MANAGEMENT SCIENCE
Vol. 50, No. 10, October 2004, pp. 1390–1406
ISSN 0025-1909 | EISSN 1526-5501 | 04 | 5010 | 1390
DOI 10.1287/mnsc.1040.0284

Efficient Diversification According to Stochastic Dominance Criteria

Timo Kuosmanen
Environmental Economics and Natural Resources Group, Wageningen University, Hollandseweg 1, 6706 KN Wageningen, The Netherlands, timo.kuosmanen@wur.nl

This paper develops the first operational tests of portfolio efficiency based on the general stochastic dominance (SD) criteria that account for an infinite set of diversification strategies. The main insight is to preserve the cross-sectional dependence of asset returns when forming portfolios by reexpressing the SD criteria in T-dimensional Euclidean space, with elements representing rates of return in T different states of nature. We characterize subsets of this state-space that dominate a given evaluated return vector by first- and second-order SD. This allows us to derive simple SD efficiency measures and test statistics, computable by standard mathematical programming algorithms. The SD tests and efficiency measures are illustrated by an empirical application that analyzes industrial diversification of the market portfolio.

Key words: stochastic dominance; portfolio choice; efficiency; diversification; mathematical programming
History: Accepted by Ravi Jagannathan, finance; received September 19, 2002. This paper was with the author 5 months for 6 revisions.

1. Introduction

Stochastic dominance (SD) criteria are well-established analytical tools for studying decision making under risk and uncertainty (see, e.g., Bawa 1982; and Levy 1992, 1998 for survey and references). The SD approach is nonparametric in the sense that its criteria do not impose explicit specification of an investor's utility function, or restrictions on function forms of probability/frequency distributions. It accounts for the entire probability distribution and employs some general conditions for an investor's risk preferences.

In finance literature, the SD approach is applied to compare the performance of alternative investment portfolios regarding their observed rates of return. Therefore, the empirical application of SD (as well as algorithmic development) has typically focused on discrete empirical distributions where each observed state of nature occurs with equal probability. Bawa et al. (1979) list three good reasons for this approach: (1) Decision making between risky alternatives is primarily based on historical data, thus, discrete distributions are, in practice, the most common objects of choice; (2) if we fully specify one of the most frequently used continuous distributions (e.g., normal), the SD criteria tend to reduce to a simpler form (e.g., the mean-variance rule) so that full-scale comparisons of empirical distributions are not needed; (3) from a Bayesian perspective, when the true distributions are unknown, the use of an empirical distribution function is justified by the von Neumann-Morgenstern-Savage axioms.

Despite the theoretical appeal of the SD criteria, the two-moment *mean-variance* (MV; Markowitz 1952, 1959) approach has been more influential in empirical portfolio analysis. The appeal of the MV model lies in its ability to test and build efficient *diversification* strategies (see the sharp note by Frankfurter and Phillips 1975). In contrast, SD currently offers no methods for identifying *all* efficient diversification strategies, or even for testing whether a given portfolio is SD efficient.[1] The current practice is to use simple enumerative crossing algorithms for revealing dominance relationships between two given portfolios (Levy and Hanoch 1970, Porter et al. 1973, Bawa et al. 1979, Aboudi and Thon 1994, among many others). The difficulty arises when diversification possibilities are introduced. Given an infinite number of alternative diversified portfolios, the standard crossing algorithms cannot confirm (in a finite time) whether or not a given portfolio is SD efficient.[2] Quoting the conclusion from the extensive survey of the SD literature by Levy (1992, p. 583):

> Ironically, the main drawback of the SD framework is found in the area of finance where it is most intensively

[1] It is not very complicated to identify *some* SD efficient diversification strategies (see, e.g., Markowitz 1977, Gavish 1977, Ziemba 1978). For example, the MV efficient portfolios are always second-order SD efficient.

[2] The crossing algorithms are useful in fields where diversification is either impossible or unimportant: Examples include comparisons of income distributions in poverty studies and crop-yield distributions in agricultural economics.

used, namely, in choosing the efficient diversification strategies. This is because as yet there is no way to find the SD efficient set of diversification strategies as prevailed by the M-V framework. Therefore, the next important contribution in this area will probably be in this direction.

It is the goal of this paper to present a first contribution in that direction.[3] The main insight encompasses the following: The conventional approach to SD efficiency ignores the cross-sectional dependence of asset returns when forming portfolios, and sorts return data asset by asset in ascending order to "translate" them into empirical distribution functions (EDFs). The standard definition of the SD is subsequently applied to these EDFs. We observe that the essential difficulties of diversification arise in the translation step: EDF of a diversified portfolio cannot be recovered directly from the marginal EDFs of the underlying assets because they do not contain the original cross-sectional information on state-specific returns. Considering this, a reverse route is adopted in this paper. Marketed portfolios are modeled as vectors of the T-dimensional Euclidean space, representing rates of return in T different states of nature. We do not translate these vectors into marginal EDFs because rates of return for diversified portfolios can only be calculated using cross-sectional information contained in the original state-specific return vectors. Instead, we reexpress the SD criteria in terms of these vectors. More specifically, we analytically characterize subsets of the T-dimensional state-space that dominate a given portfolio return vector by first- and second-order SD (henceforth FSD and SSD, respectively).[4] Interestingly, these dominating sets exhibit a relatively simple polyhedral structure. Utilizing these insights, we develop general SD efficiency tests that compare a given portfolio with an optimally diversified reference portfolio consisting of multiple assets. The test statistics are formulated in terms of standard linear programming (LP) and binary mixed integer linear programming (MILP), which enables us to compute them using standard, widely available algorithms.

This paper builds on three distinct lines of research worth acknowledging. First, the dominating sets characterized herein relate to the theorems on portfolio efficient sets by Peleg and Yaari (1975) and Dybvig and Ross (1982), among others. The main difference from the previous studies is that we build on the general probability distribution definition of the SD, which extends the scope of the analysis beyond the expected utility framework. Second, the analytical characterization of the SSD dominating set is inspired by the treatment of the linear assignment problem by von Neumann (1953) and Koopmans and Beckmann (1957). Third, our efficiency measures and test statistics draw heavily from Farrell's (1957) approach to measuring productive efficiency of the firm, also known as data envelopment analysis (DEA; Charnes et al. 1978).

The rest of the paper is organized as follows. Section 2 introduces notation, defines the key concepts, and explains why the traditional approach to the SD analysis fails to account for portfolio diversification. Section 3 derives analytical characterizations of the FSD and SSD dominating sets and briefly illustrates them using simple graphical examples. We subsequently use these results to derive *exact* SD efficiency tests for both FSD and SSD cases in §4. Section 5 analyzes and compares the computational burden associated with alternative tests. Section 6 applies the new SD techniques to analyze industrial diversification of the market portfolio. Another application to forest portfolio management is provided as an online attachment.[5] Section 8 concludes by exploring directions for further research. Appendix 1 presents formal proofs of the main theorems, and Appendix 2 presents three numerical examples.

2. Preliminaries

Consider two risky portfolios j and k with rates of return distributed according to cumulative distribution functions (CDFs) G_j and G_k, respectively.

Definition 1. Portfolio j *dominates* portfolio k by FSD (SSD), denoted by jD_1k (jD_2k) if and only if

$$\text{FSD:}\quad \begin{aligned} &G_k(r) - G_j(r) \geq 0 \quad \forall r \in \mathbb{R}, \quad \text{and} \\ &G_k(r) - G_j(r) > 0 \quad \text{for some } r \in \mathbb{R}; \end{aligned}$$

$$\text{SSD:}\quad \begin{aligned} &\int_{-\infty}^{r} [G_k(s) - G_j(s)]\, ds \geq 0 \quad \forall r \in \mathbb{R}, \quad \text{and} \\ &\int_{-\infty}^{r} [G_k(s) - G_j(s)]\, ds > 0 \quad \text{for some } r \in \mathbb{R}. \end{aligned}$$

[3] The present paper is a fully revised version of the working paper Kuosmanen (2001), which sketched the first known necessary and sufficient conditions for SD efficiency in the case of full diversification possibilities. Parallel to this research, Post (2003a) addresses the same problem from the dual perspective of the expected utility theory using Afriat's (1967) theorem. Post (2003a) originally developed the computationally tractable simplification that is discussed in §4.4 below, and also provides a detailed discussion of statistical inference in the SD framework.

[4] Extensions to higher-order SD criteria are left for further research. Even without diversification, the standard crossing algorithms become more complicated in cases of higher-order SD because violations of SD can occur outside the discrete set of observed points (see Aboudi and Thon 1994, and Levy 1998). These complications carry over to the present setup as well. The extensions suggested by Kuosmanen (2001) and Post (2003a) ignore these complications.

[5] The electronic companion is available at http://mansci.pubs.informs.org/ecompanion.html.

The SD criteria have the following well-known economic interpretation in the perspective of the standard expected utility theory (henceforth EUT). Consider the Bernoullian utility function $u: \mathbb{R} \to \mathbb{R}$. FSD condition $j\mathrm{D}_1 k$ is equivalent to the preference of portfolio j to k by all nonsatiated investors (i.e., $u'(r) \geq 0\ \forall r \in \mathbb{R}$) as first demonstrated by Quirk and Saposnik (1962). SSD dominance $j\mathrm{D}_2 k$ is, in turn, equivalent to the preference of portfolio j to k by all nonsatiated and risk-averse investors (i.e., $u'(r) \geq 0$, $u''(r) \leq 0\ \forall r \in \mathbb{R}$), as proved by Fishburn (1964) (see also Rothschild and Stiglitz 1970). It should be emphasized, however, that the SD criteria do not depend on the much-debated independence axiom of the EUT (Machina 1982). Indeed, the FSD criterion is valid for most of the alternative non-expected-utility theories for choice under uncertainty.[6] The SSD and higher-order SD criteria are not as directly compatible with alternative non-EUT approaches as FSD, but they can be modified to fit into non-EUT frameworks in a relatively straightforward manner (see, e.g., Levy and Wiener 1998 for an extension of the SD criteria to the prospect theory).

In an empirical portfolio analysis, the underlying probability distributions (G) must be estimated from available data. We thus consider a finite (and therefore discrete) sample of N assets from T states of nature, indexed as $\nu \equiv \{1, 2, \ldots, N\}$ and $\tau \equiv \{1, 2, \ldots, T\}$, respectively. This gives panel data represented by matrix $Y \equiv (Y_1 \ldots Y_N)$ with $Y_j \equiv (Y_{j1} \ldots Y_{jT})'$. States are assumed to be drawn randomly without replacement from a pool of possible states (i.e., the probability that a particular state occurs is the same in all T draws).

Portfolios can be modeled in terms of portfolio weights, or equivalently by using state-specific return vectors (henceforth, *portfolio return profiles*). Portfolio weights are denoted by a column vector $\lambda \in \Lambda$, where $\Lambda \subset \{\lambda \in \mathbb{R}^N \mid \sum_{i=1}^N \lambda_i = 1\}$ represents their feasible domain, which is assumed to be closed and bounded. The set of marketed portfolio return profiles (henceforth, *market set*) spanned by Λ is $\Psi \equiv \{y \in \mathbb{R}^T \mid y = Y\lambda;\ \lambda \in \Lambda\}$.

