경제학석사 학위논문

Testing for Almost Stochastic Dominance with Applications to Investment Decision Making

완화된 확률적 지배관계 검정과 투자 결정 문제에의 응용

2023년 8월

서울대학교 대학원

경제학부 경제학 전공

배 원 우

Testing for Almost Stochastic Dominance with Applications to Investment Decision Making

지도교수 황 윤 재

이 논문을 경제학석사 학위논문으로 제출함
2023년 6월

서울대학교 대학원
경제학부 경제학 전공
배 원 우

배원우의 경제학석사 학위논문을 인준함
2023년 6월

위 원 장 ______ 서 명 환 ______

부위원장 ______ 황 윤 재 ______

위 원 ______ 이 서 정 ______

Abstract

I propose a novel nonparametric test to assess the null hypothesis of almost stochastic dominance (ASD) in the presence of an unknown parameter. The conventional stochastic dominance (SD) rule entails ranking distributions for all utility functions within a specific class, which can be restrictive in practice. To overcome this limitation, Leshno and Levy (2002) introduced the ASD rule, which applies to most rather than all decision makers by eliminating economically pathological preferences. The ASD rule finds application in numerous empirical economic problems, including investment decisions. In this paper, I propose an integral-type test statistic that relies on empirical distribution functions and suggest bootstrap procedures to calculate critical values. I apply the test to compare return distributions of portfolios based on stock market anomalies across different investment horizons. The results of the ASD tests provide support for the effectiveness of investment strategies employing market anomalies.

Keywords: Almost Stochastic Dominance, Test Consistency, Stock Market Anomalies
Student Number: 2021-25134

Contents

List of Tables

List of Figures

1 Introduction

Numerous economic problems entail the comparison of diverse prospects. Alongside other methodologies like mean-variance analysis in finance, there has been significant attention given to ranking stochastic objects based on their distributions. The stochastic dominance (SD) rule, available in various orders, offers a consistent and weak ordering of distributions for multiple economic outcomes, such as investment strategies and welfare results. This ordering through SD encompasses a wide class of utility functions. For instance, the first-order SD rule ranks decisions made by agents with strictly increasing utility functions, while the second-order SD rule pertains to decisions of agents with strictly increasing and strictly concave utility functions.

The SD rule's *uniform* ordering property, which applies to a broad utility class, represents a key advantage. However, this advantage is not fully realized in practical applications due to the SD rule's overly stringent nature. Even a slight deviation from the SD rule renders the ordering invalid, and such deviations may arise when distributions intersect or, in other words, when an agent exhibits "extreme" preferences. For instance, consider two prospects, X_1 and X_2, where X_1 offers either \$2 or \$3 with equal probability, and X_2 provides \$1 or $\$1,000,000$ with equal probability. Intuitively, it seems likely that "most" investors would prefer X_2 over X_1. However, neither X_1 nor X_2 stochastically dominates the other because their distributions intersect, making it difficult to establish a definitive ranking based solely on the SD rule.

To address the limitation of the SD rule and eliminate extreme preferences, Leshno and Levy (2002) introduced the concept of almost stochastic dominance (ASD) as an alternative. Unlike the SD rule, which applies to all decision makers, the ASD rule is designed to be applicable to "most" decision makers. However, Tzeng, Huang, and Shih (2013) found that the original ASD did not satisfy the uniform ordering property and, as a result, modified its definition to ensure satisfaction.

Specifically, the ASD rule now excludes "economically pathological" preferences from the class of utility functions it considers. These preferences are mathematically valid but economically irrelevant. For instance, one example of pathological preferences provided by Leshno and Levy (2002) is a myopic utility function represented as $u(x) = \frac{x^\alpha}{\alpha}$, where $0 < \alpha < 1$.

The ASD rule allows for certain limited violations of the SD rule by narrowing the choice set to a group of utility functions with bounded derivatives. This means that crossings between distributions, which are often encountered in empirical examples, are permitted. By allowing such crossings, the ASD rule expands the set of prospects that can be ranked compared to the traditional SD rule. Moreover, it maintains the uniformity property for a reasonably large class of utility functions.

The ASD rule has found numerous applications within the realm of financial decision-making. For instance, Bali, Demirtas, Levy, and Wolf (2009) employ the ASD concept to provide support for the widely practiced strategy of initially allocating a larger proportion to stocks and gradually shifting funds to bonds as the investment horizon becomes shorter. The argument favoring long-term stock investment is further explored by Levy (2009).

Levy (2012) utilizes the ASD rule to devise algorithms for deriving ASD-efficient investment sets and demonstrates their effectiveness. Do (2021) employs the ASD rule to present the outperformance of socially responsible investing (SRI) portfolios when compared to market indexes. For additional applications, one can refer to Levy (2016). It is worth noting, however, that empirical analyses utilizing the ASD rule have predominantly relied on numerical computations of the violation ratio, often overlooking the consideration of sampling errors.

The objective of this study is to devise a test for the hypothesis of almost stochastic dominance in scenarios where an unknown parameter is involved. Bae and Whang (2023) have introduced an ASD test without finite dimensional parameters, and this paper extends that test to encompass situations where distributions of random variables reliant on unknown parameters are compared.

In this work, I explore an L_p- or supremum- type test statistic, which is based on empirical distribution functions. To compute critical values, I introduce bootstrap methods tailored to various sampling schemes, and demonstrate their asymptotic validity. Employing the proposed test, I empirically assess the conventional practice of investing in stock market anomalies.

This paper contributes to the extensive body of literature focused on testing stochastic dominance (SD) hypotheses. The previous literature has proposed diverse approaches to testing SD, including McFadden (1989), Klecan, McFadden, and McFadden (1991), Kaur, Rao, and Singh (1994), Anderson (1996), and Davidson and Duclos (2000). Barrett and Donald (2003) present a consistent bootstrap method for testing SD of any order between two prospects using an independent sampling scheme. Linton, Maasoumi, and Whang (2005) develop a consistent subsampling test for SD under general sampling schemes, accommodating time series dependence. Linton, Song, and Whang (2010) suggest a bootstrap-based test that enhances power performance by leveraging information from the binding part of inequality restrictions.

This study also intersects with tests for various weaker forms of the SD relation. For instance, Álvarez-Esteban, del Barrio, Cuesta-Albertos, and Matrán (2016) propose a test for the approximate SD relationship based on mixture (or contaminated) models, while Knight and Satchell (2008) present a test for the infinite order SD hypothesis. Notably, none of these previous works formulate an inference method for the almost stochastic dominance (ASD) hypothesis, despite the ASD rule being a prominent concept in empirical research. For a comprehensive overview of the literature, refer to Whang (2019).

As an empirical demonstration, I apply the developed test to assess the investment practice exploiting anomalies in the market including the value premium. Initially, I establish that the standard SD test does not consistently endorse investors' preferences for one prospect over another across the examined investment horizons. However, through the utilization of the proposed test, empirical evidence emerges supporting investors' escalating preferences for high-return stocks across various investment horizons. This test demonstrates particular relevance in the context of financial decision-making problems due to their inherent requirement for both economic and mathematical rationality.

The subsequent sections of this paper are structured as follows: In Section 2, I present the definition of ASD and introduces the pertinent hypotheses. Section 3 defines the test statistics and explores their asymptotic properties. Section 4 introduces the bootstrap inference method and establishes its asymptotic validity. Section 5 conducts an empirical analysis of the popular investment strategy using the test. Finally, I conclude the paper with final remarks in Section 6.

2 Almost Stochastic Dominance and the Hypotheses of Interest

2.1 Almost Stochastic Dominance

This paper adopts the concept of almost stochastic dominance given by Tzeng, Huang, and Shih (2013).[1]

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.

[1]The definition is a corrected version of the original definition of ASD by Leshno and Levy (2002) so that ASD can have the expected utility maximization property.

2.2 Hypotheses of Interest

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

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

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 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 $p \in \{1, 2\}$. I 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, the population quantity is defined 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

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

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^*.

Based on the test for almost stochastic dominance proposed by Bae and Whang (2023), I extend the test so that it can include the case of residual almost stochastic dominance. Let $X_k(\theta)$ be specified as

$$X_k(\theta) = \varphi_k(W, \theta),\ k = 1, 2,$$

where W is a random vector in $\mathbb{R}^{d_W}$ and $\varphi_k(\cdot, \theta)$ is a real-valued function known up to the parameter $\theta \in \Theta \subset \mathbb{R}^{d_\theta}$. Specifically, let $X_k = X_k(\theta_0)$ for some $\theta_0 \in \Theta$. One typical example of X_k is the residual from the linear regression $X_k = Y_k - Z_k^\intercal\theta_0$, where $Y_k = Z_k^\intercal\theta_0 + \epsilon_k$ with $\mathbf{E}(\epsilon_k|Z_k) = 0$ a.s. In this case, define $W = (Y, Z)$ and $\varphi_k(w, \theta) = y_k - z_k^\intercal\theta$,

$w = (y, z)$.

3 Test Statistics and Large Sample Properties

3.1 Test Statistics

Define the test statistic based on data $\{W_t : t = 1, \ldots, T_k\}$. First, estimate F_k using the empirical cumulative distribution function (ECDF)

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

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

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

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 $|a| = [a]_+ - [a]_-$, and $\hat{\theta}$ denotes a consistent estimator of θ_0.

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

$$\begin{aligned} \bar{F}_k^{(m)}(x, \theta) &:= \frac{1}{T_k} \sum_{t=1}^{T_k} \frac{(x - X_{k,t}(\theta))^{m-1} 1(X_{k,t}(\theta) \le x)}{(m-1)!} \\ &= \frac{1}{T_k} \sum_{t=1}^{T_k} h_{x,m}(X_{k,t}(\theta)),\ k = 1, 2, \end{aligned}$$

where $h_{x,m}(\varphi) := (x - \varphi)^{m-1} 1\{\varphi \le x\} / (m-1)!$. To test the null hypothesis $H_0^{(m)}$, this paper considers 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

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

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

3.2 Large Sample Properties

I present the regularity conditions to derive the asymptotic properties of S_T. In this subsection, I mainly focus on $S_{T,1}$, which corresponds to the first inequality restriction concerning the violation from the m-th order SD. Once I obtain the large sample property of this term, I can test the main hypothesis of interest since the asymptotic normality of each $S_{T,j}$ for $2 \leq j \leq m$ is straightforward.[2] To derive the asymptotic property of the test statistics, I specify conditions for the data generating process of W. Let $B_\Theta(\delta) \equiv \{\theta \in \Theta : \|\theta - \theta_0\| < \delta\}$ be the δ-neighborhood of θ_0, where $\|\cdot\|$ denotes the Euclidean norm. I assume that the observed data are generated under either of the following sampling schemes.

Assumption 1 (Type I sampling)**.**

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

(b) The union of supports of $X_{k,t}(\theta_0)$, $k = 1, 2$, is $\mathcal{X} = [\underline{x}, \overline{x}]$, $-\infty < \underline{x} < \overline{x} < \infty$, and the distribution of $X_{k,t}(\theta_0)$ 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) $\{(W_{1,t}, W_{2,t})^\intercal : t = 1, \ldots, T\}$ 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.

Type I sampling implies that, for each $k = 1, 2$, the observations are independent across t, while Type II sampling allows for dependence across k and t. I further assume the conditions for $X_{k,t}(\theta)$ and $\hat{\theta}$ in the following two assumptions. These assumptions are based on the assumptions from Linton, Song, and Whang (2010) and Lee, Linton, and Whang (2022), where conditions on the data generating process with the unknown parameter are specified in the context of stochastic dominance and time stochastic dominance, respectively.

Assumption 3.

(a) For some $\delta > 0$, $\mathbf{E}\left[\sup_{\theta \in B_\Theta(\delta)} |X_{k,t}(\theta)|^{2((m-1)\vee 1)+\delta}\right] < \infty$

(b) For some $\delta > 0$, there exists a non-random $d_\theta \times 1$ vector $\Gamma_k(x)$ such that

$$\begin{aligned}&\left|\mathbf{E}\left[h_{x,m}(X_{k,t}(\theta))\right] - \mathbf{E}\left[h_{x,m}(X_{k,t}(\theta_0))\right] - \Gamma_k(x)^\intercal(\theta - \theta_0)\right| \\ \leq& C\|\theta - \theta_0\|^2,\ k = 1, 2.\end{aligned}$$

where C is a constant.

(c) Condition (A) below holds when $m = 1$ and condition (B) holds under $m \geq 2$:

(A) There exist $\delta, C > 0$ and a subvector W_1 of W such that (i) the conditional density of W given W_1 is bounded uniformly over $\theta \in B_\Theta(\delta)$, (ii) for each θ_1

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

and θ_2 in $B_\Theta(\delta)$, $\varphi_k(W,\theta_1)-\varphi_k(W,\theta_2)$ is measurable with respect to the σ-field of W_1, and (iii) for each $\theta \in B_\Theta(\delta)$ and for each $\epsilon > 0$, $k = 1, 2$,

$$\sup_{w_1} \mathbf{E}\left[\sup_{\theta_1 \in B_\Theta(\delta)} |\varphi_k(W,\theta_1) - \varphi_k(W,\theta_2)|^2 \middle| W_1 = w_1\right] \leq C\epsilon^{2s_2} \tag{3.3}$$

for some $s_2 \in (\lambda/2, 1]$ with $\lambda = 2 \times 1\{m=1\} + 1\{m \geq 2\}$, where the supremum over w_1 runs in the support of W_1.

