Soares (2010). Let $\psi_j : \mathbb{R} \to \{0, 1\}$, $1 \leq j \leq m$, be a selection function which drops its argument whenever the argument is distant from zero in the direction of the inequality restrictions of the null hypothesis. Specifically, we define these functions as follows:

$$\psi_j(x) = 1\left(x \geq -\kappa_{T,j}\right) \text{ for } 1 \leq j \leq m, \tag{4.1}$$

where $\kappa_{T,j} = \kappa_j\sqrt{\log T}$ for $\kappa_j > 0$. Then, the following lemma holds.

Lemma 1. *Suppose that Assumption 1 or Assumption 2 holds. Suppose further that c_T is a positive sequence as $T_1, T_2 \to \infty$ or $T \to \infty$. Then,*

$$P\left\{S_T = \Lambda_p\left(\psi_1\left(\sqrt{T}d_{m,1}\right) \cdot S_{T,1}, \ldots, \psi_m\left(\sqrt{T}d_{m,m}\right) \cdot S_{T,m}\right)\right\} \to 1.$$

Lemma 1 shows that S_T is approximated by an integral with its arguments selected using the rule similar to the generalized moment selection of Andrews and Soares (2010) in large samples. This result suggests that we can consider a bootstrap procedure that mimics the representation of S_T in Lemma 1.

Under the Type I sampling, we use the standard nonparametric bootstrap procedure. Under the Type II sampling, we consider the stationary bootstrap procedure proposed by Politis and Romano (1994). The stationary bootstrap resample is strictly stationary conditional on the original sample. Let $\{L_i\}_{i\in\mathbb{N}}$ denote a sequence of i.i.d. random block lengths following the geometric distribution with a parameter $p \equiv p_T \in (0,1): \; P^*(L_i = l) = p(1-p)^{l-1}$ for each positive integer l. Here, P^* denotes the conditional probability given the original sample.

Equivalently, we can describe the stationary bootstrap procedure as follows. Let $X^*_{k,1}$ be picked at random from the original T observations, so that $X^*_{k,1} = X_{k,I_1}$, where $I_1, I_2, \ldots$ is a sequence of independent and identically distributed variables having the discrete uniform distribution on $\{1,\ldots,T\}$. With probability p, let $X^*_{k,2}$ be picked at random from the original T observations; with probability $1-p$, let $X^*_{k,2} = X_{k,I_1+1}$ so that $X^*_{k,2}$ would be the next observation in the original time series following X_{k,I_1}. In general, given that $X^*_{k,t}$ is determined by the Jth observation $X_{k,J}$ in the original time series, let $X^*_{k,t+1}$ be equal to $X_{k,J+1}$ with probability $1-p$ and picked at random from the original T observations with probability p. We assume that the parameter p satisfies the following growth condition:

Assumption 3. Under the Type II sampling, $p + \left(\sqrt{T}p\right)^{-1} \to 0$ as $T \to \infty$.

We suggest computing the bootstrap critical value for the Type I data in the following steps:

(1) For each $k = 1, 2$, draw a bootstrap sample $\mathcal{S}^*_k := \{X^*_{k,t} : t = 1,\ldots,T_k\}$, where $X^*_{k,t}$ for $t = 1,\ldots,T_k$ are independently drawn with replacement from the original sample $\mathcal{S}_k := \{X_{k,t} : t = 1,\ldots,T_k\}$.

(2) Using the bootstrap sample $\mathcal{S}^*_k$, compute the IEDFs:

$$\bar{F}_k^{(m)*}(\cdot) := \frac{1}{T_k}\sum_{t=1}^{T_k}\frac{(x - X^*_{k,t})^{m-1}1(X^*_{k,t} \le \cdot)}{(m-1)!},\; k = 1, 2. \tag{4.2}$$

(3) Compute the bootstrap test statistic

$$S^*_T = \Lambda_p\left(\psi_1\left(S_{T,1}\right)\cdot S^*_{T,1},\ldots,\psi_m\left(S_{T,m}\right)\cdot S^*_{T,m}\right), \tag{4.3}$$

where

$$S^*_{T,1} = \int_{\widehat{\mathcal{C}}_0(\widehat{c}_T)}\left\{(1-\epsilon)\left[\nu_T^{(m)*}(x)\right]_+ + \epsilon\left[\nu_T^{(m)*}(x)\right]_-\right\}dx$$

$$
\begin{aligned}
&+(1-\epsilon)\int_{\widehat{\mathcal{C}}_+(\widehat{c}_T)} \nu_T^{(m)*}(x)dx + \epsilon \int_{\widehat{\mathcal{C}}_-(\widehat{c}_T)} \nu_T^{(m)*}(x)dx \\
S_{T,j}^* &= \nu_T^{(j)*}(\overline{x}), \ \text{for } 2 \le j \le m.
\end{aligned}
$$

Here, $\nu_T^{(m)*}(x)$ denotes the bootstrap version of $\nu_T^{(m)}(x)$ and $\widehat{\mathcal{C}}_0(\widehat{c}_T)$, $\widehat{\mathcal{C}}_+(\widehat{c}_T)$, and $\widehat{\mathcal{C}}_-(\widehat{c}_T)$ are the estimated contact sets, i.e.,

$$
\nu_T^{(m)*}(x) := \sqrt{T}\left[(\bar{F}_1^{(m)*}(x) - \bar{F}_2^{(m)*}(x)) - (\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x))\right]
$$

$$
\widehat{\mathcal{C}}_0(\widehat{c}_T) := \left\{x \in \mathcal{X} : \sqrt{T}\left|\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right| \le \widehat{c}_T\right\} \tag{4.4}
$$

$$
\widehat{\mathcal{C}}_+(\widehat{c}_T) := \left\{x \in \mathcal{X} : \sqrt{T}(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)) > \widehat{c}_T\right\} \tag{4.5}
$$

$$
\widehat{\mathcal{C}}_-(\widehat{c}_T) := \left\{x \in \mathcal{X} : \sqrt{T}(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)) < -\widehat{c}_T\right\}, \tag{4.6}
$$

and $\widehat{c}_T$ is a positive sequence satisfying Assumption 4 below.

(4) Repeat the steps (1)-(3) above B-times, and compute the $(1-\alpha)$-th quantile of the bootstrap distribution of S_T^* as the bootstrap critical value $c_{T,\alpha}^*$.

Similarly, for the Type II data, the following steps can be used:

(1) Draw a bootstrap sample $\mathcal{S}^* := \{Z^* \equiv (X_{1,t}^*, X_{2,t}^*)^T : t = 1, \dots, T\}$, where $X_{k,t}^*$ for $k = 1, 2$, $t = 1, \dots, T_k$ are constructed using the stationary bootstrap of Politis and Romano (1994) from the original sample $\mathcal{S} := \{Z \equiv (X_{1,t}, X_{2,t})^T : t = 1, \dots, T\}$.

(2) Using the bootstrap sample $\mathcal{S}^*$, compute the IEDFs:

$$
\bar{F}_k^{(m)*}(\cdot) := \frac{1}{T}\sum_{t=1}^{T} \frac{(x - X_{k,t}^*)^{m-1} 1(X_{k,t}^* \le \cdot)}{(m-1)!}, \ k = 1, 2. \tag{4.7}
$$

(3) Compute the bootstrap test statistic

$$
S_T^* = \Lambda_p\left(\psi_1\left(S_{T,1}\right) \cdot S_{T,1}^*, \dots, \psi_m\left(S_{T,m}\right) \cdot S_{T,m}^*\right), \tag{4.8}
$$

where

$$
\begin{aligned}
S_{T,1}^* &= \int_{\widehat{\mathcal{C}}_0(\widehat{c}_T)} \left\{(1-\epsilon)\left[\nu_T^{(m)*}(x)\right]_+ + \epsilon\left[\nu_T^{(m)*}(x)\right]_-\right\} dx \\
&+(1-\epsilon)\int_{\widehat{\mathcal{C}}_+(\widehat{c}_T)} \nu_T^{(m)*}(x)dx + \epsilon \int_{\widehat{\mathcal{C}}_-(\widehat{c}_T)} \nu_T^{(m)*}(x)dx \\
S_{T,j}^* &= \nu_T^{(j)*}(\overline{x}), \ \text{for } 2 \le j \le m.
\end{aligned}
$$

(4) Repeat the steps (1)-(3) above B-times, and compute the $(1-\alpha)$-th quantile of the bootstrap distribution of S_T^* as the bootstrap critical value $c_{T,\alpha}^*$.

Since the test statistic S_T may have a limiting distribution degenerate to zero in some interior cases of the null hypothesis, we suggest taking the maximum of an arbitrarily small number η, say $\eta = 10^{-6}$, and the critical value from Step (4), in order to control the overall size of the test. That is, we take

$$c^*_{T,\alpha,\eta} = \max\{c^*_{T,\alpha}, \eta\} \tag{4.9}$$

as our critical value.

To establish the validity of the bootstrap procedures, we first show the consistency of contact set estimation. We need the following assumption to prove the consistency.

Assumption 4. For each $T_1, T_2 \geq 1$, there exist non-stochastic sequences $c_{T,L}, c_{T,U} > 0$ such that $c_{T,L} \leq \widehat{c}_T \leq c_{T,U}$ and

$$P\{c_{T,L} \leq \widehat{c}_T \leq c_{T,U}\} \to 1, \text{ and } c_{T,L} + \sqrt{T} c_{T,U}^{-1} \to \infty$$

Then, the following lemma holds.

Lemma 2. *Suppose that Assumption 1 or 2 holds. Then, we have*

$$\begin{aligned} P\left\{\widehat{\mathcal{C}}_0(\widehat{c}_T) = \mathcal{C}_0\right\} &\to 1 \\ P\left\{\widehat{\mathcal{C}}_+(\widehat{c}_T) = \mathcal{C}_+\right\} &\to 1 \\ P\left\{\widehat{\mathcal{C}}_-(\widehat{c}_T) = \mathcal{C}_-\right\} &\to 1 \end{aligned}$$

Based on the consistency result of the contact set estimators, we now establish the bootstrap consistency of the bootstrap test statistic. Using the bootstrap consistency, the validity of our bootstrap procedure is shown in the following theorem.

Theorem 1. *Suppose that Assumption 1 holds or Assumption 2 and Assumption 3 hold. Furthermore, Assumption 4 holds. Then, the following holds:*

$$\limsup_{T_1, T_2 \to \infty} P\{S_T > c^*_{T,\alpha,\eta}\} \leq \alpha$$

with equality holding when one of the inequality restrictions of the the null hypothesis binds.

Remark. Another crucial difference between our testing procedure and GLW's is the difference in the bootstrap test statistic. Since our tests for higher order ASD are joint tests and GLW's test is a simple test, we compare $S^*_{T,1}$ which the bootstrap test statistic for the first inequality restriction of the null hypothesis with GLW's bootstrap test statistic $S^{GLW*}_{T,1} := \int_{\mathcal{X}} \left[\nu_T^{(m)*}(x)\right]_+ dx$ for fair comparison. Then, we obtain that $S^*_{T,1} \leq S^{GLW*}_{T,1}$

because

$$
\begin{aligned}
S_{T,1}^{GLW*} - S_{T,1}^{*} &= \int_{\mathcal{X}} \left[\nu_T^{(m)*}(x)\right]_+ dx \\
&- \left(\int_{\widehat{\mathcal{C}}_0(\widehat{c}_T)} \left\{ (1-\epsilon) \left[\nu_T^{(m)*}(x)\right]_+ + \epsilon \left[\nu_T^{(m)*}(x)\right]_- \right\} dx \right. \\
&\left. + (1-\epsilon) \int_{\widehat{\mathcal{C}}_+(\widehat{c}_T)} \nu_T^{(m)*}(x) dx + \epsilon \int_{\widehat{\mathcal{C}}_-(\widehat{c}_T)} \nu_T^{(m)*}(x) dx \right) \\
&= \int_{\widehat{\mathcal{C}}_0(\widehat{c}_T)} \left\{ \epsilon \left[\nu_T^{(m)*}(x)\right]_+ - \epsilon \left[\nu_T^{(m)*}(x)\right]_- \right\} dx \\
&+ \int_{\widehat{\mathcal{C}}_+(\widehat{c}_T)} \left\{ \epsilon \left[\nu_T^{(m)*}(x)\right]_+ - (1-\epsilon) \left[\nu_T^{(m)*}(x)\right]_- \right\} dx \\
&+ \int_{\widehat{\mathcal{C}}_-(\widehat{c}_T)} \left\{ (1-\epsilon) \left[\nu_T^{(m)*}(x)\right]_+ - \epsilon \left[\nu_T^{(m)*}(x)\right]_- \right\} dx \\
&\geq 0.
\end{aligned}
$$

Thus, the test proposed in this paper is less conservative and more powerful than the test suggested by GLW.

5 Asymptotic Power Properties

We investigate power properties of our tests. We first establish consistency of our proposed test:

Theorem 2. *Suppose that Assumption 1 holds or Assumption 2 and Assumption 3 hold. Furthermore, Assumption 4 holds. Then, under a fixed alternative hypothesis $H_1^{(m)}$, we have, as $T_1, T_2 \to \infty$,*

$$P\left\{S_T > c_{T,\alpha,\eta}^*\right\} \to 1.$$

Therefore, our bootstrap test is consistent against any type of fixed alternative.

Let $\mathcal{P}$ be the collection of all the potential distributions of $\{X_1, X_2\}$ that satisfy Assumption 1 or Assumption 2 and 3. Furthermore, let $\mathcal{P}_0$ be the collection of probabilities that satisfy $H_0^{(m)}$. To investigate asymptotic local power properties, we consider a sequence of probabilities $P_T \in \mathcal{P} \setminus \mathcal{P}_0$. We confine our attention to the following sequence of local alternatives under P_T,

$$
\begin{aligned}
H_a^{(m)} : F_1^{(m)}(x) - F_2^{(m)}(x) &= \mu_m(x) + \delta(x)/\sqrt{T} \\
&\text{and} \\
F_1^{(j)}(\overline{x}) - F_2^{(j)}(\overline{x}) &= \mu_j(\overline{x}) \text{ for } 2 \leq j \leq m.
\end{aligned}
\tag{5.1}
$$

Define $\mathcal{C}_a^0 \equiv \{x \in \mathcal{X} : \mu_m(x) = 0\}$, $\mathcal{C}_a^+ \equiv \{x \in \mathcal{X} : \mu_m(x) > 0\}$, and $\mathcal{C}_a^- \equiv \{x \in \mathcal{X} : \mu_m(x) < 0\}$. We assume the following for the functions $\mu_j(\cdot)$, $1 \leq j \leq m$, and $\delta(\cdot)$.

Assumption 5. $\mu_j(\cdot)$, $1 \le j \le m$, and $\delta(\cdot)$ are bounded, real functions satisfying

(a) (i) $\int_{\mathcal{X}} \left\{ [\mu_m(x)]_+ - \epsilon \left|\mu_m(x)\right| \right\} dx = 0.$

(ii) $\mu_j(\overline{x}) \le 0$ for $2 \le j \le m$.

