Consistent Tests for Completely Monotone Stochastic Dominance*

Chris J. Bennett†

First Draft: January 27, 2007
This version: June 30, 2007

Abstract

We propose statistical tests for the hypothesis that infinite-degree stochastic dominance (ISD) holds between some pair of random variables. This work is motivated by the fact that a test for ISD can can be interpreted as a test of expected utility maximization of a completely monotone utility function—a rich class that includes logarithmic, exponential, and power utility. Thus our results allow for the testing of a fundamental assumption that, in one form or another, is pervasive in economics and finance. Infinite-degree stochastic dominance can also be viewed as a criterion for the ranking of various investment alternatives, and hence the test developed herein are useful when only empirical distributions associated with financial returns are observed. The statistical tests we propose are based on a *sup* statistic involving the difference between the one-sided empirical Laplace transforms of the distributions under consideration. We prove that the tests are consistent and use Monte Carlo methods to examine their statistical power and verify that the tests are correctly sized in finite samples. In an empirical application we use our method to test for infinite-degree stochastic dominance amongst major financial indices.

JEL Classification: C12, C15, G11
Keywords: Stochastic dominance, empirical process, bootstrap test

*I thank the members of my thesis committee, in particular John Knight and Stephen Sapp, for their comments and sugggestions on an earlier draft of this paper. Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

†Department of Economics, Social Science Centre, The University of Western Ontario, London, Ontario, N6A 5C2, Canada. E-mail: cbenne2@uwo.ca

1 Introduction

Stochastic dominance rules yield partial orderings over a set of random variables, or prospects, where these orderings are defined in terms of primitive conditions on the distribution functions of the random variables being compared. In the simplest case, due to Lehmann (1955), a prospect X is said to first-degree (weakly) stochastically dominate another prospect Y if $F(x) \leq G(x)$ for all $x \geq 0$, where F and G denote the cumulative distribution functions of X and Y, respectively. Building on this work, Hadar and Russell (1969) proposed second-degree stochastic dominance (SSD) which defines a ranking in terms of the areas under the distribution functions. Formally, X is said to second-degree (weakly) stochastically dominate Y if, for all $x \geq 0$, $\int_0^x F(x) \leq \int_0^x G(x)$. Third-, fourth-, and the general case of n-th degree stochastic dominance have since been introduced and examined (see Whitmore (1989) for the development of n-th degree stochastic dominance). Conceptually, each of the rules provides some measure of the extent to which a given distribution associates high (low) probability with a high (low) outcome of a random variable.

Economic interest in stochastic dominance is due primarily to its relationship with expected utility maximization. For example in Quirk and Saposnik (1962) it is shown that a distribution is first-degree strictly dominated if and only if it is inadmissible for all expected utility maximizers with monotonic preferences. In the case of second-degree stochastic dominance, Hadar and Russell (1969) established that SSD orders a set of probability distributions in exactly the same way as expected utility for weakly risk averse individuals with monotonic preferences. More generally, it has been shown (see, for example, Whitmore, 1989) that the ranking of distributions implied by n-th degree stochastic dominance is consistent with the ranking provided by expected utility for the class of all utility functions for which $(-1)^{(i+1)} u^{(i)}(x) > 0$, for every $i \in \{1, \ldots, n\}$, where $u^{(i)}$ denotes the i-th derivative of u. For an extensive survey of these and other results concerning stochastic dominance the reader is referred to Levy (1992).

In this paper we concentrate on the weakest[1] notion of stochastic dominance, namely the limiting case of infinite-degree stochastic dominance (ISD). Brockett and Golden (1987), Whitmore (1989), and Thistle (1993) have shown that ISD can be characterized by a simple condition involving the difference between the one-sided Laplace transforms of the distributions being compared, namely X infinite-degree (weakly) dominates Y if and only if, for all $t \geq 0$, $\int_0^\infty e^{-tx} dF(x) \leq \int_0^\infty e^{-tx} dG(x)$. These authors have also shown that ISD orders distributions in exactly the same way as expected utility for the class of all completely monotone utility functions. The completely monotone class of utility functions consists of those utility functions which are both infinitely differentiable and for which the derivatives alternate in sign. Examples include logarithmic and power utility, both of which are common in Macroeconomics and Finance.

[1] Weakest in the sense that that it is implied by any lower order of stochastic dominance

Like all orderings based of stochastic dominance, ISD suffers from the practical limitation that comparisons must often be made based on finite samples, or when only empirical distributions are observed. Having to make comparisons based on estimated Laplace transforms as opposed to their population counterparts motivates the development of a statistical test of infinite order stochastic dominance. Our contribution in this paper is the development of such a procedure. Specifically, we propose statistical tests for the hypothesis that infinite-degree stochastic dominance (ISD) holds between some pair of random variables. We allow for the observations to be serially dependent and for general dependence amongst the prospects, and we also treat the case of mutually and temporally independent samples of different sizes. Bootstrap methods are proposed to obtain the critical values for the tests, and we show that the resulting tests are consistent against general alternatives.

There are several reasons to be specifically interested in a test of infinite-degree stochastic dominance. First, a test for maximality using infinite-order stochastic dominance can be interpreted as a test of maximization of a completely monotone function.[2] Since the class of completely monotone functions includes many of the commonly assumed representations of preferences, such as the constant relative risk aversion utility functions, a test of ISD furnishes a test of the assumption that an agent is an expected (completely monotone) utility maximizer, an assumption that in one form or another is pervasive in economics.

We also argue that infinite degree stochastic dominance is a reasonable criterion for use by managers in determining the efficient set from within the feasible set of risky investment alternatives. Our argument is based on the following two points: (i) The ranking obtained under infinite degree stochastic dominance is equivalent to that obtained under expected utility for the class of completely monotonic utility functions (c.f. Brockett and Golden (1987), Whitmore (1989), and Thistle (1993))—a rich class of utility functions that includes logarithmic, exponential, and power utility. Any manager with preferences that are consistent with those implied by a completely proper utility function will therefore not exclude any expected utility efficient portfolio by ranking according to infinite-degree stochastic dominance; (ii) Ultimately, any portfolio which is selected is not going to be optimal for all investors in the fund, and without a complete specification of any investor's utility function there will be no way to determine for which particular individuals the selection is suboptimal. If, for a random participant in the fund with preferences not characterized by a completely monotone utility function, a fund manager assigns equal prior probability to this participant agreeing or disagreeing with the elimination of a dominated prospect, then it is rational in terms of an expected majority preference (which is made precise in Section 3) for the manager to select the efficient set based on infinite-degree stochastic dominance. Given that many managers are compensated by the nominal value of a fund—which is directly proportional to the number of participants in a fund—we find this to be a compelling

[2]McFadden (1989) makes this point in his paper, but leaves open the question of an appropriate test statistic.

argument for the use of ISD in practice.

Additionally, it is well-known that infinite-degree stochastic dominance is necessary for any lower order of stochastic dominance; that is, a rejection of ISD implies a rejection of stochastic dominance at any lower order. Thus, although interest often centers on direct tests for, say, first- through third-degree stochastic dominance, a single application of an ISD test can potentially rule out dominance at these orders. This point becomes more interesting when we take into account that the ISD test proposed herein jointly tests *all* of the restrictions implied by stochastic dominance, not a (compact) subset of the implied restrictions, and yet is of the same order of complexity as McFadden's (1989) test for first-order stochastic dominance. In other words, the ISD test is simple to implement relative to higher finite-order tests and is consistent when the bounded support assumption, which is necessary in most testing procedures for lower order dominance, is relaxed.

The rest of the paper is organized as follows. In the next section we define infinite-degree stochastic dominance and and review some of the findings in this literature. It is in this section that we formalize our arguments for the use of infinite degree SD in finance. In Section 3, we present our test for ISD and develop the asymptotic theory. Section 4 offers some Monte Carlo evidence concerning the finite sample properties of our test and also contains an application of the test to the ranking of various financial market indicies. Section 5 concludes.

2 Infinite-Degree Stochastic Dominance

Infinite-degree stochastic dominance has been studied by Brockett and Golden (1987), Whitmore (1989), and Thistle (1993). In this section we review some of their findings. This review serves to motivate the development of a test for ISD as well as provide the reader with the appropriate background.

We begin with a definition. Let F_n denote the nth integral of a distribution F; that is

$$F_n = \int F_{n-1}(x)dx, \quad n \geq 1,$$

with the convention that $F = F_0$. The rule for nth degree stochastic dominance is as follows:

Definition 1. *F degree-n stochastically dominates G iff*

1. $F_{n-1}(x) \leq G_{n-1}(x)$, $\forall x \in X$, *with* $F_{n-1}(x) < G_{n-1}(x)$ *for some* x*; and, for* $n \geq 3$,
2. $F_k(x^*) \leq G_k(x^*)$, $k = 1, 2, \ldots, n-2$.

Let FD_nG denote that distribution F nth degree stochastically dominates distribution G. The recursive nature of the definition makes it clear that stochastic dominance of degree n implies FD_rG for all $r > n$. Moreover, stochastic dominance of degree n for any finite n implies infinite degree stochastic dominance, where infinite degree stochastic dominance, denoted by $FD_\infty G$, is defined by letting $n \to \infty$ in the definition. Important for our purposes later on is the contrapositive statement; that is, if $\neg FD_\infty G$ then $(\forall n)\neg FD_n G$. Hence, any test that rejects infinite order stochastic dominance also rejects stochastic dominance of any other degree. We also note that in general higher order stochastic dominance does not imply lower order stochastic dominance. A partial converse is available, however, in the case of infinite degree stochastic dominance. Specifically, $FD_\infty G$ implies the existence of a finite n such that $FD_n G$ (Thistle, 1993).

In the rest of the paper we assume that utility is defined on $\mathbb{R}_+ = [0, \infty)$. This is not a restrictive assumption when our utility function is defined over, say, wealth or financial (gross) returns. We also assume that utility is differentiable and denote the derivative of u by u'.