(B) There exist $\delta, C > 0$ such that Condition (iii) above is satisfied with the conditional expectation replaced by the unconditional one.

Assumption 3 (a) is a moment condition with local boundedness. In the case of linear regression models where $Y_{k,t} = Z_{k,t}^\intercal \theta_0 + \epsilon_{k,t}$, one can write $X_{k,t}(\theta) = \epsilon_{k,t} + Z_{k,t}^\intercal(\theta_0 - \theta)$ and the condition is satisfied when $\mathbf{E}\left[|\epsilon_{k,t}|^{2((m-1)\vee 1)+\delta}\right] < \infty$ and $\mathbf{E}\left[|Z_{k,t}|^{2((m-1)\vee 1)+\delta}\right] < \infty$. Assumption 3 (b) specifies differentiablity of the functional $\int h_{x,m}(X_{k,t}(\theta))dP$ in $\theta \in B_\Theta(\delta)$. Assumption 3 (c) poses (conditional) locally uniform $L_2(P)$-continuity of $\varphi_k(W,\theta)$ in $\theta \in B_\Theta(\delta)$. See Linton, Song, and Whang (2010) and Lee, Linton, and Whang (2022), for further implications of these assumptions.

Assumption 4.

(a) For each $\epsilon > 0$, $P\left\{\|\hat{\theta} - \theta_0\| > \epsilon\right\} = o(1)$ as $T_1, T_2 \to \infty$ or $T \to \infty$.

(b) For each $\epsilon > 0$, $k = 1, 2$,

$$P\left\{\sup_{x \in \mathcal{X}} \left|\sqrt{T_k}\Gamma_k(x)^\intercal \left[\hat{\theta} - \theta_0\right] - \frac{1}{\sqrt{T_k}} \sum_{t=1}^{T_k} \psi_{x,k}(W_t, \theta_0)\right| > \epsilon\right\} \to 0, \tag{3.4}$$

where $\psi_{x,k}(\cdot)$ satisfies that there exist $\eta, \delta > 0$ such that for all $x \in \mathcal{X}$, $\mathbf{E}[\psi_{x,k}(W_t, \theta_0)] = 0$, and

$$\mathbf{E}\left[\sup_{x \in \mathcal{X}} \sup_{\theta \in B_\Theta(\delta)} |\psi_{x,k}(W,\theta)|^{2+\eta}\right] < \infty.$$

(c) There exist a bounded function V on $\mathcal{X}$ and constants $\delta, C > 0$ and $s_1 \in (1/2, 1]$ such that for each $(x_1, \theta_1) \in \mathcal{X} \times B_\Theta(\delta)$ and for each $\epsilon > 0$, $k = 1, 2$,

$$\mathbf{E}\left[\sup_{x \in \mathcal{X}: d_V(x,x_1) \leq \epsilon} \sup_{\theta \in B_\Theta(\delta): \|\theta - \theta_1\| \leq \epsilon} \left|\psi_x^\Delta(W,\theta) - \psi_x^\Delta(W,\theta_1)\right|^2\right] < C\epsilon^{2s_1},$$

where $d_V(x, x_1) := |V(x) - V(x_1)|$ and $\psi_x^\Delta(w,\theta) := \psi_{x,1}(w,\theta) - \psi_{x,2}(w,\theta)$.

Assumption 4 (a) specifies consistency of the estimator for θ_0. Assumption 4 (b) indicates that the functional Γ_k at the estimators has an asymptotic linear representation, which can be established by expanding the functional in terms of the estimator, $\hat{\theta}$ and using the asymptotic linear representation of the estimator. Assumption 4 (c) is a locally uniform L_2-continuity condition. See Linton, Song, and Whang (2010), for further implications of these assumptions.

Under either sampling type, I can show from these assumptions 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 $\nu_T^{(m)}(x) := \sqrt{T}[(\bar{F}_1^{(m)}(x,\hat{\theta}) - \bar{F}_2^{(m)}(x,\hat{\theta})) - (F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0))]$. Let $\nu_{1,2}^{(m)}$ be a mean zero Gaussian process with covariance function given by

$$C(x_1,x_2) := \lim_{T\to\infty} Cov[V_{x_1}(W_t,\theta_0), V_{x_2}(W_t,\theta_0)], \tag{3.5}$$

and

$$\begin{aligned} V_x(w,\theta_0) &:= h_{x,m}^{\Delta}(w,\theta) + \psi_x^{\Delta}(w,\theta), \\ h_{x,m}^{\Delta}(w,\theta) &:= h_{x,m}(\varphi_1(w,\theta)) - h_{x,m}(\varphi_2(w,\theta)), \end{aligned}$$

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., $d_{m,j} = 0$, $\forall j = 1,\ldots,m$.),

$$S_T \Rightarrow S_0 := \Lambda_p(S_{0,1},\ldots,S_{0,m}), \tag{3.6}$$

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

$$\mathcal{C}_0 = \{x \in \mathcal{X} : F_1^{(m)}(x,\theta_0) = F_2^{(m)}(x,\theta_0)\} \tag{3.7}$$

$$\mathcal{C}_+ = \{x \in \mathcal{X} : F_1^{(m)}(x,\theta_0) > F_2^{(m)}(x,\theta_0)\} \tag{3.8}$$

$$\mathcal{C}_- = \{x \in \mathcal{X} : F_1^{(m)}(x,\theta_0) < F_2^{(m)}(x,\theta_0)\}. \tag{3.9}$$

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

4 Bootstrap Critical Values

The asymptotic distribution of the test statistic depends on the true data generating process. Here, I 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 $\mathcal{X}$, it is necessary to estimate these contact sets. Before introducing the estimators for the contact sets, I introduce related notations. Specifically, define the r-enlargement of the contact sets for $r > 0$ as follows:

$$\mathcal{C}_0(r) := \{x \in \mathcal{X} : \sqrt{T}\left|F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0)\right| \le r\}$$

$$\mathcal{C}_+(r) := \{x \in \mathcal{X} : \sqrt{T}(F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0)) > r\}$$

$$\mathcal{C}_-(r) := \{x \in \mathcal{X} : \sqrt{T}(F_1^{(m)}(x, \theta_0) - F_2^{(m)}(x, \theta_0)) < -r\}.$$

To describe the joint testing procedure, I first introduce selection functions following the moment selection idea of Andrews and Soares (2010). Let $\xi_j : \mathbb{R} \to \{0, 1\}$, $1 \le j \le m$, 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, define these functions as follows:

$$\xi_j(x) = 1\left(x \ge -\kappa_{T,j}\right) \text{ for } 1 \le j \le m, \tag{4.1}$$

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. Then,*

$$P\left\{S_T = \Lambda_p\left(\xi_1\left(\sqrt{T}d_{m,1}\right)\cdot S_{T,1}, \ldots, \xi_m\left(\sqrt{T}d_{m,m}\right)\cdot S_{T,m}\right)\right\} \to 1.$$

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 a bootstrap procedure that mimics the representation of S_T in Lemma 1.

Under the Type I sampling, I use the standard nonparametric bootstrap procedure. Under the Type II sampling, I 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, one 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. I assume that the parameter p satisfies the following growth condition:

Assumption 5. Under the Type II sampling, $p + \left(\sqrt{T}p\right)^{-1} \to 0$ as $T \to \infty$.

I 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 $\mathcal{S}_k^* := \{W_{k,t}^* : t = 1, \ldots, T_k\}$, where $W_{k,t}^*$ for $t = 1, \ldots, T_k$ are independently drawn with replacement from the original sample $\mathcal{S}_k := \{W_{k,t} : t = 1, \ldots, T_k\}$.

(2) Using the bootstrap sample $\mathcal{S}_k^*$, compute $X_{k,t}^*(\theta) = \varphi_k(W_t^*, \theta)$, the estimate $\hat{\theta}^*$, and the (I)EDFs:

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

(3) Compute the bootstrap test statistic

$$S_T^* = \Lambda_p \left(\xi_1 \left(S_{T,1} \right) \cdot S_{T,1}^*, \ldots, \xi_m \left(S_{T,m} \right) \cdot S_{T,m}^* \right), \tag{4.3}$$

where

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

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

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

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

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

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

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

(2) Using the bootstrap sample $\mathcal{S}^*$, compute $X_{k,t}^*(\theta) = \varphi_k(W_t^*, \theta)$, the estimate $\hat{\theta}^*$, and the (I)EDFs:

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

(3) Compute the bootstrap test statistic

$$S_T^* = \Lambda_p \left(\xi_1 \left(S_{T,1} \right) \cdot S_{T,1}^*, \ldots, \xi_m \left(S_{T,m} \right) \cdot S_{T,m}^* \right), \tag{4.8}$$

where

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

(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, I 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, take

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

as the critical value.

To establish the validity of the bootstrap procedures, I introduce the following assumptions.

Assumption 6. For $\psi_{x,k}$ in Assumption 4 (b), for any $\epsilon > 0$,

$$
P\left\{\sup_{x\in\mathcal{X}}\left|\sqrt{T_k}\hat{\Gamma}_k(x) - \frac{1}{\sqrt{T_k}}\sum_{t=1}^{T_k}\left\{\psi_{x,k}(W_t^*,\hat{\theta}) - \sum_{t=1}^{T_k}\psi_{x,k}(W_t,\hat{\theta})\right\}\right| > \epsilon\middle|\mathcal{W}_T\right\} \to_p 0, \tag{4.10}
$$

where $\mathcal{W}_T$ is the σ-field generated by $\{W_{k,t} : k = 1,2; t = 1,\ldots,T_k\}$ and $\hat{\Gamma}_k(x) := (1/T_k)\sum_{t=1}^{T_k}\left[h_{x,m}(\varphi_k(W_t^*,\hat{\theta}^*)) - h_{x,m}(\varphi_k(W_t^*,\hat{\theta}))\right]$.

Assumption 7. For each $T_1, T_2 \geq 1$, there exist non-stochastic sequences $c_{T,L}, c_{T,U} > 0$ such that $c_{T,L} \leq \hat{c}_T \leq c_{T,U}$ and

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

Assumption 6 assumes the bootstrap analogue of the asymptotic linearity of $\sqrt{T_k}\Gamma_k(x)^\intercal[\hat{\theta} - \theta_0]$. This condition can be proved when one establishes the validity of bootstrap confidence sets for $\hat{\theta}^*$ under either sampling type. Assumption 7 specifies conditions for the consistency of contact set estimation.

Lemma 2. *Suppose that Assumption 1 or 2, and the other assumptions hold. Then, the followings hold*

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

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

Theorem 1. *Suppose that Assumption 1 holds or Assumption 2, and Assumption 3-7 hold. 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.

Now, I investigate power properties of the tests. Consistency of the proposed test is established in the following theorem.

Theorem 2. *Suppose that Assumption 1 holds or Assumption 2, and Assumption 3-7 hold. Then, under a fixed alternative hypothesis $H_1^{(m)}$, as $T_1, T_2 \to \infty$,*

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

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

5 Application to Investment Decision Making

Substantial evidence exists regarding stock market anomalies that remain unexplained by popular asset pricing models, including the Capital Asset Pricing Model (CAPM) introduced by Sharpe (1964), and the Fama-French three-factor model developed by Fama and French (1993). Conventional decision-making criteria such as the Mean-Variance criterion or the Stochastic Dominance (SD) rule have demonstrated limited usefulness in justifying efficient investment opportunities arising from these anomalies. Despite the theoretical underpinnings of these investment principles, many financial advisors capitalize on the abnormal returns yielded by these anomalies.

In this subsection, I explore stock return data to ascertain whether the practice of investing in stock market anomalies aligns with the expected utility framework. This exploration involves the application of the ASD test to assess the high-return assets and low-return assets associated with stock market anomalies. By investigating these anomalies through the lens of the ASD test, I seek to shed light on whether the observed investment practice can be rationalized within the expected utility paradigm.

This study focuses on five prominent stock market anomalies, as identified by Bali, Brown, and Demirtras (2011). Bali, Brown, and Demirtras (2011) investigated the anomalies using the ASD rule to support market practices. However, their approach lacks formal statistical inference for almost stochastic dominance (ASD) and their definition of ASD does not adhere to the uniform ordering property embraced in this paper. As a result, the current study aims to establish valid statistical evidence concerning ASD relationships among the assets of interest.

Specifically, this paper delves into the examination of the Size Premium, Value Premium, Momentum profits, Short-term Reversal, and Long-term Reversal anomalies. This investigation is carried out using monthly returns derived from size, book-to-market, momentum, short-term reversal, and long-term reversal portfolios, sourced from Kenneth French's online data library. By utilizing formal statistical methods and adhering to the definition of ASD with the the uniform ordering property, this study seeks to provide rigorous statistical substantiation of the ASD relations within these specific anomalies.

The Size Premium pertains to the size anomaly observed in the US equity market, initially identified byBanz (1981). Banz (1981) demonstrated that smaller stocks tend to outperform stocks with larger market capitalization. This size effect has been replicated across multiple markets, leading to the widespread adoption of investment strategies that involve buying small stocks while shorting large stocks. To examine the Size Premium, I investigate size portfolios formed at the end of each June. These portfolios are constructed utilizing the June market equity (ME) and New York Stock Exchange (NYSE) breakpoints, covering the period from July 1926 to April 2023.