(b) $\int_{\mathcal{C}_a^0} \left\{ [\delta(x)]_+ - \epsilon \left|\delta(x)\right| \right\} dx + (1-\epsilon) \int_{\mathcal{C}_a^+} \delta(x)dx + \epsilon \int_{\mathcal{C}_a^-} \delta(x)dx > 0.$

Assumption 5 ensures that the local alternative hypothesis H_a is non-void, violating the first inequality restriction of the null hypothesis. Specifically, Assumption 5 (a) makes sure that the local alternatives under the sequence of probabilities P_T converge to the null hypothesis. Let $Q(\cdot)$ denote Lebesgue measure on $\mathbb{R}$. Since the equality in Assumption 5 (a) can be written as $(1-\epsilon) \int_{\mathcal{C}_a^+} \mu_m(x)dx + \epsilon \int_{\mathcal{C}_a^-} \mu_m(x)dx > 0$ along with $Q(\mathcal{C}_0 \cup \mathcal{C}_+) = Q(\mathcal{X} \setminus \mathcal{C}_a^-) > 0$. Then, $Q(\mathcal{C}_a^+) = 0$ implies that the sequence of local alternatives deviates from the binding part of the null hypothesis where two distributions coincide. On the other hand, $Q(\mathcal{C}_a^+) > 0$ corresponds to the case where the local alternatives goes away from the binding part of the null hypothesis where there are crossings between two distributions.

We can consider testing the null hypothesis of $H_{0,1}^{(m)} : d_{m,1} \le 0$ since this is the hypothesis of interest for the GLW test. Then, the local alternatives we consider are of the form

$$H_{a,1}^{(m)} : F_1^{(m)}(x) - F_2^{(m)}(x) = \mu_m(x) + \delta(x)/\sqrt{T}. \tag{5.2}$$

where $\mu_m(\cdot)$ and $\delta(\cdot)$ are satisfying Assumption 5 (a) (i) and Assumption 5 (b). Following the same bootstrap procedure in Section 4, we define the bootstrap critical value $c_{T,1,\alpha}^*$ as the $(1-\alpha)$-th quantile of the bootstrap distribution of $S_{T,1}^*$. Let $c_{T,1,\alpha}^{GLW*}$ denote the bootstrap critical value, i.e., the $(1-\alpha)$-th quantile of the bootstrap distribution of $S_{T,1}^{GLW*}$. Alternatively, we compare our test with GLW's by considering an alternative testing procedure in which we replace the original $S_{T,1}^*$ in Step 3 of the testing procedure with $S_{T,1}^{GLW*}$. Let S_T^{GLW*} and $c_{T,\alpha,\eta}^{GLW*}$ denote the bootstrap test statistic and the critical value in this case, respectively.

Then, the following theorem implies that the GLW test is inadmissible in the sense that our bootstrap test strictly dominates their test in test power.

Theorem 3. *Suppose that Assumption 1 holds or Assumption 2 and Assumption 3 hold.*

(1) Suppose that Assumption 5 (a) (i) and Assumption 5 (b) additionally hold. Under the local alternative hypothesis $H_{a,1}^{(m)}$ *satisfying (5.2),*

$$\lim_{T_1,T_2 \to \infty} P_T\{S_{T,1} > c_{T,1,\alpha}^*\} > \lim_{T_1,T_2 \to \infty} P_T\{S_{T,1} > c_{T,1,\alpha}^{GLW*}\}. \tag{5.3}$$

(2) Suppose that Assumption 5 additionally holds. Under the local alternative hypothesis

$H_a^{(m)}$ *satisfying (5.1),*

$$\lim_{T_1,T_2\to\infty} P_T\{S_T > c^*_{T,\alpha,\eta}\} > \lim_{T_1,T_2\to\infty} P_T\{S_T > c^{GLW*}_{T,\alpha,\eta}\}. \tag{5.4}$$

The intuition behind this result is that the GLW test has weak power against the alternatives where there are crossings between two distributions. Since their bootstrap distribution is mimicking the null distribution of the test statistic under the special case of the binding null hypothesis in which the two distributions coincide, GLW's test is less sensitive to violation of the null hypothesis in this direction. Indeed, their test has actually even weaker power than the hypothetical test whose bootstrap test statistic mimics the null distribution under this special case of the binding null hypothesis. This is because the bootstrap test statistic would become $S^{Hypothetical*}_{T,1} := \int_{\mathcal{X}} \left\{ \left[\nu_T^{(m)*}(x)\right]_+ - \epsilon \left|\nu_T^{(m)*}(x)\right| \right\} dx < S^{GLW*}_{T,1} = \int_{\mathcal{X}} \left[\nu_T^{(m)*}(x)\right]_+ dx$ in this case. Thus, the GLW test has weaker local power than our test under either $Q(\mathcal{C}_a^+) = 0$ or $Q(\mathcal{C}_a^+) > 0$.

6 Monte Carlo Simulation

In this section, we conduct a set of Monte Carlo simulations to investigate the finite sample performance of the bootstrap tests. We evaluate the size and power properties of the proposed tests for the first order and the second order ASD.

To implement our testing procedure, we need to choose the allowed area violation which is specified by a fixed constant ϵ. In an experimental study by which one can determine the set of economically irrelevant utility functions, Levy, Leshno, and Leibovitch (2010) find that the minimum allowed area violations across all the subjects are $\epsilon = 0.059$ for AFSD and $\epsilon = 0.032$ for ASSD. Following this empirical guide in the literature, we fix ϵ for both AFSD and ASSD to 0.05 for the sake of simplicity.

For both AFSD and ASSD simulations, we calculate the test statistics by the trapezoidal rule using 100 grid points. The bootstrap sample size is $B = 200$ and the number of iteration is $R = 1000$. The sample size is $T = T_1 = T_2 \in \{100, 200, 300, 400, 500\}$.

The parameter $\widehat{c}_T$ for the estimation of contact sets is chosen in a data-driven way as follows:

$$\widehat{c}_T = C_{cs}(\log\log T) q_{1-\alpha_T}(R_T^*), \tag{6.1}$$

where

$$R_T^* \equiv \max\left\{\sup_x \nu_T^{(m)*}(x), \lambda\sqrt{\log T}\right\}$$

and $q_{1-\alpha_T}(R_T^*)$ is the $(1-\alpha_T)$-th quantile of the bootstrap distribution of R_T^* with $\alpha_T = 0.1/\log T$, C_{cs} is a constant, and λ is a small number.[5] For tuning parameters, we take

[5]This choice of $\widehat{c}_T$ ensures that it diverges to infinity as slowly as possible while automatically adjusting

$C_{cs} = 0.2$ for both AFSD and ASSD under $p = 1$ and $p = 2$, and $\lambda = 10^{-6}$. For ASSD only, we take $\kappa_1 = 0.05$ and $\kappa_2 = 1$. We implement sensitivity analysis to these tuning parameters.

To prevent scaling issues, we choose normalizing factors proportional to the standard deviation of each term. For AFSD, we take $\widehat{\sigma}_1^2 = \widehat{\mathrm{Var}}\left(\bar{F}_1^{(2)}(\overline{x}) - \bar{F}_2^{(2)}(\overline{x})\right)$. For ASSD, we take $\widehat{\sigma}_1^2 = \widehat{\mathrm{Var}}\left(\bar{F}_1^{(3)}(\overline{x}) - \bar{F}_2^{(3)}(\overline{x})\right)$ and $\widehat{\sigma}_2^2 = \widehat{\mathrm{Var}}\left(\bar{F}_1^{(2)}(\overline{x}) - \bar{F}_2^{(2)}(\overline{x})\right)$.

6.1 AFSD

We consider the following data generating processes to check the finite sample property of the AFSD test when the contact set $\mathcal{C}_0$ is not empty: when $x_0 \neq 0$,

$$\begin{aligned} X &\equiv U_1 \\ Y &\equiv c_0^{-1}(U_2 - a_0)1\{0 < U_2 \leq x_0\} \\ &\quad + U_2 1\{x_0 \leq U_2 \leq x_1\} + c_1^{-1}(U_2 - a_1)1\{x_1 < U_2 < 1\}, \end{aligned}$$

where U_1 and U_2 are independent generated from the uniform distribution with range $[0, 1]$, and $x_0, x_1 \in \mathbb{R}$, $a_0 = \frac{x_0}{2}$, $c_0 = (x_0 - a_0)/x_0$, $a_1 = \frac{x_1}{2}$, and $c_1 = (x_1 - a_1)/x_1$. Note that the dgps are the modified version of those from Linton, Song, and Whang (2010). Then, the distribution of Y is given by

$$P\{Y \leq y\} = \begin{cases} 1 & \text{if } y \geq (1 - a_1)/c_1 \\ c_1 y + a_1 & \text{if } x_1 < y < (1 - a_1)/c_1 \\ y & \text{if } x_0 \leq y \leq 1 \\ c_0 y + a_0 & \text{if } -a_0/c_0 < y \leq x_0 \\ 0 & \text{if } y \leq -a_0/c_0 \end{cases}$$

Let $x_0, x_1 \in \mathbb{R}$, $a_0 = \frac{x_0}{2}$, $c_0 = (x_0 - a_0)/x_0$, $a_1 = \frac{x_1}{2}$, and $c_1 = (x_1 - a_1)/x_1$. When $x_0 = 0$,

$$Y \equiv U_2 1\{x_0 \leq U_2 \leq x_1\} + c_1^{-1}(U_2 - a_1)1\{x_1 < U_2 < 1\}$$

and

$$P\{Y \leq y\} = \begin{cases} 1 & \text{if } y \geq (1 - a_1)/c_1 \\ c_1 y + a_1 & \text{if } x_1 < y < (1 - a_1)/c_1 \\ y & \text{if } x_0 \leq y \leq 1. \end{cases}$$

To investigate the size property of the AFSD test, we consider the interior and binding cases of the null hypothesis $H_0^{(1)}$, that is, $d_{1,1} < 0$ and $d_{1,1} = 0$. Specifically, we can think

to changes in scale. See Lee, Song, and Whang (2018) and Chernozhukov, Lee, and Rosen (2013) for similar ideas.

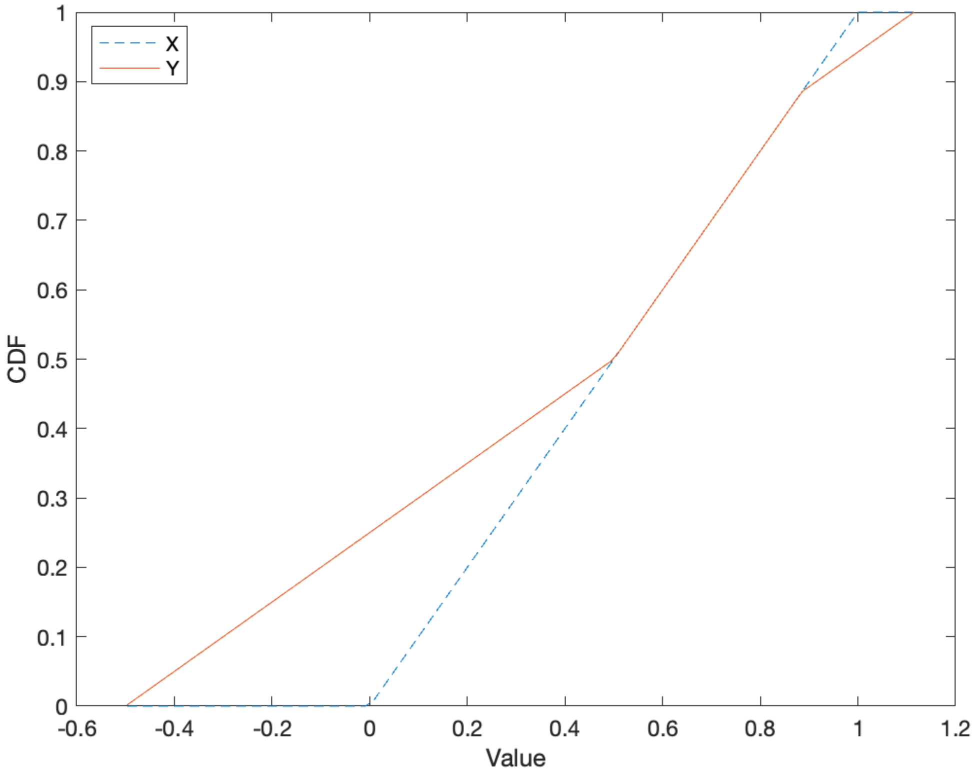
of two different scenarios for each case depending on whether there are crossings between two distributions. When $d_{1,1} < 0$ without crossings, X dominates Y by FSD and so by AFSD, which we call "Dominance" case. When $d_{1,1} < 0$ with crossings, X dominates Y by AFSD but not by FSD, which we call "Crossing (Interior)" case. When $d_{1,1} = 0$ without crossings, X trivially dominates Y by FSD and so by AFSD, which we call "Same Distribution" case. When $d_{1,1} = 0$ with crossings, X dominates Y by AFSD but not by FSD, which we call "Crossing (Boundary)" case.

Likewise, we consider the the exterior cases of the null hypothesis $H_0^{(1)}$ to investigate the power property of the AFSD test. When $d_{1,1} > 0$ without crossings, Y dominates X by FSD and so by AFSD, which we call "Reverse Dominance" case. When $d_{1,1} > 0$ with crossings, X does not dominate Y by AFSD, but Y also does not dominate X by FSD, which we call "Crossing (Exterior)" case. We summarize our parameter choice for each of these cases in Table 1. As an illustration, Figure 1 shows the cumulative distribution functions of X and Y for "Crossing (Boundary)" case.

Table 1: Simulation Design for AFSD

	Case	x_0	x_0	$d_{1,1}$
Size	Dominance	0.75	1	-0.0141
	Crossing (Interior)	0.75	0.95	-0.0129
	Same Distribution	0	0.5	0
	Crossing (Boundary)	0.5	0.8853	0
Power	Reverse Dominance	0	0.8	0.019
	Reverse Dominance	0	0.7	0.0428
	Reverse Dominance	0	0.6	0.076
	Crossing (Exterior)	0.5	0.8	0.0128
	Crossing (Exterior)	0.5	0.7	0.0365
	Crossing (Exterior)	0.5	0.6	0.0698

Figure 1: CDFs



For AFSD, we only report the results using the max type statistic for $p = 1$ since the sum type statistic or $p = 2$ variations show similar results. We provide sensitivity analysis result for C_{cs} by allowing for its variations from 0.1 to 0.3. Table 2 and 3 show the rejection ratio under the null hypothesis, while Table 4 shows that under the alternative hypothesis. In the interior case of the null hypothesis, the AFSD test shows proper size control. When the null hypothesis binds, the AFSD test has an exact size under both types of the LFC. Consistent with the theoretical prediction, the GLW test under "Crossing (Boundary)" case shows conservative size control. Under the alternative hypothesis, the AFSD test is more powerful than the GLW test under "Reverse Dominance" case. As the local power analysis shows, the AFSD test is even more powerful under "Crossing (Exterior)" case.