Let $\Im$ denote the class of bounded nondecreasing functions and define the functional $T : \Im \to \mathcal{C}([0, \infty))$ by

$$T(\psi) = \int_0^\infty e^{-tx} d\psi(t) \tag{1}$$

Further, define

$$\mathcal{M} = \{T(\psi) | \psi \in \Im\}$$

Lemma 1. *(Bernstein's Theorem) Marginal utility u' is completely monotone on $\mathbb{R}_+$ if and only if $u' \in \mathcal{M}$.*

For an arbitrary $u' \in \mathcal{M}$, say

$$u'(x) = \int_0^\infty e^{-tx} d\psi(t)$$

we can recover a utility function representing the underlying preferences simply as

$$\begin{aligned} u(x) &= \int_0^\infty u'(x)dx && (2) \\ &= \int_0^\infty \{\int_0^\infty e^{-tx} d\psi(t)\}dx && (3) \\ &= \int_0^\infty t^{-1}(1 - e^{-tx})d\psi(t) && (4) \end{aligned}$$

Denote the set of all such utility functions by U_∞. It then follows that for arbitrary distributions F and G, and arbitrary $u \in U_\infty$ that $E_F[u(X)] > E_G[u(Y)]$ if and only if

$$\int_0^\infty \int_0^\infty t^{-1}(1 - e^{-tx})d\psi(t)dF(x) - \int_0^\infty \int_0^\infty t^{-1}(1 - e^{-tx})d\psi(t)dg(x) > 0,$$

or equivalently, if and only if

$$\int_0^\infty t^{-1}[M_F(-t) - M_G(-t)]d\psi(t) < 0$$

where $M_F(t) = E_F(e^{tX})$ is the Laplace transform associated with F.[3]. It is precisely this inequality that leads to Thistle (1993)'s Proposition 4 which forms the basis of our testing procedure[4]:

Proposition 1. *For all $u \in U_\infty$, $FD_\infty G$ if and only if $M_F(-t) < M_G(-t)$, $\forall t \in \mathbb{R}_+$.*

Notice that the one-sided Laplace transform condition for ISD is well defined for all possible pairwise comparisons. Indeed, that

$$E_F[e^{-tX}] < \infty$$

for any non-negative continuous random variable X with distribution F is a direct consequence of the fact that $(\forall x, t \in \mathbb{R}_+)$ $|e^{-tx}| \leq 1$ and the Lebesgue Dominated Convergence Theorem.

2.1 Infinite-Degree Stochastic Dominance in Finance

Stochastic dominance beyond third-degree has received relatively little attention is the finance literature. In this section we discuss some of the reasons for this aversion to higher degree SD, and argue that infinite-degree SD is a reasonable criterion under certain circumstances.

2.1.1 Behaviour characterized by U_∞

The usefulness of stochastic dominance, like any other decision rule, is related the size of the efficient set implied by the ordering. Letting E_n denote the efficient set associated with n-th degree SD, it is well known that E_n is monotonically decreasing in n. It is also well known that the duality between SD and expected utility holds for more restrictive classes of utility as n increases. Thus a tradeoff occurs between our desire for smaller efficient sets and our willingness to adopt more restrictive assumptions regarding investor behaviour. In the case of infinite-degree SD we have $FD_\infty G$ if and only if $F \succcurlyeq G$ for every investor with completely monotonic utility, the set U^∞.

All utility functions in the class U^∞ are mixtures of exponential utilities (c.f. equation (2)). Mixtures of exponential utilities have many desirable properties (see Pratt 1964, Pratt and Zeckhauser 1984, Schlaifer 1969, 1971, Fishburn 1980) and include many commonly used functional forms such the family of isoelastic utility functions.

Pratt and Zeckhauser (1984) have examined the set of utility functions possessing the so-called "proper risk aversion property". This is the set of utility function having the property that whenever each of two or more lottery choices is found individually undesirable, the lotteries are also found to be undesirable when taken together. Pratt and Zeckhauser

[3] $M_F(t)$ is also referred to as the moment generating function associated with F in situations where $M_F(t)$ exists in some neighbourhood of 0. See Billingsley (1978) p. 241-243 for more details

[4] This proposition is also stated and proved independently in Brockett and Golden (1987)

(1984) show that all such "proper risk aversion" utilities enjoy the decreasing absolute risk aversion property. Additionally, they show that U^{∞} is a subset of the class of proper risk averse utilities, and, consequently inherits the decreasing absolute risk aversion property.

Motivated by the potential inadequacy of the mean-variance approximation to expected utility, Scott and Horvath (1980) obtained a characterization of U_{∞} by explicitly considering the direction of preference for higher order moments. Recall that the usual risk averse investor is assumed to have a utility function satisfying

$$U'(w) > 0,$$

and

$$U''(w) < 0,$$

for all w. The authors show that, in addition to these two assumptions, that strict consistency of preference[5] for the nth moment is a sufficient condition for $u \in U_{\infty}$. It is further shown that investors who are not strictly consistent in preference direction must exhibit on average negative preference for even central moments and positive preference for odd central moments. This last result could be used to argue that the efficient set implied by ISD approximates[6] the efficient set of an arbitrary risk-averse investor.

2.1.2 Computational Concerns

With the notable exception of Tehranian (1980), empirical work on stochastic dominance has been primarily concentrated on testing for first- through third-degree stochastic dominance. Part of the aversion to the testing down procedure employed by Tehranian (1980) is that the complexity of the testing procedure for nth degree SD tends to be an increasing function of n. The limiting case of ISD, however, is of the same complexity as testing for FDS and thus complexity is not a valid argument against using ISD. Moreover, if one is considering a testing down procedure and the definitive ranking of risky alternatives occurs at degree n, it is more efficient to consider a single application of the ISD criterion as the end result is identical but the computational cost is substantially reduced.

[5] An investor who is strictly consistent in preference direction for the n-th moment has a utility function for which:

$$\begin{aligned} U^n(w) &> 0 \quad \forall w, \\ U^n(w) &= 0 \quad \forall w, \text{ or} \\ U^n(w) &< 0 \quad \forall w. \end{aligned}$$

[6] Approximates in the sense that there is significant overlap in the efficient sets.

2.1.3 Invariance under Affine Transformations

In some situations it may be relevant to know the impact that a transformation of a random variable has upon the ranking implied by stochastic dominance. For instance, how is ISD affected when different marginal tax rates lead to different after-tax return distributions for otherwise identical individuals.

Levy and Sarnat (1971b) prove that if X dominates Y by the various SD rules, then αX also dominates αY by the corresponding SD rule where $\alpha > 0$. Hadar and Russell (1974) analyze the condition under which the transformation $aX + b$ dominates the original distribution X. Meyer (1989) and Brooks and Levy (1989) deal with very general transformations but only consider first- and second-degree stochastic dominance.

Let X and Y be random variables with distributions F and G. The following theorem shows that weak dominance is preserved under affine transformations.

Theorem 1. *Let $T(z) = a + bz$ for $(a, b) \in \mathbb{R} \times \mathbb{R}_+$. If X weakly dominates Y, then $T(X)$ weakly dominates $T(Y)$.*

Proof of Theorem 1. The proof is a straightforward consequence of the monotonicity of the exponential function.

□

2.1.4 Rationality of ISD Efficiency

The purpose of this section is to provide theoretical justification for the use of infinite-degree stochastic dominance (ISD) in the investment decision process. We accomplish this by showing that it is rational for a manager whose objective is investor-attrition minimization to pare down the feasible set of investment alternatives using the ISD criterion.

We consider a situation in which a manager is charged with selecting an investment from a predetermined feasible set of investment opportunities. This decision is made on behalf of N investors, and the manager's compensation is linked directly to the demand for participation in the investment group. We have in mind here examples such as an executive in a firm making investment decisions on behalf of the owners (shareholders), or a manager of a large mutual fund selecting an investment portfolio on behalf of the fund participants.

The manager does not observe individual utility functions, but he knows that each individual is risk averse and has smooth[7] preferences that are monotonic in terminal wealth or returns. Formally, we define U^2 to be the space of all twice differentiable and monotonically increasing (ordinal) utility functions whose first two derivatives alternate in sign. The manager's knowledge about the investors is then summarized by $u_i \in U^2$ for every $i \in \{1, \dots, N\}$.

We define $\mathcal{D}$ to be the space of distribution functions over terminal wealth (returns). That is, each distribution in $\mathcal{D}$ corresponds to the distribution of returns associated with a

[7]Here "smooth" is used in the sense that preferences are representable by a continuous utility function

particular investment opportunity from within the feasible set. For all $F, G \in \mathcal{D}$ we write FD_nG if F degree-n stochastically dominates G, and $F \succcurlyeq_A G$ if $E_F[u(X)] \geq E_G[u(Y)]$ for every $u \in A$ with strict inequality for some $u \in A$.

Before proceeding we state as lemmas some results that will be used in our argument below. Let $\{U^n\}_{n=2}^{\infty}$ be a sequence of monotonically decreasing sets, i.e. $U^{n+1} \subset U^n$. We make use of the following well-known result:

Lemma 2. *The sequence $\{U^n\}_{n=2}^{\infty}$ can be chosen such that, for every $n \geq 2$, FD_nG if and only if $F \succcurlyeq_{U^n} G$*

For a given pair of prospects, say $F, G \in \mathcal{D}$, and for a fixed $n > 2$ define the partition

$$\mathcal{P}^n_{\succcurlyeq}(F, G) = \{u \in U^2/U^n : F \succcurlyeq G\},$$

and

$$\mathcal{P}^n_{\prec}(F, G) = \{u \in U^2/U^n : G \succ F\}.$$

The partition splits the collection of individuals whose preferences are not represented by a utility function in U^n into those who weakly prefer F to G, and those who strictly prefer G to F.

Let $\mu : \mathcal{B}(U^2) \to [0, 1]$ be a probability measure on the space of functions U^2, where $\mathcal{B}(U^2)$ is the collection of Borel subsets of U^2. The measure μ represents the manager's beliefs about the distribution of preferences in the population of investors U^2. We make the following assumption about beliefs

Assumption 1. $\frac{\mu(U^{n+1})}{\mu(U^n)} \in (0, 1)$ *for every $n \geq 1$.*

We also make the following assumption about beliefs:

Assumption 2. *(Symmetric Prior) Given FD_nG and $\neg FD_{n-k}G$ for $1 \leq k \leq n-1$, $\mu(\mathcal{P}^n_{\succcurlyeq}) = \mu(\mathcal{P}^n_{\prec})$. If $\neg FD_nG$ and $\neg GD_nF$ for all n, then $\mu(\mathcal{P}^2_{\succcurlyeq}) = \mu(\mathcal{P}^2_{\prec})$.*

The first part of assumption (2) requires the manager to believe that it is equally likely that a randomly selected investor with $u \in U^0/U^n$ will agree or disagree with the ranking provided by n-th degree stochastic dominance. Since the manager knows only that a randomly selected investor's preferences can be represented by a utility function $u \in U^2$, there is no reason to believe that preferential treatment would be given to either set. In other words, we justify this assumption by appealing to the principle of insufficient reason. The second part of this assumption rules out any information content in the case of no stochastic dominance.

