TESTING FOR STOCHASTIC DOMINANCE EFFICIENCY

Olivier Scaillet
HEC Genève and FAME
Université de Genève
Bd Carl Vogt, 102
CH - 1211 Genève 4, Suisse
olivier.scaillet@hec.unige.ch

Nikolas Topaloglou
HEC Genève
Université de Genève
Bd Carl Vogt, 102
CH - 1211 Genève 4, Suisse
nikolas.topaloglou@hec.unige.ch

This version: July 2005 (first version : December 2004)

Abstract

We consider consistent tests for stochastic dominance efficiency at any order of a given portfolio with respect to all possible portfolios constructed from a set of assets. We propose and justify approaches based on simulation and the block bootstrap to achieve valid inference in a time series setting. The test statistics and the estimators are computed using linear and mixed integer programming methods. The empirical application shows that the Fama and French market portfolio is FSD and SSD efficient, although it is mean-variance inefficient.

Key words and phrases: Nonparametric, Stochastic Ordering, Dominance Efficiency, Linear Programming, Mixed Integer Programming, Simulation, Bootstrap.

JEL Classification: C12, C13, C15, C44, D81, G11.
AMS 2000 Subject Classification: 62G10, 62P05, 90C05, 91B06, 91B30.

1 Introduction

Stochastic dominance (SD) is a central theme in a wide variety of applications in economics, finance and statistics, see e.g. the review papers by Kroll and Levy (1980) and Levy (1992), the classified bibliography by Mosler and Scarsini (1993) and the books by Shaked and Shanthikumar (1994) and Levy (1998). It aims at comparing random variables in the sense of stochastic orderings expressing the common preferences of rational decision-makers. Stochastic orderings are binary relations defined on classes of probability distributions. They translate mathematically intuitive ideas like "being larger" or "being more variable" for random quantities. They extend the classical mean-variance approach to compare riskiness.

The main attractiveness of the SD approach is that it is nonparametric, in the sense that its criteria do not impose explicit specifications of an investor preferences or restrictions on the functional forms of probability distributions. Rather, they rely only on general preference and beliefs assumptions. Thus, no specification is made about the return distribution, and the observed discrete distribution is posited to represent the underlying distribution.

Traditionally, SD was tested pairwise. Only recently Kuosmanen (2004) and Post (2003) have introduced the notion of stochastic dominance efficiency (SDE). This notion is a direct extension of SD to the case where full diversification is allowed. In that setting both authors derive statistics to test for SDE of a given portfolio with respect to all possible portfolios constructed from a set of financial assets. Such a derivation relies intrinsically on using ranked observations under an i.i.d. assumption on the asset returns. [1]

The goal of this paper is to develop consistent tests for SDE at *any* order for *time-dependent* data. Serial correlation is known to pollute financial data (see the empirical section), and to alter, often severely, the size and power of testing procedures when neglected. We rely on Kolmogorov-Smirnov type tests inspired by the consistent procedures developped by Barrett and Donald (2003), hereafter BD, in testing for SD. [2] We develop general SD efficiency tests that compare a given portfolio with an optimal diversified portfolio formed from a given finite set of assets. We build on the general distribution definition of SD in contrast to the traditional expected utility framework.

Note that De Giorgi (2004) have solved a portfolio selection problem based on reward-risk measures consistent with second-order stochastic dominance (SSD). If investors have homogeneous expectations and optimally hold reward-risk efficient

[1] Remark that contrary to the initial observations, ranked observations, i.e. order statistics, are no more i.i.d.. Besides, each order statistic has a different mean since expectations of order statistics correspond to quantiles. However the approach suggested by Post (2003) based on the bootstrap is valid (see Nelson and Pope (1992) for an early use of bootstrap in SD tests). Indeed bootstrapping the ranked observations or the initial observations does not affect bootstrap distributions of test statistics, at least in an i.i.d. framework.

[2] Other SD tests has been suggested in the literature; see e.g. Anderson (1996), Beach and Davidson (1983), Davidson and Duclos (2000). However these tests rely on pairwise comparisons made at a fixed number of arbitrary chosen points. This is not a desirable feature since it introduces the possibility of test inconsistency.

portfolios, then in the absence of market frictions, the portfolio of all invested wealth, or the market portfolio, will itself be a reward-risk efficient portfolio. The market portfolio should therefore be itself efficient in the sense of SSD according to that theory (see De Giorgi and Post (2005) for a rigorous derivation of this result). This reasoning is similar to the one underlying the derivation of the CAPM (Sharpe (1964), Lintner (1965)), where investors optimally hold mean-variance efficient portfolios. A direct test for SSDE of the market portfolio can be viewed as a nonparametric way to test empirically for such a theory.

The paper is organized as follows. In Section 2, we recall the notion of SDE introduced by Kuosmanen (2004) and Post (2003), and discuss the general hypotheses for testing SDE at any order. We describe the test statistics, and analyse the asymptotic properties of the testing procedures. We follow BD, who extend and justify the procedure of McFadden (1989) (see also Abadie (2002)) leading to consistent tests of SD, and we use simulation based procedures to compute p-values. From a technical point of view, we modify their work to accommodate the presence of full diversification and time-dependent data. Note that other resampling methods such as subsampling are also available (see Linton, Maasoumi and Whang (2005) for the standard SD tests). Linton, Post and Whang (2005) follow this route, in the context of testing procedures for SDE. They use subsampling to estimate the p-values, and they also discuss power issues of the testing procedures. In their work they focus on the dominance criteria of order two and three. In our paper, we also test for FSDE, although it gives necessary and not sufficient optimality conditions, as argued in Post (2005). The FSD criterion places on the form of the utility function beyond the usual requirement that it is nondecreasing, i.e., investors prefer more to less. Thus, this criterion is appropriate for both risk averters and risk lovers since the utility function may contain concave as well as convex segments. Owing to its generality, the FSD permits a preliminary screening of investment alternatives eliminating those which no rational investor will ever choose. The SSD criterion adds the assumption of global risk aversion. This criterion is based on a stronger assumption and therefore, it permits a more sensible selection of investments. The test statistic for SSD is formulated in terms of standard linear programming (LP). Note that numerical implementation of FSD is much more difficult since we need to develop mixed integer programs (MIP). Nevertheless, standard widely available algorithms can be used to compute both test statistics.

In Section 3 we discuss two practical ways to compute the p-values for testing SDE. The first one relies on a simulation-based multiplier method while the second relies on a block bootstrap method. In Section 4 we provide an empirical illustration. We analyze whether the Fama and French market portfolio can be considered as efficient according to FSD, and SSD criteria when confronted to diversification principles made of the 6 Fama and French benchmark portfolios formed on size and book-to-market equity ratio, and the riskless asset. The fact that many institutional investors invest in mutual funds, motivate us to test the efficiency of the market portfolio. These funds track value-weighted equity indices which strongly resemble the market portfolio. We find that the market portfolio is FSD and SSD efficient. We give some concluding remarks in Section 5. Proofs and mathematical programming formulations are gathered in an Appendix.

2 Testing for stochastic dominance efficiency

We consider a strictly stationary process $\{\boldsymbol{Y}_t, t \in \mathbb{Z}\}$ taking values in a compact set $\mathbb{E} \subset \mathbb{R}^n$. The observations consist in a realization of $\{\boldsymbol{Y}_t;\ t = 1, ..., T\}$. These data correspond to observed returns of n financial assets. For inference we also need the process being strong mixing (α-mixing) with mixing coefficients α_t such that $\alpha_T = O(T^{-a})$ for some $a > 1$ as $T \to \infty$ (see Doukhan (1994) for relevant definition and examples). In particular returns generated by various stationary ARMA, GARCH and stochastic volatility models meet this requirement (Carrasco and Chen (1998)). We denote by $F(\boldsymbol{y})$, the continuous cdf of $\boldsymbol{Y} = (Y_1, ..., Y_n)'$ at point $\boldsymbol{y} = (y_1, ..., y_n)'$.

Let us consider a portfolio $\boldsymbol{\lambda} \in \mathbb{L}$ where $\mathbb{L} := \{\boldsymbol{\lambda} \in \mathbb{R}^n_+ : \boldsymbol{e}'\boldsymbol{\lambda} = 1\}$ with $\boldsymbol{e}$ for a vector made of ones. This means that short sales are not allowed and that the portfolio weights sum to one. Let us denote by $G(z, \boldsymbol{\lambda}; F)$ the cdf of the portfolio return $\boldsymbol{\lambda}'\boldsymbol{Y}$ at point z given by:

$$G(z, \boldsymbol{\lambda}; F) = \int_{\mathbb{E}} \mathbb{I}\{\boldsymbol{\lambda}'\boldsymbol{u} \leq z\} dF(\boldsymbol{u}).$$

We denote by a the largest scalar so that $G(a, \boldsymbol{\lambda}; F) = 0$ for all portfolios $\boldsymbol{\lambda} \in \mathbb{L}$, and b the smallest scalar so that $G(b, \boldsymbol{\lambda}; F) = 1$ for all portfolios $\boldsymbol{\lambda} \in \mathbb{L}$. Then the interval made of the upper bond a and the lower bound b forms the narrowest common support of all portfolio return distributions, and the nature of $\mathbb{E}$ and $\mathbb{L}$ ensures that the interval $[a, b]$ is finite.