The Value Premium denotes the positive correlation observed between a proxy for the value-to-price ratio, such as the book-to-market ratio, and security returns. This phenomenon has garnered substantial evidence dating back to Graham, Dodd, Cottle, and Tatham (1962), highlighting the persistent value premium wherein stocks with elevated book-to-market ratios (value stocks) tend to outperform those with lower ratios (growth stocks). To assess the Value Premium, I examine book-to-market portfolios. These portfolios are formulated based on the BE/ME (Book Equity to Market Equity) ratio and are constructed at the end of each June. The NYSE breakpoints are employed for portfolio construction, spanning the period from July 1926 to April 2023.

Momentum profits arise from a consistent pattern wherein stocks that have exhibited high (low) returns over the preceding 2 to 12 months tend to continue having high (low) returns in the subsequent 1 to 12 months. In other words, the momentum phenomenon signifies that stocks with positive performance continue their winning streak, while underperforming stocks persistently lag behind. The presence of momentum profits has been substantiated by Jegadeesh and Titman (1993, 2001) through the implementation of a zero-investment strategy. To assess the existence of momentum profits, I analyze momentum portfolios. These portfolios are formulated based on prior (12-months ago; 2 months ago) NYSE return decile breakpoints. The analysis encompasses the time span from January 1927 to April 2023.

The Short-term Reversal phenomenon pertains to a negative serial correlation observed in short-term stock returns. A wealth of evidence, including studies by Jegadeesh (1990) and Lehmann (1990), indicates that short-term winners tend to be outperformed by short-term losers. This finding has consistently held true over time and remains robust to date. To examine the presence of short-term reversal, I analyze short-term reversal portfolios. These portfolios are constructed based on prior (1-month ago; current) NYSE return decile breakpoints. The investigation spans the period from July 1926 to April 2023.

The Long-term Reversal phenomenon is rooted in the tendency of investors to excessively extrapolate past information into the future. In essence, stocks that have experienced recent negative news tend to be undervalued, while those with recent positive news become overvalued. De Bondt and Thaler (1985) attribute the long-term reversal to suboptimal Bayesian decision-making and empirically establish that long-term winners are outperformed by long-term losers. To assess the existence of the long-term reversal, I examine long-term reversal portfolios. These portfolios are constructed based on prior (60-months ago; 13 months ago) NYSE return decile breakpoints. The analysis covers the period from January 1931 to April 2023.

I obtain the lowest and the highest decile portfolios for each anomaly and test 5 anomalies by applying the AFSD or ASSD test. For each anomaly, I test whether the lowest decile portfolio dominates the highest one: for the size premium,

$$H_0 : Small \succeq_{AnSD} Big \text{ and } H_0 : Big \succeq_{AnSD} Small \text{ for } n = 1, 2, \tag{5.1}$$

for the value premium,

$$H_0 : Growth \succeq_{AnSD} Value \text{ and } H_0 : Value \succeq_{AnSD} Growth \text{ for } n = 1, 2, \tag{5.2}$$

and for the momentum profits, the short-term reversal, and the long-term reversal,

$$H_0 : Loser \succeq_{AnSD} Winner \text{ and } H_0 : Winner \succeq_{AnSD} Loser \text{ for } n = 1, 2. \tag{5.3}$$

I consider six different investment horizons: 1-Month, 6-Month, 12-Month, 24-Month, 48-Month, 60-Month.

Table 1 presents the descriptive statistics for the monthly value-weighted returns of these extreme decile portfolios. Except for the momentum portfolios, the mean-variance criterion cannot determine which to invest in since the mean and the standard deviation of high returns exceed those of low returns. Empirical CDFs and ICDFs in Appendix C. show crossings between the lowest and highest decile returns distributions for shorter horizons. Table 2 reports estimated degrees of violation $\hat{\theta}_1$, $\hat{\theta}_2$[3] for SD of the highest decile portfolios over the lowest decile ones, which speaks to the need for testing for ASD.

Before applying the ASD tests, I test the hypotheses using the FSD test or SSD test proposed by Lintion, Song, and Whang (2010). Here, I use stationary bootstrap with optimal block length choice following Politis and White (2004). Table 3 and 4 show results for the FSD and SSD tests. For the size portfolios, there is no FSD relation for 1-Month, 6-Month, 12-Month, and 24-Month horizons at the significance level of 1%, and no SSD relation for 1-Month horizon at the 10% significance level. For the book-to-market portfolios, there is no FSD relation for 1-Month and 6-Month horizons at the significance level of 1% and 10%, respectively. For the moment portfolios and the short-term reversal portfolios, no FSD relation is found for 1-Month horizon at the 5% significance level.

Then, I apply the ASD testing method to these cases. While Levy, Leshno, and Leibovitch (2010) suggests the empirical guide for the choice of ϵ, I run tests with $\epsilon \in \{0.05, 0.1, 0.15, 0.2, 0.25\}$ for AFSD and $\epsilon \in \{0.03, 0.05, 0.1, 0.15, 0.2\}$ for ASSD. Here, I only report the results for $\epsilon = 0.05$ for AFSD tests and $\epsilon = 0.03$ for ASSD tests. Table 5 and 6 show p-values for the AFSD and ASSD tests with selected ϵ's. Given $\epsilon = 0.05$ for AFSD and $\epsilon = 0.03$ for ASSD, I find support for all the anomalies at the significance level of 1% except for the AFSD relation for 1-Month and 6-Month size premium, 1-Month value premium, and 1-Month momentum profits.

[3] 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 $\theta_m \leq \epsilon$.

Table 1: Descriptive statistics

	Observation	Mean	Std. dev.	Min	25%	Median	75%	Max	Skewness	Kurtosis
Panel A. Size Portfolios										
Small	1162	0.013	0.097	-0.343	-0.028	0.011	0.048	1.210	3.825	40.609
Big	1162	0.009	0.050	-0.280	-0.017	0.012	0.037	0.352	0.043	6.483
Panel B. Book-to-market Portfolios										
Growth	1162	0.009	0.057	-0.307	-0.019	0.010	0.041	0.422	0.037	5.713
Value	1162	0.013	0.091	-0.440	-0.024	0.014	0.051	0.987	2.011	22.007
Panel C. Momentum Portfolios										
Loser	1156	0.003	0.099	-0.423	-0.039	0.002	0.041	0.936	1.685	14.737
Winner	1156	0.015	0.064	-0.285	-0.021	0.018	0.055	0.267	-0.497	2.018
Panel D. Short-term Reversal Portfolios										
Loser	1167	0.014	0.086	-0.345	-0.026	0.014	0.051	0.582	0.723	6.711
Winner	1167	0.005	0.069	-0.330	-0.031	0.008	0.042	0.637	0.903	12.440
Panel E. Long-term Reversal Portfolios										
Loser	1108	0.014	0.087	-0.410	-0.027	0.010	0.048	0.916	2.097	19.795
Winner	1108	0.009	0.063	-0.341	-0.024	0.013	0.047	0.308	-0.388	3.472

Table 2: Estimated degree of violation for High Return$\succeq_{SD}$Low Return

	Size		*Value* Horizon		*Momentum*		*ST Reversal*		*LT Reversal*	
	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$
1-Month	0.401	0.307	0.367	0.295	0.265	0.000	0.185	0.078	0.256	0.135
6-Month	0.312	0.101	0.248	0.064	0.100	0.000	0.002	0.000	0.119	0.009
12-Month	0.237	0.054	0.151	0.030	0.062	0.000	0.000	0.000	0.000	0.000
24-Month	0.161	0.038	0.098	0.022	0.008	0.000	0.000	0.000	0.002	0.000
48-Month	0.060	0.006	0.060	0.009	0.000	0.000	0.000	0.000	0.000	0.000
60-Month	0.046	0.004	0.048	0.007	0.000	0.000	0.000	0.000	0.000	0.000

Table 3: p-values from the SD Tests (1)

	Size Portfolios			
Horizon	S$\succeq_{1SD}$B	B$\succeq_{1SD}$S	S$\succeq_{2SD}$B	B$\succeq_{2SD}$S
1-Month	0.000	0.000	0.060	0.055
6-Month	0.000	0.000	0.280	0.045
12-Month	0.000	0.000	0.120	0.000
24-Month	0.020	0.000	0.135	0.000
48-Month	0.550	0.000	0.270	0.000
60-Month	0.495	0.000	0.260	0.000
	Book-to-market Portfolios			
Horizon	G$\succeq_{1SD}$V	V$\succeq_{1SD}$G	G$\succeq_{2SD}$V	V$\succeq_{2SD}$G
1-Month	0.000	0.005	0.025	0.115
6-Month	0.000	0.090	0.005	0.405
12-Month	0.005	0.260	0.000	0.300
24-Month	0.000	0.440	0.000	0.325
48-Month	0.000	0.490	0.000	0.235
60-Month	0.000	0.490	0.000	0.240

Table 4: p-values from the SD Tests (2)

	Momentum Portfolios			
Horizon	L$\succeq_{1SD}$W	W$\succeq_{1SD}$L	L$\succeq_{2SD}$W	W$\succeq_{2SD}$L
1-Month	0.000	0.000	0.000	1.000
6-Month	0.000	0.230	0.000	1.000
12-Month	0.000	0.980	0.000	1.000
24-Month	0.000	0.915	0.000	1.000
48-Month	0.000	1.000	0.000	1.000
60-Month	0.000	1.000	0.000	1.000
	Short-term Reversal Portfolios			
Horizon	L$\succeq_{1SD}$W	W$\succeq_{1SD}$L	L$\succeq_{2SD}$W	W$\succeq_{2SD}$L
1-Month	0.035	0.000	0.325	0.000
6-Month	0.905	0.000	0.695	0.000
12-Month	1.000	0.000	1.000	0.000
24-Month	1.000	0.000	1.000	0.000
48-Month	1.000	0.000	1.000	0.000
60-Month	1.000	0.000	1.000	0.000
	Long-term Reversal Portfolios			
Horizon	L$\succeq_{1SD}$W	W$\succeq_{1SD}$L	L$\succeq_{2SD}$W	W$\succeq_{2SD}$L
1-Month	0.115	0.000	0.270	0.025
6-Month	0.405	0.015	0.680	0.025
12-Month	0.755	0.015	0.965	0.000
24-Month	1.000	0.040	1.000	0.000
48-Month	0.990	0.015	0.685	0.000
60-Month	1.000	0.000	1.000	0.000

Table 5: p-values from the ASD Tests (1)

	Size Portfolios					
Horizon	ϵ	S$\succeq_{A1SD}$B	B$\succeq_{A1SD}$S	ϵ	S$\succeq_{A2SD}$B	B$\succeq_{A2SD}$S
1-Month	0.05	0.000	0.000	0.03	0.345	0.020
6-Month	0.05	0.000	0.000	0.03	0.490	0.010
12-Month	0.05	0.075	0.000	0.03	0.405	0.015
24-Month	0.05	0.195	0.005	0.03	0.440	0.030
48-Month	0.05	0.695	0.020	0.03	1.000	0.025
60-Month	0.05	0.835	0.035	0.03	1.000	0.010
	Book-to-market Portfolios					
Horizon	ϵ	G$\succeq_{A1SD}$V	V$\succeq_{A1SD}$G	ϵ	G$\succeq_{A2SD}$V	V$\succeq_{A2SD}$G
1-Month	0.05	0.000	0.000	0.03	0.065	0.400
6-Month	0.05	0.000	0.195	0.03	0.005	0.470
12-Month	0.05	0.000	0.575	0.03	0.020	1.000
24-Month	0.05	0.040	0.665	0.03	0.030	1.000
48-Month	0.05	0.025	0.775	0.03	0.020	1.000
60-Month	0.05	0.005	0.825	0.03	0.005	1.000

Table 6: p-values from the ASD Tests (2)

	Momentum Portfolios					
Horizon	ϵ	L$\succeq_{A1SD}$W	W$\succeq_{A1SD}$L	ϵ	L$\succeq_{A2SD}$W	W$\succeq_{A2SD}$L
1-Month	0.05	0.000	0.010	0.03	0.000	1.000
6-Month	0.05	0.000	0.420	0.03	0.000	1.000
12-Month	0.05	0.000	0.745	0.03	0.000	1.000
24-Month	0.05	0.000	1.000	0.03	0.000	1.000
48-Month	0.05	0.000	1.000	0.03	0.000	1.000
60-Month	0.05	0.000	1.000	0.03	0.000	1.000
	Short-term Reversal Portfolios					
Horizon	ϵ	L$\succeq_{A1SD}$W	W$\succeq_{A1SD}$L	ϵ	L$\succeq_{A2SD}$W	W$\succeq_{A2SD}$L
1-Month	0.05	0.175	0.000	0.03	0.475	0.000
6-Month	0.05	1.000	0.005	0.03	1.000	0.000
12-Month	0.05	1.000	0.000	0.03	1.000	0.000
24-Month	0.05	0.995	0.005	0.03	1.000	0.000
48-Month	0.05	1.000	0.005	0.03	1.000	0.000
60-Month	0.05	0.995	0.005	0.03	1.000	0.000
	Long-term Reversal Portfolios					
Horizon	ϵ	L$\succeq_{A1SD}$W	W$\succeq_{A1SD}$L	ϵ	L$\succeq_{A2SD}$W	W$\succeq_{A2SD}$L
1-Month	0.05	0.325	0.000	0.03	0.545	0.030
6-Month	0.05	0.805	0.000	0.03	1.000	0.005
12-Month	0.05	0.915	0.015	0.03	1.000	0.000
24-Month	0.05	0.995	0.010	0.03	1.000	0.005
48-Month	0.05	0.995	0.010	0.03	1.000	0.020
60-Month	0.05	1.000	0.025	0.03	1.000	0.010

6 Conclusion

This study introduces L_p-type tests for almost stochastic dominance (ASD) and establishes their asymptotic distributions within both independent and dependent sampling schemes, even in cases where the involved random variables are contingent on unknown parameters. Bootstrap procedures are proposed under both sampling schemes to emulate the null distribution of the approximated test statistic. These procedures leverage information from the binding segments of the inequality constraints. It is proven that the bootstrap tests maintain asymptotically correct size and exhibit asymptotic exactness when one of the inequality constraints is binding. The study establishes test consistency against a fixed alternative hypothesis and applies the proposed testing methodology to validate prevalent investment decisions relating to anomalies in the stock market.