Define the function $V : U^0 \times \mathcal{D}^2 \to \mathbb{R}$ by

$$V(u; F, G) = \begin{cases} -1, & \text{if } G \succ_{\{u\}} F \\ 1, & \text{if } F \succcurlyeq_{\{u\}} G \end{cases}$$

The function V can be viewed as an indirect utility function over the space U^2 that is induced by the pair (F, G). Intuitively, $V(u; F, G)$ records the vote of an individual with utility function u who is asked whether they weakly prefer F to G (+1), or strictly prefer G to F (-1).

Theorem 2 (Expected Majority Preference)**.** *Suppose assumptions (1) and (2) hold. Then*

$$FD_nG \quad \textit{for some } n \geq 2 \Rightarrow \int V(u; F, G)d\mu > 0$$

Proof of Theorem (2). First, we state as lemmas two well-known results from the stochastic dominance literature:

Lemma 3. *If FD_nG, then FD_mG for every $m > n$.*

Lemma 4 (Thistle, 1993)**.** *If $FD_\infty G$, then there exists some finite n such that FD_nG.*

Taken together, Lemmas (3) and (4) imply that it is sufficient to prove that for a finite n, FD_nG implies that $\int V(u; F, G)d\mu > 0$. It follows from Assumption (2)(i) that

$$\int V(u; F, G)d\mu > 0 = \mu(U^n).$$

That $\mu(U^n)$ has positive mass is an immediate consequence of Assumption (1). □

The theorem demonstrates that a portfolio manager with prior beliefs satisfying A1 and A2 will have an expected majority preference for the exclusion of a risky prospect if the risky prospect is stochastically dominated at some order. In other words, if a feasible investment option is stochastically dominated at any order and is discarded from consideration, the manager believes that the majority of investors will be in agreement with this action. Since, for any n, nth degree stochastic dominance implies ∞-degree SD it is suffices to consider only ∞-degree SD when forming the efficient set of investment alternatives. Note that the theorem says nothing about the ranking of distributions F and G when neither $FD_\infty G$ nor $GD_\infty F$ is true. In other words, the ordering is incomplete and hence the efficient set will consist of non-dominated return distributions as well as SD-incomparable distributions.

Constructing efficient sets according to infinite-order stochastic dominance clearly involves taking a normative stand. Under assumptions A1 and A2, the choice implied is precisely that which the manager believes makes an arbitrary member of the fund *ex-ante* least likely to be in disagreement with an exclusion from the efficient set. Since the fund manager will inevitably be forced to select a single portfolio from the efficient set—a choice which in itself requires a normative judgement since it will not be optimal for all investors in the fund—we view the use of ∞-SD to be an objective and transparent means of paring down the efficient set.

3 Statistical Tests for Infinite-Degree Stochastic Dominance

In most situations it is unreasonable to assume perfect knowledge of the distributions being compared. Instead, comparisons must be made when only empirical distributions are observed. In this section we develop statistical procedures to test the hypothesis that infinite-degree stochastic dominance holds between a pair of random variables. We first develop a test that is appropriate for prospects that exhibit temporal and mutual dependence, but have equal sample sizes. We then present a modified statistic for situations in which the prospects are mutually independent and we have different sample sizes for each prospect.

Throughout this section we consider tests based on a null of dominance; that is, if F and G are the distributions under consideration, we assume under the null either that $FD_\infty G$ or that $GD_\infty F$. Dominance under the null is the conventional approach to statistical tests of SD (see, for example, McFadden 1989 and Anderson 1996). The advantage of such a test in the case of ISD is that rejection of, say, $FD_\infty G$ leads to a rejection of $FD_n G$ for all n, therefore allowing us rule out dominance of F over G at any order. The disadvantage is that rejection of the null does not lead to any conclusive alternative. It could be the case, for instance, that rejection of the null implies that the distribution is dominated. Equally plausible, however, is that the distributions are not comparable. This would be the case whenever the Laplace transforms cross at some point. In principle, the role of F and G can be reversed and the resulting test can be combined with the original and used to distinguish between these alternatives. Alternatively, a testing procedure and decision rule can be developed to handle both cases simultaneously. Such a strategy is taken in Bishop, Formby, and Thistle (1992) and in Knight and Satchell (2006), and also discussed herein.

3.1 Testing when dominance is assumed under the null

Let X and Y be nonnegative random variables defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and let F and G denote the distributions of X and Y, respectively. From Proposition (1), the hypothesis that $FD_\infty G$ is equivalent to $M_F(-t) \leq M_G(-t)$ and $\neg F \equiv G$[8] . The null is compound since $FD_\infty G$ is true for many distributions F with G fixed; and the probability of rejecting the null hypothesis when it is true is greatest in the limiting case $F \equiv G$. Following McFadden (1989) we define the significance level of the test to be the supremum of the rejection probabilities for all cases satisfying the null. This has the effect of making the null hypothesis $H_0 : M_F(-t) \leq M_G(-t)$ for $t \in \mathbb{R}_+$, against $H_1 : M_F(-t) > M_G(-t)$, for some $t \in \mathbb{R}_+$, with the significance level equal to the probability of rejecting H_0 when $F \equiv G$.

Define the function $D : \mathbb{R}_+ \to [-1, 1]$ by $D(t) = M_F(-t) - M_G(-t)$. The definition of

[8] Throughout $M_F(-t)$ denotes the one-sided Laplace transform associated with the random variable $X \sim F$.

D permits us to reformulate the test as

$$H_0 : D(t) \leq 0 \text{ for all } t \in [0, \infty],$$
$$H_1 : D(t) > 0 \text{ for some } t \in [0, \infty].$$

Further, we note that D is continuous in t and

$$D(0) = 0 \text{ and } \lim_{t \to \infty} D(t) = 0.$$

Using these features, and letting $I = [0, \infty]$ we obtain the more convenient formulation

$$H_0 : \sup_{t \in \mathcal{I}} D(t) = 0, \tag{5}$$
$$H_1 : \sup_{t \in \mathcal{I}} D(t) > 0.$$

A natural test statistic for testing the hypothesis in (5) is the empirical analogue

$$S_n = \sup_{t \in \mathcal{I}} \sqrt{n} D_n(t), \tag{6}$$

where $D_n(t)$ is the difference between the empirical Laplace transforms, i.e.

$$D_n(t) = n^{-1} \sum_{i=1}^{n} [e^{-tX_i} - e^{-tY_i}].$$

Recalling that the supremum of $D(t)$ is nonnegative and equal to zero under the null, a large (positive) value of the statistic S_n can be interpreted as evidence in favour of rejecting the null.

It is worth noting that S_n jointly tests all of the restrictions of the hypothesis. An alternative approach which imposes a countable subset of the restrictions and for which the distributional properties are more easily characterized is to base the test statistic on the maximum over a finite grid of points. Indeed, under weak assumptions on the sampling process it is simple exercise to show that $\sqrt{n}(D_n(t_1), \ldots, D_n(t_k)) \to^d N_k(0, \Sigma)$ under the null. An application of the continuous mapping theorem (van der Vaart, 1998, p. 7) then yields $\sqrt{n} \max_{D_{t_i}} (D(t_1), \ldots, D(t_k)) \to^p \max_i (X_1, \ldots, X_k)$, where $(X_1, \ldots, X_k) \sim N_k(0, \Sigma)$. Thus we can fix $\mathbf{t} \in \mathbb{R}^n_{++}$ and then compare $\min_{t_i} D_n(\mathbf{t})$ against the critical value associated with the nth order statistic from a n-variate Normal distribution. This is the approach taken in Knight and Satchell (forthcoming). There are several drawbacks to this approach. First, if we wish to use even a moderately large dimensional vector (i.e. for moderately large n) difficulties arise in practice since the integral we are required to compute is of the same dimension. Additionally, there is no clear strategy on how to determine the points at which to evaluate the function $D_n(t)$, and this choice certainly has implications for the statistical power of the resulting test. The most important drawback, however, is that such a test introduces the possibility of test inconsistency as only a subset of the the restrictions implied by the null are actually imposed.[9]

In the remainder of the paper we concentrate solely on tests of the form (6) that impose all of the restrictions implied by the hypothesis of dominance. The challenge that remains, as usual, is the characterization of the sampling distribution of S_n. Here, some valuable insights can be gained from the existing literature on testing for lower orders of stochastic dominance. To illustrate the close parallel between testing for ISD and testing for lower orders of stochastic dominance we make use of the formulation in (5). Specifically, if we instead define $\mathcal{I} = [0, \bar{t}]$ and $D(t) = D^1(t) = F(t) - G(t)$, the hypothesis is precisely that of first-order stochastic dominance. Similarly, if we define $D^2(t) = \int_0^t F(s) - G(s)ds$, and $D^j(t) = \int_0^t D^{j-1}(s)ds$, then

$$H_0 : \sup_{t \in \mathcal{I}} D^j(t) = 0,$$

$$H_1 : \sup_{t \in \mathcal{I}} D^j(t) > 0.$$

corresponds to a test of j-th order stochastic dominance of F over G.

In the first paper to consider statistical tests for stochastic dominance when only empirical distributions are observed, McFadden (1989) develops tests for first and second-order stochastic dominance based on a Kolmogorov-Smirnov type statistic. In his work it is assumed that the samples are random and mutually independent. Exact distributional results are obtained for the sampling distribution of the test statistic in the case of first-degree dominance. Exact distributional results are not available for the KS type statistic at higher orders, and thus McFadden proposes a simulation procedure for obtaining the critical values.

Anderson (1996) proposes a test based on Pearson's chi-square test. Davidson and Duclos (2000) develop their test again based on a finite grid of points. They . Using recent developments in the empirical process literature, Barrett and Donald develop a consistent test which imposes all of the restrictions over any compact subset of $[0, \infty)$. In the context of mutually independent i.i.d. samples, they they derive the limiting distribution of the supremum statistic and establish the consistency of their test. Linton et al extend the work of Barrett and Donald by considering both mutual and temporal dependence. They propose a subsampling procedure for obtaining the critical values of the tests and also establish consistency under their assumptions.

The test statistic is computed as the supremmum over a compact subset of the difference of the empirical distribution functions, i.e. $EDF_F - EDF_G$, when F is assumed to weakly first-order stochastically dominate G under the null. The hull hypothesis is composite since the null hypothesis is true for many different G distributions with F fixed.

[9]This potential for test inconsistency also arises in the stochastic dominance tests of Anderson (1996) and Davidson and Duclos (2000)

In this paper, we consider only situations in which tests of the above hypothesis are carried out when only empirical distributions are observed. In other words, we assume that the investigator does not know (and is unwilling to assume) the parametric family of distributions to which either F or G belong. There are several key statistical issues that arise when conducting tests of this nature. First,

This test, under various assumptions about the sampling process and

which is considered in McFadden (1989), Anderson (1996), Davidson and Duclos (. Similarly,

The null imposes a continuum of restrictions, one for each $t \in \mathbb{R}_+$.