Let us define for $z \leq b$:

$$\begin{aligned}
\mathcal{J}_1(z, \boldsymbol{\lambda}; F) &= G(z, \boldsymbol{\lambda}; F), \\
\mathcal{J}_2(z, \boldsymbol{\lambda}; F) &= \int_a^z G(u, \boldsymbol{\lambda}; F) du = \int_a^z \mathcal{J}_1(u, \boldsymbol{\lambda}; F) du, \\
\mathcal{J}_3(z, \boldsymbol{\lambda}; F) &= \int_a^z \int_a^u G(v, \boldsymbol{\lambda}; F) dv du = \int_a^z \mathcal{J}_2(u, \boldsymbol{\lambda}; F) du,
\end{aligned}$$

and so on. [3]

From DD Eq. (2), we know that

$$\mathcal{J}_j(z, \boldsymbol{\lambda}; F) = \int_a^z \frac{1}{(j-1)!} (z - u)^{j-1} dG(u, \boldsymbol{\lambda}, F),$$

which can be rewritten as

$$\mathcal{J}_j(z, \boldsymbol{\lambda}; F) = \int_{\mathbb{E}} \frac{1}{(j-1)!} (z - \boldsymbol{\lambda}'\boldsymbol{u})^{j-1} \mathbb{I}(\boldsymbol{\lambda}'\boldsymbol{u} \leq z) dF(\boldsymbol{u}). \tag{2.1}$$

[3] These multiple integrals of the cdf are not well defined if a or b are infinite. This is the case for instance when the support of F is $\mathbb{R}^n$ instead of $\mathbb{E}$. Note however that we may consider any finite interval containing $[a, b]$ and define the integrals accordingly. Hence if we take such an interval sufficiently large, this should not affect the stochastic ordering in practice.

The general hypotheses for testing the stochastic dominance efficiency of order j of $\boldsymbol{\tau}$, hereafter SDE_j, can be written compactly as: [4]

$$\begin{aligned} H_0^j: &\ \mathcal{J}_j(z,\boldsymbol{\tau};F) \leq \mathcal{J}_j(z,\boldsymbol{\lambda};F) \quad \text{for all } z \in [a,b] \text{ and for all } \boldsymbol{\lambda} \in \mathbb{L}, \\ H_1^j: &\ \mathcal{J}_j(z,\boldsymbol{\tau};F) > \mathcal{J}_j(z,\boldsymbol{\lambda};F) \quad \text{for some } z \in [a,b] \text{ or for some } \boldsymbol{\lambda} \in \mathbb{L}. \end{aligned}$$

In particular we get SSDE when $j = 2$.

The empirical counterpart to (2.1) is simply obtained by integrating with respect to the empirical distribution $\hat{F}$ of F, which yields:

$$\mathcal{J}_j(z,\boldsymbol{\lambda};\hat{F}) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{(j-1)!}(z-\boldsymbol{\lambda}'\boldsymbol{Y}_t)^{j-1}\mathbb{I}(\boldsymbol{\lambda}'\boldsymbol{Y}_t \leq z), \qquad (2.2)$$

and can be rewritten more compactly for $j \geq 2$ as:

$$\mathcal{J}_j(z,\boldsymbol{\lambda};\hat{F}) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{(j-1)!}(z-\boldsymbol{\lambda}'\boldsymbol{Y}_t)_+^{j-1},$$

where $(x)_+ = \max(x,0)$.

Since $\sqrt{T}(\hat{F}-F)$ tends weakly to a mean zero Gaussian process $\mathcal{B}\circ F$ in the space of continuous functions on $\mathbb{E}$ (see e.g. the multivariate functional central limit theorem for stationary strongly mixing sequences stated in Rio (2000) [5]), we may derive the limiting behaviour of (2.2) using the Continuous Mapping Theorem (as in Lemma 1 of BD).

Lemma 2.1. *The vector* $\sqrt{T}[\mathcal{J}_j(\cdot;\hat{F}) - \mathcal{J}_j(\cdot;F)]$ *tends weakly to a Gaussian process* $\mathcal{J}_j(\cdot;\mathcal{B}\circ F)$ *with mean zero and covariance function given by:*

- for $j = 1$*:*

$$\begin{aligned} \Omega_1(z_1,z_2,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) &= E[G(z_1,\boldsymbol{\lambda}_1;\mathcal{B}\circ F)G(z_2,\boldsymbol{\lambda}_2;\mathcal{B}\circ F)] \\ &= \sum_{t\in\mathbb{Z}} E\left[\mathbb{I}\{\boldsymbol{\lambda}_1'\boldsymbol{Y}_0 \leq z_1\}\mathbb{I}\{\boldsymbol{\lambda}_2'\boldsymbol{Y}_t \leq z_2\}\right] - G(z_1,\boldsymbol{\lambda}_1;F)G(z_2,\boldsymbol{\lambda}_2;F), \end{aligned}$$

- for $j \geq 2$*:*

$$\Omega_j(z_1,z_2,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) = E\left[\mathcal{J}_j(z_1,\boldsymbol{\lambda}_1;\mathcal{B}\circ F)\mathcal{J}_j(z_2,\boldsymbol{\lambda}_2;\mathcal{B}\circ F)\right]$$

[4]Remark that the general hypotheses for testing the stochastic dominance of order j of the distribution of portfolio $\boldsymbol{\tau}$ over the distribution of portfolio $\boldsymbol{\lambda}$ take analoguous forms but for a *given* $\boldsymbol{\lambda}$ instead of several of them. Indeed the notion of stochastic dominance efficiency is a straightforward extension to the case where full diversification is allowed. Note also that saying that a portfolio dominates all other portfolios is a sufficient but not necessary condition for efficiency (being not-dominated). The last formulation is the one used by Kuosmanen (2004) and Post (2003).

[5]We may also assume a weaker, but less standard, condition, namely that the process is weakly dependent in the sense of Doukhan and Louhichi (1999) and rely on the multidimensionl functional central limit theorem in Doukhan, Fermanian and Lang (2004).

$$= \sum_{t \in \mathbb{Z}} \frac{1}{((j-1)!)^2} E\left[(z_1 - \boldsymbol{\lambda}_1' \boldsymbol{Y}_0)_+^{j-1} (z_2 - \boldsymbol{\lambda}_2' \boldsymbol{Y}_t)_+^{j-1}\right] - \mathcal{J}_j(z_1, \boldsymbol{\lambda}_1; F)\mathcal{J}_j(z_2, \boldsymbol{\lambda}_2; F),$$

with $(z_1, z_2)' \in [a, b]^2$ *and* $(\boldsymbol{\lambda}_1', \boldsymbol{\lambda}_2')' \in \mathbb{L}^2$.

For i.i.d. data the covariance kernel reduces to

$$\Omega_1(z_1, z_2, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = E\left[\mathbb{I}\{\boldsymbol{\lambda}_1' \boldsymbol{Y} \le z_1\}\mathbb{I}\{\boldsymbol{\lambda}_2' \boldsymbol{Y} \le z_2\}\right] - G(z_1, \boldsymbol{\lambda}_1; F)G(z_2, \boldsymbol{\lambda}_2; F),$$

and for $j \ge 2$:

$$\Omega_j(z_1, z_2, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \frac{1}{((j-1)!)^2} E\left[(z_1 - \boldsymbol{\lambda}_1' \boldsymbol{Y})_+^{j-1} (z_2 - \boldsymbol{\lambda}_2' \boldsymbol{Y})_+^{j-1}\right] - \mathcal{J}_j(z_1, \boldsymbol{\lambda}_1; F)\mathcal{J}_j(z_2, \boldsymbol{\lambda}_2; F).$$

Let us consider the test statistic

$$\hat{S}_j = \sqrt{T} \sup_{z, \boldsymbol{\lambda}} \left[\mathcal{J}_j(z, \boldsymbol{\tau}; \hat{F}) - \mathcal{J}_j(z, \boldsymbol{\lambda}; \hat{F})\right],$$

and a test based on the decision rule:

$$\text{“ reject } \quad H_0^j \quad \text{if} \quad \hat{S}_j > c_j \text{ ”},$$

where c_j is some critical value that will be discussed in a moment.