To derive the EDF for an arbitrary portfolio i, it is customary to rearrange elements of the portfolio return profile $y_i \in \Psi$ in nondecreasing order, and denote the resulting ranked return vector by $x_i: x_{i1} \leq x_{i2} \leq \cdots \leq x_{iT}$. Note that this operation involves a loss of potentially valuable information on the cross-sectional structure of observations; thus, we reserve y, Y for portfolio return profiles, and x for their ranked counterparts. Using ranked vectors x, the empirical distribution function (EDF) for asset j is a step function characterized as $H_n(r) \equiv \max\{t \in \tau \mid r \geq x_{nt}\}/T$. Assuming the observed states are drawn randomly without replacement from the common pool of possible states, the EDF is a nonparametric, minimum-variance, unbiased estimator of the underlying, unobservable CDF characterized by the pool of all possible states. Therefore, the standard approach is to examine SD efficiency in terms of EDF H_n (see, e.g., Bawa 1982 or Levy 1992). We follow this paradigm and apply the SD criteria to EDFs. To distinguish this from the theoretical SD conditions of Definition 1, symbol $\widehat{\mathrm{D}}_l$, $l = 1, 2$ is used for SD relations when an EDF is used to estimate the unknown CDF. The following well-known theorem forms the basis of the standard "crossing algorithm" for testing SD relationships using pairwise comparisons of asset returns (see, e.g., Levy 1992, 1998). It also forms the starting point for the developments in the following sections.[7]

Theorem 1. *The following equivalence results hold for empirical distribution functions of all portfolios j and k:*

$$FSD: \quad j\widehat{\mathrm{D}}_1 k \;\Leftrightarrow\; x_{jt} \geq x_{kt}\ \forall t \in \tau, \text{ and } x_{jt} > x_{kt} \text{ for some } t \in \tau.$$

$$SSD: \quad j\widehat{\mathrm{D}}_2 k \;\Leftrightarrow\; \sum_{i=1}^{t} x_{ji} \geq \sum_{i=1}^{t} x_{ki}\ \forall t \in \tau, \text{ and } \sum_{i=1}^{t} x_{ji} > \sum_{i=1}^{t} x_{ki} \text{ for some } t \in \tau.$$

Proof. See Appendix 1.

The notion of SD *efficiency* assumes some scarcity of investment opportunities. Focusing on the market set Ψ, SD efficiency is characterized by the following definitions:

Definition 2. Portfolio return profile $k: y_k \in \Psi$ is FSD (SSD) *efficient* in market set Ψ if and only if $j\widehat{\mathrm{D}}_1 k$ $(j\widehat{\mathrm{D}}_2 k) \Rightarrow y_j \notin \Psi$; otherwise, k is FSD (SSD) inefficient.

Unfortunately, no satisfactory method for testing the SD efficiency of a particular portfolio return profile is yet available. Therefore, a typical approach is to pick a sample of portfolio return profiles from Ψ and identify the nondominated ones by utilizing Theorem 1.[8] However, the market set Ψ contains

[6] Evidence from numerous laboratory experiments suggests that people often violate axioms of the EUT (see, e.g., Starmer 2000 for a recent survey). Consequently, alternative "non-expected-utility" theories have been (and are still being) developed. However, the EUT critique does not usually concern the SD criteria. Indeed, the experimental evidence indicates that very few people will choose a stochastically dominated option from a choice set when it is transparently obvious that the option is dominated. Quoting Starmer (2000, p. 335), "Monotonicity is the property that stochastically dominating prospects are preferred to prospects which they dominate, and it is widely held that any satisfactory theory—descriptive or normative—should embody monotonicity."

[7] A proof of this result can be found, e.g., in Levy (1998). For the sake of intuition, a simple constructive proof is also presented in Appendix 1.

[8] Bawa et al. (1979) discuss additional algorithmic insights to speed up computation in large samples.

an infinite number of portfolio return profiles; any method based on a finite number of pairwise comparisons will fail. This is a major shortcoming: We cannot verify whether a given portfolio is SD efficient or not, let alone identify all SD efficient portfolio return profiles. Consequently, we cannot even compare the relative size of the SD efficient subset of diversified portfolios to the corresponding MV efficient subset.

This paper argues that the problem with diversification is not an inherent feature of SD.[9] The complication with the conventional crossing algorithm approach arises from the impossibility of recovering the EDF of a diversified portfolio from the marginal EDFs of the underlying assets. This is illustrated in Appendix 2 by means of a simple numerical example. The real problem is the information loss incurred by ranking the return data (i.e., transforming from y to x). To be able to effectively deal with diversification, our strategy is to adopt the reverse route of characterizing the SD criteria as subsets of the state-space. To this end, we introduce the useful notion of *dominating set* associated with an arbitrary portfolio return profile y_0:

Definition 3. Set $\Delta_l(y_0) \equiv \{y \in \mathbb{R}^T \mid y\widehat{\mathrm{D}}_l y_0\}$, $l = 1, 2$, is called the *l order dominating set of* return profile y_0.

The dominating set relates to SD efficiency in the following sense:

Lemma 1. *Portfolio return profile y_0 is l order SD efficient, $l = 1, 2$, if and only if the l order dominating set of y_0 does not include any marketed portfolio return profile; that is, $\Psi \cap \Delta_l(y_0) = \varnothing$.*

Proof. Follows directly from Definitions 2 and 3. □

If we can identify the dominating set, we can use this result for testing SD efficiency by checking whether the intersection of the dominating set and the market set is empty. The next section derives the explicit characterizations of the FSD and SSD dominating sets for an arbitrary portfolio return profile y_0.

3. Dominating Sets

EDFs are based on ranked data, thus all *permutations* of y_0 have identical EDFs. Matrix $P = [P_{ij}]_{T\times T}$ is called a *permutation matrix* if its elements consist of binary integers $\{0, 1\}$ and its rows and columns sum up to unity. The set of permutation matrices can be written as

$$\Pi \equiv \left\{ [P_{ij}]_{T\times T} \;\middle|\; P_{ij} \in \{0,1\};\ \sum_{i=1}^{T} P_{ij} = \sum_{j=1}^{T} P_{ij} = 1\ \forall i, j \in \tau \right\}.$$

[9] See the exchange between Frankfurter and Phillips (1975) and Porter and Pfaffenberger (1975) on this issue.

Technically, permutation matrices allow us to sort the elements of a return vector in any arbitrary order. Thus, the set of all permutations of vector y_0 is expressed simply as $\{Py_0 \mid P \in \Pi\}$. Denote the set of vectors that are greater than or equal to some permutation of y_0 by $Q(y_0) \equiv \{y \in \mathbb{R}^T \mid \exists P \in \Pi: y \geq Py_0;\ y \neq Py_0\ \forall P \in \Pi\}$. We can now prove the following:

Theorem 2. *The FSD dominating set of portfolio return profile y_0 consists of return vectors that are greater than or equal to any permutation of y_0; that is, $\Delta_1(y_0) = Q(y_0)$.*

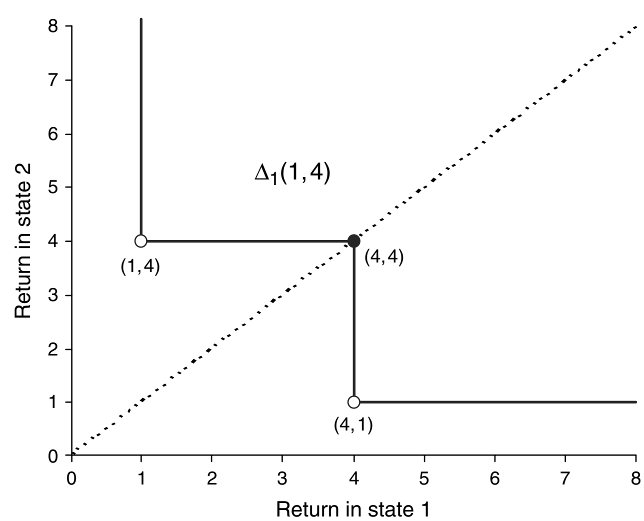
Proof. See Appendix 1.

Conveniently, $\Delta_1(y_0)$ is a positive monotonic set (i.e., $\Delta_1(y_0) = \Delta_1(y_0) + \mathbb{R}_+^T$). It is symmetric with respect to the diagonal ray $y = \alpha(1 \ldots 1)'$, $\alpha \in \mathbb{R}$, which represents the risk-free assets with a rate of return α. Unfortunately, this set is nonconvex for all risky portfolio return profiles (for which $y_{0i} \neq y_{0j}$ for some $i, j \in \tau$).

Figure 1 illustrates the FSD dominating set in the simplest conceivable case of two states of nature. Consider portfolio return profile $y_0 = (1, 4)$. Obviously, any return vector $(y_1, y_2) \geq (1, 4)$, $(y_1, y_2) \neq (1, 4)$ dominates portfolio return profile y_0 by FSD. The key insight of Theorem 2 is to consider alternative orderings of states as well. Consider the mirror image $(y_1, y_2) \geq (4, 1)$ in our example. In Figure 1, the broken line representing risk-free assets distinguishes between the two alternative orderings. In this two-dimensional case, the smallest risk-free return that dominates $(1, 4)$ by FSD is four.

Consider next the SSD dominating set. Matrix $W = [W_{ij}]_{T\times T}$ is called *doubly stochastic* if its elements are nonnegative real numbers and its rows and columns sum up to unity. Formally, the set of doubly

Figure 1 **The FSD Dominating Set of Portfolio Return Profile** (1, 4)



stochastic matrices is henceforth denoted by

$$\Xi \equiv \left\{ [W_{ij}]_{T\times T} \,\middle|\, 0 \le W_{ij} \le 1;\ \sum_{i=1}^{T} W_{ij} = \sum_{j=1}^{T} W_{ij} = 1\ \forall i, j \in \tau \right\}.$$

Observe that set Ξ only differs from Π in that the binary integer constraint is omitted. Thus, the set of permutation matrices is a subset of Ξ, but the converse is not true; that is, $\Pi \subset \Xi$. It is also worth noting that Ξ is spanned by extreme points of Π. Consequently, any $W \in \Xi$ can be expressed as a weighted average of two or more permutation matrices.

Proceeding towards the SSD criterion, consider what happens if in Theorem 2 we substitute the permutation matrix by a doubly stochastic matrix. For any $W \in \Xi - \Pi$, Wy_0 is generally "less risky" than the original portfolio return profile y_0. Indeed, Wy_0 might be viewed as a "mean-preserving antispread" of y_0 because Wy_0 and y_0 have equal mean returns, but the variance of the former tends to be smaller. For example, in the extreme case of $W_{ij} = 1/T\ \forall i, j$, we obtain a risk-free return equal to the mean return of the original portfolio return profile. Clearly, such a risk-free asset dominates any risky y_k by SSD.

We can take this line of reasoning further to derive the SSD dominating set. Denote the set of vectors that are greater than or equal to a weighed average of some permutations of y_0 by

$$S(y_0) \equiv \{y \in \mathbb{R}^T \mid \exists W \in \Xi\colon y \ge Wy_0;\ y \ne Py_0\ \forall P \in \Pi\}.$$

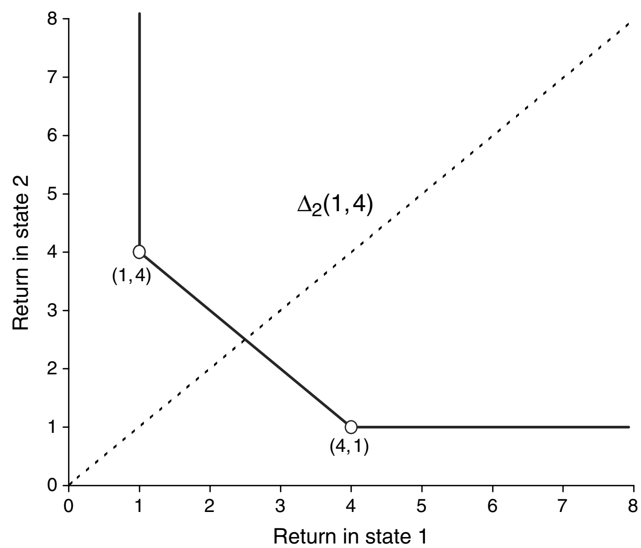
Theorem 3. *The SSD dominating set of portfolio return profile y_0 consists of return vectors that are greater than or equal to a weighted average of two or more permutations of y_0; that is, $\Delta_2(y_0) = S(y_0)$.*

Proof. See Appendix 1.

Like the FSD dominating set, set $\Delta_2(y_0)$ is positive monotonic, and symmetric with respect to the diagonal ray that represents risk-free assets. In addition, it is a convex set, which is convenient from an operational perspective. This becomes apparent in the next section.

Let us reconsider the graphical illustration of the dominating set of vector $y_0 = (1, 4)$ in Figure 2. The FSD dominating set is obviously a subset of the SSD dominating set. In addition, all return vectors that dominate convex combinations of permutations of y_0 (i.e., the mean-preserving antispread) are contained in the SSD dominating set according to Theorem 3. Thus, the triangle $[(1, 4), (4, 4), (4, 1)]$, excluding corner points $(1, 4)$ and $(4, 1)$, must be added to the FSD dominating set to obtain the SSD dominating set. The smallest risk-free rate of return that dominates portfolio return profile y_0 by SSD equals 2.5, which is also the mean return of portfolio return profile y_0. Recall that an option with the smaller mean cannot dominate by SSD, as shown by Hadar and Russel (1969).