While ASD finds empirical value across a broad spectrum of economic domains, it has predominantly seen application within financial economics, often relying on numerical simulations. Given that many empirical inquiries involving distributional comparisons or stochastic dominance could potentially benefit from ASD, the tests introduced in this paper hold relevance across diverse economic fields, including welfare economics and policy evaluation. Furthermore, this study lays the groundwork for future tests aimed at generalizing various stochastic dominance concepts, such as time stochastic dominance and Lorenz dominance.

7 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] Bae, W. and Whang, Y. J. (2023). Testing for Almost Stochastic Dominance. Working Paper.

[7] Bali, T. G., Brown, S. J., and Demirtas, K. O. (2011). Investing in Stock Market Anomalies. Available at SSRN 1806137.

[8] 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.

[9] Banz, R. W. (1981). The Relationship between Return and Market Value of Common Stocks. *Journal of financial economics*, 9(1), 3-18.

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

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

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

[13] 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.

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

[15] De Bondt, W. F. and Thaler, R. (1985). Does the Stock Market Overreact?. *The Journal of finance*, 40(3), 793-805.

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

[17] Fama, E. F. and French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of financial economics*, 33(1), 3-56.

[18] Graham, B., Dodd, D. L. F., Cottle, S., and Tatham, C. (1962). *Security Analysis: Principles and Technique* (*Vol. 5*). New York: McGraw-Hill.

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

[20] Jegadeesh, N. (1990). Evidence of Predictable Behavior of Security Returns. *The Journal of finance*, 45(3), 881-898.

[21] Jegadeesh, N. and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of finance*, 48(1), 65-91.

[22] Jegadeesh, N. and Titman, S. (2001). Profitability of Momentum Strategies: An Evaluation of Alternative Explanations. *The Journal of finance*, 56(2), 699-720.

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

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

[25] 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.

[26] Lee, K., Linton, O., and Whang, Y. J. (2023). Testing for time stochastic dominance. *Journal of Econometrics*, 235(2), 352-371.

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

[28] Lehmann, B. N. (1990). Fads, Martingales, and Market Efficiency. *The Quarterly Journal of Economics*, 105(1), 1-28.

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

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

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

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

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

[34] 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.

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

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

[37] 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.

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

[39] 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.

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

[41] Sharpe, W. F. (1964). Capital asset prices: A Theory of Market Equilibrium under Conditions of Risk. *The journal of finance*, 19(3), 425-442.

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

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

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

Appendices

Appendix A gives the proofs of Theorem 1 and 2 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, I proceed as follows. First, I establish the null distribution of the test statistic based on the representation of it given in Lemma 1. Then, I prove the bootstrap consistency result for the bootstrap test statistic. Finally, I prove the validity of the testing procedure. Since the proof under Type I sampling is straightforward, I present the proof under Type II sampling.

Null Distribution

I first derive the asymptotic distributions of $S_{T,1}, \dots, S_{T,m}$ when each inequality restriction of the null hypothesis $H_0^{(m)}$ binds. For the convenience of notation, define temporarily $\delta_T^F(x) := \sqrt{T}\left[F_1^{(m)}(x, \theta_0) - F_2^{(m)}(x, \theta_0)\right]$. When the first inequality restriction of the null hypothesis binds, one can obtain

$$\begin{aligned}
S_{T,1} = \sqrt{T}\int_{\mathcal{X}} \Bigg\{ & \left(\left[\bar{F}_1^{(m)}(x,\hat{\theta}) - \bar{F}_2^{(m)}(x,\hat{\theta})\right]_+ - \left[F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0)\right]_+\right) \\
& - \epsilon\left(\left|\bar{F}_1^{(m)}(x,\hat{\theta}) - \bar{F}_2^{(m)}(x,\hat{\theta})\right| - \left|F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0)\right|\right) \Bigg\} dx \\
= \sqrt{T}\int_{\mathcal{X}} \Bigg\{ & (1-\epsilon)\left(\left[\bar{F}_1^{(m)}(x,\hat{\theta}) - \bar{F}_2^{(m)}(x,\hat{\theta})\right]_+ - \left[F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0)\right]_+\right) \\
& + \epsilon\left(\left[\bar{F}_1^{(m)}(x,\hat{\theta}) - \bar{F}_2^{(m)}(x,\hat{\theta})\right]_- - \left[F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0)\right]_-\right) \Bigg\} dx \\
= \int_{\mathcal{C}_0} & \left\{ \left[\nu_T^{(m)}(x)\right]_+ - \epsilon\left|\nu_T^{(m)}(x)\right| \right\} dx \\
& + \int_{\mathcal{C}_+} \left((1-\epsilon)\max\left\{\nu_T^{(m)}(x), -\delta_T^F(x)\right\} + \epsilon\min\left\{\nu_T^{(m)}(x) + \delta_T^F(x), 0\right\}\right) dx \\
& + \int_{\mathcal{C}_-} \left((1-\epsilon)\max\left\{\nu_T^{(m)}(x) + \delta_T^F(x), 0\right\} + \epsilon\min\left\{\nu_T^{(m)}(x), -\delta_T^F(x)\right\}\right) dx \\
= \int_{\mathcal{C}_0} & \left\{ \left[\nu_T^{(m)}(x)\right]_+ - \epsilon\left|\nu_T^{(m)}(x)\right| \right\} dx + (1-\epsilon)\int_{\mathcal{C}_+}\left(\nu_T^{(m)}(x) + o_p(1)\right) dx + \epsilon\int_{\mathcal{C}_-}\left(\nu_T^{(m)}(x) + o_p(1)\right) dx \\
= \int_{\mathcal{C}_0} & \left\{ \left[\nu_T^{(m)}(x)\right]_+ - \epsilon\left|\nu_T^{(m)}(x)\right| \right\} dx + (1-\epsilon)\int_{\mathcal{C}_+}\nu_T^{(m)}(x)dx + \epsilon\int_{\mathcal{C}_-}\nu_T^{(m)}(x)dx + o_p(1),
\end{aligned}$$

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 $2 \le j \le m$, 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), joint convergence is obtained 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,

$$\begin{aligned} S_T &= \Lambda_p\left(\xi_1\left(\sqrt{T}d_{m,1}\right)\cdot S_{T,1},\ldots,\xi_m\left(\sqrt{T}d_{m,m}\right)\cdot S_{T,m}\right) \\ &= \Lambda_p\left(1\left(d_{m,1}\geq 0\right)\cdot S_{T,1},\ldots,1\left(d_{m,m}\geq 0\right)\cdot S_{T,m}\right) \\ &= \Lambda_p\left(1\left(d_{m,1}=0\right)\cdot\left(S_{T,1}-\sqrt{T}d_{m,1}\right),\ldots,1\left(d_{m,m}=0\right)\cdot\left(S_{T,m}-\sqrt{T}d_{m,m}\right)\right), \quad \text{(A.1)} \end{aligned}$$

where the second and last equalities hold since $\frac{\kappa_{T,j}}{\sqrt{T}}\to 0$ and the third equality holds 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,

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

Bootstrap Consistency

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

(i) I 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) $\int_{\mathcal{X}}\left\{\left[\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right]_+-\epsilon\left|\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right|\right\}dx$
$\to \int_{\mathcal{X}}\left\{\left[F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right]_+-\epsilon\left|F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right|\right\}dx.$

(b) $\bar{F}_1^{(j)}(\overline{x},\hat{\theta})-\bar{F}_2^{(j)}(\overline{x},\hat{\theta})\to F_1^{(j)}(\overline{x},\theta_0)-F_2^{(j)}(\overline{x},\theta_0)$ for $2\leq j\leq m$.

(c) For every subsequence $\{w_T := (w_{T_1},w_{T_2})\}_{T_1,T_2\geq 1}\subset\{(T_1,T_2)\}_{T_1,T_2\geq 1}$, there exists a further subsequence $\{u_T := (u_{T_1},u_{T_2})\}_{T_1,T_2\geq 1}\subset\{w_T\}_{T_1,T_2\geq 1}$ that satisfy $\mathcal{C}_0(c_{u_T,L})\subset\hat{\mathcal{C}}_0(\hat{c}_{u_T})\subset\mathcal{C}_0(c_{u_T,U})$, $\mathcal{C}_+(c_{u_T,U})\subset\hat{\mathcal{C}}_+(\hat{c}_{u_T})\subset\mathcal{C}_+(c_{u_T,L})$, and $\mathcal{C}_-(c_{u_T,U})\subset\hat{\mathcal{C}}_-(\hat{c}_{u_T})\subset\mathcal{C}_-(c_{u_T,L})$.

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

$$\begin{aligned} &\int_{\mathcal{X}}\left\{\left[\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right]_+-\epsilon\left|\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right|\right\}dx \\ &\overset{a.s.}{\to}\int_{\mathcal{X}}\left\{\left[F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right]_+-\epsilon\left|F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right|\right\}dx. \end{aligned}$$

Likewise, the strong law of large number with Assumption 4 (a) implies

$$\bar{F}_1^{(j)}(\overline{x},\hat{\theta})-\bar{F}_2^{(j)}(\overline{x},\hat{\theta})\overset{a.s.}{\to}F_1^{(j)}(\overline{x},\theta_0)-F_2^{(j)}(\overline{x},\theta_0)$$

for $2 \leq j \leq m$. Finally, Lemma 2 implies that estimated enlargements of contact sets lie in the enlargement sets of contact sets with probability approaching 1. Thus, I have

$$\begin{aligned}
&\mathcal{C}_0(c_{u_T,L}) \subset \hat{\mathcal{C}}_0(\hat{c}_{u_T}) \subset \mathcal{C}_0(c_{u_T,U}) \\
&\mathcal{C}_+(c_{u_T,U}) \subset \hat{\mathcal{C}}_+(\hat{c}_{u_T}) \subset \mathcal{C}_+(c_{u_T,L}) \\
&\mathcal{C}_-(c_{u_T,U}) \subset \hat{\mathcal{C}}_-(\hat{c}_{u_T}) \subset \mathcal{C}_-(c_{u_T,L}),
\end{aligned}$$

almost surely.

(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), $\nu_T^{(j)*}(\cdot) \Rightarrow \nu_{1,2}^{(j)}(\cdot)$ 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 $S_{T,1}^*$, I have

$$\begin{aligned}
&(1-\epsilon)\int_{\mathcal{C}_0(c_{u_T,L})}\left[\nu_T^{(m)*}(x)\right]_+ dx + \epsilon\int_{\mathcal{C}_0(c_{u_T,U})}\left[\nu_T^{(m)*}(x)\right]_- dx \\
&+ (1-\epsilon)\int_{\mathcal{C}_+(c_{u_T,U})}\nu_T^{(m)*}(x)dx + \epsilon\int_{\mathcal{C}_-(c_{u_T,L})}\nu_T^{(m)*}(x)dx \\
&\leq S_{u_T,1}^* \\
&\leq (1-\epsilon)\int_{\mathcal{C}_0(c_{u_T,U})}\left[\nu_T^{(m)*}(x)\right]_+ dx + \epsilon\int_{\mathcal{C}_0(c_{u_T,L})}\left[\nu_T^{(m)*}(x)\right]_- dx \\
&+ (1-\epsilon)\int_{\mathcal{C}_+(c_{u_T,L})}\nu_T^{(m)*}(x)dx + \epsilon\int_{\mathcal{C}_-(c_{u_T,U})}\nu_T^{(m)*}(x)dx,
\end{aligned}$$

conditional on $\mathcal{S}$, where the inequalities hold by Assumption A.1 (c). Since one can show that the upper and lower bounds of $S_{u_T,1}^*$ weakly converges to $S_{0,1}$ conditional on $\mathcal{S}$ following the similar logic deriving the null distribution of $S_{T,1}$, I obtain $S_{u_T,1}^* \Rightarrow S_{0,1}$ conditional on $\mathcal{S}$.