The following result characterizes the properties of the test, where

$$\bar{S}_j \quad := \quad \sup_{z, \boldsymbol{\lambda}} \left[\mathcal{J}_j(z, \boldsymbol{\tau}; \mathcal{B} \circ F) - \mathcal{J}_j(z, \boldsymbol{\lambda}; \mathcal{B} \circ F)\right].$$

Proposition 2.2. *Let c_j be a positive finite constant, then:*

(i) if H_0^j is true,

$$\lim_{T \to \infty} P[\textit{reject}\, H_0^j] \le P[\bar{S}_j > c_j] := \alpha(c_j),$$

with equality when $G(z, \boldsymbol{\lambda}; F) = G(z, \boldsymbol{\tau}; F)$ for all $z \in [a, b]$ and some $\boldsymbol{\lambda} \in \mathbb{L}$;

(ii) if H_0 is false,

$$\lim_{T \to \infty} P[\textit{reject}\, H_0] = 1.$$

The result provides a random variable [6] that dominates the limiting random variable corresponding to the test statistic under the null hypothesis. The inequality yield a test that will never reject more often than $\alpha(c_j)$ for any portfolio τ satisfying the null hypothesis. As noted in the result the probability of rejection will asymptotically

[6] We conjecture that a similar result holds with $\sup_{z,\boldsymbol{\lambda}} \mathcal{J}_j(z, \boldsymbol{\lambda}; \mathcal{B} \circ F)$ substituted for $\bar{S}_j$. We have not been able to show this because of the complexity of the covariance of the empirical process which impedes us to exploit the Slepian-Fernique-Marcus-Shepp inequality (see Proposition A.2.6 of VW) as in BD. Note that this second result would not bring a significant improvement in the numerical tractability of SD-tests as opposed to the context of BD.

be exactly $\alpha(c_j)$ when $G(z, \boldsymbol{\lambda}; F) = G(z, \boldsymbol{\tau}; F)$ for all $z \in [a, b]$ and some $\boldsymbol{\lambda} \in \mathbb{L}$. The first part implies that if one could find a c_j to set the $\alpha(c_j)$ to some desired probability level (say the conventional 0.05 or 0.01) then this would be the significance level for composite null hypotheses in the sense described by Lehmann (1986). The second part of the result indicates that the test is capable of detecting any violation of the full set of restrictions of the null hypothesis. Of course, in order to make the result operational, we need to find an appropriate critical value c_j. Since the distribution of the test statistic depends on the underlying distribution, this is not an easy task, and we decide hereafter to rely on two different methods to simulate p-values. The first one is based on a simulation based multiplier method and the second one on block bootstrap methods.

3 Simulating p-values

3.1 Multiplier method

In this section we consider the use of a similation based method which exploits the multiplier central limit theory discussed in VW (1996) Section 2.9. [7] The idea is to rely on artificial random numbers to simulate a process that is identical but (asymptotically) independent of $\mathcal{B} \circ F$.

To do this let $\{U_t;\ t = 1, ..., T\}$ denote a sequence of i.i.d. $N(0, 1)$ random variables that are independent of the data sample. Then the process is easily generated from:

$$\tilde{\mathcal{B}}(\boldsymbol{y}, \hat{F}) \ := \ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\mathbb{I}\{\boldsymbol{Y}_t \leq \boldsymbol{y}\} - \hat{F}(\boldsymbol{y}) \right] U_t,$$

where $\tilde{\mathcal{B}}(\boldsymbol{y}, \hat{F})$ is the value of the simulated process $\tilde{\mathcal{B}} \circ \hat{F}$ at point $\boldsymbol{y}$. We may then do inference by computing p-values of appropriate functionals of the simulated process.

Indeed if we consider

$$\tilde{S}_j := \sup_{z, \boldsymbol{\lambda}} \left[J_j(z, \boldsymbol{\tau}; \tilde{\mathcal{B}} \circ \hat{F}) - J_j(z, \boldsymbol{\lambda}; \tilde{\mathcal{B}} \circ \hat{F}) \right],$$

the p-values can be estimated from:

$$\tilde{p}_j := P_U[\tilde{S}_j > \hat{S}_j],$$

where P_U is the probability function associated with the normal random variable U and is conditional on the realized sample. The following result provides the decision rule in this environment.

Proposition 3.1. *Assuming that $\alpha < 1/2$, a test for SDE_j based on the rule:*

$$\text{“ reject } H_0^j \quad \text{if} \quad \tilde{p}_j < \alpha \text{”},$$

[7]See BD for use in stochastic dominance test, and Hansen (1996), Guay and Scaillet (2003) for other uses.

satisfies the following

$$\lim P[\textit{reject}\, H_0^j] \leq \alpha \quad \textit{if} \quad H_0^j \textit{ is true},$$
$$\lim P[\textit{reject}\, H_0^j] = 1 \quad \textit{if} \quad H_0^j \textit{ is false}.$$

The multiplier method can be justified by showing that the simulated process $\tilde{\mathcal{B}} \circ \hat{F}$ converges weakly to an identical independent copy of the Gaussian process $\mathcal{B} \circ F$. An application of the continuous mapping theorem shows then that we get a simulated copy of the bounding random variables that appears in Proposition 2.2. In practice we need to use Monte-Carlo methods to approximate the probability. The p-value is simply approximated by

$$\tilde{p}_j \approx \frac{1}{R} \sum_{r=1}^{R} \mathbb{I}\{\tilde{S}_{j,r} > \hat{S}_j\},$$

where the averaging is made on R replications. Note that the replication number can be chosen to make the approximations as accurate as one desires given time and computer constraints.

3.2 Block bootstrap methods

The second method relies on block bootstrap methods. [8] Block boostrap methods extend the nonparametric i.i.d. bootstrap to a time series context. They are based on "blocking" arguments, in which data are divided into blocks and those, rather than individual data, are resampled in order to mimick the time dependent structure of the original data. An alternative resampling technique could be subsampling, for which similar results can be shown to hold as well (see Linton, Maasoumi, and Whang (2001) for use in stochastic dominance tests).

Let b, l denote integers such that $T = bl$. We distinguish hereafter two different ways of proceeding, depending on whether the blocks are overlapping or nonoverlapping. The overlapping rule (Kunsch (1989)) produces $T - l + 1$ overlapping blocks, the kth being $\boldsymbol{B}_k = (\boldsymbol{Y}'_k, ..., \boldsymbol{Y}'_{k+l-1})'$ with $k \in \{1, ..., T - l + 1\}$. The nonoverlapping rule (Carlstein (1986)) just asks the data to be divided into b disjoint blocks, the kth being $\boldsymbol{B}_k = (\boldsymbol{Y}'_{(k-1)l+1}, ..., \boldsymbol{Y}'_{kl})'$ with $k \in \{1, ..., b\}$. In either case the block bootstrap method requires that we choose blocks $\boldsymbol{B}^*_1, ..., \boldsymbol{B}^*_b$ by resampling randomly, with replacement, from the set of overlapping or nonoverlapping blocks. If $\boldsymbol{B}^*_i = (\boldsymbol{Y}^{*\prime}_{i1}, ..., \boldsymbol{Y}^{*\prime}_{il})'$, a block bootstrap sample $\{\hat{\boldsymbol{Y}}^*_t;\ t = 1, ..., T\}$ is made of $\{\boldsymbol{Y}^*_{11}, ..., \boldsymbol{Y}^*_{1l}, \boldsymbol{Y}^*_{21}, ..., \boldsymbol{Y}^*_{2l}, ..., \boldsymbol{Y}^*_{b1}, ..., \boldsymbol{Y}^*_{bl}\}$, and we let $\hat{F}^*$ denote its empirical distribution.

Let $\hat{E}^*$ denote the expectation operator with respect to the probability measure induced by block bootstrap sampling. If the blocks are nonoverlapping, then $\hat{E}^* J_j(z, \boldsymbol{\lambda}; \hat{F}^*) = J_j(z, \boldsymbol{\lambda}; \hat{F})$. In contrast $\hat{E}^* J_j(z, \boldsymbol{\lambda}; \hat{F}^*) \neq J_j(z, \boldsymbol{\lambda}; \hat{F})$ under an overlapping scheme (Hall, Horowitz, and Jing (1996)). The resulting bias decreases the rate of convergence of the estimation errors of the block bootstrap with overlapping

[8] See BD and Abadie (2002) for use of the nonparametric i.i.d. bootstrap in stochastic dominance tests.

blocks. Fortunately this problem can be solved easily by recentering the test statistic. [9] Hence let us introduce

$$S_j^* := \sqrt{T} \sup_{z,\boldsymbol{\lambda}} \left\{ \left[J_j(z,\boldsymbol{\tau};\hat{F}^*) - \hat{E}^* J_j(z,\boldsymbol{\tau};\hat{F}^*) \right] - \left[J_j(z,\boldsymbol{\lambda};\hat{F}^*) - \hat{E}^* J_j(z,\boldsymbol{\lambda};\hat{F}^*) \right] \right\},$$

with, for the overlapping rule,

$$\hat{E}^* J_j(z,\boldsymbol{\lambda};\hat{F}^*) = \frac{1}{T-l+1} \sum_{t=1}^{T} w(t,l,T) \frac{1}{(j-1)!} (z - \boldsymbol{\lambda}'\boldsymbol{Y}_t)^{j-1} \mathbb{I}\{\boldsymbol{\lambda}'\boldsymbol{Y}_t \leq z\},$$

where

$$w(t,l,T) = \begin{cases} t/l & \text{if} \quad t \in \{1,...,l-1\}, \\ 1 & \text{if} \quad t \in \{l,...,T-l+1\}, \\ (T-t+1)/l & \text{if} \quad t \in \{T-l+2,...,T\}, \end{cases}$$

and with, for the nonoverlapping rule, $\hat{E}^* J_j(z,\boldsymbol{\lambda};\hat{F}^*) = J_j(z,\boldsymbol{\lambda};\hat{F})$.