Figure 2 **The SSD Dominating Set of Portfolio Return Profile** (1, 4)



4. Efficiency Tests and Measures

By Lemma 1, we can test SD efficiency of any given portfolio return profile y_0 simply by checking whether the dominating set and the market set share common portfolio return profiles (i.e., whether any dominating portfolio return profile is marketed). This section operationalizes this insight by presenting four efficiency measures that can be solved by standard LP techniques. The first measure provides a necessary and sufficient test statistic for FSD efficiency; the second one is a necessary test statistic for SSD efficiency; the third one complements the second statistic to form a necessary and sufficient test; and finally, the fourth one offers an alternative necessary SSD efficiency test. The tests are illustrated by three numerical examples in Appendix 2. The computational aspects are discussed in §5.

4.1. Exact FSD Test

To test FSD efficiency of portfolio return profile y_0, consider test statistic $\theta_1(y_0)$ obtained as the optimal solution to the following binary (0/1) MILP problem

$$\theta_1(y_0) = \max_{\lambda^0,\, P} \left(\sum_{t=1}^{T} \sum_{i=1}^{N+1} Y_{it}^0 \lambda_i^0 - \sum_{t=1}^{T} y_{0t} \right) \Big/ T$$
$$\text{s.t.} \quad \sum_{i=1}^{N+1} Y_{ti}^0 \lambda_i^0 \ge \sum_{j=1}^{T} P_{tj} y_{0j} \quad \forall t \in \tau \qquad (1)$$
$$P \in \Pi$$
$$\lambda^0 \in \Lambda,$$

where $Y^0 = (y_0\ \ Y)$ is the data matrix augmented by the portfolio return profile being tested for efficiency, and analogously, λ^0 is the vector of portfolio weights augmented by weight of y_0. Note that we do not require y_0 to be contained in the market set,[10] so we augment matrix Y and weights λ to guarantee feasibility. Because all elements of y_0 and Y are assumed to be finite, there always exists an optimal solution to (1).

The purpose of Problem (1) is to form a benchmark portfolio (characterized by weights λ^{0*}) that yields the highest mean return, subject to the constraint that the benchmark portfolio must dominate the evaluated portfolio (i.e., portfolio return profile y_0) by SSD. Dominance is guaranteed by the first constraint. Permutation matrix P enables us to take all possible orderings of states into account. The next theorem formally proves that θ_1 is both a necessary and sufficient FSD test statistic.

THEOREM 4. *Portfolio return profile y_0 is empirically FSD efficient in Ψ if and only if $\theta_1(y_0) = 0$.*

PROOF. See Appendix 1.

Test statistic θ_1 has an intuitive interpretation as the measure of inefficiency of the evaluated portfolio return profile. Accordant with the conventional dichotomy between return and risk, statistic $\theta_1(y_0)$ indicates the maximum increase in (mean) return obtainable without aggravating the risk exposure of the portfolio. The main difference between the SD-based measure θ_1 and the traditional MV measures lies in the treatment of risk. While the MV approach can employ variance as a measure for risk, no such quantitative measure exists for the SD. Rather, the SD approach must rely on its partial preference orderings, which offer a qualitative criterion for risk. A positive value of the SD statistic $\theta_1(y_0)$ implies that there exists portfolio $\lambda^{0*} \equiv \arg\max \theta_1(y_0)$ that dominates the evaluated portfolio return profile by FSD. While portfolio λ^{0*} may have a lower or higher variance than the evaluated portfolio, the FSD dominance implies that every nonsatiated investor, irrespective of his risk preferences, would be better off by holding λ^{0*} that yields $\theta_1(y_0)$ units higher mean return.[11] In this sense, $\theta_1(y_0)$ accounts for risk without explicitly quantifying it. Statistic $\theta_1(y_0)$ can be interpreted as the maximum loss of mean return due to FSD inefficiency.

Of course, other measures of inefficiency are also possible. For example, the minimum risk-free return that should be added to the evaluated portfolio return profile to make it efficient would be another intuitive (and perhaps more robust) measure of efficiency.[12] Formally, this measure can be defined as

$$\phi_1(y_0) \equiv \max_{\lambda^0,\, P}\left\{\phi \;\middle|\; \sum_{i=1}^{N+1} Y_{ti}^0\lambda_i^0 \geq \sum_{j=1}^{T} P_{tj}y_{0j} + \phi,\ \forall t \in \tau;\ P \in \Pi;\ \lambda^0 \in \Lambda;\ \phi \geq 0\right\}.$$

Other alternative efficiency measures are similarly obtained through straightforward modifications of Problem (1).

4.2. Necessary SSD Test

As a direct corollary of Theorems 2 and 3, we obtain a necessary test of SSD efficiency by simply relaxing the binary integer-constraint of the matrix P. Specifically,

$$\begin{aligned} \theta_2^n(y_0) = \max_{\lambda^0,\, W} & \left(\sum_{t=1}^{T}\sum_{i=1}^{N+1} Y_{it}^0\lambda_i^0 - \sum_{t=1}^{T} y_{0t}\right)\Big/ T \\ \text{s.t.}\quad & \sum_{i=1}^{N+1} Y_{ti}^0\lambda_i^0 \geq \sum_{j=1}^{T} W_{tj}y_{0j} \quad \forall t \in \tau \\ & W \in \Xi \\ & \lambda^0 \in \Lambda. \end{aligned} \tag{2}$$

THEOREM 5. *$\theta_2^n(y_0) = 0$ is a necessary condition for SSD efficiency of portfolio return profile y_0.*

PROOF. Directly analogous to the first part of the proof of Theorem 4 (Appendix 1), and thus omitted. □

The logic of this test is analogous to that of the FSD test. The test statistic $\theta_2^n(y_0)$ can be interpreted as the inefficiency measure in the mean sense, similar to the FSD case. This measure indicates the maximum increase of mean return that could be obtained by choosing an efficient portfolio from the subset that dominates the evaluated portfolio by SSD. Because the SSD dominating set is larger than the FSD dominating subset, the efficiency measures satisfy the inequality $\theta_2^n(y_0) \geq \theta_1(y_0)$. Conveniently, Problem (2) does not involve any integer variables, and can thus be solved using standard LP algorithms.

In contrast to the FSD case, however, $\theta_2^n(y_0) = 0$ is not yet a sufficient condition. In FSD, the dominating portfolio return profile yields a higher mean return by construction. This may not be the case in SSD, where a portfolio with equal mean but lower variance may dominate.

[10] This feature allows us to measure efficiency of a hypothetical, nonexisting portfolio, which might involve, for example, derivative instruments or new securities entering the market.

[11] Of course, λ^{0*} need not be the optimal, expected-utility-maximizing portfolio for all investors. Because the risk preferences differ between individuals, there generally does not exist a portfolio that would be optimal for every investor. In this respect, measuring inefficiency in terms of the mean return is as good (or as arbitrary) as any other efficiency metric.

[12] A major disadvantage of this measure, in comparison with θ_1, is that condition $\mu_1(y_0) = 0$ is not sufficient to guarantee FSD efficiency of profile y_0.

4.3. Sufficient SSD Test

It is easy to verify that any *mean-preserving antispread* Wy_0, $W \in \Xi - \Pi$ dominates the original portfolio return profile y_0 by SSD because it exhibits more equally distributed returns across states of nature. Test statistic $\theta_2^n(y_0)$ measures inefficiency in terms of foregone mean return, so it cannot identify dominance by Wy_0. Thus, to test whether any mean-preserving antispread is marketed, we calculate the following statistic:

$$\begin{aligned} \theta_2^s(y_0) = &\min_{W,\,\lambda^0,\,s^+,\,s^-} \sum_{j=1}^{T}\sum_{i=1}^{T}(s_{ij}^+ + s_{ij}^-) \\ \text{s.t.}\quad &\sum_{n=1}^{N+1} Y_{tn}^0 \lambda_i^0 = \sum_{j=1}^{T} W_{tj} y_{0j} \quad \forall t \in \tau \\ &s_{ij}^+ - s_{ij}^- = W_{ij} - \tfrac{1}{2} \quad \forall i,j \in \tau \\ &s_{ij}^+, s_{ij}^- \geq 0 \quad \forall i,j \in \tau \\ &W \in \Xi \\ &\lambda^0 \in \Lambda. \end{aligned} \tag{3}$$

This LP program tries to construct a doubly stochastic matrix W that redistributes the evaluated portfolio return profile across different states of nature to reduce the risk exposure of the resulting portfolio return profile Wy_0. The "surplus and slack variables" $s_{ij}^+, s_{ij}^- \geq 0$ (compare with Charnes et al. 1985) are computed as the positive and negative parts of $W_{ij} - \frac{1}{2}$, respectively. The sum of these variables is minimized when returns are distributed across states as equally as possible. The theoretical minimum value of the test statistic is $\frac{1}{2}T^2 - T$, which is obtained by setting $W_{ij} = 1/T$ for all i, j. The first constraint of (3) guarantees that this theoretical minimum is obtained only if the market set Ψ contains a risk-free option that yields the mean return of the evaluated portfolio return profile. Because we try to test whether any mean-preserving antispread is available in the market set, we need to compare $\theta_2^s(y_0)$ with its theoretical maximum, which will be derived next.

The theoretical maximum depends on the "ties" in the evaluated profile. A tie occurs in states of nature i and j when $y_{0i} = y_{0j}$. Of course, three-way ties (i.e., $y_{0i} = y_{0j} = y_{0k}$) or higher may also occur. In general, let d_{0k} denote the number of k-way ties in the evaluated profile. The theoretical maximum turns out to be $T^2/2 - \sum_{k=1}^{T} kd_{0k}$; that is,

$$\theta_2^s(y_0) \in \left[\frac{1}{2}T^2 - T, \frac{1}{2}T^2 - \sum_{k=1}^{T} kd_{0k}\right].$$

These properties can be used for deriving the following exact SSD test.

Theorem 6. *A necessary and sufficient condition for SSD efficiency of portfolio return profile* y_0 *is*

$$\theta_2^n(y_0) = 0 \wedge \theta_2^s(y_0) = \frac{T^2}{2} - \sum_{k=1}^{T} kd_{0k}.$$

Proof. See Appendix 1.

Together, the necessary statistic θ_2^n and the sufficient statistic θ_2^s allow us to draw the exact efficiency diagnosis. The latter statistic is only needed for verifying efficiency when $\theta_2^n(y_0) = 0$. If the necessary condition immediately shows a violation of efficiency, then the sufficiency test is unnecessary.

4.4. Special Case

Finally, consider a special case where there are no constraints on portfolio weights except that short sales are not allowed; that is, domain Λ simplifies to $\Lambda = \{\lambda \in \mathbb{R}_+^N \mid \sum_{i=1}^N \lambda_i = 1\}$. In this case, there is an alternative and more direct way of testing the necessary SSD efficiency condition. Observe that both $\Delta_2(y_0)$ and Ψ are now convex sets. This allows us to apply the theory of separating hyperplanes (e.g., Rockafellar 1970). Specifically, as SSD efficiency is equivalent to condition $\Delta_2(y_0) \cap \Psi = \varnothing$, there must exist a separating hyperplane that strongly separates set Ψ from $\Delta_2(y)$. The following necessary test statistic essentially checks if a weak separation is possible:

$$\begin{aligned} \hat{\theta}_2^n(y_0) = &\max_{w} \sum_{t=1}^{T} y_{t0} w_t \\ \text{s.t.}\quad &\sum_{t=1}^{T} Y_{tn} w_t \leq 1 \quad \forall n \in \nu \\ &w_t \leq w_s \quad \forall t, s \in \tau : y_{t0} > y_{s0} \\ &w_t \geq 0 \quad \forall t \in \tau. \end{aligned} \tag{4}$$

Theorem 7. *In the special case where there are no portfolio constraints, that is,* $\Lambda = \{\lambda \in \mathbb{R}_+^N \mid \sum_{i=1}^N \lambda_i = 1\}$, *condition* $\hat{\theta}_2^n(y_0) \geq 1$ *is a necessary but not sufficient condition for SSD efficiency of portfolio return profile* y_0.