In addition, Assumption A.1 (a) and (b) imply that $\frac{S_{T,j}}{\sqrt{T}} \to d_{m,j}$ and so $\xi_j\left(S_{T,j}\right) \to 1\left(d_{m,j} \geq 0\right) = 1\left(d_{m,j}=0\right)$ 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), one obtains joint convergence of $\xi_1\left(S_{T,1}\right) \cdot S_{T,1}^*, \ldots, \xi_m\left(S_{T,m}\right) \cdot S_{T,m}^*$ conditional on $\mathcal{S}$. The continuous mapping theorem yields (A.3) conditional on $\mathcal{S}$. Thus, one has 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 $\lim\limits_{T_1,T_2\to\infty} \sigma_{u_T} > 0$. Then, there exists a binding inequality composing the null hypothesis. Thus, one has

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

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

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

Thus, one can complete the proof by combining (A.6) and (A.7). □

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

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

for $a, b \in \mathbb{R}^m$. Then, one has

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

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 $\Lambda_p\left(\sqrt{T}d_{m,1}, \ldots, \sqrt{T}d_{m,m}\right) > 0$ under the alternative hypothesis $H_1^{(m)}$, for any constant $M_1 > 0$,

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

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. □

B Auxiliary Lemmas and Proofs of Lemmas

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

(*i*) *Then, for all $m \in \mathbb{Z}^+$, I 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 $m \in \mathbb{Z}^+$, I 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. I prove this lemma under Assumption 2. The proof of the lemma under Assumption 1 is straightforward.

(i) I prove the first result in two steps.

(1) By rearranging terms, one can write

$$\begin{aligned}
\nu_T^{(m)}(x) &= \sqrt{T}\left(\bar{F}_1^{(m)}(x,\hat{\theta}) - \bar{F}_2^{(m)}(x,\hat{\theta})\right) - \left(F_1^{(m)}(x,\theta_0) - F_2^{(m)}(x,\theta_0)\right) \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{h_{x,m}^{\Delta}(W_t,\hat{\theta}) - \mathbf{E}h_{x,m}^{\Delta}(W_t,\theta_0)\right\} \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{h_{x,m}^{\Delta}(W_t,\theta_0) - \mathbf{E}h_{x,m}^{\Delta}(W_t,\theta_0)\right\} \\
&+ \sqrt{T}(\Gamma_1 - \Gamma_2)(x)^{\intercal}[\hat{\theta} - \theta_0] + \zeta_{1T} + \zeta_{2T},
\end{aligned}$$

where

$$\begin{aligned}
\zeta_{1T} &:= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{h_{x,m}^{\Delta}(W_t,\hat{\theta}) - h_{x,m}^{\Delta}(W_t,\theta_0)\right\} \\
&- \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\mathbf{E}h_{x,m}^{\Delta}(W_t,\hat{\theta}) - \mathbf{E}h_{x,m}^{\Delta}(W_t,\theta_0)\right\}, \\
\zeta_{2T} &:= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{\mathbf{E}h_{x,m}^{\Delta}(W_t,\hat{\theta}) - \mathbf{E}h_{x,m}^{\Delta}(W_t,\theta_0)\right\} \\
&- \sqrt{T}(\Gamma_1 - \Gamma_2)(x)^{\intercal}[\hat{\theta} - \theta_0].
\end{aligned}$$

By Assumption 3 (b) and Assumption 4 (a), $\zeta_{2T} = o_p(1)$. Let

$$\mathcal{H} = \left\{h_{x,m}^{\Delta}(\cdot,\theta) - h_{x,m}^{\Delta}(\cdot,\theta_0) : (x,\theta) \in \mathcal{X} \times B_\Theta(\delta_T)\right\},$$

where $\delta_T \to 0$ is any decreasing sequence. The bracketing entropy of this class at $\epsilon \in (0,1]$ is bounded by $C\epsilon^{-\lambda/s_2}$ by Lemma A4 of Linton, Song, and Whang (2010). Assumption 3 (c) implies that the $L_2(P)$-norm of its envelope is $O(\delta_T^{2s_2}) = o(1)$. Hence, by using the maximal inequality in Theorem 2.14.2 of van der Vaart and Wellner (1996) and the fact that $\lambda/s_2 < 2$, one has $\zeta_{1T} = o_p(1)$. Now, Assumption 4 (b) gives

$$\nu_T^{(m)}(x) = \eta_T(x) + o_p(1), \text{ uniformly in } x \in \mathcal{X}, \tag{B.1}$$

where

$$\eta_T(x) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left\{V_x(W_t,\theta_0) - \mathbf{E}[V_x(W_t,\theta_0)]\right\}.$$

(2) It only remains to show $\eta_T(\cdot) \Rightarrow \nu_{1,2}^{(m)}(\cdot)$ in $l^\infty(\mathcal{X})$. Since the total boundedness of pseudometric space $(\mathcal{X},\rho)$ is clear from boundedness of $\mathcal{X}$, I 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

$$\overline{\lim_{T\to\infty}} \left\| \sup_{\rho(x_1,x_2)<\delta} |\eta_T(x_1) - \eta_T(x_2)| \right\|_q < \epsilon, \tag{B.2}$$

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

$$\rho(x_1, x_2) = \left\{ E[(h^{\Delta}_{x_1,m}(W_t,\theta_0) - \psi^{\Delta}_{x_1}(W_t,\theta_0)) - (h^{\Delta}_{x_2,m}(W_t,\theta_0) - \psi^{\Delta}_{x_2}(W_t,\theta_0))]^2 \right\}^{1/2}.$$

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

$$f_t(x) = h^{\Delta}_{x,m}(W_t,\theta_0) - \psi^{\Delta}_{x}(W_t,\theta_0).$$

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

$$\begin{aligned}
&\mathbf{E} \sup_{\substack{x_1\in\mathcal{X}\\|x_1-x|\le r}} [f_t(x_1) - f_t(x)]^2 \le 2E \sup_{\substack{x_1\in\mathcal{X}\\|x_1-x|\le r}} [h^{\Delta}_{x,m}(W_t,\theta_0)^2 + \psi^{\Delta}_{x}(W_t,\theta_0)^2] \\
&\le \frac{4}{(m-1)!}\mathbf{E} \sup_{\substack{x_1\in\mathcal{X}\\|x_1-x|\le r}} [((x_1 - X_{1,t}(\theta_0))^{m-1}1(X_{1,t}(\theta_0) \le x_1) - (x - X_{1,t}(\theta_0))^{m-1}1(X_{1,t}(\theta_0) \le x))^2 \\
&\qquad + ((x_1 - X_{2,t}(\theta_0))^{m-1}1(X_{2,t}(\theta_0) \le x_1) - (x - X_{2,t}(\theta_0))^{m-1}1(X_{2,t}(\theta_0) \le x))^2 \\
&\qquad + (m-1)!\; 2Cr^{2s_1}] \\
&\le \frac{8}{(m-1)!}\mathbf{E} \sup_{\substack{x_1\in\mathcal{X}\\|x_1-x|\le r}} [((x_1 - X_{1,t}(\theta_0))^{m-1} - (x - X_{1,t}(\theta_0))^{m-1})^2 + (1(X_{1,t}(\theta_0) \le x_1) - 1(X_{1,t}(\theta_0) \le x))^2 \\
&\qquad + ((x_1 - X_{2,t}(\theta_0))^{m-1} - (x - X_{2,t}(\theta_0))^{m-1})^2 + (1(X_{2,t}(\theta_0) \le x_1) - 1(X_{2,t}(\theta_0) \le x))^2 \\
&\qquad + (m-1)!\; Cr^{2s_1}] \\
&\le \frac{8}{(m-1)!}\mathbf{E} \sup_{\substack{x_1\in\mathcal{X}\\|x_1-x|\le r}} [C_1|x_1 - x| + 1(x < X_{1,t}(\theta_0) \le x_1) + 1(x < X_{2,t}(\theta_0) \le x_1) + (m-1)!\; Cr^{2s_1}] \\
&\le \frac{8}{(m-1)!}\mathbf{E}[C_1 r + 1(|X_{1,t}(\theta_0) - x| \le r) + 1(|X_{2,t}(\theta_0) - x| \le r) + (m-1)!\; Cr^{2s_1}] \\
&\le C_2 \cdot r \vee r^{2s_1},
\end{aligned}$$

where the first three inequalities holds by $(a+b)^2 \le 2(a^2+b^2)$ and Assumption 4 (c), the fourth 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 $N(\epsilon, \mathcal{F}) \le C_3 \cdot (1/\epsilon)^2$ for $\epsilon \in (0,1)$ and some $C_3 > 0$ 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). $\square$

(ii) Now, I prove Lemma B.1 (ii) in two steps.

(1) By rearranging terms, one can write

$$\begin{aligned}
\nu_T^{(m)*}(x) &= \sqrt{T}\left(\bar{F}_1^{(m)*}(x, \hat{\theta}^*) - \bar{F}_2^{(m)*}(x, \hat{\theta}^*)\right) - \left(\bar{F}_1^{(m)}(x, \hat{\theta}) - \bar{F}_2^{(m)}(x, \hat{\theta})\right) \\
&= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} h_{x,m}^{\Delta}(W_t^*, \hat{\theta}^*) - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} h_{x,m}^{\Delta}(W_t, \hat{\theta}) \\
&= \hat{\Gamma}_1(x) - \hat{\Gamma}_2(x) + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} h_{x,m}^{\Delta}(W_t^*, \hat{\theta}) - \frac{1}{\sqrt{T}} \sum_{t=1}^{T} h_{x,m}^{\Delta}(W_t, \hat{\theta}),
\end{aligned}$$

where $\hat{\Gamma}_k(x)$ is as defined in Assumption 6. Then, one can obtain

$$\nu_T^{(m)*}(x) = \sqrt{T}\left(\bar{F}_1^{(m)*}(x, \hat{\theta}) - \bar{F}_2^{(m)*}(x, \hat{\theta})\right) - \left(\bar{F}_1^{(m)}(x, \hat{\theta}) - \bar{F}_2^{(m)}(x, \hat{\theta})\right) + o_{p*}(1), \quad \text{(B.3)}$$

because, by Assumption 6,

$$\hat{\Gamma}_1(x) - \hat{\Gamma}_2(x) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left\{ \psi_x^{\Delta}(W_t^*, \hat{\theta}) - \frac{1}{T} \sum_{t=1}^{T} \psi_x^{\Delta}(W_t, \hat{\theta}) \right\} + o_{p*}(1).$$

For any decreasing sequence $\delta_T \to 0$, Assumption 3 (c) implies that

$$\sup_{(x,\theta) \in \mathcal{X} \times B_{\Theta}(\delta_T)} \left| \nu_T^{(m)*}(x; \theta) - \nu_T^{(m)*}(x; \theta_0) \right| = o_{p*}(1).$$

(2) Then, it only remains to show $\nu_T^{(m)*}(\cdot; \theta_0) \Rightarrow \nu_{1,2}^{(m)}(\cdot; \theta_0)$ in $l^{\infty}(\mathcal{X})$ conditional on $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2$ in probability. This follows from weak convergence results for Hilbert space valued random variables. Specifically, one can apply Theorem 3.1 of Politis and Romano (1994). First, $\{Z_t(\cdot) \equiv h_{\cdot,m}^{\Delta}(W_t, \theta_0) - \psi_{\cdot}^{\Delta}(W_t, \theta_0) : t = 1, \ldots, T\}$ is a stationary sequence of Hilbert space valued random variables which are bounded and satisfy the mixing condition $\sum_m \alpha_Z(m) < \infty$ by Assumption 2 (a). Second, Assumption 5 satisfies the condition related to the stationary resampling scheme. Thus, one obtains the desired result by applying the bootstrap central limit theorem.

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

$$P\left\{ [S_{T,j}]_+ = \left[\xi_j\left(\sqrt{T} d_{m,j}\right) \cdot S_{T,j} \right]_+ \right\} \to 1 \quad \text{(B.4)}$$

since I 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

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

and

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

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

$$\begin{aligned}[S_{T,j}]_+ &= \max\left\{S_{T,j}-\sqrt{T}d_{m,j}+\sqrt{T}d_{m,j},0\right\}\\ &\leq \max\left\{S_{T,j}-\sqrt{T}d_{m,j}-\kappa_{T,j},0\right\}\\ &\leq \max\left\{\left|S_{T,j}-\sqrt{T}d_{m,j}\right|-\kappa_{T,j},0\right\}.\end{aligned} \tag{B.5}$$

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 $1-\xi_j\left(\sqrt{T}d_{m,j}\right)=1$. Thus, I obtain

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

with probability approaching 1.