3.2 An Integral Test

While the KS type statistic is perhaps a natural starting point for developing a test of CMSD it is certainly not the only reasonable choice. An alternative approach that we consider in this section is based on computing the one-sided area between the curves under consideration. Our desire to explore the properties of this type of statistic is motivated by findings in the goodness-of-fit literature, where it is shown that tests which measure the difference over a range rather than the maximum difference at a single point often exhibit greater power.

Recalling the original statement of the null hypothesis we had

$$H_0 : D(t) \leq 0 \text{ for all } t \in [0, \infty],$$

$$H_1 : D(t) > 0 \text{ for some } t \in [0, \infty].$$

Defining the integral

$$\delta = \int_0^\infty D(t)\mathbf{1}_{D(t)\geq 0} dt$$

we can rewrite the null in the equivalent form

$$H_0 : \delta = 0,$$

$$H_1 : \delta > 0.$$

A natural test statistic in this case is

$$\delta_n = \sqrt{n} \int_0^\infty D_n(t)\mathbf{1}_{D_n(t)\geq 0} dt$$

where $D_n(t)$ is again defined as the difference between the empirical Laplace transforms.

3.3 Testing of independent samples of different sizes

There may be some situations in which it is appropriate to assume independence between the samples obtained from each distribution. This would be appropriate when comparisons

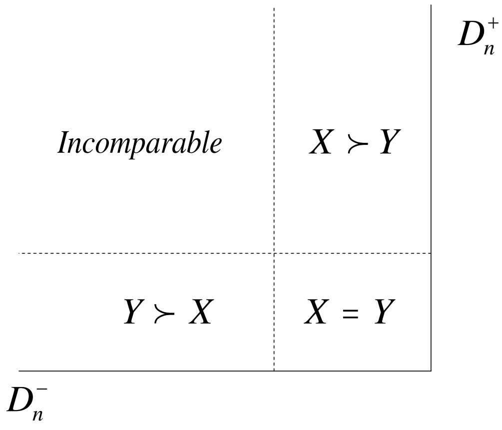

Figure 1: Decision Rule

are being made using cross-sectional data, as would be the case, for example, when income distributions are compared across countries at a given point in time. The case of independent samples can be treated without modification using the test statistic introduced above, except in situations where the sample sizes differ. In order to be able to handle such cases we now consider a simple modification of our original test statistic. Before presenting the modified statistic we first formally state our assumptions on the sampling process. These are given in assumption (3) below and are identical to the assumptions presented in Barrett and Donald (2003).

Assumption 3. *(i)* $\{X_i\}_{i=1}^N$ *and* $\{Y_i\}_{i=1}^M$ *are independent random samples from distributions with CDF's F and G;*

(ii) the sampling scheme is such that $N/(N+M) \to \alpha \in (0,1)$ *as* $N, M \to \infty$.

Define the functional $L(F)$ to be

$$L(F) = \int e^{-tx} F(dx),$$

the Laplace transform associated with distribution function F. The function $D_{n,m}(t)$ used to construct the test statistic for infinite-degree stochastic dominance can then be written

$$D_{n,m}(t) = \left(\frac{NM}{N+M}\right)^{1/2} [L(\hat{F}_N) - L(\hat{G}_M)]$$

The test is then carried out using $S_{n,m}(t) = \sup_t D_{n,m}(t)$.

3.4 Simple bidirectional tests

The test for ISD as formulated above is one in which F is assumed to (weakly) stochastically domianate G under the null. In practice, there may be no reason *a priori* to believe that a particular distribution dominates another and hence it may be more suitable to consider dominance in both directions. In principle such a test can be performed by switching the role of each distribution in the test statistic. Alternatively, we can construct a simple test that simultaneously considers dominance (or lack thereof) in both directions. The test we develop is similar to that of the previous section, although we now assume that $M_F(-t) = M_G(-t)$ for all $t \geq 0$ under the null. Rejection of this null hypothesis implies that one of the following mutually exclusive alternatives must be true: (a)$FD_\infty G$; (b) $GD_\infty F$; or (c) F and G are not comparable. It is desirable to design our test to allow us to distinguish between these alternatives. Consider the two statistics

$$D_n^+ = \sup_{t \in \mathbb{R}_+} D_n(t),$$

and

$$D_n^- = \inf_{t \in \mathbb{R}_+} D_n(t).$$

A large [resp. small] value of D_n^+ together with small [resp., large] value of D_n^- is consistent with the Laplace transform of F lying above [resp. below] G. On the other hand, observing large values in absolute terms for both D_n^+ and D_n^- is consistent with the the population Laplace transforms crossing. This decision rule is illustrated in Figure

(3).

we modify the decision rule of Knight and Satchell () and construct our test as follows:

(i) If both D_n^+ and D_n^- lead to rejection then F and G are not comparable.

(ii) If D_n^- rejects but D_n^+ does not, then $GD_\infty F$.

(iii) If D_n^+ rejects but D_n^- does not, then $FD_\infty G$.

(iv) If both D_n^- and D_n^+ do not lead to rejection then we fail to reject the null of indifference.

4 Asymptotic Theory

In this section we derive asymptotic results for the test statistics presented in the previous section. In our analysis we make use of recent developments in the empirical process literature, in particular, the empirical central limit theorem due to Arcones and Yu (1994) and the bootstrap empirical central limit theorem due to Radulović (1996), both of which apply to sequences indexed by $V - C$ subgraph classes of functions.

4.1 Sampling Process

Let $\{Z_i\}_{i=1}^n$ denote a sequence of observations with $Z_i = (X_i, Y_i)$. We continue to assume that the distributions F and G associated with X and Y, respectively, are continuous and have common support $[0, \infty)$. In addition we require the following assumptions on the sampling process in the study of the asymptotic behaviour of our test statistics:

Assumption 4. *$\{Z_i\}_{i=1}^n$ is a strictly stationary and β mixing sequence with $\beta(m) = o(m^{p/(p-2)}(\log m)^{2(p-1)/(p-2)})$ for some $2 < p < \infty$.*

Remark 1. *The mixing condition in Assumption (4) is more restrictive than α-mixing but is necessary for the bootstrap empirical CLT of Arcone of Radulović (1996).*

By requiring only stationarity and a β-mixing condition, we are able to accommodate rather general dependence structures. Moreover, we make no assumption on the form of dependence across samples.

4.2 Weak Convergence

For a fixed $t \in \mathbb{R}_+$ define the function $f : \mathbb{R}_+^2 \to \mathbb{R}$ by

$$f(x, y) = e^{-tx} - e^{-ty}.$$

We can write the test statistic in (6) compactly as

$$S_n = \sup_{t \in \mathbb{R}_+} D_n(t) \text{ with } D_n(t) = \sqrt{n} \int f(x, y) dH_n(x, y),$$

where H_n denotes the empirical distribution function. Hence, for the development of the asymptotic theory of the test statistic in (6) our interest centers on the on the class of functions

$$\mathcal{F} = \{f(\cdot, t) : t \in \mathbb{R}_+\} \tag{7}$$

Let P be the common distribution of the Z_i. Define the empirical measure as the linear combination $\mathbb{P}_n = n^{-1} \sum_{i=1}^n \delta_{Z_i}$ of the Dirac measures at the observations. The empirical measure induces a map from $\mathcal{F}$ to $\mathbb{R}$ given by

$$f \to \mathbb{P}_n f$$

where, for a given measurable function f and measure Q, $Qf = \int f dQ$. The centered and scaled version of the given map is the $\mathcal{F}$-indexed *empirical process* $\mathbb{G}_n$ given by

$$f \to \mathbb{G}_n f = \sqrt{n}(\mathbb{P}_n - P)f = \frac{1}{\sqrt{n}} \sum_{i=1}^n (f(Z_i) - Pf).$$

Definition 2. *An envelope function of a class $\mathcal{F}$ is any function $x \to F(x)$ such that $|f(x)| \leq F(x)$, for every x and $f \in \mathcal{F}$.*

Proposition 2. *For the class $\mathcal{F}$ defined in (7) there exists an envelope function $F(\mathbf{z})$ such that $PF^p < \infty$ for every $p \in (2, \infty)$.*

Proof of proposition (2). The proof is trivial since $\sup_{\mathbf{z} \in \mathbb{R}^2_+, f \in \mathcal{F}} |f(\mathbf{z})| = 1$, and so we can choose $F(x, y) = 1$. □

Before stating and proving our result establishing that (7) is a measurable VC-subgraph class we first state and prove a lemma which is required in the proof. For a fixed t define the indexed sets B_t and C_t by

$$B_t = \{(x_1, 0, y) | y < -tx_1, x_1 \in \mathbb{R}_+\},$$

and

$$C_t = \{(x, 0, y) | y < tx, x \in \mathbb{R}\}.$$

Note that each B_t is the subgraph[10] of a corresponding function $f(\cdot, t) \in \mathcal{F}$. Denote the collections of subgraphs by

$$\mathcal{S}_1 = \{B_t | t \in \mathbb{R}_+\},$$

and

$$\mathcal{S}_2 = \{C_t | t \in \mathbb{R}\}.$$

Lemma 5. *There exists a bijection $\phi : \mathcal{S}_2 \to \mathcal{S}_1$.*

Proof. Define the mapping $\phi : \mathcal{S}_2 \to \mathcal{S}_1$ by

$$\phi(C_t) = \bigcup_{x \in \mathbb{R}} \{(e^x - 1, 0, z) | z < -(e^t - 1)(e^x - 1)\}$$

Since $t_1 \neq t_2$ implies that $\phi(C_{t_1}) \neq \phi(C_{t_2})$, the map is clearly one-to-one. Let B_t be an arbitrary subgraph in $\mathcal{S}_1$. Then B_t is of the form

$$B_t = \{(x, 0, y) | y < -tx, x \in \mathbb{R}_+\}.$$

Setting $\tau = \log(t + 1)$, it follows that

$$\phi(C_\tau) = B_t.$$

Since the choice of B_t was arbitrary, we have established that the map ϕ is onto.