Proof. See Appendix 1.

This test is similar in spirit to the nonparametric consistency tests of Afriat (1972) and Varian (1983), applied to portfolio efficient sets and to Theorem 1 of Dybvig and Ross (1982), in particular. We here reinterpret the result of Dybvig and Ross from the perspective of the separating hyperplane theorem. The specific test statistic $\hat{\theta}_2^n$ is almost identical to that proposed by Post (2003a). In contrast to Post, however, we interpret (4) as a necessary condition.[13] One can show with simple numerical examples that condition $\hat{\theta}_2^n \geq 1$ does not suffice to guarantee SSD efficiency (see Appendix 2, Example 3).

[13] Difference of interpretation arises from the fact that Post (2003a) adopts a more stringent, nonstandard definition of SSD (in his Definition 1).

5. Computational Aspects

This section briefly compares the computational burden of the four tests discussed in the previous section. All test statistics were formulated as LP or MILP problems, which can be solved by standard algorithms developed for these types of problems. Sophisticated simplex, interior point, primal, dual, and barrier algorithms are widely available and can demonstrably solve complex real-world applications (see, e.g., Fourer 2001 or Bixby 2002 for review). In principle, the hardware capacity sets the limit for the size of the problem. Still, the test procedures of the previous section can become time consuming in large applications.

Table 1 reports the dimensions of the LP/MILP problems associated with the four test statistics as functions of N and T (i.e., cardinality of sets ν and τ, respectively).[14] We observe that the number of rows (i.e., constraints) is a linear function of T for all four test statistics. The number of columns (variables) is a bivariate function of N and T in all cases. For the first three statistics, the number of columns is a quadratic function, while the fourth displays a linear structure. However, the data matrix of the fourth problem is more *dense* (i.e., there are fewer zero elements) than matrices of Problems (1)–(3). This greater density partly offsets the advantageous linearity of the fourth problem. Consequently, the number of nonzero variables is a nonlinear (quadratic) bivariate polynomial in all four cases.

To provide perspective about the computational burden, Table 2 compares the dimensions of Problems (2) and (4) (the necessary SSD tests) for a selection of T, N combinations. The most critical statistic to consider is the number of nonzeros (NZ). For example, the "Performance World" homepage (2003) classifies LP problems according to the NZ as "tiny" ($0 \leq \text{NZ} \leq 2{,}000$), "small" ($2{,}000 \leq \text{NZ} \leq 50{,}000$), "medium" ($50{,}000 \leq \text{NZ} \leq 200{,}000$), "large" ($200{,}000 \leq \text{NZ} \leq 1{,}000{,}000$), and "huge" ($\text{NZ} \geq 1{,}000{,}000$). Besides the dimensions of the problem, the computation time depends on the hardware resources and the solver software used. Some impressions about computation times with different LP solvers can be obtained by comparing the NZ figures of Table 2 with benchmark cases of comparable size reported by Bixby (2002) and Mittelmann (2003), among others.

The first four cases illustrate how the size of the problem develops when T increases while N is held

Table 1 Computational Complexity of Alternative Test Statistics

Test statistic	Rows	Columns	Nonzeros	Integers
θ_1	$3T+2$	T^2+N+1	$3T^2+N(T+2)$	T^2
θ_2^n	$3T+2$	T^2+N+1	$3T^2+N(T+2)$	0
θ_2^s	$3T+2$	$3T^2+N$	$3T^2+5T+N(T+1)$	0
$\hat{\theta}_2^n$	T	$T+N$	$T(N+3)-2$	0

at the constant value of 10 (a standard choice in the finance literature). In the case of 10 assets, Problem (2) is classified as "large" when T exceeds 300 and "huge" when T exceeds 600. By contrast, Problem (4) remains "small" up to $T = 4{,}000$. In the case of $T = 2{,}000$, Problem (2) includes over 12 million NZs, which is 463 times the number of NZs in (4). The last three cases illustrate the size of the problem when T and N are unequal. Case $T = 135$, $N = 5$ corresponds to the application in the online appendix. In this case, Problem (2) is of "medium" size while Problem (4) is "tiny". Case $T = 460$, $N = 26$ corresponds to the application reported by Post (2003a). In this case, Problem (4) is a "small" one, while Problem (2) is classified as "large". Case $T = 100$, $N = 100{,}000$ serves to illustrate that Problem (4) has the computational advantage in its ability to deal with large number of states. When the large number of assets causes the computational burden, there is no notable difference in the size of the alternative problems.

The previous comparisons show that Problem (4) has the greatest computational advantage over Problem (2) in applications where T is relatively large and N is relatively small. According to Post (2003a, p. 1907), powerful statistical inference will require such a large T that solving Problem (4) becomes computationally prohibitive. It is therefore constructive to consider dimensions N and T, as well as their influence on computational burden from the statistical perspective.

Recall that the FSD and SSD tests presented above are *exact* tests for SD efficiency regarding the observed empirical distribution (EDF). While the EDF is generally a good estimator of the underlying probability distribution, it is not perfect: The SD efficiency classification based on the EDF does not always coincide with the "true" efficiency classification in terms of the underlying, unobserved CDF. The empirical efficiency diagnosis can be erroneous in two ways. A Type I error occurs when a portfolio, efficient in terms of the true CDF, is wrongly diagnosed as inefficient based on the EDF. A Type II error occurs when a truly inefficient portfolio is wrongly diagnosed as efficient. The relative frequency of Type I errors is referred to as the *statistical size* of the test. The *statistical power* of the test is one minus the relative frequency of Type II errors.

In general, there is a trade-off between the statistical size and power, which may be influenced by

[14] Computational complexity of an LP or MILP problem is usually measured by writing the problem in the *standard form* [$\max \mathbf{c}'\mathbf{x}$ s.t. $\mathbf{A}\mathbf{x}+\mathbf{s}=\mathbf{b}$, $\mathbf{x}\geq 0$, $\mathbf{s}\geq 0$] and calculating the number of rows, columns, and nonzero elements in matrix $\mathbf{A}$ augmented by the cost row $\mathbf{c}'$, ignoring the slacks $\mathbf{s}$ and the right-hand side $\mathbf{b}$.

Table 2 Necessary SSD Statistics at Different Combinations of T and N

		θ_2^n			$\hat{\theta}_2^n$		
T	N	Rows	Columns	Nonzeros	Rows	Columns	Nonzeros
100	10	302	10,011	31,020	100	110	1,298
500	10	1,502	250,011	755,020	500	510	6,498
1,000	10	3,002	1,000,011	3,010,020	1,000	1,010	12,998
2,000	10	6,002	4,000,011	12,020,020	2,000	2,010	25,998
135	5	407	18,231	55,360	135	139	1,078
460	26	1,382	211,627	646,812	460	485	13,338
100	100,000	302	110,001	10,230,000	100	100,099	10,000,298

the choice of N and T. For any given N, the empirical SD efficiency criteria are more easily met when T increases. This decreases both the size and the power of the test. Conversely, for any given T, the SD efficiency criteria become more stringent as N increases, which increases both the size and power. In conclusion, increasing N tends to improve the power but deteriorate the size, while increasing T tends to improve the size but deteriorate the power.

Traditionally, the low power has been viewed as a major problem of the SD approach (e.g., Nelson and Pope 1991). However, taking diversification into account appears to reverse this situation. In Post's (2003a) simulations, the power of the basic SSD test (referred to by Post as "naïve" for ignoring sampling errors) was found to be perfect; all inefficient portfolios were correctly diagnosed as inefficient. Unfortunately, the statistical size of the basic SSD test reveals itself as a problem: Efficient portfolios tend to be wrongly diagnosed as inefficient. While these observations aptly illustrate the effect of accounting for portfolio diversification to the statistical power of the SSD efficiency criterion, and the influence of increasing T, further evidence from the Monte Carlo simulations is needed to better understand the trade-off between dimensions N and T.

If it is possible to choose dimensions such that T becomes large and N remains small, then the statistical size of the test could be decreased. Dimension N might be kept small by aggregating assets to benchmark portfolios based on market capitalization, book-to-market ratio, earnings per price, or momentum, which can help to decrease both the computational burden and the statistical size of the test. As illustrated in Table 2, the computational advantage of Problem (4) increases when T increases and N decreases. However, Post's (2003a) simulation results suggest that the choice of T and N may not suffice to eliminate the problem of excessive statistical size: Even with $N = 25$ and $T = 2{,}000$, almost 70% of efficient portfolios were diagnosed as inefficient. Therefore, the statistical size must be reduced by other means. The relative balance between statistical power and statistical size could be adjusted in many ways. A straightforward approach is to introduce a predetermined critical tolerance level $\alpha > 0$ for the test statistic (which could be determined using a bootstrap simulation). The null hypothesis of full efficiency is maintained if the inefficiency measure does not exceed the tolerance level α. The portfolio is diagnosed as inefficient only when the SD measure exceeds this tolerance level.

Achieving a reasonable balance between power and size does not necessarily require a very large data set. The analytical test procedure proposed by Post (2003a), which is based on a nonstandard null hypothesis and an asymptotic least-favorable distribution that maximizes statistical size (and thus comes at the cost of statistical power), does require a large T (thousands of cross-sections). However, this asymptotic test may not be the most efficient solution for statistical inference. As later demonstrated by Post (2003b), the bootstrap approach can yield statistically meaningful results with a much smaller T. For $N = 26$, 100–500 cross-sections suffice to give reasonable statistical size and power. Applications of this dimensionality present no computational obstacle for Problem (2).

Because effective means to adjust the balance between statistical power and size are available, the dimensions of the problem appear to be less important for the statistical properties of the test. Thus, the dimensionality need not give such an overwhelming advantage for Problem (4) as the first rows of Table 2 suggest. Constructing the market set from asset data instead of aggregate portfolios increases dimension N and, subsequently, the computational burden of both Problems (2) and (4). Still, using disaggregate asset data offers a more detailed picture about the composition of the efficient reference portfolios, which may be of practical importance.

Besides the sampling error, it is important to note that the choice of N and T also influences the likelihood of specification errors. In reality, the probability distributions of asset returns evolve over time (e.g., due to the business cycle), which can violate our assumption that all states were drawn randomly from the same pool of possible states. As the time span T increases, the likelihood of specification errors due to the violation of this assumption increases, unless one is using high-density (i.e., tick-by-tick) data. Similarly, increasing N may involve a risk that the market set includes assets that, in reality, are not fit for the investment portfolio under evaluation (e.g., due to liquidity qualifications).

Finally, the computational cost is not the only criterion to consider when choosing the test procedure. It is worth mentioning that the complexity of Problem (2) is attributable to its generality. The benefits of this more general framework include:

(1) It is possible to model nonconvex market sets (consider, e.g., transaction costs) and nonconvex dominating sets (i.e., the FSD case). (The separating hyperplane theorem used for deriving $\hat{\theta}_2^n$ requires that both these sets are convex.[15])

(2) For every inefficient portfolio, a dominating benchmark portfolio is always explicitly identified. The identity of this benchmark enables one to assess whether all relevant constraints are modeled appropriately, or if further constraints should be added. This is especially useful in the "model-building" phase of the analysis. For computationally intensive simulations or sensitivity analyses, the identity of the benchmark portfolio is of less importance, and one can then shift to the use of the simpler statistic $\hat{\theta}_2^n$ (if applicable).

(3) It is possible to impose additional constraints on portfolio weights directly in the LP problem. Recall that test statistic $\hat{\theta}_2^n$ only applies in the special case where portfolio weights are unconstrained, nonnegative numbers that sum up to unity. Additional constraints on portfolio weights λ cannot be imposed in Problem (4) directly because Problem (4) models the market set only *implicitly* through its extreme points (the state-space tableau Y).[16] By contrast, Problems (1)–(3) model portfolio weights and their constraints *explicitly* through specification of the weight domain Λ. The ability to impose bounds on portfolio weights is very important in practical applications, as most portfolio managers face a number of explicitly specified constraints. For example, rules of mutual funds typically dictate that the investment portfolio consists of securities, bonds, options, and other assets in certain proportions; that the portfolio should be diversified across different industries and different geographic areas in a certain way; that the portfolio should include securities of small, mid-size, and large firms in certain proportions; and that the portfolio weights of individual assets should not exceed certain upper bounds.