Table 2: Size Property ($d_{1,1} < 0$)

	Dominance				Crossing (Interior)			
	$d_{1,1} = -0.0141$				$d_{1,1} = -0.0129$			
		C_{cs}				C_{cs}		
T	H_0^{GLW}	0.1	0.2	0.3	H_0^{GLW}	0.1	0.2	0.3
100	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
200	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
400	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
500	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 3: Size Property ($d_{1,1} = 0$)

	Same Distribution				Crossing (Boundary)			
	$d_{1,1} = 0$				$d_{1,1} = -0.0000$			
		C_{cs}				C_{cs}		
T	H_0^{GLW}	0.1	0.2	0.3	H_0^{GLW}	0.1	0.2	0.3
100	0.050	0.058	0.055	0.053	0.002	0.028	0.019	0.013
200	0.039	0.055	0.052	0.052	0.002	0.020	0.018	0.012
300	0.049	0.062	0.059	0.059	0.000	0.027	0.017	0.012
400	0.042	0.054	0.052	0.051	0.000	0.028	0.022	0.017
500	0.048	0.061	0.060	0.059	0.003	0.044	0.030	0.024

Table 4: Power Property ($d_{1,1} > 0$)

	Reverse Dominance				Crossing (Exterior)			
	$d_{1,1} = 0.0190$				$d_{1,1} = 0.0128$			
		C_{cs}				C_{cs}		
T	H_0^{GLW}	0.1	0.2	0.3	H_0^{GLW}	0.1	0.2	0.3
100	0.112	0.152	0.135	0.131	0.007	0.074	0.055	0.036
200	0.145	0.241	0.183	0.172	0.010	0.124	0.082	0.058
300	0.182	0.360	0.249	0.223	0.008	0.181	0.122	0.093
400	0.236	0.496	0.322	0.286	0.011	0.283	0.190	0.143
500	0.277	0.591	0.370	0.326	0.019	0.359	0.230	0.184
	Reverse Dominance				Crossing (Exterior)			
	$d_{1,1} = 0.0428$				$d_{1,1} = 0.0365$			
		C_{cs}				C_{cs}		
T	H_0^{GLW}	0.1	0.2	0.3	H_0^{GLW}	0.1	0.2	0.3
100	0.246	0.461	0.334	0.290	0.043	0.350	0.226	0.164
200	0.410	0.816	0.627	0.523	0.075	0.776	0.564	0.435
300	0.600	0.953	0.817	0.712	0.137	0.952	0.845	0.707
400	0.774	0.987	0.936	0.878	0.235	0.991	0.954	0.882
500	0.900	0.998	0.983	0.957	0.317	0.998	0.984	0.963
	Reverse Dominance				Crossing (Exterior)			
	$d_{1,1} = 0.0760$				$d_{1,1} = 0.0698$			
		C_{cs}				C_{cs}		
T	H_0^{GLW}	0.1	0.2	0.3	H_0^{GLW}	0.1	0.2	0.3
100	0.526	0.859	0.715	0.634	0.198	0.845	0.703	0.554
200	0.867	0.996	0.973	0.949	0.475	0.999	0.982	0.948
300	0.987	0.999	0.999	0.998	0.756	1.000	0.999	0.998
400	0.999	1.000	1.000	1.000	0.926	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000	0.983	1.000	1.000	1.000

6.2 ASSD

We investigate the following data generating processes:

$$X \sim F$$

$$Y \sim G$$

where

$$F(x) = \frac{-x^2 + 2mx}{m^2} 1(0 \leq x \leq m) + 1(x > m)$$
$$G(x) = \frac{x - a}{b - a} 1(a \leq x \leq b) + 1(x > b)$$

with $m = 30$ and varying $a, b \in \mathbb{R}$. The data generating processes are from Osuna (2012).

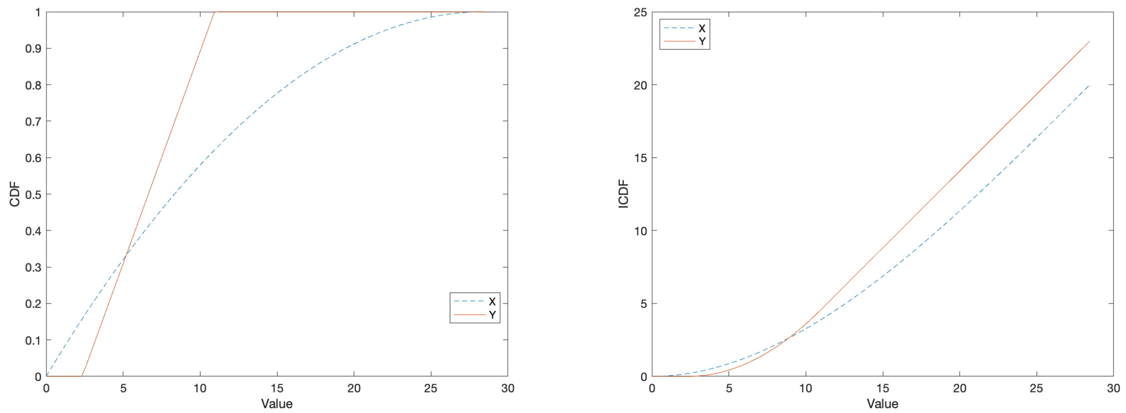
To investigate the size property of the ASSD test, we consider the interior and binding cases of the null hypothesis $H_0^{(2)}$. There are four cases: $d_{2,1} < 0$ and $d_{2,2} < 0$, $d_{2,1} = 0$ and $d_{2,2} < 0$, $d_{2,1} < 0$ and $d_{2,2} = 0$, and $d_{2,1} = 0$ and $d_{2,2} = 0$. We can further divide these cases depending on whether there are crossings between two distributions. When $d_{2,1} < 0$ and $d_{2,2} < 0$ without crossings, X dominates Y by SSD and so by ASSD, which we call "Dominance 1" case. When $d_{2,1} < 0$ and $d_{2,2} < 0$ with crossings, X dominates Y by ASSD but not by SSD, which we call "Crossing (Interior)" case. When $d_{2,1} = 0$ and $d_{2,2} < 0$, there are crossings between two distributions, so we call this "Crossing (Boundary 1)" case. When $d_{2,1} < 0$ and $d_{2,2} = 0$ without crossings, X dominates Y by SSD and so by ASSD, which we call "Dominance 2" case. When $d_{2,1} < 0$ and $d_{2,2} = 0$ with crossings, X dominates Y by ASSD but not by SSD, which we call "Crossing (Boundary 2)" case. When $d_{2,1} = 0$ and $d_{2,2} = 0$ without crossings, X trivially dominates Y by SSD and so by ASSD, which we call "Same Distribution" case. When $d_{2,1} = 0$ and $d_{2,2} = 0$ with crossings, X dominates Y by ASSD but not by SSD, which we call "Crossing (Boundary 3)" case.

Likewise, we consider the the exterior cases of the null hypothesis $H_0^{(2)}$ to investigate the power property of the ASSD test. There are five cases: $d_{2,1} > 0$ and $d_{2,2} < 0$, $d_{2,1} > 0$ and $d_{2,2} = 0$, $d_{2,1} < 0$ and $d_{2,2} > 0$, $d_{2,1} = 0$ and $d_{2,2} > 0$, and $d_{2,1} > 0$ and $d_{2,2} > 0$. When $d_{2,1} > 0$ and $d_{2,2} < 0$, there are crossings between two distributions, so we call this "Crossing (Exterior 1)" case. When $d_{2,1} > 0$ and $d_{2,2} = 0$ without crossings, Y dominates X by SSD and so by ASSD, which we call "Reverse Dominance 1" case. When $d_{2,1} > 0$ and $d_{2,2} = 0$ with crossings, X does not dominate Y by ASSD and Y does not dominate X by SSD, which we call "Crossing (Exterior 2)" case. When $d_{2,1} < 0$ and $d_{2,2} > 0$, there are crossings between two distributions, so we call this "Crossing (Exterior 3)" case. When $d_{2,1} = 0$ and $d_{2,2} > 0$, there are crossings between two distributions, so we call this "Crossing (Exterior 4)" case. When $d_{2,1} > 0$ and $d_{2,2} > 0$ without crossings, Y dominates X by SSD and so by ASSD, which we call "Reverse Dominance 2" case. When $d_{2,1} > 0$ and $d_{2,2} > 0$ with crossings, X does not dominate Y by ASSD and Y does not dominate X by SSD, which we call "Crossing (Exterior 5)" case. We summarize our parameter choice for each of these cases in Table 5. As an illustration, Figure 2 shows the (integrated) cumulative distribution functions of X and Y for "Crossing (Boundary 1)" case.

Table 5: Simulation Design for ASSD

	Case	a	b	$(d_{2,1}, d_{2,2})$
Size	Dominance 1	-12	26	$(-4.9833, -3)$
	Crossing (Interior)	1	13	$(-1.7449, -3)$
	Crossing (Boundary 1)	2.4705	11.5295	$(0, -3)$
	Dominance 2	-20	40	$(-6.25, 0)$
	Crossing (Boundary 2)	-7	27	$(-1.139, 0)$
	Same Distribution	-	-	$(0, 0)$
	Crossing (Boundary 3)	-4.2767	24.2767	$(0, 0)$
Power	Crossing (Exterior 1)	4	10	$(1.7181, -3)$
	Crossing (Exterior 1)	5	9	$(2.5097, -3)$
	Crossing (Exterior 1)	6	8	$(2.9847, -3)$
	Reverse Dominance 1	0	20	$(7.9167, 0)$
	Reverse Dominance 1	2	18	$(13.6167, 0)$
	Reverse Dominance 1	5	15	$(19.7917, 0)$
	Crossing (Exterior 2)	-1	21	$(4.8787, 0)$
	Crossing (Exterior 2)	-0.4	20.4	$(6.6476, 0)$
	Crossing (Exterior 2)	-0.1	20.1	$(7.5988, 0)$
	Crossing (Exterior 3)	-40	62	$(-9.2803, 1)$
	Crossing (Exterior 3)	-130	156	$(-72.919, 3)$
	Crossing (Exterior 3)	-230	260	$(-227.0833, 5)$
	Crossing (Exterior 4)	-23.8046	45.8046	$(0, 1)$
	Crossing (Exterior 4)	-83.1492	109.1492	$(0, 3)$
	Crossing (Exterior 4)	-144.067	174.0666	$(0, 5)$
	Reverse Dominance 2	0	22	$(23.1167, 1)$
	Reverse Dominance 2	0	26	$(49.7167, 3)$
	Reverse Dominance 2	0	30	$(71.25, 5)$
	Crossing (Exterior 5)	-13	35	$(3.6021, 1)$
	Crossing (Exterior 5)	-18	44	$(32.9038, 3)$
	Crossing (Exterior 5)	-23	53	$(82.084, 5)$

Figure 2: CDFs and ICDFs



(a) CDFs (b) ICDFs

For ASSD, we only report the results using the max type statistic for $p = 1$ since the sum type statistic or $p = 2$ variations show similar results. We relegate sensitivity analysis result for C_{cs} to the appendix. Table 6 to 9 show the rejection ratios under the null hypothesis, while Table 10 to 14 show that under the alternative hypothesis. In the interior case of the null hypothesis, the ASSD test shows proper size control. When the null hypothesis binds, the ASSD test has an exact size under every case of the LFC. Consistent with the theoretical prediction, the GLW test under "Crossing (Boundary)" case show conservative size control. Under the alternative hypothesis, the ASSD test is more powerful than the GLW test under "Reverse Dominance" case. As the local power analysis shows, the ASSD test is even more powerful under "Crossing (Exterior)" case.

Table 6: Size Property ($d_{2,1} < 0, d_{2,2} < 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(-4.9833,-3)	100	0.000	0.000	0.000	0.000	0.000	0.000
Dominance 1	200	0.000	0.000	0.000	0.000	0.000	0.000
	300	0.000	0.000	0.000	0.000	0.000	0.000
	400	0.000	0.000	0.000	0.000	0.000	0.000
	500	0.000	0.000	0.000	0.000	0.000	0.000
(-1.7449,-3)	100	0.000	0.001	0.000	0.000	0.000	0.000
Crossing (Interior)	200	0.000	0.001	0.000	0.001	0.001	0.001
	300	0.000	0.000	0.000	0.000	0.000	0.000
	400	0.000	0.000	0.000	0.000	0.000	0.000
	500	0.000	0.000	0.000	0.000	0.000	0.000

Table 7: Size Property ($d_{2,1} = 0, d_{2,2} < 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(-0.0000,-3)	100	0.000	0.001	0.000	0.006	0.000	0.000
Crossing (Boundary 1)	200	0.000	0.008	0.000	0.019	0.008	0.003
	300	0.000	0.018	0.000	0.025	0.018	0.009
	400	0.000	0.019	0.000	0.033	0.019	0.010
	500	0.000	0.021	0.000	0.029	0.021	0.017

Table 8: Size Property ($d_{2,1} < 0, d_{2,2} = 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(-6.2500,0)	100	0.002	0.002	0.051	0.051	0.051	0.051
Dominance 2	200	0.001	0.002	0.046	0.046	0.046	0.046
	300	0.000	0.002	0.049	0.049	0.049	0.049
	400	0.000	0.000	0.053	0.053	0.053	0.053
	500	0.000	0.003	0.048	0.048	0.048	0.048
(-1.1390,0)	100	0.007	0.006	0.053	0.053	0.047	0.044
Crossing (Boundary 2)	200	0.003	0.005	0.050	0.050	0.046	0.043
	300	0.002	0.003	0.054	0.054	0.054	0.049
	400	0.000	0.000	0.051	0.051	0.051	0.050
	500	0.003	0.007	0.053	0.053	0.053	0.052

Table 9: Size Property ($d_{2,1} = 0, d_{2,2} = 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(0,0)	100	0.047	0.052	0.056	0.053	0.053	0.053
Same Distribution	200	0.043	0.048	0.045	0.043	0.043	0.043
	300	0.048	0.052	0.049	0.050	0.050	0.050
	400	0.040	0.049	0.054	0.050	0.050	0.050
	500	0.041	0.047	0.058	0.052	0.052	0.052
(-0.0000,0)	100	0.019	0.019	0.057	0.053	0.049	0.049
Crossing (Boundary 3)	200	0.015	0.018	0.049	0.048	0.046	0.046
	300	0.019	0.022	0.055	0.054	0.051	0.050
	400	0.008	0.013	0.051	0.051	0.049	0.048
	500	0.017	0.019	0.050	0.050	0.049	0.046

Table 10: Power Property ($d_{2,1} > 0, d_{2,2} < 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(1.7181,-3)	100	0.000	0.073	0.000	0.135	0.063	0.016
Crossing (Exterior 1)	200	0.000	0.255	0.000	0.337	0.255	0.174
	300	0.000	0.419	0.000	0.472	0.419	0.348
	400	0.000	0.533	0.000	0.597	0.533	0.489
	500	0.000	0.648	0.000	0.700	0.648	0.599
(2.5097,-3)	100	0.000	0.231	0.000	0.347	0.221	0.091
Crossing (Exterior 1)	200	0.000	0.576	0.000	0.658	0.576	0.472
	300	0.000	0.777	0.000	0.821	0.777	0.709
	400	0.000	0.882	0.000	0.917	0.882	0.856
	500	0.000	0.947	0.000	0.964	0.947	0.937
(2.9847,-3)	100	0.000	0.379	0.000	0.523	0.369	0.200
Crossing (Exterior 1)	200	0.000	0.781	0.000	0.825	0.781	0.692
	300	0.000	0.931	0.000	0.944	0.931	0.905
	400	0.000	0.975	0.000	0.979	0.975	0.967
	500	0.000	0.994	0.000	0.996	0.994	0.992