Let us define

$$p_j^* := P[S_j^* > \hat{S}_j].$$

Then the block bootstrap method is justified by the next statement. [10]

Proposition 3.2. *Assuming that $\alpha < 1/2$, a test for SDE_j based on the rule:*

$$\text{“ } reject\, H_0^j \quad if \quad p_j^* < \alpha \text{ ”},$$

satisfies the following

$$\begin{aligned} \lim P[reject\, H_0^j] &\leq \alpha \quad if \quad H_0^j \text{ is true,} \\ \lim P[reject\, H_0^j] &= 1 \quad if \quad H_0^j \text{ is false.} \end{aligned}$$

Again we need to rely on Monte-Carlo methods to approximate the probability in a manner analogous to the one of the previous subsection.

4 Empirical application

In this section we present the results of an empirical application. To illustrate the potential of the proposed test statistics, we test whether different SDE criteria (FSDE, and SSDE) rationalize the market portfolio. Although we focus the analysis on testing the SSD efficiency of the market portfolio, we additionally test FSD to examine the degree of the subjects' rationality (in the sense that they prefer more to less).

[9] See also Hall and Horowitz (1996), Andrews (2002) for a discussion of the need for recentering to avoid excessive bias in tests based on extremum estimators.

[10] Note that deleting the expectations, i.e. setting $\hat{E}^* J_j(z,\boldsymbol{\tau};\hat{F}^*) = 0$ and $\hat{E}^* J_j(z,\boldsymbol{\lambda};\hat{F}^*) = 0$ under both rules, would not change the statement.

4.1 Description of the data

We use the 6 Fama and French benchmark portfolios constructed as the intersections of 2 ME portfolios and 3 BE/ME portfolios. The portfolios are constructed at the end of June. ME is market cap at the end of June. BE/ME is book equity at the last fiscal year end of the prior calendar year divided by ME at the end of December of the prior year. Firms with negative BE are not included in any portfolio. The annual returns are from January to December. We use data on monthly returns (month-end to month-end) from July 1963 to October 2001 (460 monthly observations) obtained from the data library on the homepage of Kenneth French[11]. Excess returns are computed from the raw return observations by subtracting the return on the one-month U.S. Treasury bill. We also use the Fama and French market portfolio, which is the value-weighted average of all non-financial common stocks listed on NYSE, AMEX, and Nasdaq, and covered by CRSP and COMPUSTAT. Further, we use the one-month U.S. Treasury bill as the riskless asset.

First we analyze the statistical characteristics of the data covering the period from July 1963 to October 2001 (460 monthly observations) that are used in the test statistics. As we can see from Table 4.1, portfolio returns exhibit considerable variance in comparison to their mean. Moreover, the skewness and kurtosis indicate that normality cannot be accepted for the majority of them. These observations suggest to adopt the FSD and SSD tests which account for the full return distribution and not only the mean and the variance.

Descriptive Statistics

No.	Mean	Std. Dev.	Skewness	Kurtosis	Minimum	Maximum
Market Portfolio	0.462	4.461	-0.498	2.176	-23.09	16.05
1	0.316	7.07	-0.337	-1.033	-32.63	28.01
2	0.726	5.378	-0.512	0.570	-28.13	26.26
3	0.885	5.385	-0.298	1.628	-28.25	29.56
4	0.323	4.812	-0.291	-1.135	-23.67	20.48
5	0.399	4.269	-0.247	-0.706	-21.00	16.53
6	0.581	4.382	-0.069	-0.929	-19.46	20.46

Table 4.1: Descriptive statistics of monthly percentage excess returns (in %) from July 1963 to October 2001 (460 monthly observations) for the Fama and French market portfolio and the 6 Fama and French benchmark portfolios formed on size and book-to-market equity ratio. Portfolio 1 has Low BE/ME and Small size, portfolio 2 has Medium BE/ME and Small Size, portfolio 3 has High BE/ME and small size...portfolio 6 has High BE/ME and Large size.

[11] http://mba.turc.dartmouth.edu/pages/faculty/ken.french

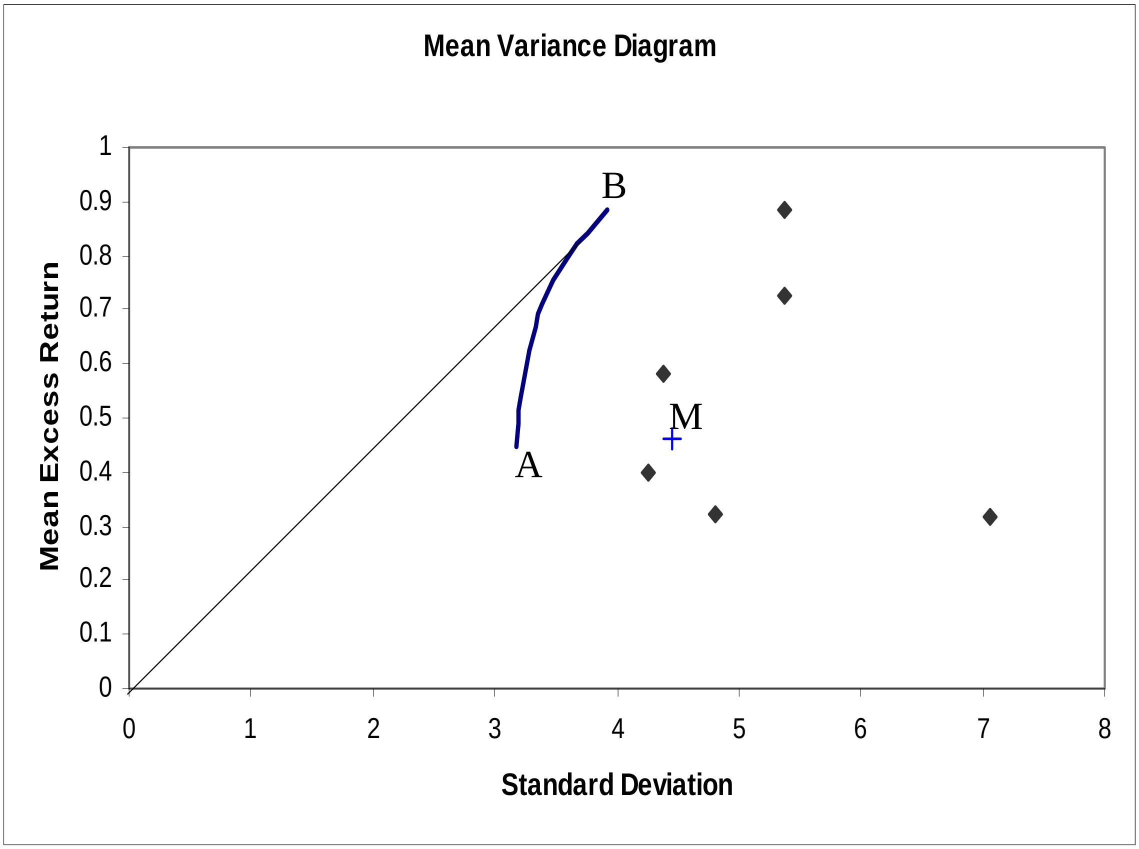

Figure 4.1: Mean-Standard Deviation efficient frontier of the 6 Fama and French risky benchmark portfolios. The diagram also shows the mean excess returns and standard deviations of the individual benchmark portfolios as well as the Fama and French Market (M) portfolio, which is the test portfolio.

One interesting feature is the comparison of the behavior of the market portfolio with that of the individual portfolios. Figure 4.1 shows the Mean-Standard Deviation efficient frontier of the 6 Fama and French risky benchmark portfolios. This Figure diagram also shows the mean excess returns and standard deviations of the individual benchmark portfolios as well as the Fama and French Market (M) portfolio, which will be the test portfolio. We observe that the test portfolio (M) is Mean-Standard Deviation inefficient. It is clear that we can construct portfolios that achieve a higher expected return for the same level of standard deviation, and a lower standard deviation for the same level of expected return. If the investor utility function is not quadratic, then the risk profile of the benchmark portfolios cannot be totally captured by the variance of these portfolios. Generally, the variance is not a satisfactory measure. It is a symmetric measure that penalizes gains and losses in the same way. Moreover, the variance is inappropriate to describe the risk of low probability events. Finally, the mean-variance approach is not consistent with second-order stochastic dominance. This is well illustrated by the mean-variance Paradox and motivates us to test whether the market portfolio is FSD and SSD efficient.