□

Proposition 3. *The class $\mathcal{F}$ defined in (7) is a measurable V-C subgraph class.*

[10]The subgraph of a function $f : X \subset \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\{(\mathbf{x}, y) : f(\mathbf{x}) < y, \mathbf{x} \in X\}$$

Proof. We prove the proposition by building the class $\mathcal{F}$ from VC classes and then appealing to results concerning the closure of VC classes under various operations. First, S_2 is a finite-dimensional vector space of measurable functions and, hence, by lemma (6) is VC subgraph. Since, from lemma (7), $\phi(\mathcal{S}_2)$ is also VC if ϕ is one-to-one , we have from lemma (5) above that $\mathcal{S}_1$ is VC. Since the exponential function is monotonic, we can apply lemma 2.6.18(viii) of VW to prove that $\mathcal{G}_1 = \{\exp(h)|h \in \mathcal{S}_1\}$ is VC. Defining the class $\mathcal{G}_2 = \{-\exp(-tz_2) : t \in \mathbb{R}_+\}$ and employing an analogous argument to that used for $\mathcal{G}_1$ in conjuction with lemma (2.6.18) implies that $\mathcal{G}_2$ is VC. Define ψ to be the correspondence which maps real-valued functions to their associated subgraph. It then follows from lemma 2.6.17 that $\psi(\mathcal{G}_1) \times \psi(\mathcal{G}_2)$ is VC. Finally, it is a well known fact in mathematics (see, for example..) that the cardinality of $\mathbb{R}^2_+$ is the same as the cardinality of $\mathbb{R}$. This fact implies the existence of a bijection $\varphi : \mathbb{R}^2_+ \to \mathbb{R}_+$. Choosing φ to be such a bijection, consider the map defined by

$$\Phi(\psi(g_1(\cdot, t_1)) \times \psi(g_2(\cdot, t_2))) = \psi(g_1(\cdot, \varphi(t_1, t_2)) + g_2(\cdot, \varphi(t_1, t_2)))$$

for all $(g_1, g_2) \in \mathcal{G}_1 \times \mathcal{G}_2$. The fact that Φ as defined above is a bijection and

$$\Phi(\psi(\mathcal{G}_1) \times \psi(\mathcal{G}_2)) = \psi(\mathcal{F})$$

completes the proof. □

We are now in a position to establish the weak convergence of the stochastic process $\mathcal{G}_n$ to a mean zero Brownian Bridge process, and this is precisely the content of Theorem (3) stated below.

Theorem 3. *Suppose Assumption (4) is satisfied. Then $\{\mathbb{G}_n(t) : t \in \mathbb{R}_+\}$ converges weakly to a mean zero Brownian Bridge $\{\mathbb{B}(t) : t \in \mathbb{R}_+\}$ in $l_\infty(\mathcal{F})$ with covariance function*

$$\begin{aligned}
\Gamma(s,t) &= \lim_{n\to\infty} \{n^{-1}\mathcal{L}_{X_1}(t+s) + 2n^{-2}\sum_{i=2}^{n}\mathcal{L}_{X_1,X_i}(t,s) + n^{-2}\sum_i\sum_j \mathcal{L}_{X_i,Y_j}(t,s) \\
&+ n^{-2}\sum_i\sum_j \mathcal{L}_{Y_i,X_j}(t,s) + n^{-1}\mathcal{L}_{Y_1}(t+s) + 2n^{-2}\sum_{i=2}^{n}\mathcal{L}_{Y_1,Y_i}(t,s)\} \\
&- (\mathcal{L}_X(t) - \mathcal{L}_Y(t))(\mathcal{L}_X(s) - \mathcal{L}_Y(s))
\end{aligned}$$

Proof of Theorem (3). Propositions (2) and (3), together with Assumption (4) establish that the necessary conditions for the empirical central limit theorem of Arcones and Yu (1994) are met. Theorem (3) then follows as an application of their CLT, which, for completeness, is stated as Theorem (13) in the technical appendix. □

This result establishes the existence of a limiting distribution for the empirical process $\{\mathbb{G}_n(f), f \in \mathcal{F}\}$. That the sequences $\sup_t \mathbb{G}_n(t)$ and $\int_0^\infty \mathbb{G}_n(t)\mathbf{1}_{\{D_n(t)>0\}}dt$ have a limiting distribution is then a direct consequence of the continuous mapping theorem and Theorem (3) above. We state these results as theorems (4) and (5), respectively.

Theorem 4. *Suppose Assumption (4) is satisfied. Then* $\sup_t \mathbb{G}_n(t)$ *converges weakly to* $\sup_t \mathbb{B}(t)$*, where* $\mathbb{B}(t)$ *is the Gaussian process defined in (3) above.*

Theorem 5. *Suppose Assumption (4) is satisfied. Then* $\int_0^\infty \mathbb{G}_n(t)\mathbf{1}_{\{D_n(t)>0\}}dt$ *converges weakly to* $\int_0^\infty \mathbb{B}(t)dt$*, where* $\mathbb{B}(t)$ *is the Gaussian process defined in (3) above.*

4.3 Test Consistency

We now establish the consistency of our proposed testing procedures. Let $c(\alpha)$ denote the critical value at significance level α associated with the limiting distribution of the given statistic.

Theorem 6 (Consistency of the Supremum Test)**.** *Given Assumption (4), then:*

(i) if H_0 *is true,*

$$\lim_{n\to\infty} P(rejectH_0) \leq P\left(\sup_{t\geq 0}\mathbb{B}(t) > c(\alpha)\right) \equiv \alpha$$

with equality when $\int e^{-tx}dF(x) = \int e^{-ty}dG(y)$ *for all* $t \in \mathbb{R}_+$

(ii) if H_0 *is false,*

$$\lim_{n\to\infty} P(rejectH_0) = 1$$

Proof. (i) Suppose H_0 is true and define $\delta(t) = \text{plim}\, D_n(t)$. From theorem (3) we have that $S_n(t)$ converges in distribution to $\sup_{t\geq 0}\mathbb{B}(t) + \sup_t \delta(t)$. Since $\sup_t \delta(t) \leq 0$ by assumption and $P(\sup_{t\geq 0}\mathbb{B}(t) + \sup_t \delta(t) > c(\alpha)) \leq P(\sup_{t\geq 0}\mathbb{B}(t) > c(\alpha))$, the result follows. (ii) Suppose H_1 is true. Then there exists some $t_0 \in \mathbb{R}_+ +$ such that $D(t_0) > 0$. Since

$$\begin{aligned}\sqrt{n}\sup_t D_n(t) &= \sqrt{n}\sup_t (D_n(t) - D(t) + D(t)) \\ &\geq \sqrt{n}\sup_t (D_n(t) - D(t)) + \sqrt{n}\sup_t D(t),\end{aligned}$$

where $\sqrt{n}\sup_t(D_n(t) - D(t)) \to^d \sup_{t\geq 0}\mathbb{B}(t)$ and $\sqrt{n}\sup_t D(t) \to \infty$, it follows that

$$P(\sqrt{n}\sup_t D_n(t) > c) \to 1$$

for all $c > 0$. □

Theorem 7 (Consistency of the Integral Test)**.** *Given Assumption (4), then:*

(i) if H_0 *is true,*

$$\lim_{n\to\infty} P(rejectH_0) \leq P\left(\int_0^\infty \mathbb{B}(t)dt > c(\alpha)\right) \equiv \alpha$$

with equality when $\int e^{-tx}dF(x) = \int e^{-ty}dG(y)$ *for all* $t \in \mathbb{R}_+$

(ii) if H_0 is false,

$$\lim_{n\to\infty} P(reject H_0) = 1$$

Proof. (i) Suppose H_0 is true and define $\delta(t) = \text{plim}\, D_n(t)$. From theorem (3) we have that $S_n(t)$ converges in distribution to $\int \mathbb{B}(t)dt + \int \delta(t)dt$. Since $\int \delta(t)dt \leq 0$ by assumption and $P(\int \mathbb{B}(t)dt + \int \delta(t)dt > c(\alpha)) \leq P(\int \mathbb{B}(t)dt > c(\alpha))$, the result follows. (ii) Suppose H_1 is true. Then there exists some region $\mathcal{R}$ such that $D(t) > 0$ for all $t \in \mathcal{R}$. The Glivenko-Cantelli property implies that $\mathbf{1}_{\{D_n(t)\geq 0\}}$ converges uniformly in probability to $\mathcal{R} \cup \mathcal{O}$, where $\mathcal{O}$ contains all t such that $D(t) = 0$ holds. Since

$$\begin{aligned}\sqrt{n}\int D_n(t)\mathbf{1}_{\{D_n(t)\geq 0\}}dt &= \sqrt{n}\int (D_n(t) - D(t) + D(t))\mathbf{1}_{\{D_n(t)\geq 0\}}dt \\ &= \sqrt{n}\int (D_n(t) - D(t))\mathbf{1}_{\{D_n(t)\geq 0\}}dt + \sqrt{n}\int D(t)\mathbf{1}_{\{D_n(t)\geq 0\}}dt,\end{aligned}$$

where $\sqrt{n}\int (D_n(t) - D(t))\mathbf{1}_{\{D_n(t)\geq 0\}}dt \to^d \int_0^\infty \mathbb{B}(t)dt$ and

$$\sqrt{n}\int D(t)\mathbf{1}_{\{D_n(t)\geq 0\}}dt \to^p \sqrt{n}\int_{\mathcal{R}} D(t)dt \to \infty,$$

it follows that

$$P\left(\sqrt{n}\int D_n(t)\mathbf{1}_{\{D_n(t)\geq 0\}}dt > c\right) \to 1$$

for all $c > 0$. □

5 Obtaining the Critical Values

Although we have established consistency of our proposed test statistics, we must recognize that these results depend crucially on our ability to consistently estimate the appropriate critical values. In this section we present a technique, in particular bootstrap procedure, for obtaining the critical values of our proposed tests. Due to the assumed presence of serial dependence the naive application of the IID bootstrap of Efron (1979) is inappropriate. Singh (1981) provides an example to illustrate the inadequacy of this procedure for dependent data. In this paper we have chosen to concentrate on a nonparametric block bootstrap procedure. There are several block bootstrap alternatives, all of which attempt to reproduce the dependence structure of the underlying process asymptotically. Carlstein (1996) proposes a bootstrap procedure whereby the observations are divided into nonoverlapping blocks and bootstrap samples are formed by uniformly sampling the blocks with replacement. The moving block bootstrap (MBB), which is developed independently in Künsch (1989) and Liu and Singh (1992), forms bootstrap samples by uniformly sampling overlapping blocks from the original sample. The circular block bootstrap of Politis and Romano (1992) resamples overlapping and periodically extended blocks of a given length. This has the effect of ensuring that every observation has an equal probability of being included in a bootstrap sample, thereby removing the "edge effects" associated with the nonoverlapping and MBB procedures. Politis and Romano (1994) also proposed the stationary block bootstrap (SB) that uses blocks of random lengths rather than blocks of a fixed length. Conditional on the original sample, the bootstrap observations generated by the SB method are then also stationary.