Table 11: Power Property ($d_{2,1} > 0, d_{2,2} = 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(7.9167,0)	100	0.138	0.147	0.050	0.113	0.113	0.113
Reverse Dominance 1	200	0.162	0.180	0.054	0.150	0.150	0.150
	300	0.216	0.236	0.053	0.196	0.196	0.196
	400	0.260	0.287	0.049	0.236	0.236	0.236
	500	0.282	0.314	0.047	0.269	0.269	0.269
(13.6167,0)	100	0.256	0.279	0.047	0.231	0.231	0.231
Reverse Dominance 1	200	0.373	0.399	0.051	0.343	0.343	0.343
	300	0.488	0.533	0.052	0.465	0.462	0.462
	400	0.611	0.712	0.048	0.590	0.590	0.589
	500	0.714	0.860	0.045	0.710	0.702	0.697
(19.7917,0)	100	0.515	0.616	0.044	0.483	0.479	0.475
Reverse Dominance 1	200	0.772	0.952	0.054	0.757	0.754	0.753
	300	0.950	1.000	0.055	0.948	0.948	0.948
	400	0.985	1.000	0.048	0.989	0.989	0.989
	500	0.999	1.000	0.044	1.000	1.000	1.000
(4.8787,0)	100	0.093	0.097	0.050	0.080	0.080	0.080
Crossing (Exterior 2)	200	0.101	0.117	0.054	0.083	0.083	0.083
	300	0.109	0.124	0.056	0.099	0.099	0.099
	400	0.139	0.150	0.050	0.123	0.123	0.123
	500	0.136	0.151	0.050	0.127	0.127	0.127
(6.6476,0)	100	0.117	0.126	0.048	0.099	0.099	0.099
Crossing (Exterior 2)	200	0.140	0.153	0.054	0.123	0.123	0.123
	300	0.163	0.191	0.052	0.155	0.155	0.155
	400	0.198	0.218	0.049	0.182	0.182	0.182
	500	0.210	0.240	0.048	0.198	0.198	0.198
(7.5988,0)	100	0.128	0.139	0.048	0.110	0.110	0.110
Crossing (Exterior 2)	200	0.156	0.170	0.055	0.146	0.146	0.146
	300	0.204	0.220	0.052	0.187	0.187	0.187
	400	0.245	0.269	0.049	0.213	0.213	0.213
	500	0.269	0.294	0.047	0.246	0.246	0.246

Table 12: Power Property ($d_{2,1} < 0, d_{2,2} > 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(-9.2803,1)	100	0.002	0.010	0.097	0.097	0.097	0.097
Crossing (Exterior 3)	200	0.001	0.006	0.115	0.115	0.115	0.115
	300	0.001	0.005	0.131	0.131	0.131	0.131
	400	0.000	0.002	0.150	0.150	0.150	0.150
	500	0.000	0.006	0.165	0.165	0.165	0.165
(-72.9190,3)	100	0.003	0.011	0.092	0.092	0.092	0.092
Crossing (Exterior 3)	200	0.001	0.007	0.123	0.123	0.123	0.123
	300	0.001	0.007	0.152	0.152	0.152	0.152
	400	0.000	0.005	0.164	0.164	0.164	0.164
	500	0.000	0.007	0.180	0.180	0.180	0.180
(-227.0833,5)	100	0.004	0.011	0.089	0.089	0.089	0.089
Crossing (Exterior 3)	200	0.001	0.007	0.124	0.124	0.124	0.124
	300	0.001	0.007	0.149	0.149	0.149	0.149
	400	0.000	0.005	0.166	0.166	0.166	0.166
	500	0.000	0.007	0.175	0.175	0.175	0.175

Table 13: Power Property ($d_{2,1} = 0, d_{2,2} > 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(0.0000,1)	100	0.003	0.011	0.128	0.128	0.128	0.127
Crossing (Exterior 4)	200	0.002	0.011	0.149	0.149	0.149	0.149
	300	0.002	0.010	0.201	0.201	0.201	0.201
	400	0.000	0.010	0.235	0.235	0.235	0.235
	500	0.003	0.015	0.260	0.260	0.260	0.260
(0.0000,3)	100	0.006	0.016	0.131	0.131	0.131	0.131
Crossing (Exterior 4)	200	0.001	0.010	0.168	0.168	0.168	0.168
	300	0.001	0.009	0.243	0.243	0.243	0.243
	400	0.000	0.011	0.273	0.273	0.273	0.273
	500	0.003	0.017	0.307	0.307	0.307	0.307
(0.0000,5)	100	0.008	0.016	0.138	0.138	0.138	0.138
Crossing (Exterior 4)	200	0.001	0.013	0.168	0.168	0.168	0.168
	300	0.001	0.008	0.245	0.245	0.245	0.245
	400	0.000	0.011	0.271	0.271	0.271	0.271
	500	0.002	0.017	0.309	0.309	0.309	0.309

Table 14: Power Property ($d_{2,1} > 0, d_{2,2} > 0$)

		$H_{0,1}^{(2)} : d_{2,1} \leq 0$		$H_{0,2}^{(2)} : d_{2,2} \leq 0$	$H_0^{(2)}$		
						C_{cs}	
$(d_{2,1}, d_{2,2})$	T	$H_{0,1}^{GLW}$	$H_{0,1}^{(2)}$	$H_{0,2}^{(2)}$	0.1	0.2	0.3
(23.1167,1)	100	0.397	0.420	0.282	0.372	0.372	0.372
Reverse Dominance 2	200	0.589	0.612	0.430	0.567	0.567	0.567
	300	0.721	0.747	0.559	0.699	0.699	0.699
	400	0.830	0.857	0.658	0.819	0.819	0.819
	500	0.910	0.924	0.757	0.900	0.900	0.900
(49.7167,3)	100	0.857	0.869	0.877	0.876	0.876	0.876
Reverse Dominance 2	200	0.988	0.989	0.995	0.994	0.994	0.994
	300	1.000	1.000	1.000	1.000	1.000	1.000
	400	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
(71.2500,5)	100	0.987	0.988	0.997	0.997	0.997	0.997
Reverse Dominance 2	200	1.000	1.000	1.000	1.000	1.000	1.000
	300	1.000	1.000	1.000	1.000	1.000	1.000
	400	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
(3.6021,1)	100	0.012	0.013	0.161	0.161	0.160	0.155
Crossing (Exterior 5)	200	0.005	0.021	0.212	0.212	0.212	0.211
	300	0.005	0.026	0.306	0.306	0.306	0.306
	400	0.001	0.032	0.352	0.352	0.352	0.352
	500	0.007	0.032	0.416	0.416	0.416	0.416
(32.9038,3)	100	0.059	0.096	0.474	0.474	0.474	0.474
Crossing (Exterior 5)	200	0.086	0.208	0.686	0.686	0.686	0.686
	300	0.126	0.370	0.829	0.829	0.829	0.829
	400	0.171	0.484	0.913	0.913	0.913	0.913
	500	0.223	0.595	0.964	0.964	0.964	0.964
(82.0840,5)	100	0.156	0.253	0.694	0.694	0.694	0.693
Crossing (Exterior 5)	200	0.274	0.541	0.911	0.911	0.911	0.911
	300	0.452	0.736	0.985	0.985	0.985	0.985
	400	0.591	0.873	0.998	0.998	0.998	0.998
	500	0.695	0.952	1.000	1.000	1.000	1.000

7 Empirical Examples

7.1 Stocks to Bonds Ratio Adjustment

One of the most popular financial market practices is that as the investment horizon shrinks, investors' portfolio's allocation should shift from primarily equities, to a balanced portfolio, and then to a primarily bond portfolio. Conventional decision rules such as the SD rule or mean-variance criterion are not consistent with this popular investment advice. Bali, Demirtas, Levy, and Wolf (2009) use the ASD rule to find support for the practitioners' view within the expected utility paradigm. However, their approach is not based on formal statistical inference and their definition of ASSD does not satisfy the uniform ordering property.

We investigate whether the return of stocks dominates that of bonds by AFSD or ASSD for different investment horizons using our testing procedure. We use monthly stock index returns (value-weighted returns on the NYSE/AMEX/Nasdaq index) and monthly returns on bonds with a maturity of 30 years from the Center for Research for Securities Prices (CRSP). The sample period is from November 1941 to December 2021. We consider six different investment horizons: 1-Month, 6-Month, 12-Month, 24-Month, 48-Month, 60-Month.

To motivate the reason for adopting the ASD criterion, we first report descriptive statistics for the 1 month, 6 months, 12 months, 24 months, 48 months, and 60 months ahead returns on bond and stock portfolios in Table 15. The mean-variance criterion cannot determine which to invest in since the mean and the standard deviation of stock returns exceed those of bond returns. Note that the geometric mean of stocks is greater than that of bonds and so the distribution of stocks will shift to the right relative to that of bonds as the investment horizon increases. Empirical CDFs and ICDFs also show crossings between stock and bond returns distributions for shorter horizons. Table 16 reports estimated degrees of violation $\widehat{\theta}_1$, $\widehat{\theta}_2$[6] for SD of stocks over bonds, which speaks to the need for testing ASD.

Before applying the ASD tests, we verify that stocks and bonds do not dominate each other by FSD or SSD using the existing test proposed by Lintion, Song, and Whang (2010). We consider

$$H_0 : Stock \succeq_{nSD} Bond \text{ and } H_0 : Bond \succeq_{nSD} Stock \text{ for } n = 1, 2 \tag{7.1}$$

Here, we use stationary bootstrap with optimal block length choice following Politis and White (2004). Table 17 shows results for the FSD and SSD tests. At the significance level of 1%, there is no FSD relation between these two prospects for 1-Month, 6-Month,

[6]See Section 2.1 for definition of the measure of deviation from m-th order SD. Here, we estimate them using (intergrated) empirical distribution functions.

and 12-Month horizons. There is no FSD relation for 1-Month, 6-Month, 12-Month, and 24-Month horizons, and no SSD relation for 1-Month horizon at the significance level of 10%.

We apply our testing method to these cases. We consider

$$H_0 : Stock \succeq_{AnSD} Bond \text{ and } H_0 : Bond \succeq_{AnSD} Stock \text{ for } n = 1, 2 \tag{7.2}$$

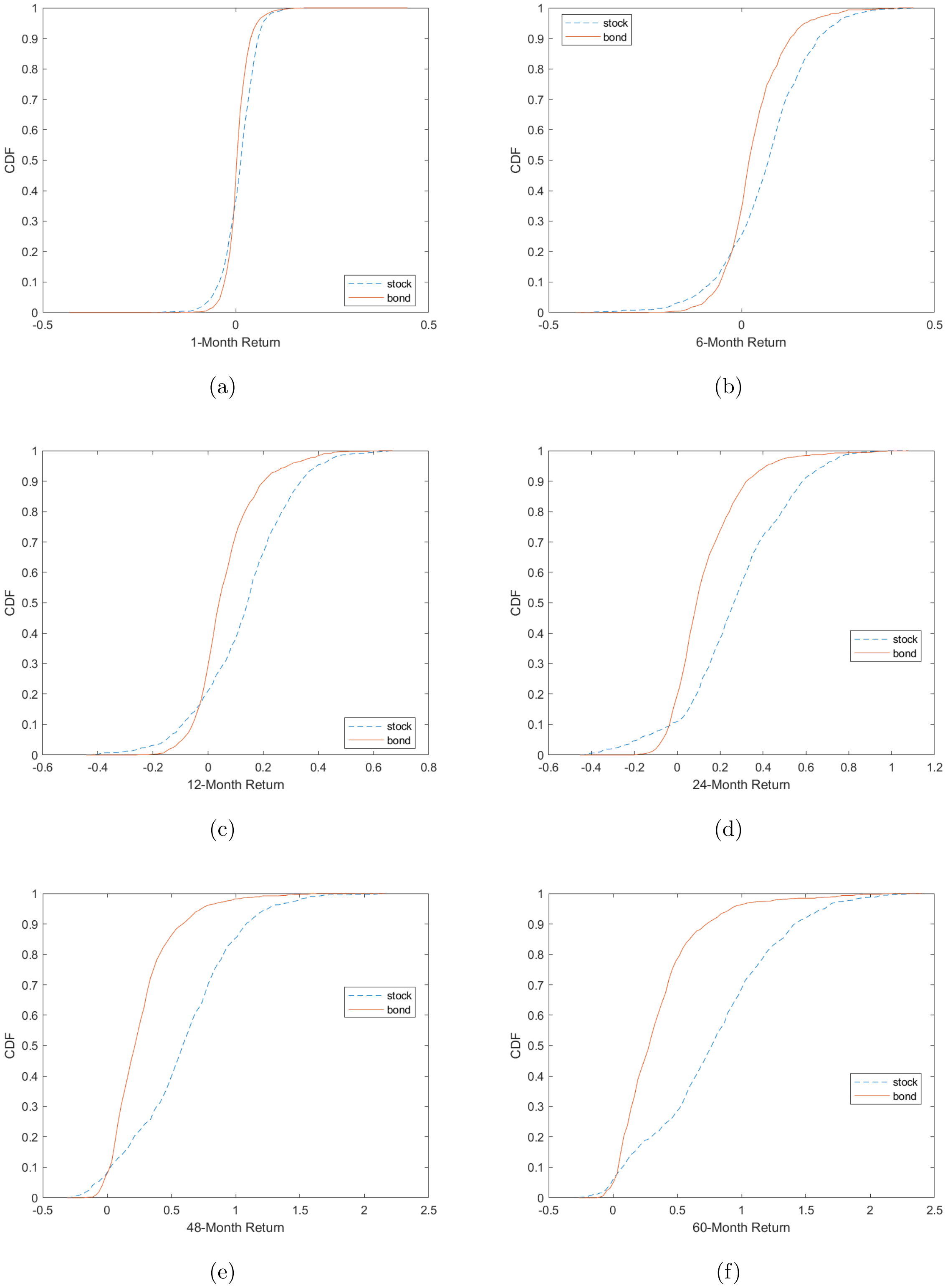
Rather than using the empirical guide suggested by Levy, Leshno, and Leibovitch (2010), the allowed area violation represented by ϵ was selected through experiments varying its values so that we can find one where there is an ASD relation of one prospect over the other. Table 18 and 19 show p-values for the AFSD and ASSD tests with selected ϵ's. We find that stocks dominate bonds by AFSD for 1-Month, 6-Month, and 12-Month horizons and by ASSD for all horizons at the significance level of 1%.

Even if we only accept the allowed area variation less than 0.05, it can be argued that most investors with nonpathological preferences would prefer stocks to bonds when the investment horizon is longer than 6-Month. Furthermore, these different values of epsilon needed to specify AFSD relations provide the rationale for the popular practice of changing portfolio allocation as the investment horizon varies.