4.2 The SDE test results

We use the block bootstrap method (proposed in section 3) to compute the p-values for testing SDE. [12] They are computed by comparing the test statistics with a number of recentered bootstrap samples. We have found a significant autocorrelation of order one at a 5% significance level in benchmark portfolios 1 to 3, while ARCH effects are present in benchmark portfolio 4 at a 5% significance level. This indicates that a block bootstrap approach should be favoured over a standard i.i.d. bootstrap approach. The original data are divided into blocks. Since the autocorrelations die out quickly, we may take a block of small size to compute the test statistics. We have chosen a size of 10 observations. The blocks are overlapping, to avoid excessive bias in the tests [13]. The p-values are approximated by

$$\tilde{p}_j \approx \frac{1}{R}\sum_{r=1}^{R} \mathbb{I}\{\tilde{S}_{j,r} > \hat{S}_j\}, \qquad j = 1, 2,$$

where the averaging is made on 300 replications. This number guarantees that the approximations are accurate enough, given time and computer constraints. We reject the null hypothesis if the p-value is lower than the significance level of 5%.

For the FSD efficiency, we found that the market portfolio is FSD highly and significantly efficient. The p-value of the test statistic is $\tilde{p}_j \approx 0.55$. The market portfolio is found to be efficient in 167 out of 300 replications.

Our test results suggest also that the market portfolio is highly and significantly SSD efficient. The p-value of the test statistic is $\tilde{p}_j \approx 0.59$. The market portfolio is found to be efficient in 179 out of 300 replications. Although Figure 4.1 shows that the market portfolio is inefficient compared to the benchmark portfolios in the mean-variance scheme, the FSD and SSD efficiency of the market portfolio proves the opposite under more general schemes. These results indicate that the whole distribution rather than the mean and the variance plays an important role in comparing portfolios. This efficiency of the market portfolio is interesting for investors. If the market portfolio is not efficient, individual investors could diversify across diverse asset portfolios and outperform the market.

The fact that the market portfolio is found to be SSDE cannot be attributed to the lack of power of the test statistic. Indeed, we use a long enough time-series of 460 return observations, and a relatively narrow cross-section of 6 benchmark portfolios.

The fact also that our test is only a necessary and not sufficient condition for the efficiency of the market portfolio, does not influence the results. Indeed, the market portfolio SD dominates all other portfolios, thus, it is not dominated by any other portfolio, and therefore it is efficient.

[12] Our experience indicates that using a bootstrap method or a multiplier method do not affect dramatically the outcome of the tests.

[13] We use the overlapping rule because in the nonoverlapping rule we need to recenter the test statistic, and the new S_j^* test is very difficult to optimize, since it involves a large number of binary variables.

4.3 Rolling window analysis

We carried out an additional test to validate the SSD efficiency of the market portfolio and the stability of the model results. It is possible that the efficiency of the market portfolio could change over time, as the risk and preferences of investors change. Therefore, the market portfolio could be efficient in the total sample, but inefficient in some subsamples. Moreover, the degree of efficiency may change over time, as pointed by Post (2003). To control for that, we perform a rolling window analysis, using a window width of 10 years. The test statistic is calculated separately for 340 overlapping 10-year periods, (July 1963-June 1973), (August 1963-July 1973),...,(November 1991-October 2001).

Figure 4.2 shows the corresponding p-values. Interestingly, we observe that the market portfolio is SSD efficient in the total sample period. The SSD efficiency is not rejected for all subsamples. The p-values are always greater than 15%, and in some cases they reach the $80\% - 90\%$. This result confirms the SSD efficiency that was found in the previous subsection, for the whole period. This means that we cannot form an optimal portfolio from the set of the 6 benchmark portfolios that dominates the market portfolio by SSD. The line exhibit large fluctuations, thus the degree of efficiency is changing over time, but remains always above the critical value.

Note that the computational complexity and the associated large solution time of the FSD test are prohibitive for a rolling window analysis, which involves a large number of optimization models (340 rolling windows times 300 bootstraps for each one times 460 programs. 460 is the number of discrete values of z, a discretization that reduces the solution time, as it is explained in the Appendix).

This is a strong indication of the SSD efficiency of the market portfolio, as the resulting p-values remain always above the critical level of 5%.

5 Concluding remarks

In this paper we develop *consistent* tests for stochastic dominance efficiency at *any* order for *time-dependent* data. We propose tests for stochastic dominance efficiency of a given portfolio with respect to all possible portfolios constructed from a set of assets. We propose and justify approaches based on simulation and the block bootstrap to achieve valid inference in a time series setting. Linear as well as mixed integer programs are developed to compute the test statistics and the estimators.

To illustrate the potential of the proposed test statistics, we test whether different SDE criteria (FSDE, and SSDE) rationalize the Fama and French market portfolio over the 6 Fama and French benchmark portfolios constructed as the intersections of 2 ME portfolios and 3 BE/ME portfolios. Empirical results indicate that the market portfolio is SSD efficient. This result is also confirmed in a rolling window analysis. In contrast, the market portfolio is Mean-Variance inefficient, indicating the weakness of the variance to capture the risk.

The next step in this work is to develop estimators of efficiency lines as suggested by Davidson and Duclos (2000) for poverty lines in stochastic dominance. For the

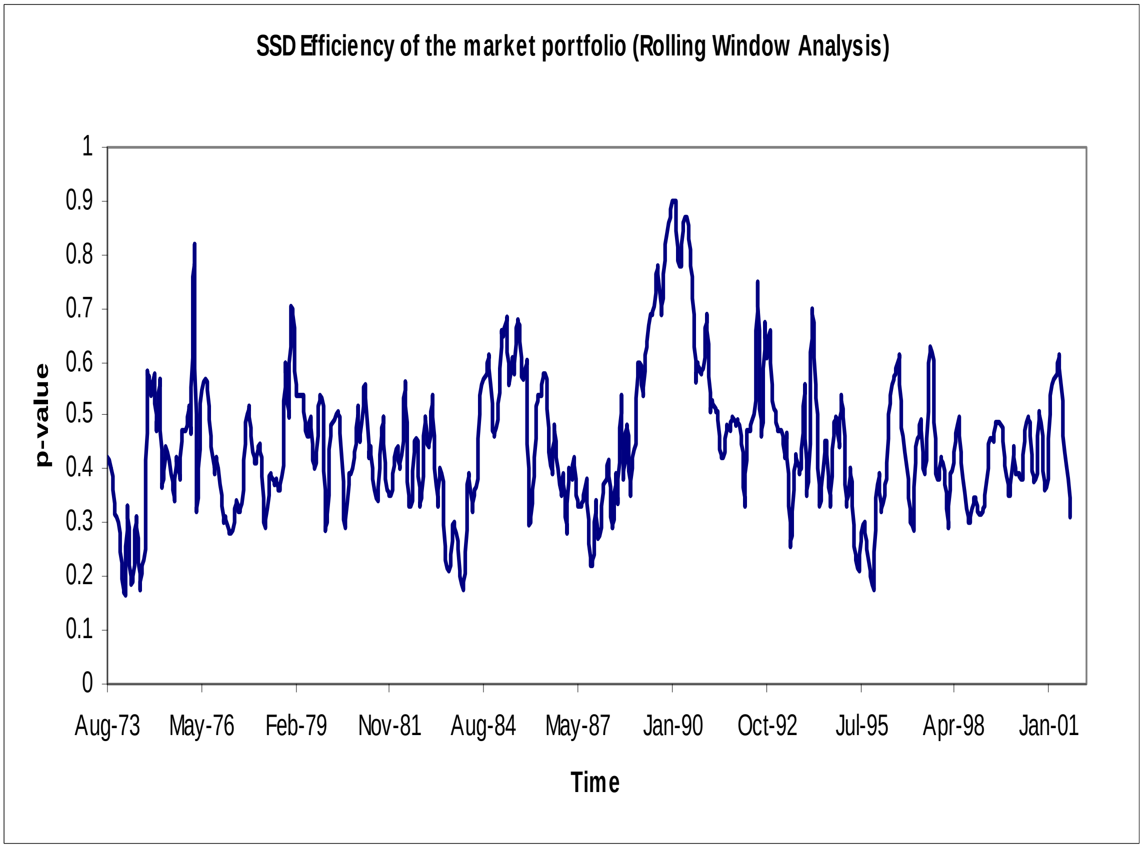

Figure 4.2: p-values for the SSD efficiency test using a rolling window of 120 months. The test statistic is calculated separately for 340 overlapping 10-year periods, (July 1963-June 1973), (August 1963-July 1973),...,(November 1991-October 2001). The SSD efficiency is not rejected.

first order we should estimate the smallest return at which the distribution associated with the portfolio under test and the smallest distribution generated by any portfolios built from the same set of assets intersect. Similarly we could rely on an intersection between integrals of these distributions to determine efficiency line at higher orders.