6. Application: Industrial Diversification of the Market Portfolio

One of the key functions of stock markets is to allocate scarce resources of the economy to different industries according to their profitability prospects and risks. To illustrate the potential of the proposed SD approach, we examine whether the U.S. stock markets have been efficiently diversified across industries. Specifically, we use the SD tests of §4 to gauge efficiency of the market portfolio relative to the market set spanned by 48 industry portfolios. That is, the portfolio weights of stocks *within* each industry are fixed, but weights can be adjusted *across* industries. Efficiency of the industrial diversification of the market portfolio is interesting for both individual investors and society as a whole. Inefficient diversification across industries implies that the resources of the economy are suboptimally allocated.[17] Moreover, identification of systematic inefficiencies in industrial diversification of the market portfolio would allow individual investors and fund managers to exploit this information in order to outperform the market.

The market portfolio was constructed as the value-weighted average of all NYSE, AMEX, and NASDAQ stocks. The industry portfolios were based on the industry classification according to the four-digit SIC code. Monthly data of returns exceeding the risk-free rate (one-month treasury bill) were obtained from the data library of Kenneth French (French 2003). After excluding periods with missing values, the data set covers periods 7/1963 to 12/2002. Table 3 reports the summary statistics of the portfolios over the entire study period.

The specification of the planning horizon involves a trade-off (as usual). On the one hand, it is meaningless to fix the industry weights of the benchmark portfolio for a long period of time because the underlying probability distribution is likely to change over time (i.e., some states become more likely to occur, while other states become less likely). On the other hand, a sufficiently large sample of return observations is

[15] Consequently, the FSD extension suggested by Post (2003a), which is based on the separating hyperplane theorem, gives a necessary but not sufficient efficiency condition.

[16] Post (2003a, Endnote 5) suggests in his approach a possibility of modeling constraints by first identifying the extreme points of the constrained market set by enumeration and then including the extreme points in the state-space tableau Y. This requires the constrained market set to be a convex polyhedron, which is not a general property of constrained market sets. Even if the constrained market set is convex, identifying the vertices of this set is not trivially easy when a large number of side constraints are present.

[17] Allocative efficiency of the stock market should not be confused with the *informational efficiency*. In an incomplete market where information is costly, informational efficiency does not necessarily guarantee allocative efficiency. See Stiglitz (1981) for an insightful discussion of the resource allocation function of the stock market.

Table 3 Summary Statistics of the 48 Industry Portfolios and the Market Portfolio

	Mean	St.dev	Skewness	Kurtosis	Min	Max		Mean	St.dev	Skewness	Kurtosis	Min	Max
Agric	0.46	6.78	0.30	4.21	−25.78	40.79	Guns	0.76	7.00	−0.03	1.54	−30.36	29.79
Food	0.62	4.64	0.15	1.83	−18.49	18.60	Gold	0.61	9.84	0.36	1.54	−31.68	45.31
Soda	0.82	5.69	−0.21	2.97	−26.49	25.60	Mines	0.33	6.65	−0.30	1.68	−34.13	20.40
Beer	0.63	5.27	0.10	2.05	−18.74	25.18	Coal	0.49	7.87	0.49	3.20	−30.74	45.92
Smoke	0.94	6.16	−0.11	2.09	−21.66	28.19	Oil	0.52	5.24	0.09	1.66	−19.31	23.41
Toys	0.58	8.46	−0.03	1.14	−33.88	34.47	Util	0.28	4.12	0.15	1.11	−12.95	18.36
Fun	0.76	7.60	−0.20	2.37	−32.50	35.60	Telecm	0.36	4.97	−0.12	1.99	−19.19	21.63
Books	0.66	5.75	−0.22	1.37	−26.70	22.55	PerSv	0.25	7.25	−0.14	2.40	−32.94	30.41
Hshld	0.52	4.94	−0.25	1.46	−22.25	16.80	BusSv	0.63	7.14	−0.15	1.10	−28.15	25.86
Clothes	0.40	6.62	−0.07	2.80	−32.14	31.84	Comps	0.42	6.76	−0.09	1.05	−27.79	25.07
Health	0.73	10.75	0.43	3.27	−44.09	48.29	Chips	0.56	7.82	−0.25	1.78	−30.57	28.94
MedEq	0.79	5.51	−0.16	0.70	−19.92	20.56	LabEq	0.68	8.11	0.00	0.49	−30.69	28.86
Drugs	0.68	5.30	0.21	2.67	−19.71	31.33	Paper	0.48	5.77	0.03	1.78	−26.89	24.02
Chems	0.39	5.31	−0.08	2.48	−28.60	20.76	Boxes	0.44	5.18	−0.06	3.00	−28.86	21.59
Rubbr	0.56	6.36	−0.25	2.69	−30.86	27.38	Transp	0.41	6.12	−0.26	1.23	−28.40	18.66
Txtls	0.41	6.21	−0.53	2.73	−33.22	22.38	Whlsl	0.62	5.95	−0.37	2.36	−31.59	17.32
BldMt	0.48	5.69	−0.19	2.60	−28.43	26.27	Retail	0.60	5.72	−0.16	2.13	−29.59	26.53
Cnstr	0.41	7.30	−0.01	1.04	−31.83	23.70	Meals	0.65	6.68	−0.49	2.25	−32.10	28.10
Steel	0.20	6.38	−0.11	2.18	−31.76	25.91	Banks	0.55	5.88	−0.01	1.68	−25.34	24.51
FabPr	0.23	6.91	−0.39	1.29	−27.45	24.40	Insur	0.53	5.57	0.18	1.46	−17.52	26.06
Mach	0.39	6.00	−0.36	2.18	−32.13	22.16	RlEst	0.48	7.03	0.21	4.70	−26.65	46.73
ElecEq	0.54	6.70	0.37	3.56	−32.69	37.57	Trade	0.59	5.22	−0.37	1.10	−20.84	16.99
Autos	0.34	6.10	−0.17	1.69	−28.70	22.15	Other	0.16	7.35	−0.47	1.30	−29.06	20.45
Aero	0.68	7.00	−0.22	1.63	−30.79	24.58							
Ships	0.59	6.66	−0.11	1.47	−32.81	19.24	Market	0.41	4.51	−0.49	1.96	−23.09	16.05

needed for approximating the underlying probability distributions with sufficient accuracy. To strike a balance, we resort to the *rolling-window* approach, using a window width of five years. Specifically, we calculate our test statistics separately for 36 overlapping five-year periods (1963–1967), (1964–1968), …, (1997–2001), (1998–2002). The annual efficiency measures are calculated as the average of the five-year periods that include the year in question.

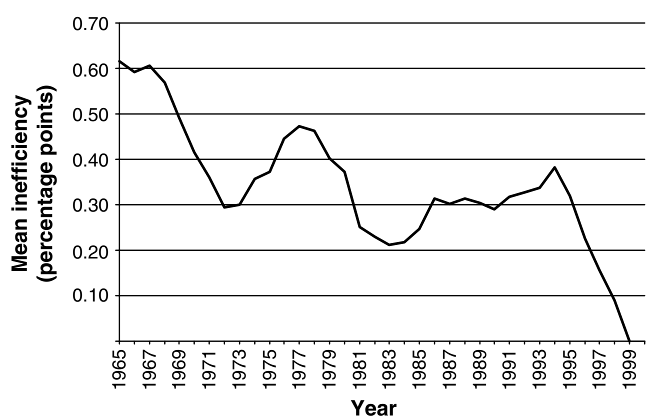
Notably, the FSD and SSD efficiency measures θ_1 and θ_2^n yielded identical values for all time periods. This means that risk preferences do not play a role in our efficiency analysis.[18] The sufficient SSD test θ_2^s and the alternative necessary SSD test $\hat{\theta}_2^n$ confirms the efficiency classification, but does not offer additional information, so we focus on interpreting the FSD measure θ_1.

[18] We conjecture that this somewhat surprising outcome arises from the fact that the evaluated market portfolio is a convex combination of all industry portfolios. Intuitively, the average market portfolio tends to be compared with reference portfolios that lie on the frontier segment of the market set, where increasing the mean return also increases the risk, which belongs to both the FSD and SSD efficient sets. Observe that the difference between the FSD and the SSD criteria only occurs in the frontier segments where increasing risk is associated with decreasing mean return. Alternatively interpreted, the assumption of risk aversion (imposed by SSD) can make a substantial difference in evaluation of high-risk, weakly diversified portfolios, while its effect tends to be minimal (or nonexistent) in evaluation of well-diversified portfolios (such as the market portfolio).

The development of FSD efficiency of the market portfolio over time is illustrated by Figure 3, where the line represents the values of the five-period moving averages of the inefficiency measures calculated for five-year windows. Three observations are worth noting.

First, we observe that the FSD measure θ_1 shows a declining trend, meaning the efficiency of the market portfolio has improved over time. In the mid-1960s the calculated inefficiency measures were as high as 0.6 percentage points. This means that we can form a benchmark portfolio that dominated the market portfolio by FSD and yielded 0.6 percentage points higher mean return. By contrast, in the last four five-year periods (from the late 1990s onward) the market port-

Figure 3 FSD Inefficiency of the Market Portfolio 1965–2000



folio proved efficient; that is, no FSD or SSD dominating portfolio existed.

Second, the FSD measure θ_1 reached relatively low levels in the early 1970s, early 1980s, and late 1990s. A rapid increase occurred in the late 1970s, and more gradual growth took place from the mid-1980s through the mid-1990s. Notably, the high-efficiency periods correspond to years of relatively stable GDP growth, while the low-efficiency periods are evidenced during more turbulent times. The inefficiencies of the late 1970s can be associated with the energy crises of 1973 and 1979, which had major repercussions for many industries, increasing uncertainties in future prospects. Similarly, the "Black Monday" stock crash in 1987 corresponded with an increasing part of the inefficiency curve. The inefficiencies during the early 1990s are more difficult to explain by particular historical events. As with catastrophic events, times of rapid growth are also difficult to predict. For example, restructuring towards the ICT lead information society in the late 1980s and early 1990s could partly explain the ex post inefficiencies during that period.

Third, the dominating benchmark portfolios (i.e., λ^* obtained as the optimal solution to Problem (1)) proved relatively well diversified across industries; imposing additional constraints for portfolio weights was unnecessary. As the data do not reveal the original industry weights, we compared our benchmark industry weights with the GDP shares of industries during the 1987–1996 period.[19] Overall, the differences were relatively small, suggesting the weights of our benchmark portfolios lie at reasonable levels. For all years there were some industries (e.g., electrical equipment) that were assigned a zero weight in the benchmark portfolio, which seems somewhat unrealistic. Still, the GDP shares of those industries were less than 2%, thus, the zero weight does not present a major deviation from reality; enforcing some strictly positive weight would not radically change the result. Regarding specific industries, the benchmark portfolios consistently gave a five to six percentage points higher weight on the trading, oil, and food products sectors than the corresponding GDP shares of these industries. On the other hand, the real estate sector that contributed approximately 25%–27% share of the GDP was allocated only marginal weights consisting of less than 1% in the dominating benchmark portfolios.[20] Other sectors underrepresented in our benchmark portfolios included transportation, wholesale, and construction, with four to seven percentage points lower portfolio weights than their GDP share.

To assess the robustness of the previous analysis to sampling error, we selected three periods for a more detailed sensitivity analysis: (1) 1964–1968, (2) 1981–1985, and (3) 1998–2002. Following the bootstrap approaches of Nelson and Pope (1991) and Post (2003a, b), we generated 500 pseudosamples of 60 periods for each three five-year periods by drawing randomly, with replacement from the empirical cross-sections observed in the corresponding period.[21] Calculating the efficiency measures for each randomly drawn pseudosample, we can get an impression about the sampling variation associated with our efficiency measures. For simplicity, here we restrict attention to the SSD efficiency criterion (recall that the original FSD and SSD efficiency measures were identical for all periods).