Table 15: Descriptive statistics of stocks and bonds

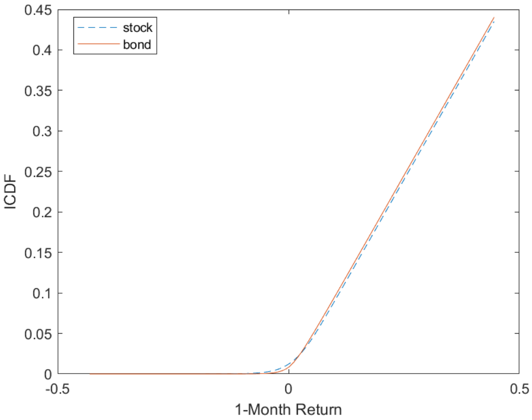
	1-Month	6-Month	12-Month	24-Month	48-Month	60-Month
			Panel A. Bond portfolio			
Arithmetic mean	0.005	0.029	0.059	0.120	0.257	0.335
Geometric mean	0.010	0.031	0.047	0.080	0.163	0.200
Std. dev.	0.027	0.070	0.106	0.160	0.267	0.344
Skewness	0.795	1.036	1.375	1.909	1.721	2.048
Kurtosis	4.243	3.290	3.428	6.148	4.057	5.759
Jarque-Bera	689.357	504.607	639.810	1708.540	895.436	1554.641
			Panel B. Stock portfolio			
Arithmetic mean	0.010	0.062	0.131	0.276	0.612	0.807
Geometric mean	0.023	0.074	0.126	0.232	0.466	0.567
Std. dev.	0.042	0.114	0.170	0.247	0.410	0.505
Skewness	-0.564	-0.223	-0.151	-0.101	0.283	0.221
Kurtosis	1.987	0.547	0.103	0.009	0.122	-0.277
Jarque-Bera	175.420	16.619	3.387	1.338	10.618	8.489

Figure 3: CDFs of returns of stock and bond portfolios

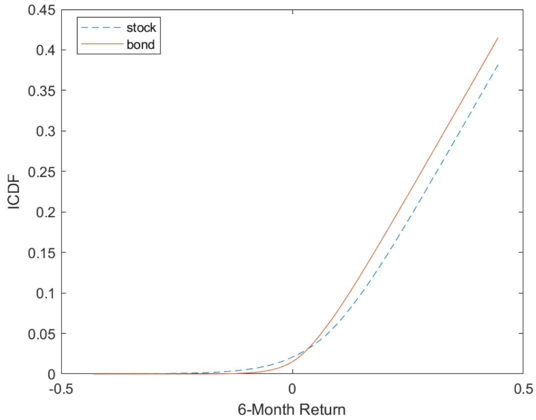


(a) (b) (c) (d) (e) (f)

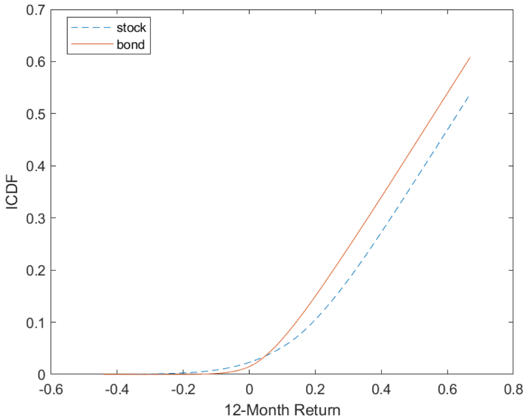
Figure 4: ICDFs of returns of stock and bond portfolios



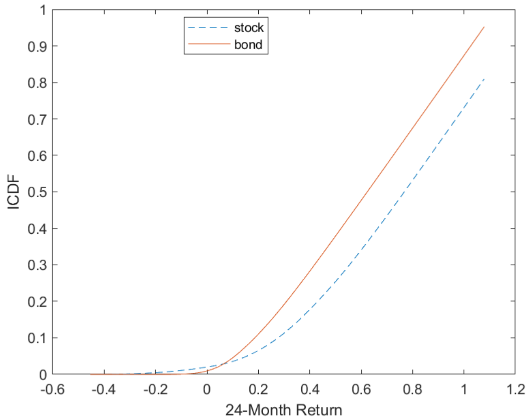
(a)



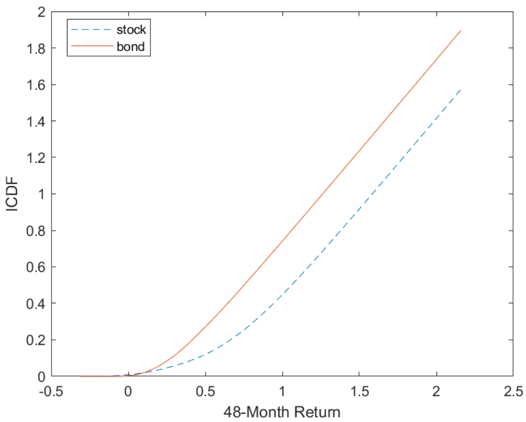
(b)



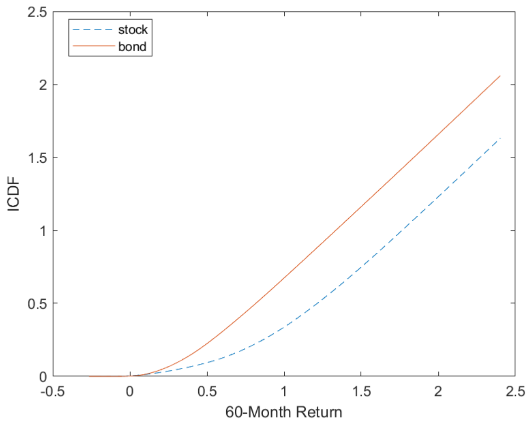
(c)



(d)



(e)



(f)

Table 16: Estimated degree of violation for Stock$\succeq_{SD}$Bond

Horizon	$\widehat{\theta}_1$	$\widehat{\theta}_2$
1-Month	0.304	0.280
6-Month	0.141	0.087
12-Month	0.104	0.048
24-Month	0.075	0.023
48-Month	0.016	0.002
60-Month	0.005	0.001

Table 17: p-values from the FSD and SSD Test

Horizon	Stock$\succeq_{1SD}$Bond	Bond$\succeq_{1SD}$Stock	Stock$\succeq_{2SD}$Bond	Bond$\succeq_{2SD}$Stock
1-Month	0.000	0.000	0.070	0.010
6-Month	0.000	0.000	0.285	0.000
12-Month	0.010	0.000	0.300	0.000
24-Month	0.085	0.000	0.365	0.000
48-Month	0.555	0.000	0.505	0.000
60-Month	0.710	0.000	0.550	0.000

Table 18: p-values from the AFSD Test

Horizon	ϵ	Stock$\succeq_{A1SD}$Bond	Bond$\succeq_{A1SD}$Stock
1-Month	0.220	0.140	0.000
6-Month	0.030	0.105	0.000
12-Month	0.010	0.165	0.000
24-Month	0.001	0.145	0.000
48-Month	0.000	0.815	0.000
60-Month	0.000	0.700	0.000

Table 19: p-values from the ASSD Test

Horizon	ϵ	Stock$\succeq_{A2SD}$Bond	Bond$\succeq_{A2SD}$Stock
1-Month	0.000	0.230	0.005
6-Month	0.000	0.440	0.000
12-Month	0.000	0.335	0.000
24-Month	0.000	0.230	0.000
48-Month	0.000	0.490	0.010
60-Month	0.000	0.575	0.015

8 Conclusion

This paper proposed L_p-type tests for ASD and attained their asymptotic distributions under both the independent and dependent sampling schemes. We proposed bootstrap procedures under both sampling schemes that mimic the null distribution of the approximation to our test statistic using the information of the binding parts of the inequality restrictions. We showed that our bootstrap tests have asymptotically correct size and are asymptotically exact when one of the inequality restrictions binds. We proved test consistency and provided local power asymptotics of our tests which dominate the previous literature. Monte Carlo simulations were conducted to verify finite sample properties of the test and sensitivity of our tests to the selection of tuning parameters. Generally, simulation results are consistent with the theoretical predictions. We applied our testing to verify popular investment decisions regarding stocks to bonds ratio.

Despite its empirical usefulness in wide areas of economics, ASD has been used mainly in financial economics and the applications in this area have been based mainly on numerical simulations. Because many empirical questions involving comparison between distributions or SD can also be potential applications of ASD, the proposed ASD tests in this paper can be utilized in various fields of economics such as welfare economics and policy evaluation. In addition, this paper will serve as a foundation for further tests for generalization of various SD concepts. Since the presence of pathological functions in a class of functions is not just a problem of the original SD, our ASD tests can be extended to almost prospect stochastic dominance (APSD) or almost time stochastic dominance (ATSD) tests as future work, each of which is a generalization of prospect stochastic dominance (PSD) and time stochastic dominance (TSD).

Appendices

Appendix A gives the proofs of Theorem 1, 2, and 3 in the main text. Appendix B contains auxiliary lemmas with their proof and the proofs of Lemma 1 and 2.

A Proofs of Main Theorems

Proof of Theorem 1. In this proof, we proceed as follows. First, we establish the null distribution of our test statistic based on the representation of it given in Lemma 1. Then, we prove the bootstrap consistency result for our bootstrap test statistic. Finally, we prove the validity of our testing procedure. Since the proof under Assumption 1 is straightforward, we present the proof under Assumption 2 and 3.

Null Distribution

We first derive the asymptotic distributions of $S_{T,1}, \ldots, S_{T,m}$ when each inequality restriction of the null hypothesis $H_0^{(m)}$ binds. For the convenience of notation, define temporarily $\delta_T^F(x) := \sqrt{T}\left[F_1^{(m)}(x) - F_2^{(m)}(x)\right]$. When the first inequality restriction of the null hypothesis binds, we have

$$
\begin{aligned}
S_{T,1} = \sqrt{T}\int_{\mathcal{X}} \Bigg\{ & \left(\left[\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right]_+ - \left[F_1^{(m)}(x) - F_2^{(m)}(x)\right]_+\right) \\
& - \epsilon\left(\left|\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right| - \left|F_1^{(m)}(x) - F_2^{(m)}(x)\right|\right)\Bigg\}dx \\
= \sqrt{T}\int_{\mathcal{X}} \Bigg\{ & (1-\epsilon)\left(\left[\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right]_+ - \left[F_1^{(m)}(x) - F_2^{(m)}(x)\right]_+\right) \\
& + \epsilon\left(\left[\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right]_- - \left[F_1^{(m)}(x) - F_2^{(m)}(x)\right]_-\right)\Bigg\}dx \\
= \int_{\mathcal{C}_0} & \left\{\left[\nu_T^{(m)}(x)\right]_+ - \epsilon\left|\nu_T^{(m)}(x)\right|\right\}dx \\
& + (1-\epsilon)\int_{\mathcal{C}_+}\left(\max\left\{\nu_T^{(m)}(x), -\delta_T^F(x)\right\} + \min\left\{\nu_T^{(m)}(x) + \delta_T^F(x), 0\right\}\right)dx \\
& + \epsilon\int_{\mathcal{C}_-}\left(\max\left\{\nu_T^{(m)}(x) + \delta_T^F(x), 0\right\} + \min\left\{\nu_T^{(m)}(x), -\delta_T^F(x)\right\}\right)dx \\
= \int_{\mathcal{C}_0} & \left\{\left[\nu_T^{(m)}(x)\right]_+ - \epsilon\left|\nu_T^{(m)}(x)\right|\right\}dx + (1-\epsilon)\int_{\mathcal{C}_+}\left(\nu_T^{(m)}(x) + o_p(1)\right)dx + \epsilon\int_{\mathcal{C}_-}\left(\nu_T^{(m)}(x) + o_p(1)\right)dx \\
= \int_{\mathcal{C}_0} & \left\{\left[\nu_T^{(m)}(x)\right]_+ - \epsilon\left|\nu_T^{(m)}(x)\right|\right\}dx + (1-\epsilon)\int_{\mathcal{C}_+}\nu_T^{(m)}(x)dx + \epsilon\int_{\mathcal{C}_-}\nu_T^{(m)}(x)dx + o_p(1),
\end{aligned}
$$

since $\sup_{x\in\mathcal{X}}\left|\nu_T^{(m)}(x)\right| = O_p(1)$ by Lemma B.1 (1). Thus, $S_{T,1} \Rightarrow S_{0,1}$ when $d_{m,1} = 0$. For

$2 \leq j \leq m$, Lemma B.1 (1) implies that when the j-th inequality binds,

$$S_{T,j} = \nu_T^{(j)}(\overline{x}) \Rightarrow \nu_{1,2}^{(j)}(\overline{x}) = S_{0,j}.$$

By Theorem 1.4.8 of van der Vaart and Wellner (1996), we have joint convergence since the asymptotic processes of $S_{T,j}$'s are separable.

Under the null hypothesis $H_0^{(m)}$, Lemma 1 implies, with probability approaching 1,

$$\begin{aligned}
S_T &= \Lambda_p\left(\psi_1\left(\sqrt{T}d_{m,1}\right) \cdot S_{T,1}, \ldots, \psi_m\left(\sqrt{T}d_{m,m}\right) \cdot S_{T,m}\right) \\
&= \Lambda_p\left(1\left(d_{m,1} \geq 0\right) \cdot S_{T,1}, \ldots, 1\left(d_{m,m} \geq 0\right) \cdot S_{T,m}\right) \\
&= \Lambda_p\left(1\left(d_{m,1} = 0\right) \cdot \left(S_{T,1} - \sqrt{T}d_{m,1}\right), \ldots, 1\left(d_{m,m} = 0\right) \cdot \left(S_{T,m} - \sqrt{T}d_{m,m}\right)\right) \\
&= \Lambda_p\left(1\left(d_{m,1} \geq 0\right) \cdot \left(S_{T,1} - \sqrt{T}d_{m,1}\right), \ldots, 1\left(d_{m,m} \geq 0\right) \cdot \left(S_{T,m} - \sqrt{T}d_{m,m}\right)\right) \\
&= \Lambda_p\left(\psi_1\left(\sqrt{T}d_{m,1}\right) \cdot \left(S_{T,1} - \sqrt{T}d_{m,1}\right), \ldots, \psi_m\left(\sqrt{T}d_{m,m}\right) \cdot \left(S_{T,m} - \sqrt{T}d_{m,m}\right)\right),
\end{aligned} \tag{A.1}$$

where the second and last equalities hold since $\frac{\kappa_{T,j}}{\sqrt{T}} \to 0$ and the third and fourth equalities hold due to the restriction of the null hypothesis. Thus, the null distribution of the test statistic S_T can be approximated with the distribution of (A.1), which implies, by the continuous mapping theorem,

$$S_T \Rightarrow \Lambda_p\left(1\left(d_{m,1} = 0\right) \cdot S_{0,1}, \ldots, 1\left(d_{m,m} = 0\right) \cdot S_{0,m}\right). \tag{A.2}$$

Bootstrap Consistency

The distribution of our bootstrap test statistic mimics the null distribution of (A.1), which is given in (A.2). Here, we establish its bootstrap consistency. To this end, (i) we first introduce assumptions such that for a fixed (nonrandom) sequence $\mathcal{S}$ the desired asymptotic results hold. Then, (ii) we show that these assumptions hold with probability approaching 1. Lastly, (iii) we prove that these results for fixed covariates hold.