Another future step in this direction is to further extend these inference procedures to prospect stochastic dominance efficiency (PSDE) and Markowitz stochastic dominance efficiency (MSDE). These concepts allows to take into account that investors react differently to gains and losses and have S-shaped or reverse S-shaped utility functions. The development of necessary and optimal conditions for PSDE and MSDE is of great interest.

Acknowledgements

The authors would like to thank professors Thierry Post and Martine Labbé for their helpful suggestions. Research was financially supported by the Swiss National Science Foundation through NCCR FINRISK. We would also like to thank the participants of the XXXVI EWGFM Meeting on Financial Optimization, Brescia Italy, May 5-7, 2005, and of the Workshop on Optimization in Finance, Coimbra Portugal, July 4-8, 2005, for their helpful comments.

APPENDIX

In the proofs we will often use the shorter notation: $D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};F)=\mathcal{J}_j(z,\boldsymbol{\tau};F)-\mathcal{J}_j(z,\boldsymbol{\lambda};F)$, while all limits are taken as T goes to infinity.

A Proof of Proposition 2.2

1. Proof of Part (i):

By the definition of $\hat{S}_j$ as well as the fact that under H_0^j, $D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};F)\le 0$ for all z and for all $\boldsymbol{\lambda}$, we get:

$$\begin{aligned}\hat{S}_j &\le \sup_{z,\boldsymbol{\lambda}}\sqrt{T}[D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};\hat{F})-D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};F)]+\sup_{z,\boldsymbol{\lambda}}\sqrt{T}D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};F)\\ &\le \sup_{z,\boldsymbol{\lambda}}\sqrt{T}[D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};\hat{F})-D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};F)].\end{aligned}$$

Therefore Part $(A)(ii)$ results from the weak convergence of $\sqrt{T}[D_j(\cdot;\hat{F})-D_j(\cdot;F)]$ and the fact that $\bar{S}_j=\sup_{z,\boldsymbol{\lambda}}D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};\mathcal{B}\circ F)$.

2. Proof of Part (ii):

If the alternative is true, then there exists some z and some $\boldsymbol{\lambda}$, say $\bar{z}\in[a,b]$ and $\bar{\boldsymbol{\lambda}}\in\mathbb{L}$, for which $D_j(\bar{z},\boldsymbol{\tau},\bar{\boldsymbol{\lambda}};F):=\delta>0$. Then the result follows using the inequality $\hat{S}_j\ge\sqrt{T}D_j(z,\boldsymbol{\tau},\boldsymbol{\lambda};\hat{F})$, and the weak convergence of $\sqrt{T}[D_j(\cdot;\hat{F})-D_j(\cdot;F)]$.

B Proof of Proposition 3.1

Let us write

$$\tilde{\mathcal{B}}(\boldsymbol{y},\hat{F}) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\mathbb{I}\{\boldsymbol{Y}_t\le\boldsymbol{y}\}-F(\boldsymbol{y})]\,U_t-\left[\hat{F}(\boldsymbol{y})--F(\boldsymbol{y})\right]\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_t.$$

First consider the second term, and recall that $\sup_{\boldsymbol{y}}|\hat{F}(\boldsymbol{y})-F(\boldsymbol{y})|\longrightarrow 0$ for almost every sample (see e.g. Rio (2000)). Then since the U_i are i.i.d. $N(0,1)$, we have that conditional on the sample:

$$\begin{aligned}&P_U\left[(\sup_{\boldsymbol{y}}|\left[\hat{F}(\boldsymbol{y})-F(\boldsymbol{y})\right]\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_t|>\epsilon)\right]\\ &\qquad=\left[P_U(\sup_{\boldsymbol{y}}|\hat{F}(\boldsymbol{y})-F(\boldsymbol{y})||\frac{1}{\sqrt{T}}\sum_{t=1}^{T}U_t|>\epsilon)\right]\\ &\qquad\le\frac{\left(\sup_{\boldsymbol{y}}|\hat{F}(\boldsymbol{y})-F(\boldsymbol{y})|\right)^2E[\frac{1}{T}\sum_{t=1}^{T}U_t^2]}{\epsilon^2}\longrightarrow 0.\end{aligned}$$

Consequently for this sample we have that $\left[\hat{F}(\boldsymbol{y}) - F(\boldsymbol{y})\right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t \stackrel{p}{\longrightarrow} 0$, which implies $\left[\hat{F}(\boldsymbol{y}) - F(\boldsymbol{y})\right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t \Longrightarrow 0$, and even, since this holds for almost all samples, $\left[\hat{F}(\boldsymbol{y}) - F(\boldsymbol{y})\right] \frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t \stackrel{a.s}{\Longrightarrow} 0$.

For the first term since $\|U\|_1 \equiv E|U| < \infty$, $\|U\|_{2,1} \equiv \int_0^\infty \sqrt{P[|U| > x]}dx < \infty$, and $E \max_{1 \le i \le n} |U_i|/\sqrt{n} \longrightarrow 0$, we deduce from the multiplier inequalities of Lemma 2.9.1 of VW that the asymptotic equicontinuity conditions for the empirical and the multiplier processes are equivalent. Indeed these inequalities derive from triangle inequalities which are unaffected by the time dependence in the data. This means that the first term converges weakly to an independent copy $\mathcal{B}' \circ F$ of $\mathcal{B} \circ F$. As in Theorem 2.9.7 of VW this leads to the almost sure conditional convergence $\sup_{h \in BL_1} |E_U h(\mathcal{B}' \circ F) - Eh(\mathcal{B} \circ F)| \stackrel{a.s.}{\longrightarrow} 0$ where BL_1 is the set of bounded Lipschitz functions on $l^\infty([0,1]^2)$.

Now to show the result concerning the asymptotic behavior of the p-values, let $\tilde{P}_j(t)$ be the c.d.f. of the process (conditional on the original sample) generated by $\tilde{S}_j$. Since $\tilde{\mathcal{B}} \circ \hat{F} \stackrel{a.s}{\Longrightarrow} \mathcal{B}' \circ F$, the CMT gives that

$$\tilde{S}_j \stackrel{a.s.}{\Longrightarrow} \sup_{z, \boldsymbol{\lambda}} D_j(z, \boldsymbol{\tau}, \boldsymbol{\lambda}; \mathcal{B}' \circ F), \tag{B.1}$$

where the latter random variable is an independent copy of $\bar{S}_j$. Note that the median of the distribution $P_j^0(t)$ of $\sup_{z, \boldsymbol{\lambda}} D_j(z, \boldsymbol{\tau}, \boldsymbol{\lambda}; \mathcal{B}' \circ F)$ is strictly positive and finite. Since $D_j(z, \boldsymbol{\tau}, \boldsymbol{\lambda}; \mathcal{B}' \circ F)$ is a Gaussian process indexed by parameters living in a compact set, P_j^0 is absolutely continuous (Tsirel'son (1975)), while $c_j(\alpha)$ defined by $P(\bar{S}_j > c_j(\alpha)) = \alpha$ is finite and positive for any $\alpha < 1/2$ (Proposition A.2.7 of VW). The event $\{\tilde{p}_j < \alpha\}$ is equivalent to the event $\{\hat{S}_j > \tilde{c}_j(\alpha)\}$ where

$$\inf\{t : \tilde{P}_j(t) > 1 - \alpha\} = \tilde{c}_j(\alpha) \stackrel{a.s.}{\longrightarrow} c_j(\alpha), \tag{B.2}$$

by (B.1) and the aforementioned properties of P_j^0. Then:

$$\begin{aligned} \lim P[\text{reject } H_0^j | H_0^j] &= \lim P[\hat{S}_j > \tilde{c}_j(\alpha)] \\ &= \lim P[\hat{S}_j > c_j(\alpha)] + \lim(P[\hat{S}_j > \tilde{c}_j(\alpha)] - P[\hat{S}_j > c_j(\alpha)]) \\ &\le P[\bar{S}_j > c_j(\alpha)] := \alpha, \end{aligned}$$

where the last statement comes from (B.2), part i) of Proposition 2.2 and $c_j(\alpha)$ being a continuity point of the distribution of $\bar{S}_j$. On the other hand part ii) of Proposition 2.2 and $c_j(\alpha) < \infty$ ensure that $\lim P[\text{reject } H_0^j | H_1^j] = 1$.

C Proof of Proposition 3.2

Conditionally on the sample, we have that

$$\sqrt{T}(\hat{F}^* - \hat{F}) \stackrel{p}{\Longrightarrow} \mathcal{B}^* \circ F, \tag{C.1}$$

where $\mathcal{B}^* \circ F$ is an independent copy of $\mathcal{B} \circ F$ (Bühlmann (1994), Peligrad (1998)).