The original inefficiency measures were the following: (1) 0.632, (2) 0, and (3) 0 (note that these differ from the five-year moving averages presented in Figure 3). The average inefficiencies in the pseudosamples were found to be considerably higher: (1) 1.010, (2) 0.612, and (3) 0.856 (with standard deviations (1) 0.224, (2) 0.229, and (3) 0.380). In all three periods, the original inefficiency measures were found in the bottom five percentiles of the pseudoinefficiency distributions. Considering the qualitative efficiency classification, all pseudosamples were diagnosed as inefficient in Period 1, while 2% of the pseudosamples were found efficient in Period 2 and 7% in Period 3, respectively.

At least three conclusions can be drawn from this limited sensitivity analysis. First, both the distribution of pseudoinefficiencies and the qualitative efficiency classifications support the view that the market portfolio has exhibited substantial inefficiency in Period 1, and that efficiency has improved considerably in Periods 2 and 3, in line with Figure 3. Second, the sensitivity analysis shows that it is highly unlikely to draw random pseudosamples from the observed EDF in which the market portfolio is diagnosed as efficient. The frequency of efficient pseudosamples increases to some extent in Periods 2 and 3, in which the original efficiency measure diagnosed the market portfolio as fully efficient. Third, the exact levels of inefficiency are highly uncertain given the large variances in the pseudoinefficiencies. In this respect, the results should be interpreted with sufficient caution.

The large variations in the pseudoefficiencies are due to the small T in combination with a relatively

[19] Recall that the market portfolio was efficient from the 1995–1999 period onwards.

[20] This disparity might be because the real estate markets tend to be highly localized, with large numbers of small firms not publicly traded in the stock market. This same remark applies, to some extent, to the transportation and construction sectors.

[21] In all three cases the mean and standard deviation of the efficiency distribution stabilize at their reported levels after the first 50 pseudosamples. Thus, 500 pseudosamples should suffice for the purposes of a sensitivity analysis.

large N. Increasing T (i.e., the window length) would probably give more robust results but would also increase the risk of violating our basic assumption that all observed states were drawn randomly from the same pool of possible states. Obviously, the return distributions change over time, so it is likely that some (if not all) of the inefficiencies measured reflect changes in the market environment. Indeed, we observed that inefficiencies peaked in time periods known to be turbulent in the stock market. To alleviate this problem, we have kept the window length to the minimum and analyzed the annual market efficiency by means of the average efficiency of those periods in which the particular year is included. Still, these measures do not fully eliminate the specification errors resulting from possible changes in the underlying probability distribution (CDF). A more systematic, full-fledged empirical analysis could attempt to alleviate the "nonstationarity" problem by utilizing high-frequency financial data (i.e., daily or tick-by-tick data instead of monthly returns), as well as more sophisticated econometric methods for identifying trends and structural breaks in the time series.

A second application of the SD tools to forest portfolio management is available as an online supplement.

7. Concluding Remarks

We developed a series of operational tests for portfolio efficiency that are based on the general stochastic dominance criteria and account for infinite numbers of diversification strategies. The key idea was to preserve the cross-sectional dependence of asset returns when forming portfolios. Instead of arranging data in the form of empirical distribution functions, we reexpressed the SD criteria in T-dimensional Euclidean space spanned by return vectors representing rates of return in T different states of nature. We derived explicit analytical characterizations for the FSD and SSD dominating sets as subsets of this T-dimensional state-space. Using these results, we further derived operational SD efficiency measures and test statistics that can be computed using standard mathematical programming algorithms and readily available software packages. The proposed SD tests and efficiency measures can be directly applied to a wide variety of practical portfolio selection problems, as exemplified by the illustrative application presented in §6.

The lack of tools for dealing with diversification has traditionally been the most serious weakness of the SD criteria in empirical portfolio analysis. We believe the new perspectives presented in this paper could contribute to a greater diffusion of the theoretically appealing SD criteria to empirical applications. In this respect, the SD framework may prove a fruitful application area for mathematical programming techniques of operations research and management sciences.

In conjunction with the problem of diversification, another serious weakness of the SD approach has been its low discriminatory power (i.e., efficient sets are large and thus uninformative). However, the existing evidence about the discriminatory power (e.g., Kroll and Levy 1980, Nelson and Pope 1991) refers to comparisons of the SD and MV efficient subsets in cases where MV accounts for diversification, while SD does not. Compared to the earlier, "approximative" necessary tests, the exact tests developed above improve the power of SD criteria decisively. Indeed, the tentative evidence from Post's (2003a) simulations seems to suggest that the statistical power of the SD tests is no longer a problem when we account for portfolio diversification. Rather, we should be more concerned with the statistical size, that is, classifying a truly efficient portfolio wrongly as inefficient based on an empirical distribution.

To find an optimal balance between the statistical power and size, we need to expand the analysis towards statistical inference, a topic that we have disregarded in this paper. In this respect, it is useful to note that the LP structure of our efficiency measures enables one to combine the SD efficiency measures with the bootstrapping approach (see Nelson and Pope 1991 and especially Post 2003a for further discussion). Other promising tools for statistical inference that are consistent with the "nonparametric" nature of SD criteria include the nonparametric Kolmogorov-Smirnov (McFadden 1989) or Wilcoxon-Mann-Whitney (Schmid and Trede 1996) tests. Although Post (2003a) presents a very impressive start towards statistical inference in the SD framework with full diversification possibilities, it warrants further development. For example, it would be interesting to see more systematic simulation evidence about the performance of the SD efficiency measures and tests in alternative conditions involving different combinations of assets and cross-sections, nonnormal true distributions, and specification errors. Such simulations could also shed light on the performance of the FSD criterion in comparison to the SSD and higher-order SD criteria. Moreover, developing operational techniques for statistical inference that do not require excessively demanding computations, or massive amounts of data, remains a challenge for further research.

An online supplement to this paper is available at http://mansci.pubs.informs.org/ecompanion.html.

Acknowledgments

The author is indebted to the editors, Professor Ravi Jagannathan and an anonymous associate editor, and to two anonymous referees for their constructive comments and useful suggestions. In addition, the author is grateful to Mrs. Jasmine van Wijgerden for proofreading the text and providing helpful suggestions regarding the style. All errors remain, of course, the author is responsibility. This study forms part of the *NOMEPRE*-research program; see http://www.sls.wau.nl/enr/staff/kuosmanen/program1/ for further information. Financial support from the Emil Aaltonen Foundation, Finland, is gratefully acknowledged.

Appendix 1. Proofs of Mathematical Theorems

Proof of Theorem 1. *FSD case*: Note that the EDFs are monotonically increasing step functions with the step height $1/T$, and width characterized by x_{jt} and x_{kt}, $t \in \tau$. Therefore, the dominance relation of Definition 1 can be determined by comparing the values of EDFs in the T vertices (i.e., steps x_{jt} and x_{kt}, $t \in \tau$) of the EDF.

SSD case: The integral of the step function H is a monotonically increasing piecewise linear function, with vertices located in $\sum_{i=1}^{t} x_{ji}$ and $\sum_{i=1}^{t} x_{ki}$, $t \in \tau$. Thus, the dominance relation of Definition 1 can be determined, analogous to the FSD case, by comparing the values of the EDFs in these T vertices.[22] □

Proof of Theorem 2. Consider an arbitrary portfolio return profile $y_1 \in \mathbb{R}^T$. The proof is presented in two parts:

(1) $y_1 \in Q(y_0) \Rightarrow y_1 \widehat{\mathrm{D}}_1 y_0$: Obviously, if $y_1 \geq P y_0$, for some $P \in \Pi$, we can safely rearrange elements of both vectors in nondecreasing order, i.e., use vectors x_1 and x_0. Clearly, the inequality $x_1 \geq x_0$ must still hold. By Theorem 1 this immediately implies

(i) $H_1(r) \leq H_0(r)\ \forall r \in \mathbb{R}$.

Because $y_1 \neq P y_0$ for all permutations, (i) must hold as a strict inequality for some $r \in \mathbb{R}$. By Definition 1, $y_1 \widehat{\mathrm{D}}_1 y_0$.

(2) $y_1 \widehat{\mathrm{D}}_1 y_0 \Rightarrow y \in Q(y_0)$: By Theorem 1, $x_1 \geq x_0$. Because there exist permutation matrices $P_h \in \Pi$: $x_h = P_h y_h$, $h = 1, 2$, there must also exist permutation $P^* \in \Pi$ such that $y_1 \geq P^* y_0$. Moreover, because $y_1 \widehat{\mathrm{D}}_1 y_0$ directly implies $y_1 \neq P y_0$, we see that $y_1 \in Q(y_0)$. □

Proof of Theorem 3. This follows from the theorem in Hardy et al. (1934),[23] which proved the following two conditions to be equivalent:

(1) $\sum_{i=1}^{l} x_{1i} \geq \sum_{i=1}^{l} x_{0i}\ \forall l = 1, \ldots, T$.

(2) There exists a doubly stochastic matrix $W \in \Xi$ such that $x_1 \geq W x_0$.

Observe that there exist permutation matrices $P_0, P_1 \in \Pi$ such that $y_h = P_h x_h$, $h = 0, 1$. By Theorem 1, $y_1 \widehat{\mathrm{D}}_1 y_0$ is equivalent to Condition (1) above. Therefore, we have the inequality $y_1 = P_1 x_1 \geq W(P_0 x_0) = \widetilde{W} y_0$; $W P_0 = \widetilde{W} \in \Xi$. To complete the proof, we note that a portfolio return profile cannot dominate "itself," or its permutation, so we must have $y_1 \neq P y_0\ \forall P \in \Pi$. □

Proof of Theorem 4. *The "only if" part*: If $\theta_1(y_0) > 0$, then there exists a marketed portfolio return profile $y^* = Y\lambda^{0*} \in \Psi$ that satisfies $y^* \geq P y_0$ due to the first constraint of (1). Moreover, $\theta_1(y_0) > 0$ implies $y^* \neq P y_0$. Thus, we see that y_0 cannot be FSD efficient because there exists $y^* \in \Delta_1(y_0)$.

The "if" part: Suppose $\theta_1(y_0) = 0$ and assume there exists $y^* \in \Psi$ such that $y^* \in \Delta_1(y_0)$. The last condition directly implies $y^* \geq P y_0$, with the strict inequality in at least one dimension. However, this implies the optimal solution to (1) must be strictly positive, which contradicts the initial assumption $\theta_1(y_0) = 0$. □

Proof of Theorem 6. The necessary test follows directly from Theorem 5, so we focus solely on the sufficient test, adopting the constructive style for the sake of intuition. $\theta_2^n(y_0) = 0$ implies that any dominating portfolio return profile with higher mean return does not exist. As a consequence, the dominating portfolio return profile, if existent, must be of form $W y_0$, $W \in \Xi - \Pi$. Problem (3) tries to construct such a portfolio return profile, minimizing departures of elements W_{ij} from the value $\frac{1}{2}$.

Suppose first that there are no ties in the evaluated profile. If $W \in \Pi$, then all elements of nonnegative $T \times T$ matrices s^+, s^- satisfy $s_{ij}^+ = \frac{1}{2}$, $s_{ij}^- = 0$ or $s_{ij}^+ = 0$, $s_{ij}^- = \frac{1}{2}$, implying $\theta_2^s(y_0) = T^2/2$. Next, consider the possibility of ties. We can mix up tied elements of y_0 without changing the portfolio return profile. Thus, only those elements of W that correspond to a pair of unequal elements of y_0 need to be binary valued. Suppose a k-way tie occurs in y_0. The sum of s^+, s^- variables is minimized by setting $W_{ij} = 1/k$ for all pairs i, j corresponding to the tied elements. Thus, the sum of s^+, s^- variables for these elements becomes $k^2(1/2 - 1/k) = (1/2)k^2 - k$. The analogous sum for k nonequal elements is $(1/2)k^2$, so a k-way tie decreases the sum of s^+, s^- variables by k. Thus, the theoretical maximum is lower by $\sum_{k=1}^{T} k d_{0k}$ when the ties are accounted for.