In the remainder of this section we focus on the moving block bootstrap procedure. Our choice of the MBB method is based on the following. First, in comparing various block bootstrap variance estimators, Lahiri (1999) demonstrates that the MBB performs favourably with respect to the nonoverlapping bootstrap, and also that fixed length blocking procedures generally have better properties than their random block length counterparts. Second, to the best of our knowledge very few block bootstrap empirical CLT results are available, but an MBB empirical central limit theorem, due to Radulović (1996), has been established for VC classes.

5.1 The Moving Block Bootstrap

Conditional on the sample $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$, let $\mathcal{B}_i = (Z_i, \ldots, Z_{i+l-1})$ denote the block of length l starting with Z_i, $1 \leq i \leq N$ where $N = n - l + 1$. Assume, temporarily, that the block length is given (we discuss methods for selecting l later in the section) and that $N = kl$. Generating moving block bootstrap samples consists of resampling uniformly with replacement from the collection of overlapping blocks $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$. Let $\mathcal{Z}* = \{\mathcal{B}_1^*, \ldots, \mathcal{B}_k^*\}$ denote an arbitrary bootstrap sample generated by this procedure. Denote by $H_n^*(x, y)$ the bivariate empirical distribution function computed from a given bootstrap sample $\mathcal{Z}^*$. The MBB version $D_n^*(t)$ of $D_n(t)$ is defined as

$$D_n^*(t) = \sqrt{n}\left[\int f(x, y) dH_n^*(x, y)\right]$$

Some care needs to be taken when computing bootstrap p-values. The null distribution imposes the restriction $E[f(X, Y)] = 0$, which is not necessarily replicated by the above resampling scheme. The bootstrap approximation to the null distribution will therefore have a random bias which may render the approximation useless.[11] In order to remove the random bias we re-center the test statistic as

$$D_n^*(t) = \sqrt{n}\left\{\int [f(x, y) - E^* f(X^*, Y^*)] dH_n^*(x, y)\right\}$$

where $E^* f(X^*, Y^*)$ is the mean of the bootstrap distribution which can be calculated explicitly as

$$E^* f(X^*, Y^*) = \frac{1}{Nl}\sum_{i=1}^{N}\sum_{j=1}^{l} \exp(-tX_{i+j-1}) - \exp(-tY_{i+j-1}).$$

[11] A similar problem is addressed in Freedman (1981), Shorack (1982), and Lahiri (1992).

The integrand in the expression above is then mean zero by construction. The MBB procedure for obtaining p-values can now be summarized as follows:

1. Conditional on $\mathcal{Z}_i$, generate B bootstrap samples $\{Z_1^*, \ldots, Z_B^*\}$ where B is chosen such that $\alpha(B+1)$ is integer-valued.

2. Compute $d_i^* = \sup_t D_{ni}^*(t)$ for $i = 1, \ldots B$.

3. Compute the $p - value$ as

$$P^*(d_i^* > D_n^+) = B^{-1} \sum_{i=1}^{B} I\{d_i^* > D_n^+\}$$

Define the bootstrap empirical measure as $\hat{\mathbb{P}}_n^* f = n^{-1} \sum_{i=1}^{k} W_{Ni} f(X_i)$, where $W_N = (W_{1N}, \ldots, W_{NN})$ is a multinomial vector with probabilities $(1/N, \ldots, 1/N)$ and number of trials k, and where W_N is independent of the sequence $(\mathcal{B}_1, \ldots, \mathcal{B}_N)$. Define $\hat{\mathbb{G}}_n^* = \sqrt{n}(\hat{\mathbb{P}}_n^* - \mathbb{P}_n)$, and let $\mathbb{G}_P$ denote a centered, tight Gaussian process. The following theorem establishes the weak convergence of $\hat{\mathbb{G}}_n^*$ to the limiting Gaussian process $\mathbb{G}_P$.

Theorem 8. *Define the class $\mathcal{F}$ as in (7)and assume that Assumption (4) holds. Let Z_i^* be the MBB sequence with block size $l = O(n^\rho)$ and $0 < \rho < \frac{1}{2}$. Then*

$$\hat{\mathbb{G}}_n^* \to \mathbb{G}_P(f)_{f \in \mathcal{F}} \text{ in probability}$$

Proof. Proposition (3) establishes that $\mathcal{F}$ is VC, a condition that is sufficient to ensure that the uniform entropy integral is finite (see Appendix). This, together with Proposition (2) establish that the conditions of Theorem (11) are met and hence that $\mathcal{F}$ is a Donsker class. The proof then follows from Theorem (12). □

Theorem (8) provides the foundation for establishing consistency of the bootstrap p-value approach, which is the content of our next theorem.

Theorem 9.

$$\lim_{n \to \infty} \mathbb{P}_n^*(d_i^* > c) \to \mathbb{P}(D_n^+ > c)$$

6 Simulation Experiments

In this section we examine the finite sample properties our proposed test statistics.

6.1 Klecan process

In each of the simulation experiments we consider some variation of the process

$$X_{ki} = (1-\lambda)\left[\alpha_k + \beta_k(\sqrt{p}Z_{0i} + \sqrt{1-\rho}Z_{ki})\right] + \lambda X_{k,i-1} \tag{8}$$

for $k = 1, 2$, where $(Z_{0i}, Z_{1i}, Z_{2i}) \sim i.i.d.N(0, I)$. This process was introduced in Klecan et al. (1991) and subsequently used in Linton et al (1993). The basic properties of the process $\{X_1, X_2\}$ are summarized in Proposition (4) below.

Proposition 4. *Suppose $\{X_k\}$ is generated according to (8). Then, the moment generating function associated with X_k, denoted by ϕ_k, is given by*

$$\phi_k(t) = \exp\left\{\alpha_k t + \frac{1}{2}t^2\frac{(1-\lambda)}{1+\lambda}\beta_k^2\right\}.$$

Additionally, the temporal and mutual dependence properties of the process (X_{1i}, X_{2i}) are characterized by

$$Cov(X_{k,i}, X_{k,i-l}) = \lambda^l\frac{(1-\lambda)}{1+\lambda}\beta_k^2,$$

and

$$Cov(X_{1,i}, X_{2,i-l}) = \lambda^l\frac{(1-\lambda)}{1+\lambda}\rho\beta_1\beta_2,$$

respectively.

The moment generating function associated with X_k is that of a normally distributed random variable. This two-parameter family of distributions allows for a simple characterization of dominance, i.e.

$$X_1 \succcurlyeq_\infty X_2 \Leftrightarrow (\alpha_1 - \alpha_2) - \frac{1}{2}\frac{(1-\lambda)}{(1+\lambda)}(\beta_1^2 - \beta_2^2)t \leq 0$$

for all $t > 0$. We note that the parameter ρ is controls the degree of mutual dependence between the two random processes, and in the case where $\rho = 0$ the processes are independent of one another. Also, the parameter λ directly influences the temporal dependence of the processes. When $\lambda = 0$, both processes are *i.i.d.* and in the case where $\lambda = \rho = 0$ the processes are *i.i.d* and mutually independent. The simple form of the m.g.f. and the ease through which the dependence structure can be manipulated makes the Klecan et al process convenient for our analysis. We point out, however, that the marginal distributions do violate our assumed support condition. In all of the simulations using this process we therefore restrict the choice parameters to ensure that the probability of observing a negative outcome is negligible.

The simulation results suggest that the bootstrap test is correctly sized when there is no temporal dependence in the data. When temporal dependence is introduced the test appears to be slightly oversized in small samples, with the magnitude of the difference between the actual and nominal size being sensitive to the choice of block length. As the sample size is increased from 100 to 500, the test appears to be correctly sized, at least for the appropriate choice of block length.

Test Size

n=100	$\lambda = 0, \ell = 1$	0.0495	0.0494	0.0520	0.0496	0.0502
		($\rho = 0$)	($\rho = 0.1$)	($\rho = 0.2$)	($\rho = 0.3$)	($\rho = 0.4$)
n=100	$\lambda = 0.1, \rho = 0.1$	0.0772	0.0692	0.0658	0.0646	0.0738
		($\ell = 1$)	($\ell = 3$)	($\ell = 5$)	($\ell = 7$)	($\ell = 9$)
n=500	$\lambda = 0.1, \rho = 0.1$	0.0700	0.0552	0.0546	0.0513	0.0532
		($\ell = 1$)	($\ell = 5$)	($\ell = 10$)	($\ell = 15$)	($\ell = 20$)

Table 1: Actual size versus nominal size of 5%

7 Empirical Applications

We now illustrate briefly the application of the above methodology using real data.

7.1 Financial Market Indicies

We now consider the application of the ISD test to the comparison of financial market indices. The data we consider consists of weekly returns for various indices for the period January 1, 1996 through December 31, 2005. This amounts to 518 observations for each index. The particular market indices we consider include the Bovespa, IPC, Hang Seng, Nikkei, FTSE, DAX, Nasdaq, and the S&P.

Table (2) contains summary statistics for the indices considered. Several indices exhibit a high degree of mutual dependence as measured by the correlation coefficient. Also of note is that the IPC index has a relatively high sample mean and low sample standard deviation making it a plausible candidate for inclusion in the ISD efficient set.

Before proceeding with the tests we also present graphical illustrations of the differences in Laplace transforms of several pairs of indices. Figure (2) contains a plot of the difference between the empirical Laplace transforms of the IPC and Nasdaq markets. The fact that the difference is negative for all values of t and, hence, that the IPC curve is always below the Nasdaq curve suggests that the IPC market dominates the Nasdaq market. In Figure (3) we have plotted the difference in the Hang Seng and Dax empirical Laplace transforms. In this case, the difference is always positive and implies that, at least in-sample, the Laplace transform of the Hang Seng is always above that of the Dax—a finding that is consistent with the hypothesis that the DAX dominates the Hang Seng. In the final graph, Figure (4), we have plotted the difference between the Bovespa and FTSE transforms. Here the difference switches signs suggesting that the Bovespa and FTSE ranking may not be resolved by infinite-degree stochastic dominance.