We can see that the functional $D_j(\cdot; F)$ is Hadamard differentiable at F by induction. Indeed $D_1(\cdot; F)$ is a linear functional, while $D_j(\cdot; F)$ is also a linear functional of a Hadamard differentiable mapping $D_{j-1}(\cdot; F)$. The delta method (VW Chapter 3.9), the CMT and (C.1) then yields:

$$S_j^* \overset{p}{\Longrightarrow} \sup_{z, \boldsymbol{\lambda}} D_j(z, \boldsymbol{\tau}, \boldsymbol{\lambda}; \mathcal{B}^* \circ F),$$

where the latter random variable is an independent copy of $\bar{S}_j$, and we can pursue as in the proof of Proposition 3.1 but using convergence in probability instead of almost sure convergence to get the final result.

D Mathematical Programming formulations

D.1 Formulation for FSDE

The test statistic $\hat{S}_1$ for FSDE is derived using MIP. The following is the full formulation of the model:

$$\begin{aligned}
\max_{\mathbf{z}, \boldsymbol{\lambda}} \quad & \hat{S}_1 = \sqrt{T} \frac{1}{T} \sum_{t=1}^{T} (L_t - W_t) && \text{(D.1a)} \\
\text{s.t.} \quad & M(L_t - 1) \leq z - \boldsymbol{\tau}' \boldsymbol{Y}_t \leq M L_t, \quad \forall t && \text{(D.1b)} \\
& M(W_t - 1) \leq z - \boldsymbol{\lambda}' \boldsymbol{Y}_t \leq M W_t, \quad \forall t && \text{(D.1c)} \\
& \boldsymbol{e}' \boldsymbol{\lambda} = 1, && \text{(D.1d)} \\
& \boldsymbol{\lambda} \geq 0, && \text{(D.1e)} \\
& a \leq z \leq b, && \text{(D.1f)} \\
& W_t \in \{0, 1\},\ L_t \in \{0, 1\}, \quad \forall t && \text{(D.1g)}
\end{aligned}$$

with M being a large constant.

The model is a Mixed Integer Program maximizing the distance between two binary variables L_t and W_t which represent the $\mathcal{J}_1(z, \boldsymbol{\tau}; \hat{F})$ and $\mathcal{J}_1(z, \boldsymbol{\lambda}; \hat{F})$, respectively. According to Inequalities (D.1b), L_t equals 1 for each scenario $t \in T$ for which $z \geq \boldsymbol{\tau}' \boldsymbol{Y}_t$ and 0 otherwise. Analogously, Inequalities (D.1c) ensure that W_t equals 1 for each scenario for which $z \geq \boldsymbol{\lambda}' \boldsymbol{Y}_t$. Equation (D.1d) defines the sum of all portfolio weights to be unity, while Inequality (D.1e) disallows for short positions in the available assets.

Under this form, this is a very difficult problem to solve. It takes more that two days to find the optimal solution. We reformulate the problem in order to reduce the solving time and obtaining a tractable formulation. The steps are the following:

1) The factor $\sqrt{T}/T$ can be left out in the objective function, since T is fixed. The goal of this is to avoid numerical problems.

2) We can see that there is a set of at most T values, say $R = \{r_1, r_2, ..., r_T\}$ containing the optimal value of the variable z.

Proof: The vectors $\boldsymbol{\tau}$ and $\boldsymbol{Y}_t$, $t = 1, ..., T$ being given, we can rank the values of $\boldsymbol{\tau}'\boldsymbol{Y}_t, t = 1, ..., T$ by order of increasing values. Let us call $r_1, ..., r_T$ the possible different values of $\boldsymbol{\tau}'\boldsymbol{Y}_t$, with $r_1 < r_2 < ... < r_T$ (actually there may be less than T different values). Now, for any z such that $r_i \leq z \leq r_i + 1, \sum_{t=1,...,T} L_t$ is constant (it is equal to the number of t such that $\boldsymbol{\tau}'\boldsymbol{Y}_t \leq r_i$). Further, when $r_i \leq z \leq r_i + 1$, the maximum value of $-\sum_{t=1,...,T} W_t$ is reached for $z = r_i$. Hence, we can restrict z to belong to the set R.

3) A direct consequence of the above is that we can solve FSDE by solving the smaller problems $P(r)$, $r \in R$, in which z is fixed to r, and then to take the value for z that yields the best total result. The advantage is that the optimal values of the L_t variables are known in $P(r)$. Precisely, $\sum_{t=1,...,T} L_t$ is equal to the number of t such that $\boldsymbol{\tau}'\boldsymbol{Y}_t \leq r$. Hence problem $P(r)$ boils down to:

$$\begin{aligned} \min \quad & \sum_{t=1}^{T} W_t \\ \text{s.t.} \quad & M(W_t - 1) \leq r - \boldsymbol{\lambda}'\boldsymbol{Y}_t \leq MW_t, \quad \forall t \in T \\ & \boldsymbol{e}'\boldsymbol{\lambda} = 1, \\ & \boldsymbol{\lambda} \geq 0, \\ & W_t \in \{0, 1\}, \quad \forall t \in T. \end{aligned} \tag{D.2a}$$

Note that this becomes a minimization problem.

Problem $P(r)$ amounts to find the largest set of constraints $\boldsymbol{\lambda}'\boldsymbol{Y}_t \geq r$ consistent with $\boldsymbol{e}'\boldsymbol{\lambda} = 1$ and $\boldsymbol{\lambda} \geq 0$.

Let $M_t = \min \boldsymbol{Y}_{t,i}, i = 1, ..., n$, i.e. the smallest entry of vector $\boldsymbol{Y}_t$.

Clearly, for all $\boldsymbol{\lambda} \geq 0$ such that $\boldsymbol{e}'\boldsymbol{\lambda} = 1$, we have that $\boldsymbol{\lambda}'\boldsymbol{Y}_t \geq M_t$. Hence, Problem $P(r)$ can be rewritten in an even reduced form:

$$\begin{aligned} \min \quad & \sum_{t=1}^{T} W_t \\ \text{s.t.} \quad & \boldsymbol{\lambda}'\boldsymbol{Y}_t \geq r - (r - M_t)W_t, \quad \forall t \in T \\ & \boldsymbol{e}'\boldsymbol{\lambda} = 1, \\ & \boldsymbol{\lambda} \geq 0, \\ & W_t \in \{0, 1\}, \quad \forall t \in T. \end{aligned} \tag{D.3a}$$

We further reduce $P(r)$ by fixing the following variables:

For all t such that $r \leq M_t$, the optimal value of W_t is equal to 0 since the halfspace defined by the t-th inequality contains the simplex.

For all t such that $r \geq M_t$, the optimal value of W_t is equal to 1 since the halfspace defined by the t-th inequality has an empty intersection with the simplex.

The computational time for this MIP formulation is significantly reduced. For the optimal solution (which involves 460 MIP optimization programs, one for each discrete value of z) it takes less than two hours. This solution time reflects an average for typical solvings of the respective problems. The problems were solved with IBM's CPLEX solver on an Intel Xeon workstation (with a 2*2.4 GHz Power, 6Gb of RAM). We note the almost exponential increase in solution time with the increasing number of observations. Clearly, this solution time is by no means prohibitive with today's available computing technologies. The problems could probably be solved much more efficiently by developing specialized algorithms that exploit the structure of MIP models. However, issues of computational efficiency are not of primary concern in this study.

D.2 Formulation of SSDE

The test statistic $\hat{S}_2$ for SSDE is derived using MIP. The full formulation of the problem is the following:

$$\max_{\mathbf{z},\boldsymbol{\lambda}} \quad \hat{S}_2 = \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(L_t - W_t) \tag{D.4a}$$

$$\text{s.t.} \quad M(F_t - 1) \leq z - \boldsymbol{\tau}'\boldsymbol{Y}_t \leq MF_t, \qquad \forall\, t \tag{D.4b}$$

$$-M(1 - F_t) \leq L_t - (z - \boldsymbol{\tau}'\boldsymbol{Y}_t) \leq M(1 - F_t), \qquad \forall\, t \tag{D.4c}$$

$$-MF_t \leq L_t \leq MF_t, \qquad \forall\, t \tag{D.4d}$$

$$W_t \geq z - \boldsymbol{\lambda}'\boldsymbol{Y}_t, \qquad \forall\, t \tag{D.4e}$$

$$\boldsymbol{e}'\boldsymbol{\lambda} = 1, \tag{D.4f}$$

$$\boldsymbol{\lambda} \geq 0, \tag{D.4g}$$

$$a \leq z \leq b, \tag{D.4h}$$

$$W_t \geq 0,\ F_t \in \{0, 1\}, \qquad \forall\, t \tag{D.4i}$$

with M being a large constant.

The model is a Mixed Integer Program maximizing the distance between two variables L_t and W_t which represent the $\mathcal{J}_2(z, \boldsymbol{\tau}; \hat{F})$ and $\mathcal{J}_2(z, \boldsymbol{\lambda}; \hat{F})$, respectively. This is a difficult problem to solve, since it is the maximization of the difference of two convex functions. We use a binary variable F_t, which, according to Inequalities (D.4b), equals 1 for each scenario $t \in T$ for which $z \geq \boldsymbol{\tau}'\boldsymbol{Y}_t$ and 0 otherwise. Then, Inequalities (D.4c) and (D.4d) ensure that the variable L_t equals $z - \boldsymbol{\tau}'\boldsymbol{Y}_t$ for the scenarios for which this difference is positive, and 0 for all the other scenarios. Inequalities (D.4e) and (D.4i) ensure that W_t equals exactly the difference $z - \boldsymbol{\lambda}'\boldsymbol{Y}_t$ for the scenarios for which this difference is positive and 0 otherwise. Equation (D.4f) defines the sum of all portfolio weights to be unity, while Inequality (D.4g) disallows for short positions in the available assets.