(i) We provide assumptions under which S_T^* has the desired asymptotic null distribution for the case of fixed covariates. Assume first that $\mathcal{S}$ is fixed, i.e., nonrandom.

Assumption A.1.

(a) $\int_{\mathcal{X}}\left\{\left[\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right]_+ - \epsilon\left|\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right|\right\}dx$
$\to \int_{\mathcal{X}}\left\{\left[F_1^{(m)}(x) - F_2^{(m)}(x)\right]_+ - \epsilon\left|F_1^{(m)}(x) - F_2^{(m)}(x)\right|\right\}dx.$

(b) $\bar{F}_1^{(j)}(\overline{x}) - \bar{F}_2^{(j)}(\overline{x}) \to F_1^{(j)}(\overline{x}) - F_2^{(j)}(\overline{x})$ for $2 \leq j \leq m$.

(c) $Q\left(\widehat{\mathcal{C}}_0(\widehat{c}_T)\right) \to Q\left(\mathcal{C}_0\right)$, $Q\left(\widehat{\mathcal{C}}_+(\widehat{c}_T)\right) \to Q\left(\mathcal{C}_+\right)$, and $Q\left(\widehat{\mathcal{C}}_-(\widehat{c}_T)\right) \to Q\left(\mathcal{C}_-\right)$.

(ii) To prove the bootstrap consistency of the proposed test, we show that Assumption A.1 holds with probability approaching 1. First, Assumption A.1 (a) holds with probability approaching 1 by Lemma B.1 (i) because the continuous mapping theorem applied to the weak convergence of $\nu_T^{(m)}(\cdot)$ to $\nu_{1,2}^{(m)}$ implies $S_{T,1} - \sqrt{T} d_{m,1} \Rightarrow S_{0,1}$ and so

$$\begin{aligned}
&\int_{\mathcal{X}} \left\{ \left[\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x) \right]_+ - \epsilon \left| \bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x) \right| \right\} dx \\
&\xrightarrow{p} \int_{\mathcal{X}} \left\{ \left[F_1^{(m)}(x) - F_2^{(m)}(x) \right]_+ - \epsilon \left| F_1^{(m)}(x) - F_2^{(m)}(x) \right| \right\} dx.
\end{aligned}$$

Likewise, Lemma B.1 (i) implies the weak convergence result $\nu_T^{(j)}(\overline{x}) \Rightarrow \nu_{1,2}^{(j)}$ and so

$$\bar{F}_1^{(j)}(\overline{x}) - \bar{F}_2^{(j)}(\overline{x}) \xrightarrow{p} F_1^{(j)}(\overline{x}) - F_2^{(j)}(\overline{x})$$

for $2 \leq j \leq m$. Finally, Lemma 2 implies that estimated enlargements of contact sets coincide with the original contact sets with probability approaching 1. Thus, we have

$$\begin{aligned}
Q\left(\widehat{\mathcal{C}}_0(\widehat{c}_T)\right) &\xrightarrow{p} Q\left(\mathcal{C}_0\right) \\
Q\left(\widehat{\mathcal{C}}_+(\widehat{c}_T)\right) &\xrightarrow{p} Q\left(\mathcal{C}_+\right) \\
Q\left(\widehat{\mathcal{C}}_-(\widehat{c}_T)\right) &\xrightarrow{p} Q\left(\mathcal{C}_-\right).
\end{aligned}$$

(iii) It only remains to show that

$$S_T^* \Rightarrow \Lambda_p \left(1\left(d_{m,1} = 0\right) \cdot S_{0,1}, \ldots, 1\left(d_{m,m} = 0\right) \cdot S_{0,m}\right), \tag{A.3}$$

conditional on $\mathcal{S}$, which is nonrandom. By Lemma B.1 (ii), $\nu_T^{(j)*} \Rightarrow \nu_{1,2}^{(j)}$ for $1 \leq j \leq m$ conditional on $\mathcal{S}$. This implies that $S_{T,j}^* \Rightarrow S_{0,j}$ for $2 \leq j \leq m$ conditional on $\mathcal{S}$. For $S_{T,1}^*$, we have

$$\begin{aligned}
S_{T,1}^* &= \int_{\widehat{\mathcal{C}}_0(\widehat{c}_T)} \left\{ (1-\epsilon) \left[\nu_T^{(m)*}(x) \right]_+ + \epsilon \left[\nu_T^{(m)*}(x) \right]_- \right\} dx \\
&\quad + (1-\epsilon) \int_{\widehat{\mathcal{C}}_+(\widehat{c}_T)} \nu_T^{(m)*}(x) dx + \epsilon \int_{\widehat{\mathcal{C}}_-(\widehat{c}_T)} \nu_T^{(m)*}(x) dx \\
&= \int_{\mathcal{C}_0} \left\{ (1-\epsilon) \left[\nu_T^{(m)*}(x) \right]_+ + \epsilon \left[\nu_T^{(m)*}(x) \right]_- \right\} dx \\
&\quad + (1-\epsilon) \int_{\mathcal{C}_+} \nu_T^{(m)*}(x) dx + \epsilon \int_{\mathcal{C}_-} \nu_T^{(m)*}(x) dx + o(1) \\
&\Rightarrow S_{0,1},
\end{aligned}$$

conditional on $\mathcal{S}$, where the second equality holds by Assumption A.1 (c). In addition, Assumption A.1 (a) and (b) imply that $\frac{S_{T,j}}{\sqrt{T}} \to d_{m,j}$ and so $\psi_j\left(S_{T,j}\right) \to 1\left(d_{m,j} \geq 0\right) =$

$1\left(d_{m,j}=0\right)$ for $1 \leq j \leq m$, where the equality holds under the null hypothesis. By Theorem 1.4.8 of van der Vaart and Wellner (1996), we have joint convergence of $\psi_1\left(S_{T,1}\right) \cdot S^*_{T,1}, \ldots, \psi_m\left(S_{T,m}\right) \cdot S^*_{T,m}$ conditional on $\mathcal{S}$. The continuous mapping theorem yields (A.3) conditional on $\mathcal{S}$. Thus, we have the desired result.

Asymptotic Size Control

Let $c_{0,\alpha}$ denote the $(1-\alpha)$-th quantile of the null distribution of S_T. In addition, let $\sigma_T := \mathbf{Var}\left(S^*_T\right)$ denote the variance of S^*_T. Then, the bootstrap consistency result implies

$$c^*_{T,\alpha} \xrightarrow{p} c_{0,\alpha}. \tag{A.4}$$

when $\lim_{T_1,T_2\to\infty} \sigma_T > 0$. There exists a subsequence $\{w_T := (w_{T_1}, w_{T_2})\}_{T_1,T_2\geq 1} \subset \{(T_1,T_2)\}_{T_1,T_2\geq 1}$ such that

$$\limsup_{T_1,T_2\to\infty} P\{S_T > c^*_{T,\alpha,\eta}\} = \lim_{T_1,T_2\to\infty} P\{S_{w_T} > c^*_{w_T,\alpha,\eta}\}, \tag{A.5}$$

where S_{w_T} and $c^*_{w_T,\alpha,\eta}$ are the same as S_T and $c^*_{T,\alpha,\eta}$, except that the sample size T is replaced by w_T. By Assumption 2 (b), $\{\sigma_T\}_{T_1,T_2\geq 1}$ is a bounded sequence. Therefore, there exists a further subsequence $\{u_T := (u_{T_1}, u_{T_2})\}_{T_1,T_2\geq 1} \subset \{w_T\}_{T_1,T_2\geq 1}$ such that σ_{u_T} converges.

Consider the case $\lim_{T_1,T_2\to\infty} \sigma_{u_T} > 0$. Then, there exists a binding inequality composing the null hypothesis. Thus, we have

$$P(S_{u_T} > c^*_{u_T,\alpha,\eta}) = P(S_{u_T} > c^*_{u_T,\alpha}) = P(S_{u_T} + o(1) > c_{0,\alpha}) \to \alpha. \tag{A.6}$$

Now, consider the other case $\lim_{T_1,T_2\to\infty} \sigma_{u_T} = 0$. Then, there are no binding inequalities, i.e., all inequalities comprising the null hypothesis hold strictly. Thus, we have

$$P(S_{u_T} > c^*_{u_T,\alpha,\eta}) = P(S_{u_T} > \eta) \to 0. \tag{A.7}$$

Thus, we complete the proof by combining (A.6) and (A.7). $\square$

Proof of Theorem 2. Note that the map Λ_p is a convex function on $\mathbb{R}^m$. By Jensen's inequality, we have

$$\Lambda_p\left(\frac{b}{2}\right) \leq \frac{\Lambda_p\left(a+b\right) + \Lambda_p\left(-a\right)}{2}$$

for $a, b \in \mathbb{R}^m$. Then, we have

$$\begin{aligned} S_T &= \Lambda_p\left(\left(S_{T,1} - \sqrt{T}d_{m,1}\right) + \sqrt{T}d_{m,1}, \ldots, \left(S_{T,m} - \sqrt{T}d_{m,m}\right) + \sqrt{T}d_{m,m}\right) \\ &\geq \frac{1}{2^{p-1}}\Lambda_p\left(\sqrt{T}d_{m,1}, \ldots, \sqrt{T}d_{m,m}\right) \\ &\quad - \Lambda_p\left(-\left(S_{T,1} - \sqrt{T}d_{m,1}\right), \ldots, -\left(S_{T,m} - \sqrt{T}d_{m,m}\right)\right). \end{aligned} \tag{A.8}$$

By Lemma B.1 (i) and the proof of Lemma 1, (A.8) is $O_p(1)$. Since $T \to \infty$ as $T_1, T_2 \to \infty$ and $\Lambda_p\left(\sqrt{T}d_{m,1}, \dots, \sqrt{T}d_{m,m}\right) > 0$ under the alternative hypothesis $H_1^{(m)}$, we have for any constant $M_1 > 0$,

$$P\left\{\frac{1}{2^{p-1}}\Lambda_p\left(\sqrt{T}d_{m,1}, \dots, \sqrt{T}d_{m,m}\right) > M_1\right\} \to 1.$$

Therefore, this implies that for any constant $M_2 > 0$,

$$P\left\{S_T > M_2\right\} \to 1.$$

The result of Theorem 2 now holds because $c^*_{T,\alpha,\eta} = O_p(1)$ by the bootstrap consistency result in the proof of Theorem 1. □

Proof of Theorem 3. 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 of the first inequality restriction of the null hypothesis. Since the bootstrap test statistics $S^{GLW*}_{T,1}, S^*_{T,1}$ are recentered, the former converges in distribution conditional on $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2$ to the distribution of $S^{GLW}_{a,1} := \int_{\mathcal{X}}\left[\nu^{(m)}_{1,2}(x)\right]_+ dx$ and the latter converges to that of
$S_{a,1} := \int_{\mathcal{C}^0_a}\left\{(1-\epsilon)\left[\nu^{(m)}_{1,2}(x)\right]_+ + \epsilon\left[\nu^{(m)}_{1,2}(x)\right]_-\right\}dx + (1-\epsilon)\int_{\mathcal{C}^+_a}\nu^{(m)}_{1,2}(x)dx + \epsilon\int_{\mathcal{C}^-_a}\nu^{(m)}_{1,2}(x)dx$.
Since

$$\begin{aligned}
S^{GLW}_{a,1} - S_{a,1} = &\int_{\mathcal{C}^0_a}\left\{\epsilon\left[\nu^{(m)}_{1,2}(x)\right]_+ - \epsilon\left[\nu^{(m)}_{1,2}(x)\right]_-\right\}dx \\
&+ \int_{\mathcal{C}^+_a}\left\{\epsilon\left[\nu^{(m)}_{1,2}(x)\right]_+ - (1-\epsilon)\left[\nu^{(m)}_{1,2}(x)\right]_-\right\}dx \\
&+ \int_{\mathcal{C}^-_a}\left\{(1-\epsilon)\left[\nu^{(m)}_{1,2}(x)\right]_+ - \epsilon\left[\nu^{(m)}_{1,2}(x)\right]_-\right\}dx \geq 0
\end{aligned}$$

and $Q(\mathcal{C}^0_a \cup \mathcal{C}^+_a) > 0$ by Assumption 5 (a), we have

$$\begin{aligned}
\int_{\mathcal{X}}\left[\nu^{(m)}_{1,2}(x)\right]_+ dx > &\int_{\mathcal{C}^0_a}\left\{(1-\epsilon)\left[\nu^{(m)}_{1,2}(x)\right]_+ + \epsilon\left[\nu^{(m)}_{1,2}(x)\right]_-\right\}dx \\
&+ (1-\epsilon)\int_{\mathcal{C}^+_a}\nu^{(m)}_{1,2}(x)dx + \epsilon\int_{\mathcal{C}^-_a}\nu^{(m)}_{1,2}(x)dx.
\end{aligned}$$

almost everywhere. Hence, $c^{GLW*}_{T,1,\alpha} > c^*_{T,1,\alpha} + o_p(1)$.

Under the local alternatives $H^{(m)}_{a,1}$, we have

$$\lim_{T_1,T_2\to\infty} P_T\{S_{T,1} > c^*_{T,1,\alpha}\}$$

$$
\begin{aligned}
&= \lim_{T_1,T_2\to\infty} P_T\Bigg\{ \int_{\mathcal{C}_a^0} \Big\{(1-\epsilon)\left[\nu_T^{(m)}(x)+\delta(x)\right]_+ + \epsilon\left[\nu_T^{(m)}(x)+\delta(x)\right]_-\Big\}\,dx \\
&\qquad\qquad + (1-\epsilon)\int_{\mathcal{C}_a^+}\left\{\nu_T^{(m)}(x)+\delta(x)\right\}dx + \epsilon\int_{\mathcal{C}_a^-}\left\{\nu_T^{(m)}(x)+\delta(x)\right\}dx > c^*_{T,1,\alpha}\Bigg\} \\
&= P\Bigg\{ \int_{\mathcal{C}_a^0} \Big\{(1-\epsilon)\left[\nu_{1,2}^{(m)}(x)+\delta(x)\right]_+ + \epsilon\left[\nu_{1,2}^{(m)}(x)+\delta(x)\right]_-\Big\}\,dx \\
&\qquad + (1-\epsilon)\int_{\mathcal{C}_a^+}\left\{\nu_{1,2}^{(m)}(x)+\delta(x)\right\}dx + \epsilon\int_{\mathcal{C}_a^-}\left\{\nu_{1,2}^{(m)}(x)+\delta(x)\right\}dx > c^*_{T,1,\alpha}\Bigg\} \\
&> P\Bigg\{ \int_{\mathcal{C}_a^0} \Big\{(1-\epsilon)\left[\nu_{1,2}^{(m)}(x)+\delta(x)\right]_+ + \epsilon\left[\nu_{1,2}^{(m)}(x)+\delta(x)\right]_-\Big\}\,dx \\
&\qquad + (1-\epsilon)\int_{\mathcal{C}_a^+}\left\{\nu_{1,2}^{(m)}(x)+\delta(x)\right\}dx + \epsilon\int_{\mathcal{C}_a^-}\left\{\nu_{1,2}^{(m)}(x)+\delta(x)\right\}dx > c^{GLW*}_{T,1,\alpha}\Bigg\}.
\end{aligned}
\tag{A.9}
$$

Thus, we have the result in (1) of Theorem 3.