If any portfolio return profile $W y_0$, $W \in \Xi - \Pi$ is marketed, then it must be obtained as a linear combination of existing assets, i.e.,

$$\sum_{n=1}^{N+1} Y_{tn}^0 \lambda_i^0 = \sum_{j=1}^{T} W_{tj} y_{0j} \quad \forall t \in \tau.$$

Now, consider the subproblem of optimizing matrices s^+, s^- for a given W. Suppose, $W_{ij} \in [\frac{1}{2}, 1)$. Clearly, the optimal s^{+*}, s^{-*} satisfy $0 \leq s_{ij}^{+*} < \frac{1}{2}$, $s_{ij}^{-*} = 0$. Similarly, if $W_{ij} \in (0, \frac{1}{2}]$, we must have that $s_{ij}^{+*} = 0$, $0 \leq s_{ij}^{-*} < \frac{1}{2}$. Therefore, if any $W y_0$, $W \in \Xi - \Pi$ is marketed, then

$$\sum_{j=1}^{T} \sum_{i=1}^{T} (s_{ij}^+ + s_{ij}^-) < \frac{T^2}{2} - \sum_{k=1}^{T} k d_{0k}.$$

In conclusion, if we set weights λ^0, W, s^+, s^- to minimize the sum $\sum_{j=1}^{T} \sum_{i=1}^{T} (s_{ij}^+ + s_{ij}^-)$ and the optimal solution equals $T^2/2 - \sum_{k=1}^{T} k d_{0k}$, then this implies the market set Ψ does not contain any mean-preserving antispread $W y_0$, $W \in \Xi - \Pi$. As a consequence, y_0 must be SSD efficient. □

Proof of Theorem 7. Consider necessity first. Recall that SSD efficiency is equivalent to condition $\Delta_2(y_0) \cap \Psi = \varnothing$. Thus, if portfolio return profile y_0 is efficient, there must exist a separating hyperplane $H = \{y \in \mathbb{R}^T \mid w'y = w'y_0;\ w \in \mathbb{R}_+^T\}$ that separates the market set Ψ from the dominating $\Delta_2(y)$. Note that in this special case the polyhedron Ψ is simply the convex hull of the underlying assets Y_n. To separate y_0 from the relative interior of Ψ, weights w must satisfy

(i) $w'y_0 \geq w'Y_n\ \forall n \in \nu$.

Also note that set $\Delta_2(y_0)$ is essentially the convex monotonic hull of permutations of vector y_0. To separate y_0 from the relative interior of $\Delta_2(y_0)$, weights w must satisfy the following two conditions:

(ii) By monotonicity of $\Delta_2(y_0)$, weights w must be *nonnegative*, i.e., $w \in \mathbb{R}_+^T$.

[22] The SSD case can also be proved using the result of Karlin and Novikoff (1963) in majorization theory; see Aboudi and Thon (1994, pp. 509–510) for discussion.

[23] See Schmeidler (1979) for discussion of the economic significance of this theorem.

(iii) By convexity of $\Delta_2(y_0)$, permutations of y_0 must lie above the separating hyperplane; i.e., $y_0 w \le (y_0 P)w$ $\forall P \in \Pi: y_0 P \ne y_0$.

Condition (iii) implies (iiib) $y_{0j} > y_{0k} \Rightarrow w_j \le w_k\ \forall j, k \in T$.

Combining these observations, SSD efficiency of y_0 implies there exist weights w that satisfy Conditions (i), (ii), and (iiib). The last two constraints of Problem (4) guarantee that Conditions (ii) and (iiib) hold. Thus, Condition (i) determines efficiency. The weights are scaled such that $\sum_{t=1}^{T} Y_{tn} w_t \le 1\ \forall n \in \nu$ (the first constraint of (4)). Therefore, Condition (i) holds if and only if the optimal solution satisfies $\hat{\theta}_2(y_0) \ge 1$.

The fact that $\hat{\theta}_2(y_0) \ge 1$ is not a sufficient condition can be verified by means of a numerical example; see, e.g., Example 2 in Appendix 2. □

Appendix 2. Three Numerical Examples

Example 1. Why can the EDF of a diversified portfolio not be recovered from the marginal EDFs of the underlying assets?

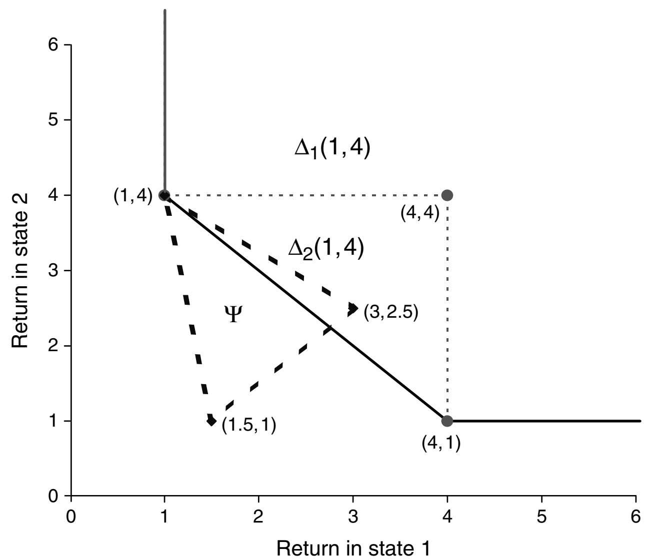
Consider assets A and B with EDFs

$$H_A(r) = \begin{cases} 0 & \text{for } r < 1 \\ 1/2 & \text{for } 1 \le r < 4 \\ 1 & \text{for } r \ge 4. \end{cases} \qquad H_B(r) = \begin{cases} 0 & \text{for } r < 0 \\ 1/2 & \text{for } 0 \le r < 3 \\ 1 & \text{for } r \ge 3. \end{cases}$$

Clearly, A dominates B by both FSD and SSD. However, what is the EDF of the portfolio $(\lambda_A, \lambda_B) = (0.5, 0.5)$? Based on these two EDFs, we cannot tell. Suppose the EDFs were constructed from observations of only two states of nature. There are two possibilities for both assets: The states underlying the EDFs could be $y_A = (1, 4)$ or $y_A = (4, 1)$ for Asset A, and $y_B = (0, 3)$ or $y_B = (3, 0)$ for Asset B, respectively. If we have $y_A = (1, 4)$, $y_B = (0, 3)$ or alternatively $y_A = (4, 1)$, $y_B = (3, 0)$, the EDF of the diversified portfolio is

$$H_\lambda(r) = \begin{cases} 0 & \text{for } r < 0.5 \\ 1/2 & \text{for } 0.5 \le r < 3.5 \\ 1 & \text{for } r \ge 3.5. \end{cases}$$

Figure A1 Illustration of the SD Tests



However, if the state-specific portfolio return profiles are $y_A = (1, 4)$, $y_B = (3, 0)$ or, alternatively, $y_A = (4, 1)$, $y_B = (0, 3)$, then diversification completely eliminates risk and the portfolio distribution becomes

$$\widetilde{H}_\lambda(r) = \begin{cases} 0 & \text{for } r < 2 \\ 1 & \text{for } r \ge 2. \end{cases}$$

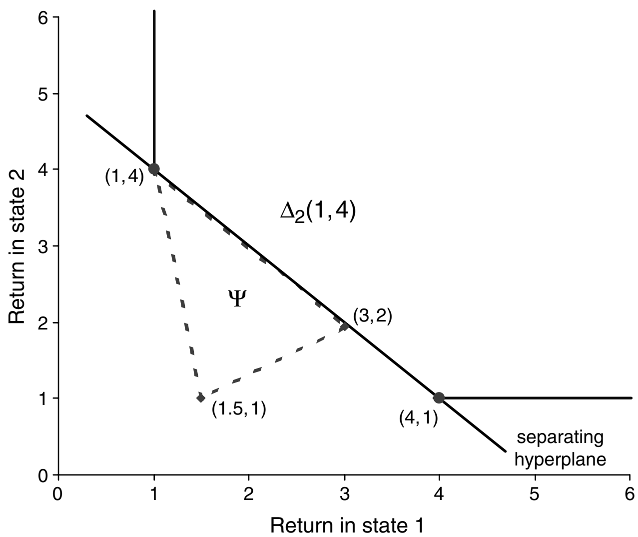
This simple example demonstrates that it is impossible to tell whether H_λ or $\widetilde{H}_\lambda$ is correct without knowing which states of nature the returns are associated with. By construction, the marginal EDFs cannot keep track of diversification possibilities. Certainly, the situation does not become any simpler if additional states and/or assets are introduced.

Example 2: Illustration of SD Tests When the Evaluated Portfolio Is FSD Efficient but SSD Inefficient. Consider a case of two states of nature, and three marketed assets (numbered as 0, 1, and 2) with portfolio return profiles $y_0 = (1, 4)$, $y_1 = (3, 2.5)$, and $y_2 = (1.5, 1)$. Let the set of portfolio weights include all nonnegative weights that sum up

Table A1 SD Statistics, Portfolio Return Profiles $y_0 = (1, 4)$, $y_1 = (3, 2.5)$, **and** $y_2 = (1.5, 1)$

$\theta_1(1,4) = \max_{\lambda^0, P} (5\lambda_0^0 + 5.5\lambda_1^0 + 2.5\lambda_2^0 - 5)/2$	$\theta_2^n(1,4) = \max_{\lambda^0, W} (5\lambda_0^0 + 5.5\lambda_1^0 + 2.5\lambda_2^0 - 5)/2$	$\hat{\theta}_2^n(y_0) = \max_{w} (w_1 + 4w_2)$
s.t. $1\lambda_0^0 + 3\lambda_1^0 + 1.5\lambda_2^0 \ge P_{11} + 4P_{12}$	s.t. $1\lambda_0^0 + 3\lambda_1^0 + 1.5\lambda_2^0 \ge W_{11} + 4W_{12}$	s.t. $3w_1 + 2.5w_2 \le 1$
$4\lambda_0^0 + 2.5\lambda_1^0 + \lambda_2^0 \ge P_{21} + 4P_{22}$	$4\lambda_0^0 + 2.5\lambda_1^0 + \lambda_2^0 \ge W_{21} + 4W_{22}$	$1.5w_1 + w_2 \le 1$
$P_{11} + P_{12} = P_{21} + P_{22} = 1$	$W_{11} + W_{12} = W_{21} + W_{22} = 1$	$w_1 \ge w_2 \ge 0$
$P_{11} + P_{21} = P_{12} + P_{22} = 1$	$W_{11} + W_{21} = W_{12} + W_{22} = 1$	
$P_{ij} \in \{0, 1\}$	$W_{11}, W_{12}, W_{21}, W_{22} \ge 0$	
$\lambda_0^0 + \lambda_1^0 + \lambda_2^0 = 1$	$\lambda_0^0 + \lambda_1^0 + \lambda_2^0 = 1$	
$\lambda_0^0, \lambda_1^0, \lambda_2^0 \ge 0$	$\lambda_0^0, \lambda_1^0, \lambda_2^0 \ge 0$	
Optimal solution:	Optimal solution:	Optimal solution:
$\theta_1(1,4) = 0$	$\theta_2^n(1,4) = 0.25$	$\hat{\theta}_2^n(1,4) = \frac{10}{11}$
$\lambda_0^{0*} = 1,\quad \lambda_1^{0*} = \lambda_2^{0*} = 0$	$\lambda_0^{0*} = 0,\quad \lambda_1^{0*} = 1,\quad \lambda_2^{0*} = 0$	$w_1^* = w_2^* = \frac{2}{11}$
$P_{11}^* = P_{22}^* = 1,\quad P_{12}^* = P_{21}^* = 0$	$W_{ij}^* = \frac{1}{2}$ (not unique)	

Figure A2 SSD Test—Necessary vs. Sufficient Conditions



to unity, so that the market set Ψ is simply the convex hull of the three return vectors. What do our tests reveal about efficiency of investing all wealth in Asset 0 in this example?

Figure A1 illustrates the example graphically: The dominating sets are identical to those introduced in Figures 1 and 2 in §3; the market set is the triangle spanned by the three return vectors. From this figure we observe that the market set overlaps with the SSD dominating set but not with the FSD dominating set. This implies the evaluated portfolio is FSD efficient but not SSD inefficient. This is confirmed by numerical calculations.

Table A1 rephrases Problems (1), (2), and (4) in this specific numerical example and reports the optimal solutions. Problem (3) is disregarded as the necessary tests already reject efficiency. The example illustrates the key difference between Problems (1) and (2): in the leftmost column of Table A1, variables P of the permutation matrix are binary integers, while in the midcolumn, variables W of the doubly stochastic matrix are real numbers in the interval $[0, 1]$. The difference is notable in this example. Observe that the evaluated portfolio return profile offers the highest possible return of 4 in State 2, so reordering the states does not help. In the SSD case we can redistribute a part of the high return of Asset 0 in State 2 to State 1 through weights W. Doing so, we see that Asset 1 offers higher returns in both states, and hence dominates Asset 0.