Again, this is a very difficult problem to solve. It takes more than a day for the optimal solution. We reformulate the problem, following the exact steps as for the FSDE. Then the model is transformed to a linear one, which is very easy to solve.

We solve SSDE by solving again smaller problems $P(r)$, $r \in R$, in which z is fixed to r, before taking the value for z that yields the best total result. The advantage is that the optimal values of the L_t variables are known in $P(r)$. Precisely, $L_t = r - \boldsymbol{\tau}'\boldsymbol{Y}_t$, for the scenarios for which this difference is positive, and zero otherwise. Hence problem $P(r)$ boils down to the linear problem:

$$\begin{aligned} \min \quad & \sum_{t=1}^{T} W_t \\ \text{s.t.} \quad & W_t \geq r - \boldsymbol{\lambda}'\boldsymbol{Y}_t, \quad \forall t \in T \\ & \boldsymbol{e}'\boldsymbol{\lambda} = 1, \\ & \boldsymbol{\lambda} \geq 0, \\ & W_t \geq 0, \quad \forall t \in T. \end{aligned} \tag{D.5a}$$

The computational time for this LP formulation is very small. For the optimal solution (which involves 460 linear optimization programs, one for each discrete value of z) using the CPLEX solver, it takes three minutes on average. We can have an even better formulation of this latter model. Instead of solving it for each discrete time of z, we can reformulate the model in order to solve for all discrete values $r_i, i = 1, ..., T$ simultaneously. The new model is the following:

$$\begin{aligned} \min \quad & \sum_{i=1}^{T} \sum_{t=1}^{T} W_{i,t} \\ \text{s.t.} \quad & W_{i,t} \geq r_i - \boldsymbol{\lambda}_i'\boldsymbol{Y}_t, \quad \forall i \in T, \quad \forall t \in T \\ & \boldsymbol{e}'\boldsymbol{\lambda}_i = 1, \quad \forall i \in T, \\ & \boldsymbol{\lambda}_i \geq 0, \quad \forall i \in T, \\ & W_{i,t} \geq 0, \quad \forall i \in T, \quad \forall t \in T. \end{aligned} \tag{D.6a}$$

The optimal portfolio $\boldsymbol{\lambda}_i$ and the optimal value of variable z, r_i are for that i, that gives the $\min \sum_{t=1}^{T} W_{i,t}$. Now, the computational time for this formulation of the problem is less than a minute.

References

[1] Abadie, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97, 284-292.

[2] Anderson, G. (1996): "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica*, 64, 1183-1193.

[3] Andrews, D. (2002): "Higher-Order Improvements of a Computationally Attractive k-Step Bootstrap for Extremum Estimators," *Econometrica*, 64, 891-916.

[4] Barrett, G., and S. Donald (2003): "Consistent Tests for Stochastic Dominance," *Econometrica*, 71, 71-104.

[5] Beach, C., and R. Davidson (1983): "Distribution-free Statistical Inference with Lorenz Curves and Income Shares," *Review of Economic Studies*, 50, 723-735.

[6] Bühlmann, P. (1994): "Blockwise Bootstrapped Empirical Process for Stationary Sequences," *Annals of Statistics*, 22, 995-1012.

[7] Carlstein, E. (1986): "The Use of Subseries Methods for Estimating the Variance of a General Statistic from a Stationary Time Series," *Annals of Statistics*, 14, 1171-1179.

[8] Carrasco, M., and X. Chen (1998): "Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models," *Econometric Theory*, 18, 17-39.

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

[10] De Giorgi, E. (2004): "Reward-Risk Portfolio Selection and Stochastic Dominance," forthcoming in *Journal of Banking and Finance*.

[11] De Giorgi, E. and T. Post (2005): "Second order stochastic dominance, reward-risk portfolio selection and the CAPM," Working paper, NCCR.

[12] Doukhan, P. (1994): *Mixing: Properties and Examples.* Lecture Notes in Statistics, 85. Berlin: Springer-Verlag.

[13] Doukhan, P., J.-D. Fermanian, and G. Lang (2004): "Copulas of a Vector-Valued Stationary Weakly Dependent Process," CREST working paper.

[14] Doukhan, P., and S. Louhichi (1999): "A New Weak Dependence Condition and Applications to Moment Inequalities," *Stochastic Processes and their Applications*, 84, 313-342.

[15] Guay, A. and O. Scaillet (2003): "Indirect Inference, Nuisance Parameter and Threshold Moving Average Models," *Journal of Business and Economic Statistics*, 21, 122-132.

[16] Hall, P. and J. Horowitz (1996): "Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moment Estimators," *Econometrica*, 64, 891-916.

[17] Hall, P., Horowitz, J. and B.-Y. Jing (1995): "On Blocking Rules for the Bootstrap with Dependent Data," *Biometrika*, 82, 561-574.

[18] Hansen, B. (1996): "Inference when a Nuisance Parameter is not Identified under the Null Hypothesis," *Econometrica*, 64, 413-430.

[19] Kahneman, D. and A. Tversky (1979): "Prospect Theory of Decisions under Risk," *Econometrica*, 47, 263-291.

[20] Kroll, Y. and H. Levy (1980): "Stochastic Dominance Criteria : a Review and Some New Evidence," in *Research in Finance*, Vol. II, 263-227, JAI Press, Greenwich.

[21] Kunsch, H. (1989): "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217-1241.

[22] Kuosmanen, T. (2004): "Efficient Diversification According to Stochastic Dominnace Criteria," *Management Science*, 50, 1390-1406.

[23] Lehmann, E. (1986) : *Testing Statistical Hypotheses*, 2nd Ed., John Wiley & Sons, New York.

[24] Levy, H. (1992): "Stochastic Dominance and Expected Utility : Survey and Analysis," *Management Science*, 38, 555-593.

[25] Levy, H. (1998) : *Stochastic Dominance*, Kluwer Academic Publisher, Boston.

[26] Levy, M. and H. Levy (2002): "Prospect Theory: Much Ado about Nothing," *Management Science*, 48, 1334-1349.

[27] Levy, H. and Z. Wiener (1998): "Stochastic Dominance and Prospect Dominance with Subjective Weighting Functions," *Journal of Risk and Uncertainty*, 16, 147-163.

[28] Lintner, J. (1965): "Security Prices, Risk and Maximal Gains from Diversification," *Journal of Finance*, 20, 587-615.

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

[30] Linton, O., T. Post and Y.-J. Whang (2005): "Testing for stochastic dominance efficiency," LSE working paper.

[31] McFadden, D. (1989): "Testing for Stochastic Dominance," in *Studies in the Economics of Uncertainty*, eds Fomby T. and Seo T., 113-134, Springer-Verlag, New York.

[32] Mosler, K. and M. Scarsini (1993): *Stochastic Orders and Applications, a Classified Bibliography*, Springer-Verlag, Berlin.

[33] Nelson, R. and R. Pope (1992): "Bootstrapped Insights into Empirical Applications of Stochastic Dominance," *Management Science*, 37, 1182-1194.

[34] Newey, W. and D. McFadden (1994) : "Large Sample Estimation and Hypotheses Testing," in *Handbook of Econometrics*, Vol. IV, eds R. Engle and D. MacFadden, 2111-2245, North Holland, Amsterdam.

[35] Peligrad, M. (1998): "On the Blockwise Bootstrap for Empirical Processes for Stationary Sequences," *Annals of Probability*, 26, 877-901.

[36] Post, T. (2003): "Empirical Tests for Stochastic Dominance Efficiency," *Journal of Finance*, 58, 1905-1031.

[37] Post, T. (2005): "Wanted: A test for FSD optimality of a given portfolio," Erasmus University working paper.

[38] Post, T. and H. Levy (2003): "Does Risk Seeking Drive Stock Prices? A Stochastic Dominance Analysis of Aggregate Investor Preferences and Beliefs," forthcoming in *Review of Financial Studies*.

[39] Rio, E. (2000): *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*. Mathématiques et Applications, 31. Berlin: Springer-Verlag.

[40] Shaked, M. and J. Shanthikumar (1994): *Stochastic Orders and their Applications*, Academic Press, New York.

[41] Sharpe, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425-442.

[42] Tsirel'son, V. (1975): "The Density of the Distribution of the Maximum of a Gaussian Process," *Theory of Probability and their Applications*, 16, 847-856.

[43] van der Vaart, A. and J. Wellner (1996): *Weak convergence and empirical processes*, Springer Verlag, New York.