Likewise, S_T^{GLW*} and S_T^* converge in distribution conditional on $\mathcal{S}$ to the distribution of $\Lambda_p\big(1\,(d_{m,1}\geq 0)\cdot S_{a,1}^{GLW}, 1\,(d_{m,2}\geq 0)\cdot S_{0,2},\ldots,1\,(d_{m,m}\geq 0)\cdot S_{0,m}\big)$ and $\Lambda_p\big(1\,(d_{m,1}\geq 0)\cdot S_{a,1}, 1\,(d_{m,2}\geq 0)\cdot S_{0,2},\ldots,1\,(d_{m,m}\geq 0)\cdot S_{0,m}\big)$, respectively. Since $S_{a,1}^{GLW} > S_{a,1} > 0$ almost everywhere under the local alternatives,

$$
\begin{aligned}
&\Lambda_p\left(1\,(d_{m,1}\geq 0)\cdot S_{a,1}^{GLW}, 1\,(d_{m,2}\geq 0)\cdot S_{0,2},\ldots,1\,(d_{m,m}\geq 0)\cdot S_{0,m}\right) \\
&\quad > \Lambda_p\left(1\,(d_{m,1}\geq 0)\cdot S_{a,1}, 1\,(d_{m,2}\geq 0)\cdot S_{0,2},\ldots,1\,(d_{m,m}\geq 0)\cdot S_{0,m}\right)
\end{aligned}
$$

almost everywhere and so $c^{GLW*}_{T,\alpha,\eta} > c^*_{T,\alpha,\eta} + o_p(1)$ under the local alternatives.

Then, under the local alternatives $H_a^{(m)}$, we also have

$$
\begin{aligned}
&\lim_{T_1,T_2\to\infty} P_T\{S_T > c^*_{T,\alpha,\eta}\} \\
&\ = \lim_{T_1,T_2\to\infty} P_T\Big\{\Lambda_p\left(S_{T,1},S_{T,2},\ldots,S_{T,m}\right) > c^*_{T,\alpha,\eta}\Big\} \\
&\ = \lim_{T_1,T_2\to\infty} P_T\Big\{\Lambda_p\left(S_{T,1},\nu_T^{(2)}(\overline{x})+\sqrt{T}\mu_2(\overline{x}),\ldots,\nu_T^{(m)}(\overline{x})+\sqrt{T}\mu_m(\overline{x})\right) > c^*_{T,\alpha,\eta}\Big\} \\
&\ = P\Bigg\{\Lambda_p\Bigg(\int_{\mathcal{C}_a^0} \Big\{(1-\epsilon)\left[\nu_{1,2}^{(m)}(x)+\delta(x)\right]_+ + \epsilon\left[\nu_{1,2}^{(m)}(x)+\delta(x)\right]_-\Big\}\,dx \\
&\qquad\qquad + (1-\epsilon)\int_{\mathcal{C}_a^+}\left\{\nu_{1,2}^{(m)}(x)+\delta(x)\right\}dx + \epsilon\int_{\mathcal{C}_a^-}\left\{\nu_{1,2}^{(m)}(x)+\delta(x)\right\}dx, \\
&\qquad\qquad 1\,(\mu_2(\overline{x})=0)\cdot\nu_{1,2}^{(2)}(\overline{x}),\ldots,1\,(\mu_m(\overline{x})=0)\cdot\nu_{1,2}^{(m)}(\overline{x})\Bigg) > c^*_{T,\alpha,\eta}\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
> P\Bigg\{ \Lambda_p \Bigg(& \int_{\mathcal{C}_a^0} \left\{ (1-\epsilon) \left[\nu_{1,2}^{(m)}(x) + \delta(x) \right]_+ + \epsilon \left[\nu_{1,2}^{(m)}(x) + \delta(x) \right]_- \right\} dx \\
& + (1-\epsilon) \int_{\mathcal{C}_a^+} \left\{ \nu_{1,2}^{(m)}(x) + \delta(x) \right\} dx + \epsilon \int_{\mathcal{C}_a^-} \left\{ \nu_{1,2}^{(m)}(x) + \delta(x) \right\} dx, \\
& 1\left(\mu_2(\overline{x}) = 0\right) \cdot \nu_{1,2}^{(2)}(\overline{x}), \ldots, 1\left(\mu_m(\overline{x}) = 0\right) \cdot \nu_{1,2}^{(m)}(\overline{x}) \Bigg) > c_{T,\alpha,\eta}^{GLW*} \Bigg\}.
\end{aligned} \tag{A.10}
$$

Thus, we obtain the desired result in (2) of Theorem 3.

□

B Auxiliary Lemmas and Proofs of Lemmas

Lemma B.1. *Suppose that Assumption 1 or 2 holds.*

(i) Then, for all $m \in \mathbb{Z}^+$, we have

$$\nu_T^{(m)}(\cdot) \Rightarrow \nu_{1,2}^{(m)}(\cdot) \text{ in } l^\infty(\mathcal{X}).$$

(ii) Suppose further that Assumption 3 holds. Then, for all $m \in \mathbb{Z}^+$, we have

$$\nu_T^{(m)*}(\cdot) \Rightarrow \nu_{1,2}^{(m)}(\cdot) \text{ in } l^\infty(\mathcal{X})$$

conditional on $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2$ in probability.

Proof of Lemma B.1. We prove this lemma under Assumption 2. The proof of the lemma under Assumption 1 is straightforward.

(1) Since the total boundedness of pseudometric space $(\mathcal{X}, \rho)$ is clear from boundedness of $\mathcal{X}$, we only need to verify (a) finite dimensional convergence and (b) the stochastic equicontinuity result: that is, for each $\epsilon > 0$, there exists $\delta > 0$ such that

$$\overline{\lim_{T\to\infty}} \left\| \sup_{\rho(x_1,x_2)<\delta} \left| \nu_T^{(m)}(x_1) - \nu_T^{(m)}(x_2) \right| \right\|_q < \epsilon, \tag{B.1}$$

where the pseudo-metric on $\mathcal{X}$ is given by

$$
\begin{aligned}
\rho(x_1, x_2) = \Big\{ & E[((x_1 - X_{1,t})^{m-1} 1(X_{1,t} \le x_1) - (x_1 - X_{2,t})^{m-1} 1(X_{2,t} \le x_1)) \\
& - ((x_2 - X_{1,t})^{m-1} 1(X_{1,t} \le x_2) - (x_2 - X_{2,t})^{m-1} 1(X_{2,t} \le x_2))]^2 \Big\}^{1/2}.
\end{aligned}
$$

The finite dimensional convergence result holds by the Cramer-Wold device and a CLT for bounded random variables of Corollary 5.1. of Hall and Heyde (1980) since $\{(X_{1,t}, X_{2,t})^T : t = 1, \ldots, T\}$ is strictly stationary and α-mixing with $\sum_{m=1}^\infty \alpha(m) < \infty$ by Assumption 2 (a).

To show the stochastic equicontinuity condition, let

$$\mathcal{F} = \{f_t(x) : x \in \mathcal{X}\},$$

where

$$f_t(x) = (x - X_{1,t})^{m-1}1(X_{1,t} \le x) - (x - X_{2,t})^{m-1}1(X_{2,t} \le x).$$

Then, $\mathcal{F}$ is a class of uniformly bounded functions that satisfy the L^2-continuity condition since, for some $C_1, C_2 > 0$,

$$\begin{aligned}
&E \sup_{\substack{x_1 \in \mathcal{X} \\ |x_1 - x| \le r}} [f_t(x_1) - f_t(x)]^2 \le 2E \sup_{\substack{x_1 \in \mathcal{X} \\ |x_1 - x| \le r}} [((x_1 - X_{1,t})^{m-1}1(X_{1,t} \le x_1) - (x - X_{1,t})^{m-1}1(X_{1,t} \le x))^2 \\
&\qquad\qquad + ((x_1 - X_{2,t})^{m-1}1(X_{2,t} \le x_1) - (x - X_{2,t})^{m-1}1(X_{2,t} \le x))^2] \\
&\le 4E \sup_{\substack{x_1 \in \mathcal{X} \\ |x_1 - x| \le r}} [((x_1 - X_{1,t})^{m-1} - (x - X_{1,t})^{m-1})^2 + (1(X_{1,t} \le x_1) - 1(X_{1,t} \le x))^2 \\
&\qquad + ((x_1 - X_{2,t})^{m-1} - (x - X_{2,t})^{m-1})^2 + (1(X_{2,t} \le x_1) - 1(X_{2,t} \le x))^2] \\
&\le 4E \sup_{\substack{x_1 \in \mathcal{X} \\ |x_1 - x| \le r}} [C_1|x_1 - x| + 1(x < X_{1,t} \le x_1) + 1(x < X_{2,t} \le x_1)] \\
&\le 4E[C_1 r + 1(|X_{1,t} - x| \le r) + 1(|X_{2,t} - x| \le r)] \le C_2 r,
\end{aligned}$$

where the first two inequalities hold by $(a + b)^2 \le 2(a^2 + b^2)$, the third inequality holds by $a^n - b^n \le (a - b)\sum_{k=0}^{n-1} a^{n-k-1}b^k$ and boundedness of random variables, and the last inequality holds by absolute continuity with respect to Lebesgue measure of Assumption 2 (b). Then, the bracketing number satisfies $N(\epsilon, \mathcal{F}) \le C \cdot (1/\epsilon)^2$ for some $C > 0$ and so $\int_0^1 \epsilon^{-1/2} N(\epsilon, \mathcal{F})^{1/q} dx < \infty$. Furthermore, Assumption 2 (a) implies that $\sum_{m=1}^{\infty} m^{q-2}\alpha(m)^{2/(q+2)} = \sum_{m=1}^{\infty} O(m^{q-2-A \cdot 2/(q+2)}) < \infty$. Thus, the stochastic equicontinuity condition holds by Theorem 2.2. of Andrews and Pollard (1994) with $Q = q$ and $\gamma = 2$. This yields Lemma B.1. (i). □

(2) Lemma B.1 (ii) follows from weak convergence results for Hilbert space valued random variables. Specifically, we can apply Theorem 3.1 of Politis and Romano (1994). First, $\{Z_t \equiv (\cdot - X_{1,t})^{m-1}1(X_{1,t} \le \cdot) - (\cdot - X_{2,t})^{m-1}1(X_{2,t} \le \cdot) : t = 1, \ldots, T\}$ is a stationary sequence of Hilbert space valued random variables which are bounded and satisfy the mixing condition $\sum_j \alpha_Z(j) < \infty$ by Assumption 2 (a). Second, Assumption 3 satisfies the condition related to the stationary resampling scheme. Thus, we have the desired result by applying the bootstrap central limit theorem.

Proof of Lemma 1. It suffices to show that for $1 \le j \le m$,

$$P\left\{[S_{T,j}]_+ = \left[\psi_j\left(\sqrt{T}d_{m,j}\right) \cdot S_{T,j}\right]_+\right\} \to 1 \tag{B.2}$$

since we consider a function Λ_p of the form (2.2) or (2.3) which satisfies $\Lambda_p(a_1, \ldots, a_j, \ldots, a_m) = \Lambda_p(a_1, \ldots, [a_j]_+, \ldots, a_m)$ for $a_j \in \mathbb{R}$, $1 \le j \le m$. Note that

$$\begin{aligned} S_{T,j} &= \psi_j\left(\sqrt{T} d_{m,j}\right) \cdot S_{T,j} + \left(1 - \psi_j\left(\sqrt{T} d_{m,j}\right)\right) \cdot S_{T,j} \\ &= \begin{cases} S_{T,j}, & \text{if } \sqrt{T} d_{m,j} \ge -\kappa_{T,j} \\ S_{T,j}, & \text{otherwise} \end{cases} \end{aligned}$$

and

$$\psi_j\left(\sqrt{T} d_{m,j}\right) \cdot S_{T,j} = \begin{cases} S_{T,j}, & \text{if } \sqrt{T} d_{m,j} \ge -\kappa_{T,j} \\ 0, & \text{otherwise.} \end{cases}$$

Suppose $\psi_j\left(\sqrt{T} d_{m,j}\right) = 0$, i.e., $\sqrt{T} d_{m,j} < -\kappa_{T,j}$. Then, we have

$$\begin{aligned} [S_{T,j}]_+ &= \max\left\{S_{T,j} - \sqrt{T} d_{m,j} + \sqrt{T} d_{m,j}, 0\right\} \\ &\le \max\left\{S_{T,j} - \sqrt{T} d_{m,j} - \kappa_{T,j}, 0\right\} \\ &\le \max\left\{\left|S_{T,j} - \sqrt{T} d_{m,j}\right| - \kappa_{T,j}, 0\right\}. \end{aligned} \tag{B.3}$$

Since $\left|S_{T,j} - \sqrt{T} d_{m,j}\right| = O_p(1)$ by Lemma B.1 (i) and $\kappa_{T,j}$ goes to infinity, the upper bound of $[S_{T,j}]_+$ is $o_p(1)$ when $1 - \psi_j\left(\sqrt{T} d_{m,j}\right) = 1$. Thus, we obtain

$$\begin{aligned} [S_{T,j}]_+ &= \psi_j\left(\sqrt{T} d_{m,j}\right) \cdot [S_{T,j}]_+ + \left(1 - \psi_j\left(\sqrt{T} d_{m,j}\right)\right) \cdot [S_{T,j}]_+ \\ &= \left[\psi_j\left(\sqrt{T} d_{m,j}\right) \cdot S_{T,j}\right]_+ + 1\left(\sqrt{T} d_{m,j} < -\kappa_{T,j}\right)[S_{T,j}]_+ \\ &= \left[\psi_j\left(\sqrt{T} d_{m,j}\right) \cdot S_{T,j}\right]_+, \end{aligned}$$

with probability approaching 1.

It only remains to show $\left|S_{T,j} - \sqrt{T} d_{m,j}\right| = O_p(1)$ for $1 \le j \le m$. Since $S_{T,1}$ takes a different from the other $S_{T,j}$'s, we consider two cases: $j = 1$ and $2 \le j \le m$. When $j = 1$, we have

$$\begin{aligned} \left|S_{T,1} - \sqrt{T} d_{m,1}\right| = &\left| \int_{\mathcal{X}} \sqrt{T}\left\{\left[\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right]_+ - \left[F_1^{(m)}(x) - F_2^{(m)}(x)\right]_+\right\} dx \right. \\ &\left. - \epsilon \int_{\mathcal{X}} \sqrt{T}\left\{\left|\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right| - \left|F_1^{(m)}(x) - F_2^{(m)}(x)\right|\right\} dx \right| \\ \le &\int_{\mathcal{X}} \left|\sqrt{T}\left[\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) - \left(F_1^{(m)}(x) - F_2^{(m)}(x)\right)\right]_+\right| \end{aligned}$$

$$
\begin{aligned}
&+\epsilon\int_{\mathcal{X}}\left|\sqrt{T}\left|\left(\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right)-\left(F_1^{(m)}(x)-F_2^{(m)}(x)\right)\right|\right| \\
&\leq (1+\epsilon)\cdot Q(\mathcal{X})\cdot\sup_{x\in\mathcal{X}}\left|\nu_T^{(m)}(x)\right| = O_p(1), \qquad \text{(B.4)}
\end{aligned}
$$

where the inequality holds by $[a]_+ + [b]_+ \leq [a-b]_+$ and $|a|-|b| \leq |a-b|$ for $a, b \in \mathbb{R}$, and the last equality holds by Lemma B.1 (i).