Also, test statistic $\hat{\theta}_2(y_0)$ applies to this example. In Problem (4) we try to find a tangent line for the SSD dominating set at point $(1, 4)$, which separates the market set Ψ from the SSD dominating set. Such a tangent line does not exist because the two sets overlap. The nonexistence reveals itself in the value of the test statistic being less than one.

Example 3: Illustration of SSD Inefficiency That Is Not Revealed by the Necessary SSD Conditions. Modify the previous example so that the rate of return for y_1 decreases by 0.5 in State 2; that is, $y_1 = (3, 2)$. Figure A2 illustrates this new situation. Of course, FSD efficiency does not change when the market set is contracted, so we ignore the FSD case. Interestingly, Assets 0 and 1 now yield an equal mean return. From Figure A2 we see that there exists a separating hyperplane that weakly separates the market set and the dominating set. As a consequence, the necessary SSD tests classify portfolio return profile y_0 as efficient. This is confirmed by the algebraic solution of Test Problems (2) and (4) reported in the two leftmost columns of Table A2.

Although portfolio return profile $y_0 = (1, 4)$ satisfies both alternative necessary SSD conditions, it is not SSD efficient.

Table A2 SSD Statistics, Portfolio Return Profiles $y_0 = (1, 4)$, $y_1 = (3, 2)$, and $y_2 = (1.5, 1)$

$\theta_2^n(1,4) = \max_{\lambda^0, W}(5\lambda_0^0 + 5.5\lambda_1^0 + 2\lambda_2^0 - 5)/2$	$\hat{\theta}_2^n(y_0) = \max_{w}(w_1 + 4w_2)/2$	$\theta_2^s(y_0) = \min_{W,\lambda^0,s^+,s^-}(s_{11}^+ + s_{12}^+ + s_{21}^+ + s_{22}^+ + s_{11}^- + s_{12}^- + s_{21}^- + s_{22}^-)$
s.t. $1\lambda_0^0 + 3\lambda_1^0 + 1.5\lambda_2^0 \ge W_{11} + 4W_{12}$	s.t. $3w_1 + 2w_2 \le 1$	s.t. $\lambda_0^0 + 3\lambda_1^0 + 1.5\lambda_2^0 = W_{11} + 4W_{12}$
$4\lambda_0^0 + 2\lambda_1^0 + \lambda_2^0 \ge W_{21} + 4W_{22}$	$1.5w_1 + w_2 \le 1$	$4\lambda_0^0 + 2\lambda_1^0 + 1\lambda_2^0 = W_{21} + 4W_{22}$
$W_{11} + W_{12} = W_{21} + W_{22} = 1$	$w_1 \ge w_2 \ge 0$	$W_{11} = \frac{1}{2} + s_{11}^+ - s_{11}^-$
$W_{11} + W_{21} = W_{12} + W_{22} = 1$		$W_{12} = \frac{1}{2} + s_{12}^+ - s_{12}^-$
$W_{11}, W_{12}, W_{21}, W_{22} \ge 0$		$W_{21} = \frac{1}{2} + s_{21}^+ - s_{21}^-$
$\lambda_0^0 + \lambda_1^0 + \lambda_2^0 = 1$		$W_{22} = \frac{1}{2} + s_{22}^+ - s_{22}^-$
$\lambda_0^0, \lambda_1^0, \lambda_2^0 \ge 0$		$W_{11} + W_{12} = W_{21} + W_{22} = 1$
		$W_{11} + W_{21} = W_{12} + W_{22} = 1$
		$W_{11}, W_{12}, W_{21}, W_{22} \ge 0$
		$\lambda_0^0 + \lambda_1^0 + \lambda_2^0 = 1$
		$\lambda_0^0, \lambda_1^0, \lambda_2^0 \ge 0$
Optimal solution:	Optimal solution:	Optimal solution:
$\theta_2^n(1,4) = 0$	$\hat{\theta}_2^n(1,4) = 1$	$\theta_2^s(1,4) = 0$
$\lambda_0^{0*} = 1$, $\lambda_1^{0*} = 0$, $\lambda_2^{0*} = 0$, $W_{ij}^* = 1$ (not unique)	$w_1^* = w_2^* = \frac{1}{5}$	$\lambda_0^{0*} = \frac{1}{4}$, $\lambda_1^{0*} = \frac{3}{4}$, $\lambda_2^{0*} = 0$
		$W_{ij}^* = \frac{1}{2}$
		$s_{ij}^{+*} = s_{ij}^{-*} = 0$

We can obtain the same mean return with lower risk by diversifying between Assets 0 and 1. We see this in Figure 4 in that the boundary of the dominating set includes return vectors that are also contained in the market set. In other words, the tangent line weakly separates the two sets, but a strong separation is impossible. To identify this type of inefficiency, we need to examine the sufficient SSD test statistic θ_2^s. The rightmost column of Table A2 rephrases Problem (3) in this example. Minimizing the sum of surplus and slack variables s^+, s^-, we note that in this example risk can be fully eliminated through diversification: By setting $\lambda_0^0 = \frac{1}{4}$, $\lambda_1^0 = \frac{3}{4}$, $\lambda_2^0 = 0$ we obtain a risk-free return of 2.5.

References

Aboudi, R., D. Thon. 1994. Efficient algorithms for stochastic dominance tests based on financial market data. *Management Sci.* **40**(4) 508–515.

Afriat, S. 1967. The construction of a utility function from expenditure data. *Internat. Econom. Rev.* **8** 67–77.

Bawa, V. S. 1982. Stochastic dominance: A research bibliography. *Management Sci.* **28** 698–712.

Bawa, V. S., E. Lindenberg, L. Rafsky. 1979. An algorithm to determine stochastic dominance admissible sets. *Management Sci.* **25**(7) 609–622.

Bixby, R. E. 2002. Solving real-world linear programs: A decade and more of progress. *Oper. Res.* **50**(1) 3–15.

Charnes, A., W. W. Cooper, E. Rhodes. 1978. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **2**(6) 429–444.

Charnes, A., W. W. Cooper, B. Golany, L. Seiford, J. Stutz. 1985. Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *J. Econometrics* **30**(1) 91–107.

Dybvig, P. H., S. A. Ross. 1982. Portfolio efficient sets. *Econometrica* **50**(6) 1525–1546.

Farrell, M. J. 1957. The measurement of productive efficiency. *J. Roy. Statist. Soc. Ser. A* **120**(3) 253–290.

Fishburn, P. C. 1964. *Decision and Value Theory*. John Wiley & Sons, New York.

Fourer, R. 2001. Linear programming: 2001 software survey. *OR/MS Today* **28**(4) 58–59, 68.

Frankfurter, G. M., H. E. Phillips. 1975. Efficient algorithms for conducting stochastic dominance tests on large numbers of portfolios, a comment. *J. Financial Quant. Anal.* **10** 177–179.

French, K. R. 2003. *Kenneth R. French—Data Library*. Accessed September 2, 2003. <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html>.

Gavish, B. 1977. A relaxation algorithm for building undominated portfolios. *J. Banking Finance* **1** 143–150.

Hadar, J., W. R. Russel. 1969. Rules for ordering uncertain prospects. *Amer. Econom. Rev.* **59** 25–34.

Hardy, G. H., J. E. Littlewood, G. Polya. 1934. *Inequalities*. Cambridge University Press, Cambridge, U.K.

Karlin, S., A. Novikoff. 1963. Generalized convex inequalities. *Pacific J. Math.* 1251–1279.

Koopmans, T. C., M. Beckmann. 1957. Assignment problems and the location of economic activities. *Econometrica* **25** 53–76.

Kroll, Y., H. Levy. 1980. Sampling errors and portfolio efficient analysis. *J. Financial Quant. Anal.* **15** 655–688.

Kuosmanen, T. 2001. Stochastic dominance efficiency tests under diversification. Working paper W-283, Helsinki School of Economics and Business Administration, Helsinki, Finland.

Levy, H. 1992. Stochastic dominance and expected utility: Survey and analysis. *Management Sci.* **38**(4) 555–593.

Levy, H. 1998. *Stochastic Dominance: Investment Decision Making Under Uncertainty*. Kluwer Academic Publishers, Norwell, MA.

Levy, H., G. Hanoch. 1970. Relative effectiveness of efficiency criteria for portfolio selection. *J. Financial Quant. Anal.* **5** 63–76.

Levy, H., Z. Wiener. 1998. Stochastic dominance and prospect dominance with subjective weighting functions. *J. Risk Uncertainty* **16**(2) 147–163.

Machina, M. J. 1982. Expected utility analysis without the independence axiom. *Econometrica* **50**(2) 277–324.

Markowitz, H. M. 1952. Portfolio selection. *J. Finance* **12** 77–91.

Markowitz, H. M. 1959. *Portfolio Selection*. John Wiley & Sons, New York.

Markowitz, H. M. 1977. An algorithm for finding undominated portfolios. H. Levy, M. Sarnat, eds. *Financial Decision Making Under Uncertainty*. Academic Press, New York, 3–10.

McFadden, D. 1989. Testing for stochastic dominance. T. B. Fomby, T. K. Seo, eds. *Studies in the Economics of Uncertainty. In Honor of Josef Hadar*. Springer Verlag, Berlin and New York, 113–134.

Mittelmann, H. 2003. Benchmarks for optimization software. Accessed September 1, 2003. http://plato.la.asu.edu/bench.html.

Nelson, R. D., R. Pope. 1991. Bootstrapping insights into empirical applications of stochastic dominance. *Management Sci.* **37** 1182–1194.

Neumann, J. von. 1953. A certain zero-sum two-person game equivalent to the optimal assignment problem. H. W. Kuhn, A. W. Tucker, eds. *Contributions to the Theory of Games*, Vol. II. Princeton University Press, Princeton, NJ, 5–12.

Neumann, J. von, O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Peleg, B., M. E. Yaari. 1975. A price characterization of efficient random variables. *Econometrica* **43**(2) 283–292.

Performance World. 2003. GAMS Development Corp. Accessed August 31, 2003. <http://www.gamsworld.org/performance/plib/size.htm>.

Porter, R. B., R. C. Pfaffenberger. 1975. Efficient algorithms for conducting stochastic dominance tests on large numbers of portfolios: Reply. *J. Financial Quant. Anal.* **10** 181–185.

Porter, R. B., J. R. Wart, D. L. Ferguson. 1973. Efficient algorithms for conducting stochastic dominance tests on large numbers of portfolios. *J. Financial Quant. Anal.* **8** 71–82.

Post, G. T. 2003a. Empirical tests for stochastic dominance efficiency. *J. Finance* **58**(5) 1905–1931.

Post, G. T. 2003b. Asset pricing and omitted moments: A stochastic dominance analysis of market efficiency. Erasmus Research Institute of Management (ERIM), Rotterdam, The Netherlands.

Quirk, J., R. Saposnik. 1962. Admissibility and measurable utility functions. *Rev. Econom. Stud.* **29** 140–146.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press, Princeton, NJ.

Rothschild, M., J. E. Stiglitz. 1970. Increasing risk I: A definition. *J. Econom. Theory* **2** 225–243.

Schmeidler, D. 1979. A bibliographical note on a theorem of Hardy, Littlewood, and Polya. *J. Econom. Theory* **20** 125–128.

Schmid, F., M. Trede. 1996. Testing for first-order stochastic dominance: A new distribution free test. *Statistician* **45**(3) 371–380.

Starmer, C. 2000. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *J. Econom. Lit.* **38**(2) 332–382.

Stiglitz, J. 1981. The allocation role of the stock market: Pareto optimality and competition. *J. Finance* **36**(2) 235–251.

Varian, H. 1983. Nonparametric tests of models of investor behavior. *J. Financial Quant. Anal.* **18** 269–278.

Ziemba, W. T. 1978. Portfolio applications: Computational aspects. G. A. Whitmore, M. C. Findlay, eds. *Stochastic Dominance: An Approach to Decision-Making Under Risk*. D.C. Heath, Lexington, MA, 199–260.