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

$$\begin{aligned}\left|S_{T,1}-\sqrt{T}d_{m,1}\right| = &\left|\int_{\mathcal{X}}\sqrt{T}\left\{\left[\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right]_+-\left[F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right]_+\right\}dx\right.\\ &\left.-\epsilon\int_{\mathcal{X}}\sqrt{T}\left\{\left|\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right|-\left|F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right|\right\}dx\right|\\ \leq &\int_{\mathcal{X}}\left|\sqrt{T}\left[\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right]_+\right|\\ &+\epsilon\int_{\mathcal{X}}\left|\sqrt{T}\left|\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right|\right|\\ \leq &(1+\epsilon)\cdot Q(\mathcal{X})\cdot\sup_{x\in\mathcal{X}}\left|\nu_T^{(m)}(x)\right|=O_p(1),\end{aligned} \tag{B.6}$$

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 $2\leq j\leq m$, one has

$$\begin{aligned}\left|S_{T,j}-\sqrt{T}d_{m,j}\right| &= \left|\sqrt{T}\left[\bar{F}_1^{(j)}(\overline{x},\hat{\theta})-\bar{F}_2^{(j)}(\overline{x},\hat{\theta})\right]-\left[F_1^{(j)}(\overline{x},\theta_0)-F_2^{(j)}(\overline{x},\theta_0)\right]\right|\\ &\leq \left|\nu_T^{(j)}(\overline{x})\right|=O_p(1),\end{aligned} \tag{B.7}$$

where the last equality holds by Lemma B.1 (i). Thus, I obtain the desired result. □

Proof of Lemma 2. Since the empirical processes $\nu_T^{(m)}(\cdot)$ is asymptotically tight by

Lemma B.1 (i), one has

$$P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right|>\hat{c}_T-c_{T,L}\right)\to 0$$
$$P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right|>c_{T,U}-\hat{c}_T\right)\to 0$$

by Assumption 4. Equivalently, one has

$$P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right|\le\hat{c}_T-c_{T,L}\right)\to 1$$
$$P\left(\sqrt{T}\sup_{x\in\mathcal{X}}\left|\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right|\le c_{T,U}-\hat{c}_T\right)\to 1.$$

Here, I only show the result for $\hat{\mathcal{C}}_0(\hat{c}_T)$. Let $x\in\mathcal{C}_0(c_{T,L})$. Then, by the triangular inequality,

$$\begin{aligned}\sqrt{T}\left|\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right|&\le\sqrt{T}\left|\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right|\\&\quad+\sqrt{T}\left|F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right|\\&\le(\hat{c}_T-c_{T,L})+c_{T,L}=\hat{c}_T\end{aligned}$$

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

$$\begin{aligned}\sqrt{T}\left|F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right|&\le\sqrt{T}\left|\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right|\\&\quad+\sqrt{T}\left|\left(\bar{F}_1^{(m)}(x,\hat{\theta})-\bar{F}_2^{(m)}(x,\hat{\theta})\right)-\left(F_1^{(m)}(x,\theta_0)-F_2^{(m)}(x,\theta_0)\right)\right|\\&\le\hat{c}_T+(c_{T,U}-\hat{c}_T)=c_{T,U}\end{aligned}$$

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

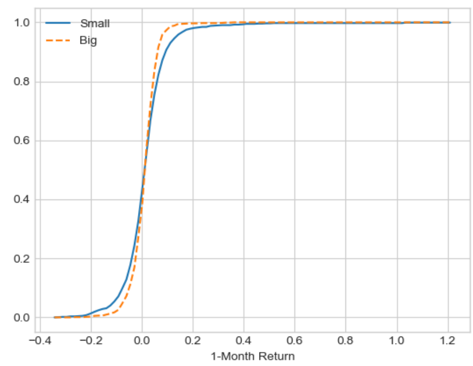
Similiarly, one can obtain

$$\begin{aligned}&P\left\{\mathcal{C}_0(c_{T,L})\subset\hat{\mathcal{C}}_0(\hat{c}_T)\subset\mathcal{C}_0(c_{T,U})\right\}\to 1\\&P\left\{\mathcal{C}_+(c_{T,U})\subset\hat{\mathcal{C}}_+(\hat{c}_T)\subset\mathcal{C}_+(c_{T,L})\right\}\to 1\\&P\left\{\mathcal{C}_-(c_{T,U})\subset\hat{\mathcal{C}}_-(\hat{c}_T)\subset\mathcal{C}_-(c_{T,L})\right\}\to 1.\end{aligned}\tag{B.8}$$

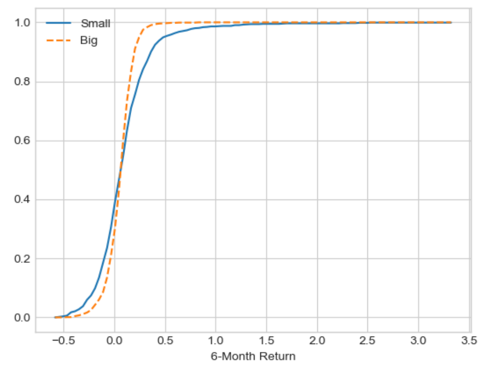
□

C EDFs and IEDFs of the Portfolio Returns

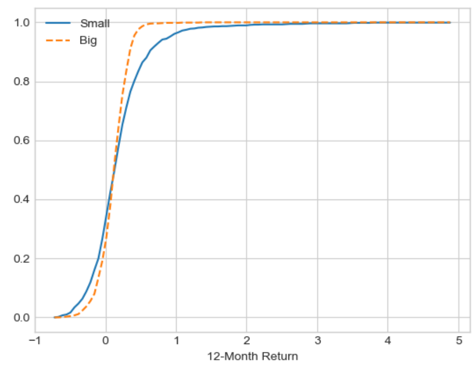
Figure C.1: EDFs of returns of small and big portfolios



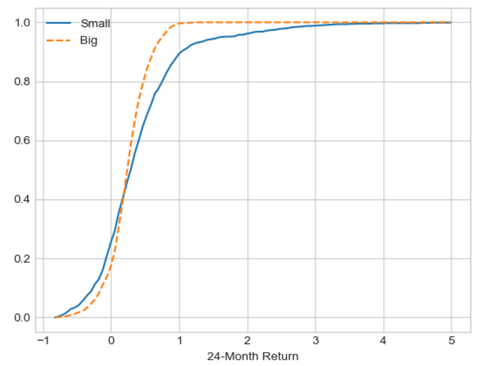
(a)



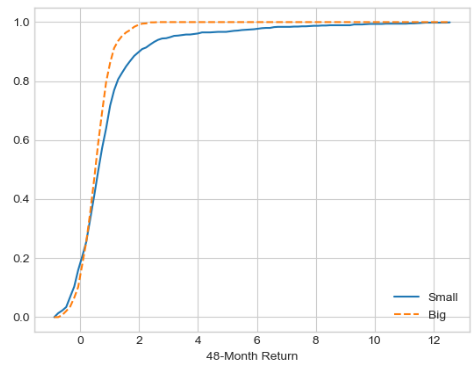
(b)



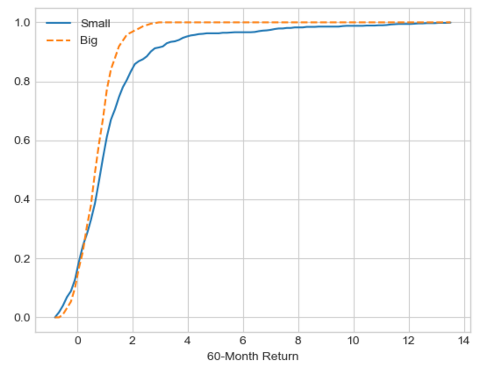
(c)



(d)

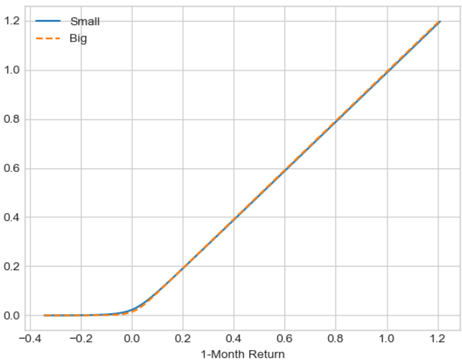


(e)

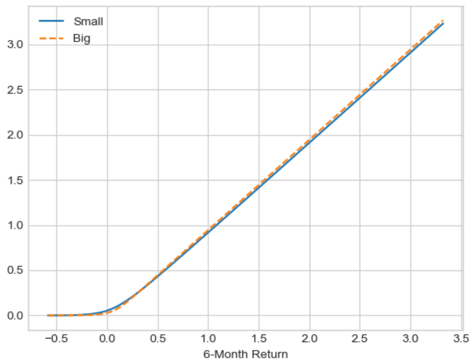


(f)

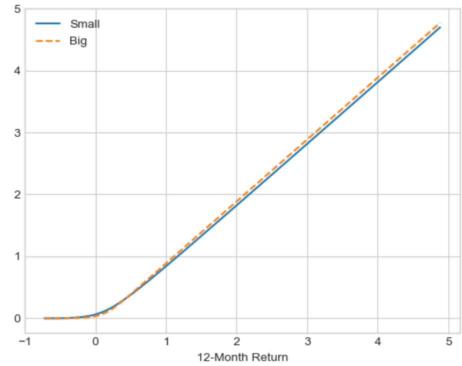
Figure C.2: IEDFs of returns of small and big portfolios



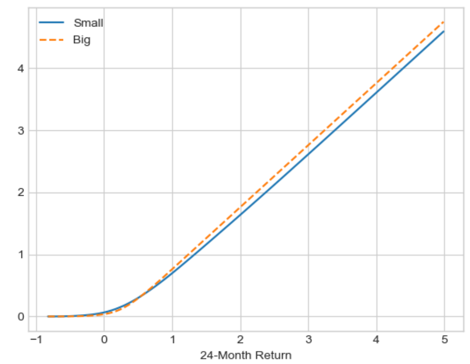
(a)



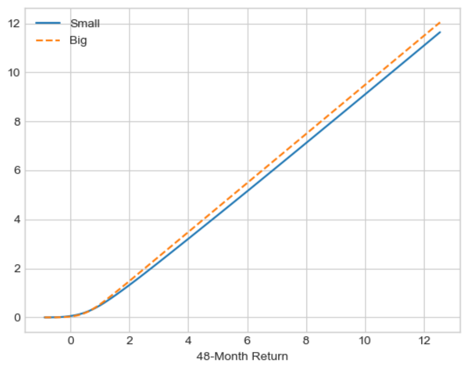
(b)



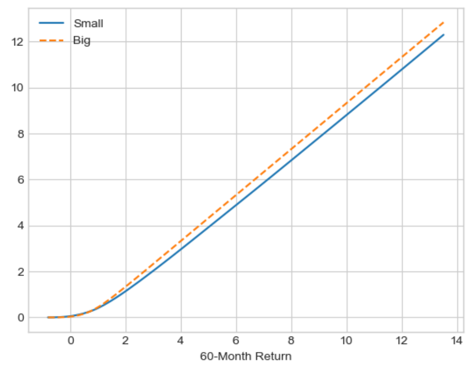
(c)



(d)

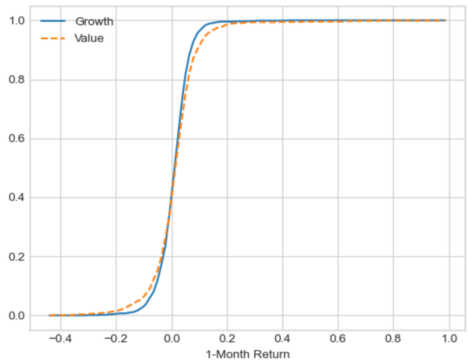


(e)

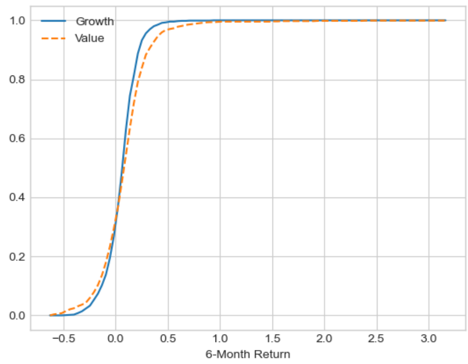


(f)

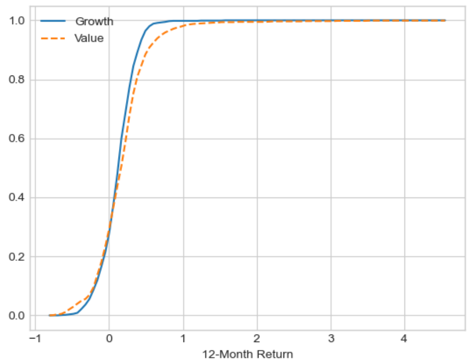
Figure C.3: EDFs of returns of growth and value portfolios



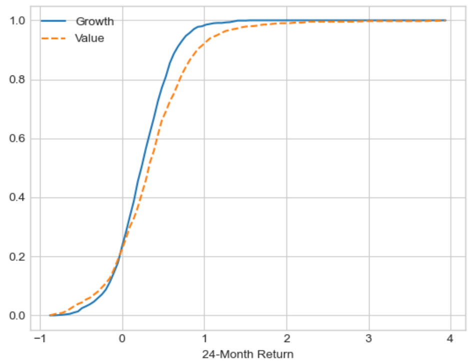
(a)



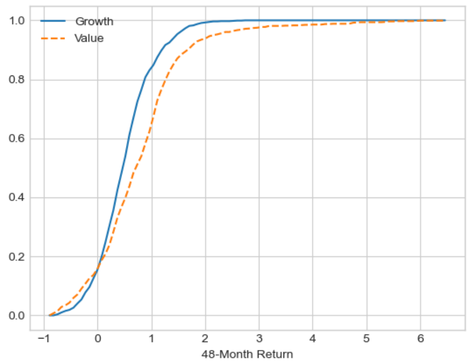
(b)