When $2 \leq j \leq m$, we have

$$
\begin{aligned}
\left|S_{T,j}-\sqrt{T}d_{m,j}\right| &= \left|\sqrt{T}\left[\bar{F}_1^{(j)}(\overline{x})-\bar{F}_2^{(j)}(\overline{x})\right]-\left[F_1^{(j)}(\overline{x})-F_2^{(j)}(\overline{x})\right]\right| \\
&\leq \left|\nu_T^{(j)}(\overline{x})\right| = O_p(1), \qquad \text{(B.5)}
\end{aligned}
$$

where the last equality holds by Lemma B.1 (i). Thus, we have the desired result. $\square$

Proof of Lemma 2. Since the empirical processes $\nu_T^{(m)}(\cdot)$ is asymptotically tight by Lemma B.1 (i), we have

$$
\begin{aligned}
&P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right)-\left(F_1^{(m)}(x)-F_2^{(m)}(x)\right)\right| > \widehat{c}_T - c_{T,L}\right) \to 0 \\
&P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right)-\left(F_1^{(m)}(x)-F_2^{(m)}(x)\right)\right| > c_{T,U} - \widehat{c}_T\right) \to 0
\end{aligned}
$$

by Assumption 4. Equivalently, we have

$$
\begin{aligned}
&P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right)-\left(F_1^{(m)}(x)-F_2^{(m)}(x)\right)\right| \leq \widehat{c}_T - c_{T,L}\right) \to 1 \\
&P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right)-\left(F_1^{(m)}(x)-F_2^{(m)}(x)\right)\right| \leq c_{T,U} - \widehat{c}_T\right) \to 1.
\end{aligned}
$$

(1) Let $x \in \mathcal{C}_0(c_{T,L})$. Then, by the triangular inequality,

$$
\begin{aligned}
\sqrt{T}\left|\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right| &\leq \sqrt{T}\left|\left(\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right)-\left(F_1^{(m)}(x)-F_2^{(m)}(x)\right)\right| \\
&\quad + \sqrt{T}\left|F_1^{(m)}(x)-F_2^{(m)}(x)\right| \\
&\leq (\widehat{c}_T - c_{T,L}) + c_{T,L} = \widehat{c}_T
\end{aligned}
$$

with probability approaching 1. Thus, we have $P\left(\mathcal{C}_0(c_{T,L}) \subset \widehat{\mathcal{C}}_0(\widehat{c}_T)\right) \to 1$. Now, let $x \in \widehat{\mathcal{C}}_0(\widehat{c}_T)$. The triangular inequality implies

$$
\begin{aligned}
\sqrt{T}\left|F_1^{(m)}(x)-F_2^{(m)}(x)\right| &\leq \sqrt{T}\left|\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right| \\
&\quad + \sqrt{T}\left|\left(\bar{F}_1^{(m)}(x)-\bar{F}_2^{(m)}(x)\right)-\left(F_1^{(m)}(x)-F_2^{(m)}(x)\right)\right| \\
&\leq \widehat{c}_T + (c_{T,U} - \widehat{c}_T) = c_{T,U}
\end{aligned}
$$

with probability approaching 1. Thus, we have $P\left(\widehat{\mathcal{C}}_0(\widehat{c}_T) \subset \mathcal{C}_0(c_{T,U})\right) \to 1$.

(2) Let $x \in \mathcal{C}_+(c_{T,U})$. Then, we have

$$\begin{aligned}
\sqrt{T}\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) &= \sqrt{T}\left[\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) - \left(F_1^{(m)}(x) - F_2^{(m)}(x)\right)\right] \\
&\quad + \sqrt{T}\left(F_1^{(m)}(x) - F_2^{(m)}(x)\right) \\
&> (\widehat{c}_T - c_{T,U}) + c_{T,U} = \widehat{c}_T
\end{aligned}$$

with probability approaching 1. Thus, we have $P\left(\mathcal{C}_+(c_{T,U}) \subset \widehat{\mathcal{C}}_+(\widehat{c}_T)\right) \to 1$. Now, let $x \in \widehat{\mathcal{C}}_+(\widehat{c}_T)$. The triangular inequality implies

$$\begin{aligned}
\sqrt{T}\left(F_1^{(m)}(x) - F_2^{(m)}(x)\right) &= \sqrt{T}\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) \\
&\quad + \sqrt{T}\left[\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) - \left(F_1^{(m)}(x) - F_2^{(m)}(x)\right)\right] \\
&> \widehat{c}_T - (\widehat{c}_T - c_{T,L}) = c_{T,L}
\end{aligned}$$

with probability approaching 1. Thus, we have $P\left(\widehat{\mathcal{C}}_+(\widehat{c}_T) \subset \mathcal{C}_+(c_{T,L})\right) \to 1$.

(3) Let $x \in \mathcal{C}_-(c_{T,U})$. Then, we have

$$\begin{aligned}
\sqrt{T}\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) &= \sqrt{T}\left[\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) - \left(F_1^{(m)}(x) - F_2^{(m)}(x)\right)\right] \\
&\quad + \sqrt{T}\left(F_1^{(m)}(x) - F_2^{(m)}(x)\right) \\
&< (c_{T,U} - \widehat{c}_T) - c_{T,U} = -\widehat{c}_T
\end{aligned}$$

with probability approaching 1. Thus, we have $P\left(\mathcal{C}_-(c_{T,U}) \subset \widehat{\mathcal{C}}_-(\widehat{c}_T)\right) \to 1$. Now, let $x \in \widehat{\mathcal{C}}_-(\widehat{c}_T)$. The triangular inequality implies

$$\begin{aligned}
\sqrt{T}\left(F_1^{(m)}(x) - F_2^{(m)}(x)\right) &= \sqrt{T}\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) \\
&\quad + \sqrt{T}\left[\left(\bar{F}_1^{(m)}(x) - \bar{F}_2^{(m)}(x)\right) - \left(F_1^{(m)}(x) - F_2^{(m)}(x)\right)\right] \\
&< -\widehat{c}_T + (\widehat{c}_T - c_{T,L}) = -c_{T,L}
\end{aligned}$$

with probability approaching 1. Thus, we have $P\left(\widehat{\mathcal{C}}_-(\widehat{c}_T) \subset \mathcal{C}_-(c_{T,L})\right) \to 1$. Therefore, we obtain

$$\begin{aligned}
&P\left\{\mathcal{C}_0(c_{T,L}) \subset \widehat{\mathcal{C}}_0(\widehat{c}_T) \subset \mathcal{C}_0(c_{T,U})\right\} \to 1 \\
&P\left\{\mathcal{C}_+(c_{T,U}) \subset \widehat{\mathcal{C}}_+(\widehat{c}_T) \subset \mathcal{C}_+(c_{T,L})\right\} \to 1 \\
&P\left\{\mathcal{C}_-(c_{T,U}) \subset \widehat{\mathcal{C}}_-(\widehat{c}_T) \subset \mathcal{C}_-(c_{T,L})\right\} \to 1.
\end{aligned} \tag{B.6}$$

Next, we claim

$$
\begin{aligned}
&P\left\{\mathcal{C}_0(c_{T,U}) \subset \mathcal{C}_0\right\} \to 1 \\
&P\left\{\mathcal{C}_+ \subset \mathcal{C}_+(c_{T,U})\right\} \to 1 \\
&P\left\{\mathcal{C}_- \subset \mathcal{C}_-(c_{T,U})\right\} \to 1.
\end{aligned} \tag{B.7}
$$

(1) Suppose $x \in \mathcal{C}_0(c_{T,U})$. Then, for large enough T, we have $|F_1(x) - F_2(x)| \leq \frac{c_{T,U}}{\sqrt{T}} < \epsilon$ for arbitrary $\epsilon > 0$ by definition of $c_{T,U}$. Thus, $F_1(x) = F_2(x)$ for large enough T, which implies that $x \in \mathcal{C}_0$ with probability approaching 1. We have the first part of the claim.

(2) Suppose $x \in \mathcal{C}_+$. Then, $F_1(x) - F_2(x) = c > 0$. By definition of $c_{T,U}$, we have $\frac{c_{T,U}}{\sqrt{T}} < c$ for large enough T. Thus, $F_1(x) - F_2(x) = c > \frac{c_{T,U}}{\sqrt{T}}$ for large enough T, which implies that $x \in \mathcal{C}_+(c_{T,U})$ with probability approaching 1. We have the second part of the claim.

(2) Suppose $x \in \mathcal{C}_-$. Then, $F_1(x) - F_2(x) = c < 0$. By definition of $c_{T,U}$, we have $-\frac{c_{T,U}}{\sqrt{T}} > -c$ for large enough T. Thus, $F_1(x) - F_2(x) = c < -\frac{c_{T,U}}{\sqrt{T}}$ for large enough T, which implies that $x \in \mathcal{C}_-(c_{T,U})$ with probability approaching 1. We have the last part of the claim.

Since we trivially have $\mathcal{C}_0 \subset \mathcal{C}_0(c_{T,L})$, $\mathcal{C}_+(c_{T,L}) \subset \mathcal{C}_+$, and $\mathcal{C}_-(c_{T,L}) \subset \mathcal{C}_-$ almost everywhere, combining (B.6) and (B.7) yields the desired result of the lemma. $\square$

References

[1] Álvarez-Esteban, P. C., del Barrio, E., Cuesta-Albertos, J. A., and Matrán, C. (2016). A Contamination Model for the Stochastic Order. *Test*, 25(4), 751-774.

[2] Anderson, G. (1996). Nonparametric Tests of Stochastic Dominance in Income Distributions. *Econometrica*, 1183-1193.

[3] Anderson, G., Post, T., and Whang, Y. J. (2020). Somewhere between Utopia and Dystopia: Choosing from Multiple Incomparable Prospects. *Journal of Business and Economic Statistics*, 38(3), 502-515.

[4] Andrews, D. W. and Pollard, D. (1994). An Introduction to Functional Central Limit Theorems for Dependent Stochastic Processes. *International Statistical Review*, 119-132.

[5] Andrews, D. W. and Soares, G. (2010). Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection. *Econometrica*, 78(1), 119-157.

[6] Bali, T. G., Demirtas, K. O., Levy, H., and Wolf, A. (2009). Bonds versus Stocks: Investors' Age and Risk Taking. *Journal of Monetary Economics*, 56(6), 817-830.

[7] Barrett, G. F. and Donald, S. G. (2003). Consistent Tests for Stochastic Dominance. *Econometrica*, 71(1), 71-104.

[8] Chernozhukov, V., Lee, S., and Rosen, A. M. (2013). Intersection Bounds: Estimation and Inference. *Econometrica*, 81(2), 667-737.

[9] Davidson, R. (2009). Testing for Restricted Stochastic Dominance: Some Further Results. *Review of Economic Analysis*, 1(1), 34-59.

[10] Davidson, R. and Duclos, J. Y. (2000). Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality. *Econometrica*, 68(6), 1435-1464.

[11] Davidson, R. and Duclos, J. Y. (2013). Testing for Restricted Stochastic Dominance. *Econometric Reviews*, 32(1), 84-125.

[12] Do, T. K. (2021). Socially Responsible Investing Portfolio: An Almost Stochastic Dominance Approach. *International Journal of Finance and Economics*, 26(1), 1122-1132.

[13] Guo, X., Levy, H., and Wong, W. K. (2015). Consistent Tests for Almost Stochastic Dominance. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2592205

[14] Hall, P. and Heyde, C. C. (2014). *Martingale Limit Theory and its Application*. Academic press.

[15] Kaur, A., Rao, B. P., and Singh, H. (1994). Testing for Second-Order Stochastic Dominance of Two Distributions. *Econometric Theory*, 10(5), 849-866.

[16] Klecan, L., McFadden, R., and McFadden, D. (1991). A Robust Test for Stochastic Dominance. Working Paper, Department of Economics, MIT.

[17] Knight, J. and Satchell, S. (2008). Testing for infinite order stochastic dominance with application to finance, risk and income inequality/ *Journal of Economics and Finance*, 32(1), 35-46.

[18] Lee, S., Song, K., and Whang, Y. J. (2018). Testing for a General Class of Functional Inequalities. *Econometric Theory*, 34(5), 1018-1064.

[19] Leshno, M. and Levy, H. (2002). Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance. *Management Science*, 48(8), 1074-1085.

[20] Levy, H. (2016). *Stochastic Dominance: Investment Decision Making under Uncertainty*. Springer.

[21] Levy, M. (2009). Almost Stochastic Dominance and Stocks for the Long Run. *European Journal of Operational Research*, 194(1), 250-257.

[22] Levy, M. (2012). Almost Stochastic Dominance and Efficient Investment Sets. *American Journal of Operations Research*, 2, 313-321.

[23] Levy, H., Leshno, M., and Leibovitch, B. (2010). Economically Relevant Preferences for All Observed Epsilon. *Annals of Operations Research*, 176(1), 153-178.

[24] Linton, O., Maasoumi, E., and Whang, Y. J. (2005). Consistent Testing for Stochastic Dominance under General Sampling Schemes. *The Review of Economic Studies*, 72(3), 735-765.

[25] Linton, O., Song, K., and Whang, Y. J. (2010). An Improved Bootstrap Test of Stochastic Dominance. *Journal of Econometrics*, 154(2), 186-202.

[26] McFadden, D. (1989). Testing for Stochastic Dominance. In *Studies in the Economics of Uncertainty* (pp. 113-134). Springer, New York, NY.

[27] Osuna, E. E. (2012). Crossing Points of Distributions and a Theorem that Relates Them to Second Order Stochastic Dominance. *Statistics and Probability Letters*, 82(4), 758-764.

[28] Politis, D. N. and Romano, J. P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association*, 89(428), 1303-1313

[29] Politis, D. N. and Romano, J. P. (1994). Limit Theorems for Weakly Dependent Hilbert Space Valued Random Variables with Application to the Stationary Bootstrap. *Statistica Sinica*, 461-476.

[30] Politis, D. N. and White, H. (2004). Automatic Block-Length Selection for the Dependent Bootstrap. *Econometric reviews*, 23(1), 53-70.

[31] Tzeng, L. Y., Huang, R. J., and Shih, P. T. (2013). Revisiting Almost Second-Degree Stochastic Dominance. *Management Science*, 59(5), 1250-1254.

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

[33] Whang, Y. J. (2019). *Econometric Analysis of Stochastic Dominance: Concepts, Methods, Tools, and Applications*. Cambridge University Press.