In Table (3) we have compiled the p-values associated with all of the possible pairwise tests of the hypothesis that market A dominates market B at infinite order. The (i, j)-th entry in the matrix corresponds to the p-value associated with the test market i stochastically dominates market j. Thus, if the p-value in (i, j) is significant but the p-value (j, i) is not, then we conclude that market i dominates market j. Alternative, if both of the p-values

Summary Statistics

	Bovespa	Dax	FTSE	HangSeng	IPC	Nikkei	Nasdaq	S&P
Mean	0.4897	0.2102	-0.0921	0.1208	0.4023	-0.0660	-0.1006	-0.0745
St.D.	0.4850	0.3341	0.1534	0.3580	0.3553	0.2879	0.2352	0.1490

Correlation Matrix

	Bovespa	Dax	FTSE	HangSeng	IPC	Nikkei	Nasdaq	S&P
Bovespa	1.0000	0.453	0.001	0.322	0.564	0.175	0.019	0.027
Dax	0.453	1.0000	-0.018	0.501	0.491	0.112	0.070	0.040
FTSE	0.001	-0.018	1.0000	0.006	0.021	0.013	0.164	0.204
Hang Seng	0.322	0.501	0.006	1.0000	0.422	0.123	0.051	0.018
IPC	0.564	0.491	0.021	0.422	1.0000	0.129	0.058	0.016
Nikkei	0.175	0.112	0.013	0.123	0.129	1.0000	0.054	0.050
Nasdaq	0.019	0.070	0.164	0.051	0.058	0.054	1.0000	0.878
S&P	0.027	0.040	0.204	0.018	0.016	0.050	0.878	1.0000

Table 2: Summary statistics for the weekly returns on the financial indices

were found to be significant then we would conclude that the population Laplace transforms cross and therefore the ranking cannot be resolved by stochastic dominance. Similarly, if both of the p-values are insignificant then we cannot reject that the population empirical Laplace transforms are equal and again the ranking cannot be resolved by stochastic dominance. Based on the results in the table, we are able to conclude that the DAX and FTSE dominate the S&P, and that the IPC market dominates every market except possibly the Bovespa and Nikkei.

ISD Comparison Tests of Various Financial Market Indicies

	Bovespa	Dax	FTSE	HangSeng	IPC	Nasdaq	Nikkei	S&P
Bovespa	1.0000	0.4975	0.4875	0.5355	0.1701	0.5375	0.4434	0.4965
Dax	0.2793	1.0000	0.4775	1.0000	**0.0040**	1.0000	0.2502	0.6066
FTSE	0.2432	0.2943	1.0000	1.0000	**0.0020**	1.0000	0.2142	1.0000
Hang Seng	0.0871	0.1301	0.2092	1.0000	**0.0060**	1.0000	0.1421	0.5205
IPC	1.0000	0.5746	0.5405	1.0000	1.0000	1.0000	0.5826	0.5796
Nasdaq	0.1091	0.0140	0.1011	0.4074	**0.0000**	1.0000	0.0761	0.3684
Nikkei	0.4586	0.5617	0.5245	1.0000	0.1471	1.0000	1.0000	0.5796
S&P	0.1021	**0.0091**	**0.0200**	0.4705	**0.0000**	1.0000	0.0851	1.0000

Table 3: Bootstrap p-values associated with testing the null $r \not\succcurlyeq c$

8 Conclusion

To be completed...

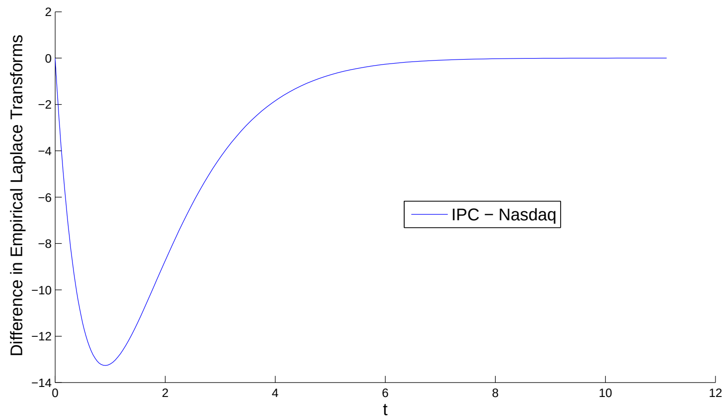

Figure 2: Graph corresponding to a test of the hypothesis that IPC dominates Nasdaq

9 Technical Appendix

In this section we collect some of the major results from modern empirical process theory. We borrow heavily from Kosorok (2006) and van der Vaart and Wellner (1996).

Theorem 10. *The empirical process X_n converges weakly to a tight stochastic process X in $l(\mathcal{F})$ if and only if:*

1. *for all finite $\{t_1, \ldots, t_k\} \subset T$, the multivariate distribution of $\{X_n(t_1), \ldots, X_n(t_k)\}$ converges to that of $\{X(t_1), \ldots, X(t_n)\}$.*

2. *There exists a semimetric ρ for which T is totally bounded and*

$$\lim_{\delta \downarrow 0} \lim\sup\nolimits_{n\to\infty} P^* \left\{ \sup_{s,t \in T \text{ with } \rho(s,t)<\delta} |X_n(s) - X_n(t)| > \epsilon \right\} = 0,$$

for

all $\epsilon > 0$.

A class of functions for which Theorem (10) holds is called a *Donsker* class, or more precisely, a P-Donsker class. Condition (i) is convergence of all finite dimensional distributions. This condition is typically easy to verify through the use of a multivariate central limit theorem, or a univariate central limit theorem coupled with the Cramer-Wold device. Condition (ii), which is referred to as stochastic equicontinuity, can be regarded as a probabilistic and asymptotic generalization of uniform continuity (see Andrews ()). This condition which is often extremely difficult to verify directly is nevertheless an essential ingredient in any proof

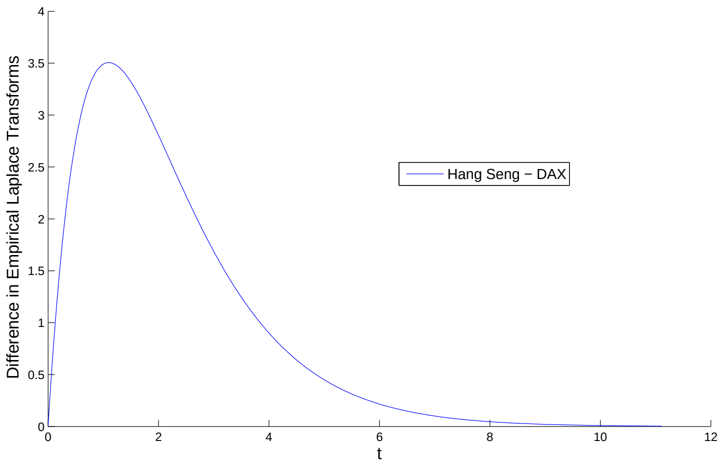

Figure 3: Graph corresponding to a test of the hypothesis that Hang Seng dominates DAX

of weak convergence. The difficulties associated with verifying condition (ii) of the theorem have been the primary motivating factor behind a search for more easily verifiable sufficient conditions.

The complexity, or *entropy*, of the class $\mathcal{F}$ has been shown to play a prominent role in whether $\mathcal{F}$ is a Donsker class. There are two major entropy measures which are known as entropy with bracketing and entropy based on covering numbers. For the sake of brevity, we will focus on the latter and refer to reader to section 2.7 of VW () for definitions and results concerning bracketing. The covering number $N(\epsilon, \mathcal{F}, L_r(Q))$ is the minimum number of $L_r(Q)$ ϵ-balls needed to cover $\mathcal{F}$, where and $L_r(Q)$ ϵ-ball around a function $g \in L_r(Q)$ is the set $\{h \in L_r(Q) : ||h - g||_{Q,r} < \epsilon\}$. For a collection of balls to cover $\mathcal{F}$, all elements of $\mathcal{F}$ must be included in at least one of the balls, but it is not necessary that centers of the balls be contained in $\mathcal{F}$. The entropy is the logarithm of the covering number. Define the uniform covering number

$$\sup_Q N(\epsilon||F||_{Q,r}, \mathcal{F}, L_r(Q)),$$

where $F : X \to \mathbb{R}$ is an envelope for $\mathcal{F}$, and where the supremum is taken over all finitely discrete probability measures Q with $||F||_{Q,r} > 0$. The uniform entropy integral is

$$J(\delta, \mathcal{F}, L_r) = \int_0^\delta \sqrt{\log \sup_Q N(\epsilon||F||_{Q,r}, \mathcal{F}, L_r(Q))} d\epsilon. \tag{9}$$

The following theorem provides a Donsker result for uniform entropy:

Theorem 11. *Let $\mathcal{F}$ be a measurable class of measurable functions with $J(1, \mathcal{F}, L_2) < \infty$. If $PF^2 < \infty$, then $\mathcal{F}$ is P-Donsker.*

Theorem (11) shows that convergence of the integral in (9) together with a moment condition on the envelope is sufficient for weak convergence of the empirical process. Vap-

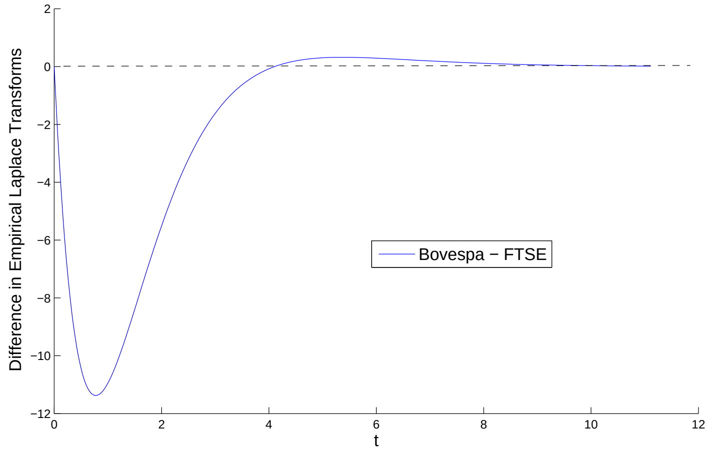

Figure 4: Graph corresponding to a test of the hypothesis that Bovespa dominates FTSE

nik and Červonenkis (), showed that for classes of sets satisfying certain combinatorial conditions

$$\sup_Q N(\epsilon||F||_{Q,r}, \mathcal{F}, L_r(Q)) < K\left(\frac{1}{\epsilon}\right)^V, \tag{10}$$

for some finite number V, which certainly implies $J(1, \mathcal{F}, L_2) < \infty$. A class of sets for which (10) holds is referred to as a VC class, or simply VC. We now state several useful results concerning VC classes. The proofs of these results can be found in VW ().

Lemma 6. *(2.6.15 of VW) Any finite-dimensional vector space mathcal$F$ of measurable functions $f : X \to \mathbb{R}$ is VC of index smaller than or equal to dim$(\mathcal{F}) + 2$.*

The next two results concern closure under various operations on VC classes.