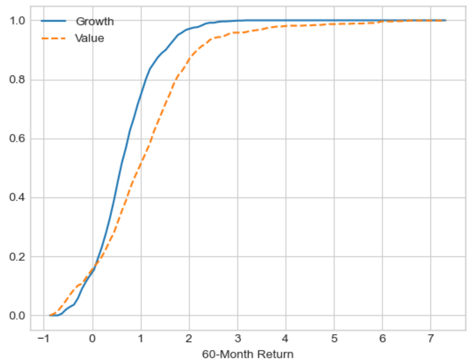
(c)



(d)

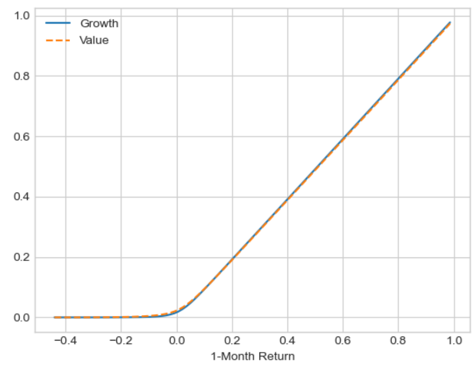


(e)

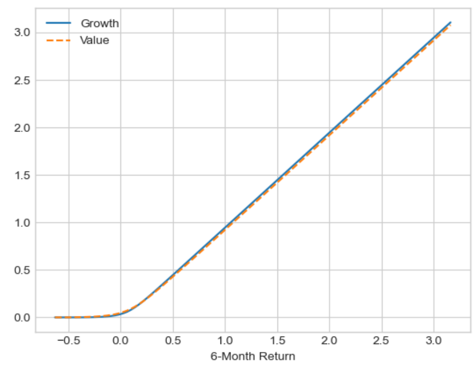


(f)

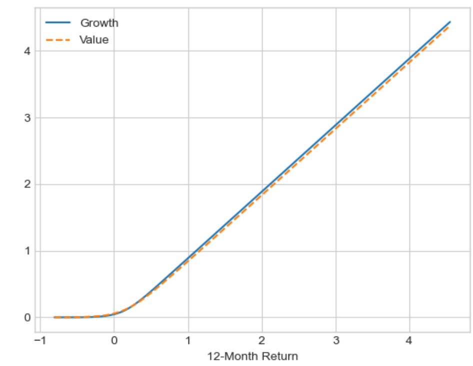
Figure C.4: IEDFs of returns of growth and value portfolios



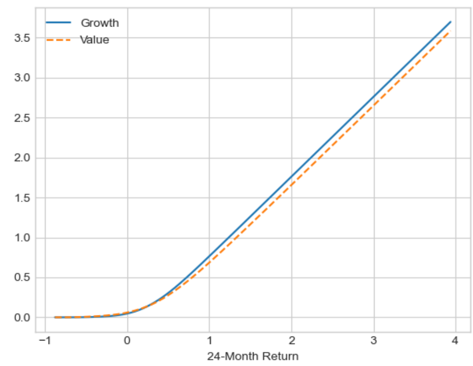
(a)



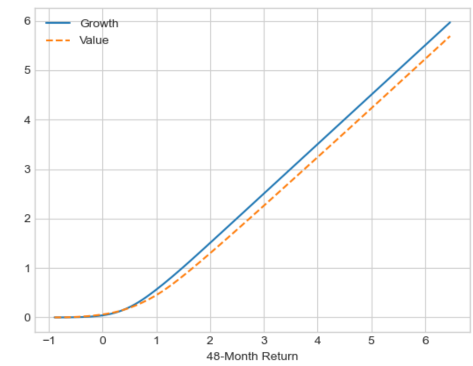
(b)



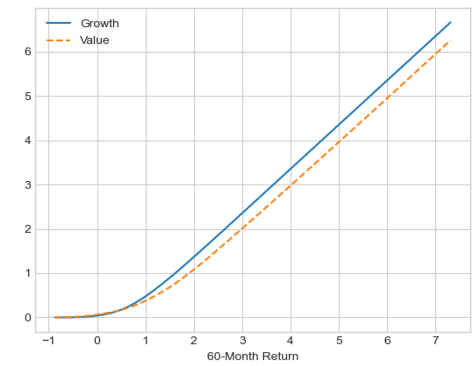
(c)



(d)

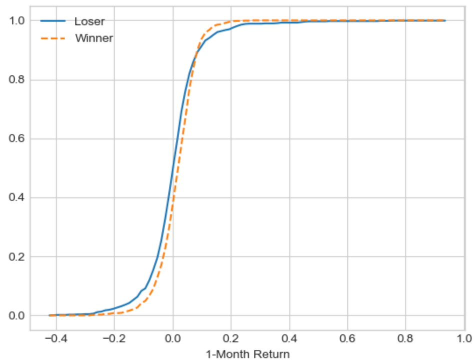


(e)

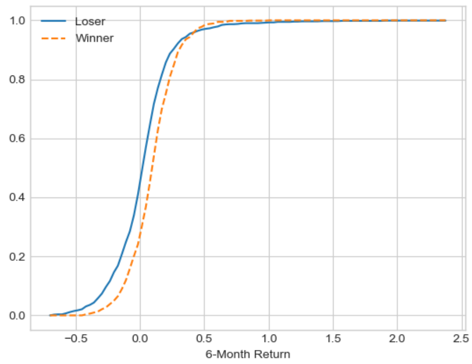


(f)

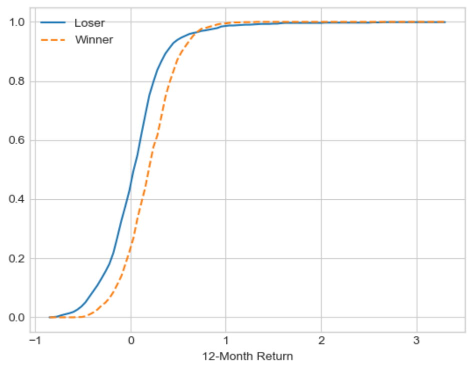
Figure C.5: EDFs of returns of (momentum) loser and winner portfolios



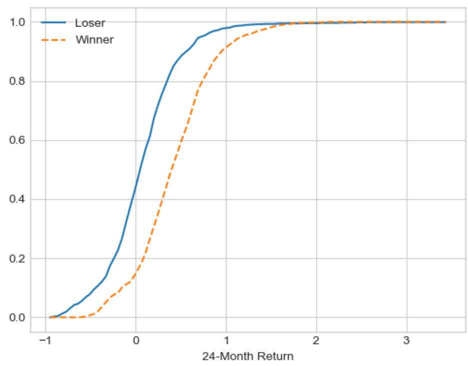
(a)



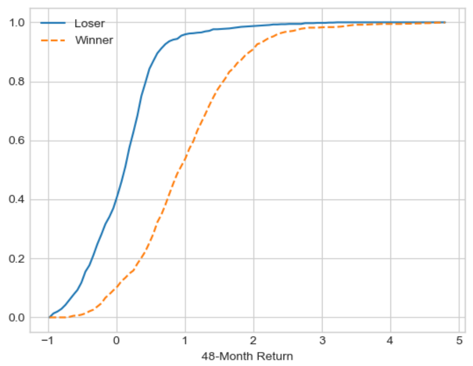
(b)



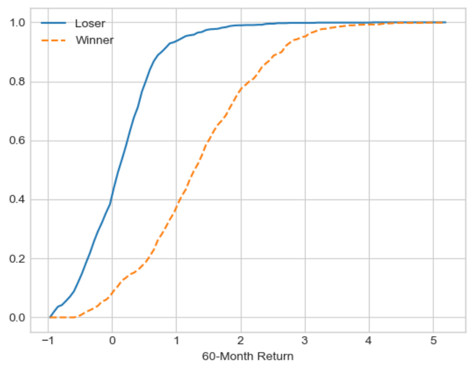
(c)



(d)

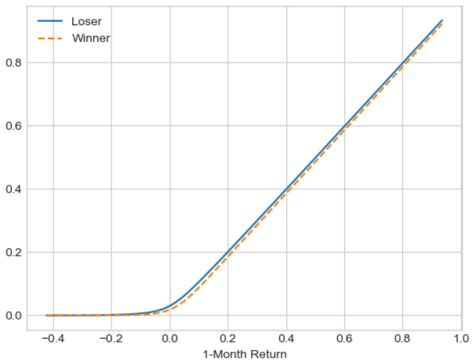


(e)

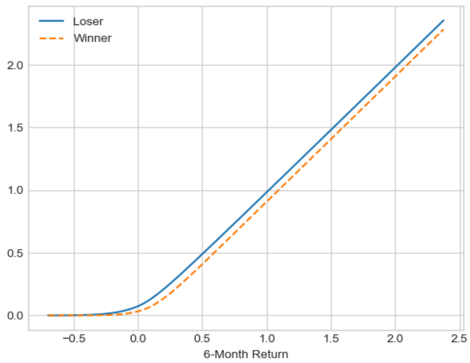


(f)

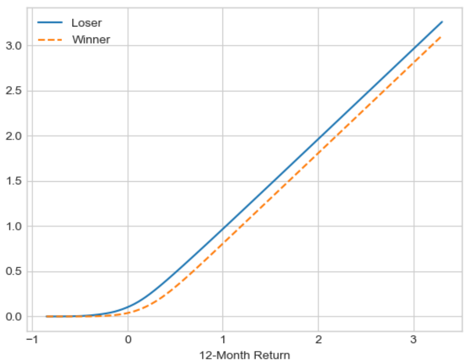
Figure C.6: IEDFs of returns of (momentum) loser and winnerg portfolios



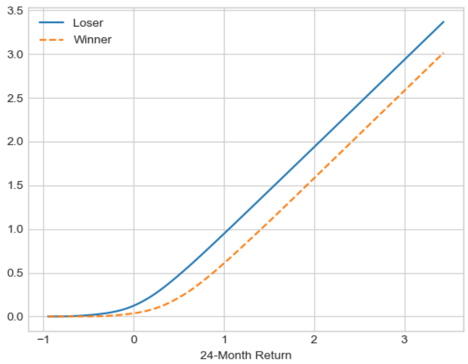
(a)



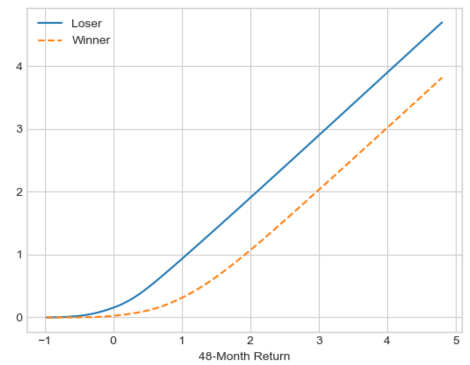
(b)



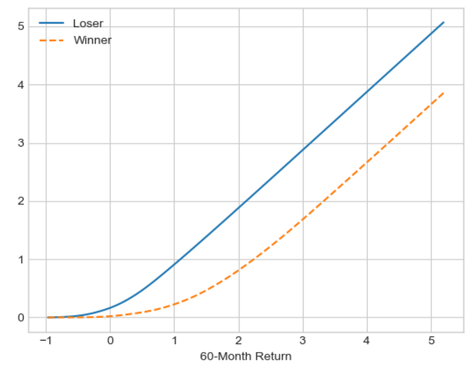
(c)



(d)

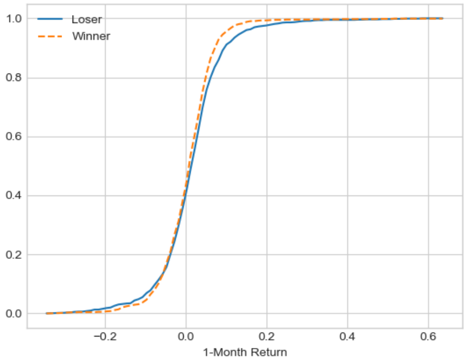


(e)

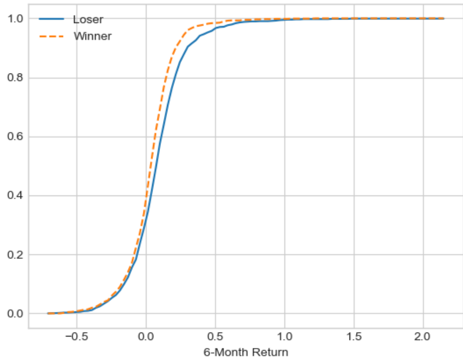


(f)

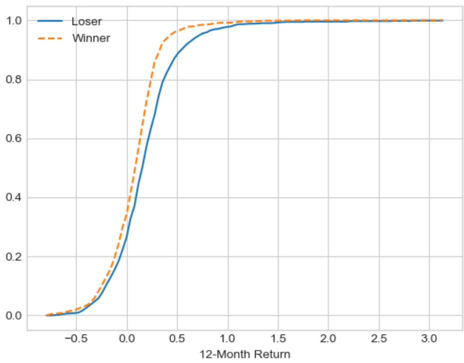
Figure C.7: EDFs of returns of (short-term reversal) loser and winner portfolios



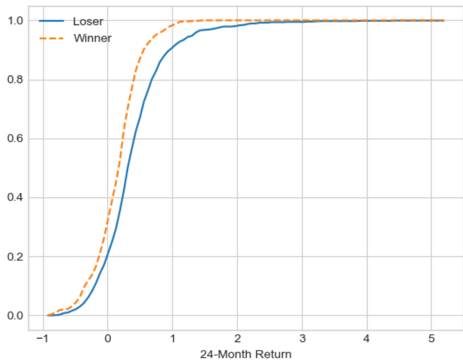
(a)



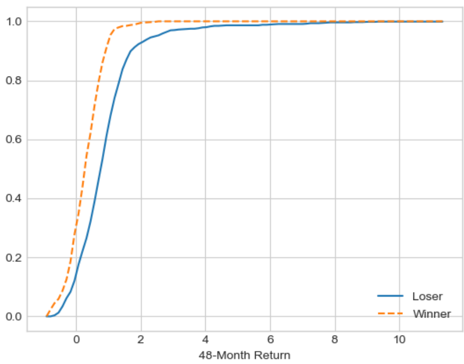
(b)



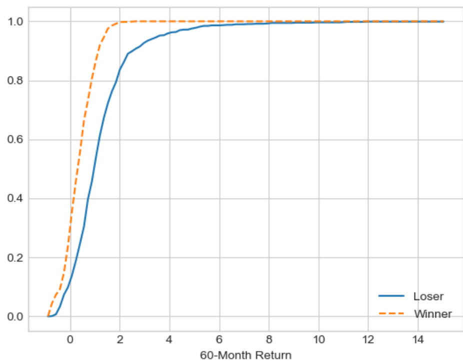
(c)



(d)

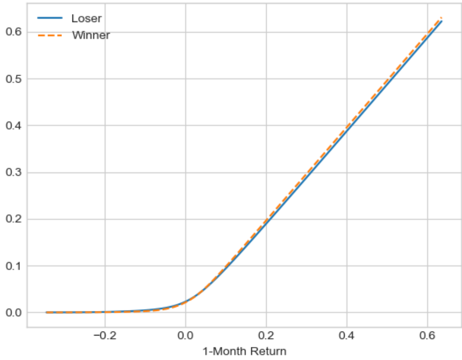


(e)

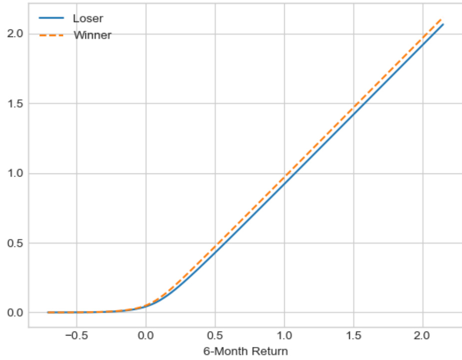


(f)

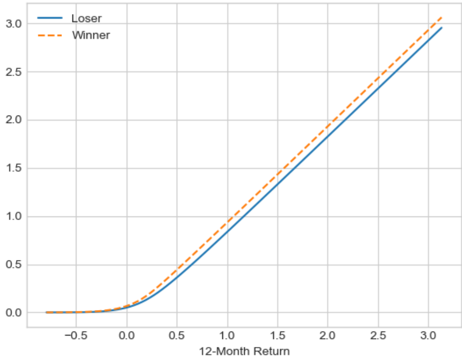
Figure C.8: IEDFs of returns of (short-term reversal) loser and winnerg portfolios



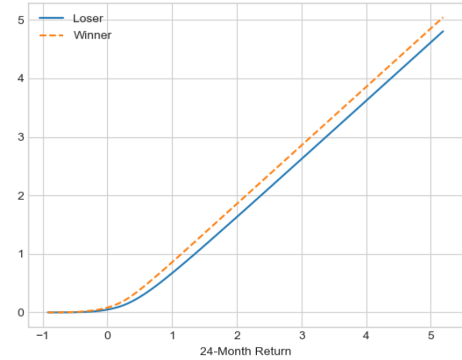
(a)



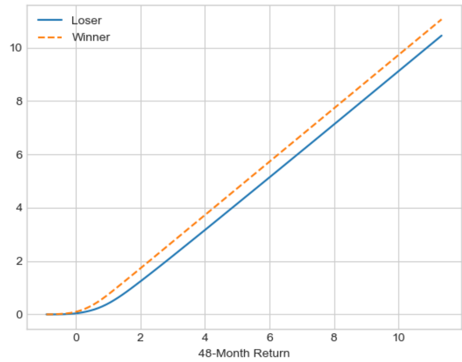
(b)



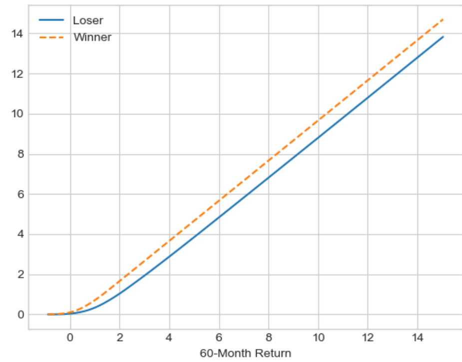
(c)



(d)

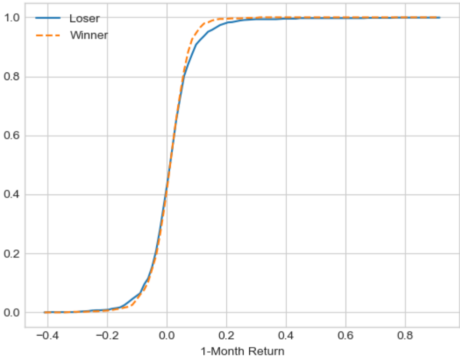


(e)

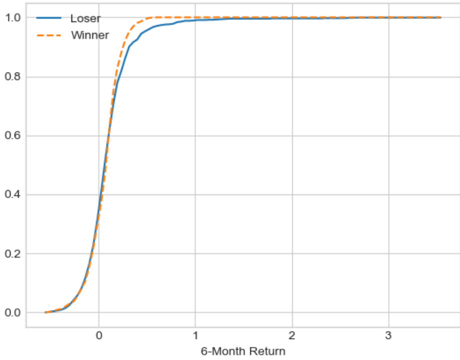


(f)

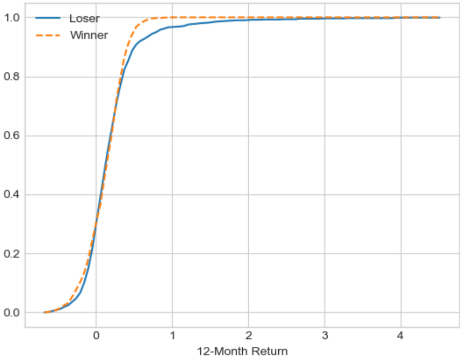
Figure C.9: EDFs of returns of (long-term reversal) loser and winner portfolios



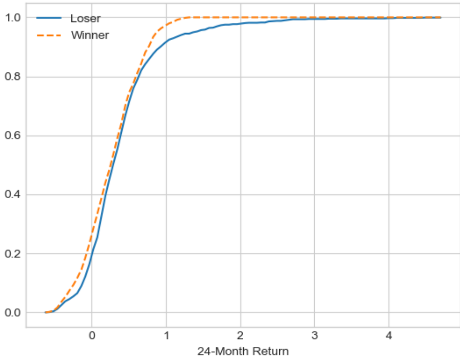
(a)



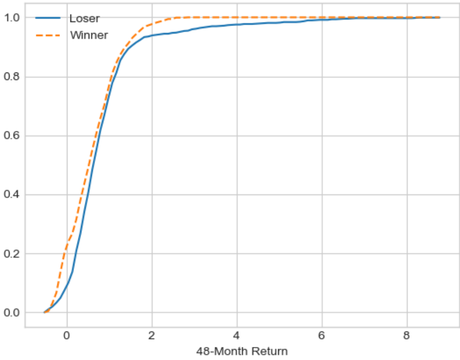
(b)



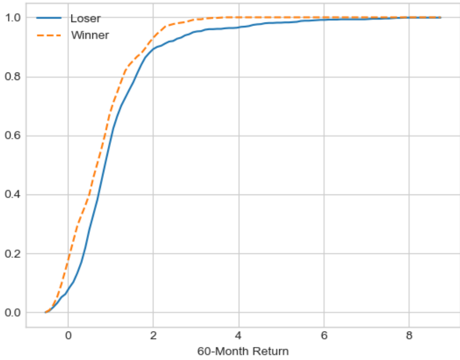
(c)



(d)

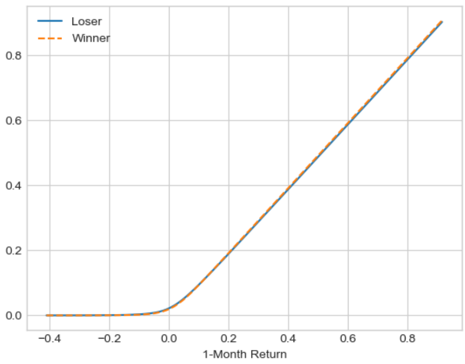


(e)

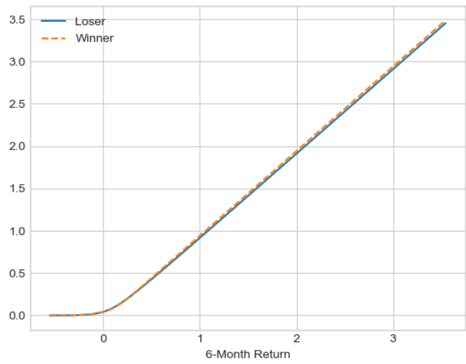


(f)

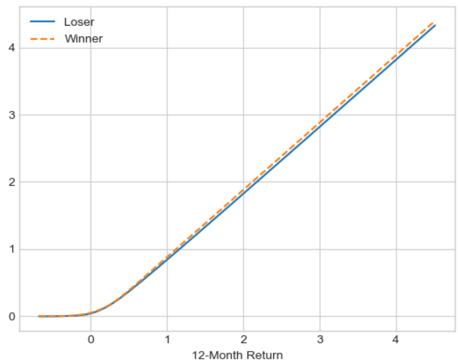
Figure C.10: IEDFs of returns of (long-term reversal) loser and winnerg portfolios



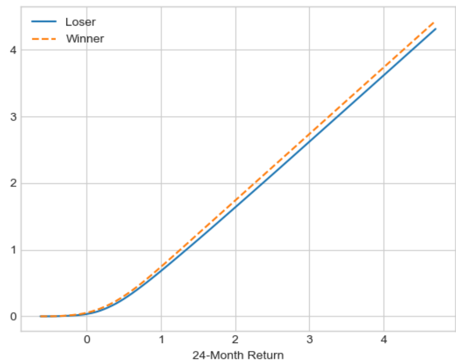
(a)



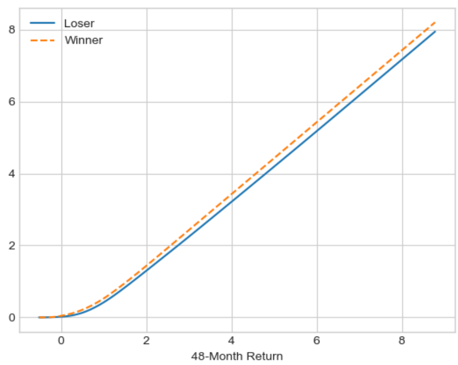
(b)



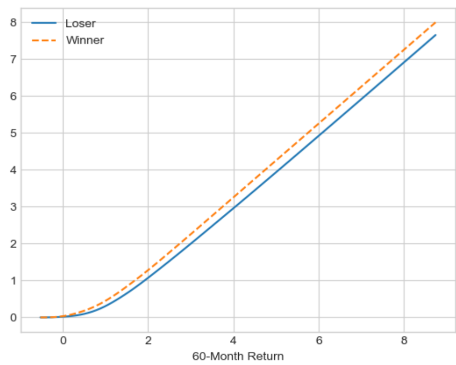
(c)



(d)



(e)



(f)

국문초록

본 논문에서는 알려지지 않은 모수의 존재하에서 완화된 확률적 지배관계 귀무가설을 검정하기 위한 새로운 비모수적 추론 방법을 제안한다. 전통적인 확률적 지배관계 규칙은 특정 효용함수 집단에 속하는 모든 의사 결정자에 대해 동일한 분포 순위를 매기는 것을 요구하는데, 이는 실제 경제 분석에서 제한적일 수 있다. 이러한 한계를 극복하기 위해 Leshno와 Levy (2002)는 효용함수 집단 내의 경제적으로 병적인 선호를 제거하여 모든 의사 결정자가 아닌 대부분의 의사 결정자에게 적용되는 완화된 확률적 지배관계 규칙을 도입했다. 완화된 확률적 지배관계 규칙은 투자 결정을 포함한 다양한 경제 문제에 응용된다. 이 논문에서는 경험적 분포 함수에 기반한 검정 통계량과 붓스트랩 방법을 이용하여 임계값을 계산할 수 있는 방법을 제시한다. 주식 시장 이상현상을 기반으로 한 포트폴리오의 수익 분포 간의 비교에 본 논문의 검정 방법을 적용한 결과, 주식 시장 이상현상을 활용하는 투자 전략이 기대 효용 극대화의 결과임이 나타났다.

주요어: 완화된 확률적 지배관계, 검정 일치성, 주식 시장 이상현상
학번: 2021-25134