Lemma 7. *(2.6.17 of VW) Let $\mathcal{C}$ and $\mathcal{D}$ be VC-classes of sets in a set $\mathcal{X}$ and $\phi : \mathcal{X} \to \mathcal{Y}$. Then*

1. *$\phi(\mathcal{C})$ is VC if ϕ is one-to-one; For VC classes $\mathcal{C}$ and $\mathcal{D}$ in sets $\mathcal{X}$ and $\mathcal{Y}$,*
2. *$\mathcal{C} \times \mathcal{D}$ is VC is $\mathcal{X} \times \mathcal{Y}$.*

The Donsker property of class $\mathcal{F}$ leads directly to consistency results concerning bootstrapping empirical processes. Following Kosorok (2006), we define the bootstrap empirical measure as $\hat{\mathbb{P}}_n f = n^{-1}\sum_{i=1}^n W_{ni} f(X_i)$, where $W_n = (W_{1n}, \ldots, W_{nn})$ is a multinomial vector with probabilities $(1/n, \ldots, 1/n)$ and number of trials n, and where W_n is independent of the data sequence $(X_1, \ldots, X_n)$. Let $\hat{\mathbb{G}}_n = \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P})$, and $\mathbb{G}$ be the standard Brownian bridge in $l^\infty(\mathcal{F})$.

Theorem 12. *The following are equivalent*

1. $\mathcal{F}$ *is* P*-Donsker*

2. $\hat{\mathbb{G}}_n \to^P \mathbb{G}$ *in* $l^\infty(\mathcal{F})$.

Theorem 13 (Theorem 2.1 of Arcones and Yu (1994))**.** *Suppose that* $\mathcal{F}$ *is a measurable VC subgraph class of functions satisfying*

$$PF^p < \infty$$

for some $2 < p < \infty$*, where* $F(x) = \sup_{f\in\mathcal{F}} |f(x)|$*. If the* β *mixing coefficient of the stationary sequence satisfies*

$$k^{p/(p-2}(\log m)^{2(p-1)/(p-2)}\beta_m \to 0$$

then

$$\left\{n^{-1/2}\sum_{j=1}^{n}(f(X_j) - Pf) : f \in \mathcal{F}\right\}$$

converges in law to a Gaussian process $\{G(f)\}_{f\in\mathcal{F}}$ *which has a version with uniformly bounded and uniformly continuous paths with respect to the* $||\cdot||_2$ *norm.*

Theorem 14. *Let* $\mathcal{F}$ *be a VC class of functions with envelope function* F *satisfying* $EF^p < \infty$ *for some* $p > 2$*. Let* X_i^* *be the MBB sequence with block size* $b = O(n^\rho)$ *and* $0 < \rho < \frac{p-2}{2(p-1)}$*. If the original sequence* X_i *is* β*-mixing with* $\beta_i = O(i^{-q})$ *and* $q > p/(p-2)$*, then*

$$\sqrt{n}(P_n^* f - P_n f)_{f\in\mathcal{F}} \to G_P(f)_{f\in\mathcal{F}} \text{ in probability.}$$

Appendix: Proofs

Proof of ().

$$cov(D_n(t), D_n(s)) = E[D_n(t)D_n(s) - E(D_n(t))E(D_n(s))] \tag{11}$$

The second term consists of the expectation of the difference of the empirical Laplace transforms and so by stationarity it follows immediately that

$$E(D_n(t))E(D_n(s)) = D(t)D(s),$$

where $D(t) = \mathcal{L}_F(t) - \mathcal{L}_G(t)$, the difference in the population Laplace transforms. Focusing on the first term in (11) we have

$$\begin{aligned}
E[D_n(t)D_n(s)] &= E[(\mathcal{L}_{\hat{F}}(t) - \mathcal{L}_{\hat{F}}(t))(\mathcal{L}_{\hat{F}}(s) - \mathcal{L}_{\hat{G}}(s))] \\
&= E[\mathcal{L}_{\hat{F}}(t)\mathcal{L}_{\hat{F}}(s) - \mathcal{L}_{\hat{F}}(t)\mathcal{L}_{\hat{G}}(s) - \mathcal{L}_{\hat{G}}(t)\mathcal{L}_{\hat{F}}(s) + \mathcal{L}_{\hat{G}}(t)\mathcal{L}_{\hat{G}}(s)] \\
&= \frac{1}{n}\mathcal{L}_X(s+t) + \frac{2}{n^2}\sum_{j=1}^{n-1}(n-j)\mathcal{L}_{X_1,X_{1+j}}(s,t) - \frac{4}{n^2}\sum_{j=0}^{n-1}(n-j)\mathcal{L}_{X_1,Y_{1+j}}(s,t) \\
&+ \frac{1}{n}\mathcal{L}_Y(s+t) + \frac{2}{n^2}\sum_{j=1}^{n-1}(n-j)\mathcal{L}_{Y_1,Y_{1+j}}(s,t)
\end{aligned}$$

□

Properties of the integral test.

$$\int_a^b D_n(t)dt = \frac{1}{n}\sum_{i=1}^{n}[\frac{1}{X_i}(e^{-aX_i} - e^{bX_i}) - \frac{1}{Y_i}(e^{-aY_i} - e^{bY_i})] \tag{12}$$

Note that

$$\lim_{n\to\infty}\int_a^b D_n(t)dt = \int_a^b \lim_{n\to\infty} D_n(t)dt$$

by the Lebesgue Dominated Convergence Theorem, and thus it follows that under the null

$$\lim_{n\to\infty}\int_a^b D_n(t)dt = 0$$

□

References

[1] Fred D. Arditti. A utility function depending on the first three moments: Reply. *The Journal of Finance*, 24(4):720, Sep. 1969. GR: Comment.

[2] Philippe Barbe and Patrice Bertail. *The weighted bootstrap*, volume 98. Springer-Verlag, New York, 1995.

[3] O. E. Barndorff-Nielsen and D. R. Cox. *Asymptotic techniques for use in statistics.* Chapman and Hall, London ;; New York, 1989.

[4] Vijay S. Bawa. Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, 2(1):95–121, Jan. 1985.

[5] P. Billingsley. *Probability and Measure.* John Wiley & Sons, New York, 1978.

[6] Karl Borch. The rationale of the mean-standard deviation analysis: Comment. *The American Economic Review*, 64(3):428–430, Jun. 1974.

[7] Patrick L. Brockett and Yehuda Kahane. Risk, return, skewness and preference. *Management Science*, 38(6):851–866, Jun. 1992.

[8] Russell Davidson and James MacKinnon. *Bootstrap Methods in Econometrics*, volume 1 of *Chapter 23 Palgrave Handbook of Economics*, pages 812–838. Palgrave Macmillan, 2006.

[9] Bradley Efron. *The jackknife, the bootstrap, and other resampling plans*, volume 38. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1982.

[10] Bradley Efron and Robert Tibshirani. *An introduction to the bootstrap.* Chapman and Hall, New York, 1993.

[11] W.E. Frank. A likelihood ratio test for stochastic ordering. *Journal of American Statistical Association*, 70(2):686, 1984.

[12] P. Hall. *The bootstrap and Edgeworth expansion.* Springer-Verlag, New York, 1992.

[13] P. Hall. *Chapter 39 Methodology and theory for the bootstrap*, volume 4 of *Handbook of Econometrics*, pages 2341–2381. Elsevier, 1994.

[14] Joel L. Horowitz. *Chapter 52 The Bootstrap*, volume 5 of *Handbook of Econometrics*, pages 3159–3228. Elsevier, 2001.

[15] W.H. Jean. Comparison of moment and stochastic dominance ranking methods. *Journal of Financial and Quantitative Analysis*, 10(3):151, 1975.

[16] O. Maurice Joy and R. Burr Porter. Stochastic dominance and mutual fund performance. *The Journal of Financial and Quantitative Analysis*, 9(1):25–31, Jan. 1974.

[17] Alex Kane. Skewness preference and portfolio choice. *The Journal of Financial and Quantitative Analysis*, 17(1):15–25, Mar. 1982.

[18] J. Knight and S. Satchell. Testing for infinite order stochastic dominance with applications to finance, risk and income inequality. *Forthcoming*, 2007.

[19] S. N. Lahiri. *Resampling Methods for Dependent Data.* Springer, New York, 2003.

[20] H. Levy. Stochastic dominance and expected utility: Survey and analysis. *Management Science*, 38(4):555, 1992.

[21] H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77, 1952.

[22] D. McFadden. Testing for Stochastic Dominance. In Thomas B. Fomby and Tae Kun Seo, editors, *Studies in the Economics of Uncertainty. In Honour of Josef Hadar*, pages 113–134. Springer Verlag, New York, 1989.

[23] D. McFadden. *Testing for Stochastic Dominance*, volume 1 of *Studies in the Economics of Uncertainty*, pages 81–98. Elsevier, 2006.

[24] Robb J. Muirhead. *Aspects of multivariate statistical theory.* Wiley, New York ;; Toronto, 1982.

[25] J.W. Pratt and R.J. Zeckhauser. Proper risk aversion. *Econometrica*, 55(1):143–154, 1980.

[26] D. Radulović. On the Bootstrap and Empirical Processes for Dependent Sequences. In Herold Dehling, Thomas Mikosch, and Michael Sorensen, editors, *Empirical Process Techniques for Dependent Data*, pages 345–364. Birkhauser, Berlin, 2002.

[27] Andrzej Ruszczynski and Robert J. Vanderbei. Frontiers of stochastically nondominated portfolios. *Econometrica*, 71(4):1287–1297, Jul. 2003.

[28] Robert C. Scott and Philip A. Horvath. On the direction of preference for moments of higher order than the variance. *The Journal of Finance*, 35(4):915–919, Sep. 1980.

[29] R. J. Serfling. *Approximation theorems of mathematical statistics.* Wiley, New York ;; Toronto, 1980.

[30] Jun Shao and Dongsheng Tu. *The jackknife and bootstrap.* Springer Verlag, New York, 1995.

[31] Hassan Tehranian. Empirical studies in portfolio performance using higher degrees of stochastic dominance. *The Journal of Finance*, 35(1):159–171, Mar. 1980.

[32] P.D. Thistle. Negative moments, risk aversion, and stochastic dominance. *Journal of Financial and Quantitative Analysis*, 4:3, 1993.

[33] James K. Weeks. Stochastic dominance: A methodological approach to enhancing the conceptual foundations of operations management theory. *The Academy of Management Review*, 10(1):31–38, Jan. 1985.

[34] G.A. Whitmore. Stochastic Dominance for the Class of Completely Monotonic Utility Functions. In Thomas B. Fomby and Tae Kun Seo, editors, *Studies in the Economics of Uncertainty. In Honour of Josef Hadar*, pages 77–88. Springer Verlag, New York, 1989.